Expectations, Learning and Business Cycle Fluctuations*

Stefano Eusepi† Federal Reserve Bank of New York

Bruce Preston‡ Columbia University and NBER

July 11, 2008

Abstract

This paper develops a theory of expectations-driven business cycles based on learning. Agents have incomplete knowledge about how market prices are determined and shifts in expectations of future prices affect dynamics. In a real business cycle model, the theoretical framework amplifies and propagates technology shocks. Improved correspondence with data arises from dynamics in beliefs being themselves persistent and because they generate strong intertemporal substitution effects in consumption and leisure. Output volatility is comparable with a rational expectations analysis with a standard deviation of technology shock that is 20 percent smaller, and has substantially more volatility in investment and hours. Persistence in these series is captured, unlike in standard models. Inherited from real business cycle theory, the benchmark model suffers a comovement problem between consumption, hours, output and investment. An augmented model that is consistent with expectations-driven business cycles, in the sense of Beaudry and Portier (2006), resolves these counterfactual predictions.

JEL Classifications: E32, D83, D84

Keywords: Expectations-Driven Business Cycles, Learning Dynamics, Pigou Cycles

---

*The authors thank seminar participants at the Australasian Workshop in Macroeconomic Dynamics 2008, Australian National University, Columbia University, ESSIM 2008, Federal Reserve Bank of Atlanta, Northwestern University, Society of Economic Dynamics 2007, Society of Computational Economics 2008, Universitat Automoma and CREI-Universitat Pompeu Fabra. We thank Albert Marcet, Kristoffer Nimark, Jaume Ventura and particularly Florin Bilbiie for useful conversations and comments. The views expressed in the paper are those of the authors and are not necessarily reflective of views at the Federal Reserve Bank of New York or the Federal Reserve System. The usual caveat applies.

†Federal Reserve Bank of New York. E-mail: stefano.eusepi@ny.frb.org.

‡Department of Economics, Columbia University, 420 West 118th St. New York NY 10027. E-mail: bp2121@columbia.edu

1 Introduction

Recently there has been a renewed interest in shifting expectations as a source of business cycle fluctuations. A range of models have been explored that rely variously on multiple equilibria, exogenous news about future productivity and imperfect information — see, for example, Benhabib and Farmer (1994), Schmitt-Grohe (2000), Beaudry and Portier (2006), Jaimovich and Rebelo (2008) and Lorenzoni (2008). These frameworks seek not only to explain business cycles fluctuations with changes in expectations but also to resolve comovement problems that arise in real business cycle theory.

This paper explores an alternative theory based on learning dynamics. In the context of an otherwise standard stochastic growth model, we consider an environment in which households and firms have an incomplete model of the macroeconomy, knowing only their own objectives, constraints and beliefs. Agents are optimizing, have a completely specified belief system but do not know the equilibrium mapping between primitive disturbances, the aggregate state of the economy and market clearing prices. By extrapolating from historical patterns in observed data they approximate this mapping to forecast exogenous variables relevant to their decision problems, such as the rental rate of capital and the real wage. This belief structure has the property that beliefs affect the true data generating process of the economy which in turn affects belief formation. The economy is therefore self-referential: shifts in beliefs about future returns to labor and capital affect current market clearing prices which in turn can reinforce beliefs. In this environment, current prices can become less informative about future economic conditions generating fluctuations in real activity.

This kind of mechanism driving business cycle fluctuations is found in early writings on macroeconomic dynamics. For example, Pigou (1927), on page 122, writes:

> "[...] a rise in prices, however brought about, by creating some actual and some counterfeit prosperity for business man, is liable to promote an error of optimism, and a fall in prices an error of pessimism, and this mutual stimulation of errors and price movements may continue in a vicious spiral until it is checked by some interference from outside."

Hence, shifts in expectations, whether in part due to changes in fundamentals or in part due to error are a source of business cycle fluctuation that may be self-fulfilling. Our model is very much in the spirit of this quote.[1] Learning breaks the tight link between fundamentals and, through expectations formation, equilibrium prices and allocations, giving rise to additional volatility relative to a rational expectations analysis of the model. Furthermore, shifts in expectations occur not because of *exogenous* "news shocks" about the future state of the economy — as proposed in the recent literature on expectations-driven business cycles — or "sunspots" but because of the agents' learning process, which depends on current available data. Revisions in agents' beliefs about future returns to their capital holdings generate *endogenous* aggregate demand shocks which amplify the effects of exogenous changes to productivity. Moreover, learning might be thought to improve the internal propagation mechanisms of the model, since beliefs are a function of historical data, introducing an additional state variable.

Calibrating the model to match properties of post-war U.S. quarterly data, the central results of the paper are as follows. First, learning amplifies technology shocks. Relative to a rational expectations analysis of the model, a 20 percent smaller standard deviation of technology shocks is required to match the standard deviation of HP-filtered output data. Moreover, the volatility of investment and hours relative to output is 40 and 25 percent greater than under rational expectations. Second, the persistence properties of our model bear much closer resemblance to observed data. The first order autocorrelation properties of output, hours and investment growth are well matched despite shocks being identically and independently distributed over time — hump-shaped impulse responses are observed. These features of the data are typically problematic for real business cycle theory as documented by Cogley and Nason (1993) and Rotemberg and Woodford (1996). In general, the learning model provides a superior characterization of second order moments of observed data than does the model under rational expectations.

The improvement in fit can be traced to shifting beliefs acting as endogenous demand shocks, interpretable as news and having similar effects to government expenditure or invest-

[1] The model will not have the property of "vicious spirals".

ment specific technology shocks. The latter will be given specific emphasis when interpreting results. The only source of exogenous variation are technology shocks, which have two effects. First, as in standard real business cycle theory, a temporary technology shock shifts the production frontier with well understood implications. Second, in subsequent periods, households revise their beliefs in response to changed market opportunities. In particular, *relative* to rational expectations, households are more optimistic about the future path of returns to capital and more pessimistic about future returns to labor. The former leads to substitution of current for future consumption and a high marginal utility of income, an effect reinforced by lower projected wages. Combined, these expectations effects induce a larger fall in consumption and consequently a larger shift in labor supply and investment in the period after the shock. This amplification of substitution and income effects in response to a technology shock relative to rational expectations explains the increased volatility in these variables. The delayed adjustment in beliefs explains the persistence. Furthermore, these observations highlight our connection to Pigou (1927): erroneous optimism or pessimism about future returns to capital and wages are in part validated by the data. Moreover, shifts in expectations about future returns to labor and capital are for a given technological frontier and endogenous to the technology shock. In this sense they have similarity to demand shocks in so far as hours and consumption negatively comove in response to a revision in expectations.

As there is additional endogenous variation in the marginal utility of income for a given production frontier, the model suffers a comovement problem. Hours and consumption display negative correlation. Furthermore, while hours growth exhibits positive autocorrelation, consumption growth has negative autocorrelation. The third result of the paper shows that in a model of the kind proposed by Beaudry and Portier (2006) this comovement problem can be resolved. That paper explores primitive assumptions on technology and preferences that are consistent with so called expectations-driven business cycles — in response to an expectational shock, output, hours, investment and consumption display positive comovement. We propose a new pairing of assumptions that delivers this property. They are a small degree of increasing returns combined with non-separability in utility between consumption and hours.[2]

[2]The results do not rely on having an upward sloping demand for labor schedule or indeterminacy. Moreover, indentical results obtain if increasing returns are replaced by a model of endogenous entry and exit as

The first assumption provides an endogenous shift in the production frontier from external economies so that consumption does not crowd out investment, while the latter assists in capturing the comovement between hours and consumption. Under these assumptions, which introduce no additional states variables, our model provides an even better characterization of observed data. Moreover, the interaction of learning with these model features provides additional amplification and propagation relative to a rational expectations model with these characteristics and our baseline model under learning. The modified model implies some 30 percent greater volatility in output for a given technology shock.

The results are robust to a range of alternative assumptions. In particular, our analysis could be criticized on the ground that it is well understood that real business cycle theory fails to account for various properties of observed data without augmenting the model with additional frictions such as variable capital utilization and investment adjustment costs. We show that our benchmark model performs well when compared to rational expectations models with these features.

Finally, we compare our analysis with earlier explorations of learning as a source of amplification and propagation. In particular, we revisit the analysis of Williams (2003) which also looked at learning in a standard real business cycle model. That paper concludes that learning based on extrapolating historical patterns in observed data, as considered here, is unlikely to help improve the performance of real business cycle models. Reproducing that analysis in the context of our model shows that this is indeed the case. The difference in conclusions stems from the failure in Williams (2003) to model optimal decisions conditional on maintained beliefs as done in Marcet and Sargent (1989) and Preston (2005). Williams (2003) also considers a related but distinct class of learning models in which agents face uncertainty about structural parameters rather than the statistical relation between prices and state variables that obtains in equilibrium. This conceptually distinct exercise is argued to be more promising in generating economic fluctuations. The results of this paper indicate this is not necessarily true.

This paper belongs to a long literature reconciling the predictions of real business cycle

considered by Portier (1995) and Jaimovich (2007).

theory with observed data — see, inter alia, Hansen (1985), Rogerson (1988), Christiano and Eichenbaum (1992), Benhabib and Farmer (1994), Andolfatto (1996), Burnside and Eichenbaum (1996), Carlstrom and Fuerst (1997) and Schmitt-Grohe (2000). These papers introduce various frictions including indeterminacy of rational expectations equilibrium to benchmark theory to improve the amplification and propagation of technology shocks. Our paper furthers this line of inquiry by considering learning dynamics as an alternative friction.

The introduction of imperfect information and learning in the real business cycle framework dates back to Kydland and Prescott (1982). In their model, the stochastic process for technology is composed of two unobserved shocks with different persistence. Agents face a signal extraction problem when predicting future productivity. More recently, Edge, Laubach, and Williams (2007) show in a similar model that learning has important effects in the response of the economy to a persistent shift in productivity growth. In this class of models learning is not an endogenous source of propagation because changes in endogenous variables do not affect the agents' learning process. On the contrary, gradual recognition of the productivity changes generates a gradual response to the shock — a property determined by the specified signal-to-noise ratio in the exogenous process.

In addition to the above mentioned learning literature, our paper relates to other recent contributions by Milani (2006), Carceles-Poveda and Giannitsarou (2007) and Huang, Liu, and Zha (2008). Milani (2006) considers whether learning dynamics improve the fit of a simple estimated New Keynesian model. Carceles-Poveda and Giannitsarou (2007) analyze the consequences of learning dynamics for asset pricing in a real business cycle model. Huang, Liu, and Zha (2008) explore the amplification and propagation of technology shocks. Like Williams (2003), these papers consider models in which only one-period-ahead forecasts matter for household and firm behavior — decisions are not optimal given the maintained beliefs. Moreover, and in further contrast to these papers, the analysis here studies model properties at the steady state distribution of beliefs so that initial conditions are not a source of amplification and propagation.

Finally, this paper connects with recent work on imperfect information and business cycle dynamics. Nieuwerburgh and Veldkamp (2006) consider a model where agents have a noisy signal about aggregate and idiosyncratic productivity and explore implications for generating long expansions and short contractions in economic activity. Lorenzoni (2008) develops a theory of demand shocks in a model with heterogeneous productivity shocks and diverse private information. Agents' signal extraction problem lead to expectational errors relative to a full information model, which generate dynamics that are qualitatively like demand shocks when the only primitive disturbances are technology shocks.

The paper proceeds as follows. Section 2 lays out a simple real business cycle model. Section 3 discusses the assumed belief structure. Section 4 details the data and calibration. Section 5 presents the core results under our benchmark assumptions. Section 6 gives results for a model consistent with expectations driven business cycles in the sense of Beaudry and Portier (2006). Section 7 provides some robustness exercises. Section 8 concludes.

2 A Simple Model

The following section details a stochastic growth model similar in spirit to Kydland and Prescott (1982), Prescott (1986) and King, Plosser, and Rebelo (1988). A continuum of households faces a canonical consumption allocation problem and decides how much to consume of the economy's single available good, how much to invest, and how much labor to supply to firms in the production of the available good. A continuum of competitive firms produces goods using labor and capital as inputs. The major difference to this earlier literature is the incorporation of near-rational beliefs, delivering an anticipated utility model of the kind discussed by Kreps (1998) and Sargent (1999). The analysis follows Marcet and Sargent (1989) and Preston (2005), solving for optimal decisions conditional on current beliefs.

Various mechanisms of persistence, such as investment adjustment costs and variable capital utilization, are abstracted from. This facilitates identification of key mechanisms operating in our model that would be present in more richly specified environments and provides pellucid results on the ability of near-rational expectations to replicate salient features of the data.[3] The sequel demonstrates that frictions of this kind tend to amplify further the effects

[3]This should not be taken to suggest that the benchmark real business cycle model is the best competing model. Our approach seeks to elucidate a new theoretical mechanism for expectations-driven business cycles, while at the same time showing it is consistent with empirical regularities.

identified in our benchmark analysis.

2.1 Microfoundations

Households. Households maximize their intertemporal utility derived from consumption and leisure

$$\hat{E}_t^j \sum_{T=t}^{\infty} \beta^{T-t} \left[\ln C_T^j - \nu \left(H_T^j \right) \right] \tag{1}$$

subject to the flow budget constraint

$$C_t^j + K_{t+1}^j = R_t^K K_t^j + W_t H_t^j + (1-\delta) K_t^j \tag{2}$$

where C_t^j denotes household $j's$ consumption, K_t^j the holdings of the aggregate capital stock available at the beginning of period t, with $K_0^j > 0$ given; and H_t^j represents the fraction of the available time (normalized to one unit per period) spent on non-leisure activities. The function $v(\cdot)$ is convex. The functional forms are chosen to be consistent with a balanced growth path — see King, Plosser and Rebelo (1988). Households supply labor and capital in perfectly competitive markets. R_t^K is the rental rate of capital and W_t is the real wage. The household's discount factor and the depreciation rate of capital satisfy $0 < \beta, \delta < 1$.

The expectation operator $\hat{E}_t^j$ denotes agent j's subjective beliefs. In forming expectations, households and firms observe only their own objectives, constraints and realizations of aggregate variables that are exogenous to their decision problems and beyond their control. The agent's problem is to choose sequences of consumption, hours worked, and capital in order to maximize (1) subject to (2), taking as given prices and the capital stock available at the beginning of the period. Beliefs are specified in the next section.

Household optimization yields the conditions

$$W_t = C_t^j \nu_H \left(H_t^j \right) \tag{3}$$

from equating the marginal rate of substitution between an additional unit of consumption and additional unit labor supply to their relative prices and

$$\beta \hat{E}_t^j \left[\frac{C_t^j}{C_{t+1}^j} \left(R_{t+1}^K + (1-\delta) \right) \right] = 1 \tag{4}$$

the Euler equation from equating the marginal rate of substitution between consumption today and tomorrow to the real interest rate.

The paper's primary goal is the quantitative evaluation of the model. Following Kydland and Prescott (1982), a log-linear approximation of the model around a balanced growth path is employed. For any variable G_t define $g_t = G_t/X_t$ as the corresponding normalized variable, where X_t is the level of technology in period t described further below. The model is then studied in log deviation from a non-stochastic steady state in these transformed variables so that $\hat{g}_t = \ln(g_t/\bar{g})$, with $\bar{g}$ denoting the steady state value of g_t. Details of the steady state and the log-linear approximation are confined to the appendix. Here it suffices to note that consumption, investment, output, the capital stock and real wages grow at the rate of technological progress in the balanced growth path so that

$$y_t = \frac{Y_t}{X_t};\ c_t = \frac{C_t}{X_t};\ i_t = \frac{I_t}{X_t};\ w_t = \frac{W_t}{X_t} \text{ and } k_t = \frac{K_t}{X_{t-1}}$$

are stationary. Hours and the rental rate of capital are stationary on the balanced growth path. Studying the approximated model also facilitates economic interpretation of later results.

Log-linearizing, solving the flow budget constraint forward, imposing the transversality condition and substituting for hours using a log-linear approximation to (3) gives the intertemporal budget constraint

$$\epsilon_c \hat{E}_t \sum_{T=t}^{\infty} \beta^{T-t} \hat{c}_T^j = \beta^{-1} \hat{k}_t^j + \hat{E}_t \sum_{T=t}^{\infty} \beta^{T-t} \left[\epsilon_w \hat{w}_T + \bar{R} \hat{R}_T^K - \beta^{-1} \hat{\gamma}_T \right].$$

The coefficients ϵ_c and ϵ_w are constants that are composites of model primitives, $\bar{R} > 0$ the gross rental rate, and $\hat{\gamma}_t = \ln(X_t/(X_{t-1}\bar{\gamma}))$ the log deviation of the growth rate in technological progress relative to steady state growth. This relation states the expected present value of consumption must be equal to the capital stock available at the beginning of the period plus the expected present value of wage and rental income. These latter variables are outside the control of the household, given the assumption of competitive markets.

To determine optimal consumption decisions, combine the intertemporal budget constraint with a log-linear approximation to (4) to yield

$$\begin{aligned}
\hat{c}_t^j &= \frac{1-\beta}{\epsilon_c}\left[\beta^{-1}\left(\hat{k}_t^j - \hat{\gamma}_t\right) + \bar{R}\hat{R}_t^K + \epsilon_w \hat{w}_t\right] \\
&\quad + \hat{E}_t^j \sum_{T=t}^{\infty} \beta^{T-t}\left[\frac{(1-\beta)}{\epsilon_c} - \beta\right]\beta\bar{R}\hat{R}_{T+1}^K + \\
&\quad \hat{E}_t^j \sum_{T=t}^{\infty} \beta^{T-t}\frac{(1-\beta)}{\epsilon_c}\beta\epsilon_w\hat{w}_{T+1}.
\end{aligned} \tag{5}$$

The consumption decision rule comprises three terms. The first shows the impact that the current level of the capital stock and current prices have on consumption. The second and third terms show how expected variations in permanent income affect current consumption. The former has two parts corresponding to the positive income effect and the negative substitution effect of higher returns to capital on current consumption. The latter has only one part as the income and substitution effects of a wage increase both increase current consumption.

Firms. There is a continuum of identical competitive firms of mass one. Each produces the economy's only good using capital and labor as inputs according to the production function

$$Y_t^i = \left(K_t^i\right)^{\alpha}\left(X_t H_t^i\right)^{1-\alpha}$$

where $0 < \alpha < 1$. Labor augmenting technological progress, X_t, satisfies the stochastic process

$$\ln\left(X_{t+1}/X_t\right) = \ln\bar{\gamma} + a_{t+1}$$

where a_t is an independent, identically distributed random variable with zero mean and standard deviation σ_A. $\bar{\gamma} > 0$ is the steady state rate of technology growth. This aggregate disturbance is the only source of exogenous variation in the model. Each firm chooses labor and capital inputs to maximize profits

$$\Pi_t^i = Y_t^i - R_t^K K_t^j - W_t H_t^j$$

taking factor prices as given. The first-order conditions to a firm's optimization problem provide

$$\begin{aligned}
W_t &= (1-\alpha)\left(K_t^i\right)^{\alpha}\left(X_t\right)^{1-\alpha}\left(H_t^i\right)^{-\alpha} \\
R_t^K &= \alpha\left(K_t^i\right)^{\alpha-1}\left(X_t H_t^i\right)^{1-\alpha}
\end{aligned}$$

which equate factor prices with their real marginal products.

2.2 Market clearing and aggregate dynamics

We are interested in studying the behavior of macroeconomic aggregates. As households have the same preferences and constraints; firms the same technology; and beliefs are assumed homogeneous across all agents (although they are assumed not to be aware of that) the analysis considers a symmetric equilibrium in which $\hat{k}_t^i = \hat{k}_t^j = \hat{k}_t$; $\hat{H}_t^j = \hat{H}_t^i = \hat{H}_t$; $\hat{\imath}_t^i = \hat{\imath}_t^j = \hat{\imath}_t$ for all i, j, t.

Integrating over the continuum provides aggregate consumption demand

$$\begin{aligned} \hat{c}_t \quad = \quad & \frac{1-\beta}{\epsilon_c}\left[\beta^{-1}\hat{k}_t + \bar{R}\hat{R}_t^K - \beta^{-1}\hat{\gamma}_t + \epsilon_w \hat{w}_t\right] \\ & + \hat{E}_t \sum_{T=t}^{\infty} \tilde{\beta}^{T-t}\left[\frac{(1-\beta)}{\epsilon_c} - \beta\right]\beta\bar{R}\hat{R}_{T+1}^K + \\ & \hat{E}_t \sum_{T=t}^{\infty} \beta^{T-t}\frac{(1-\beta)}{\epsilon_c}\beta\epsilon_w \hat{w}_{T+1} \end{aligned} \tag{6}$$

where $\int \hat{E}_t^j dj = \hat{E}_t$ denotes average expectations in the population. Aggregate consumption dynamics inherit the properties of individual decision rules. This is the only model equation that depends on expectations, and therefore of central focus. If near-rational expectations are to be a source of amplification and propagation of primitive shocks, the effects must originate here.

A log-linear approximation yields the remaining model equations. Aggregate capital dynamics are given by the accumulation equation

$$\hat{k}_{t+1} = \left[\alpha\frac{\bar{y}}{\bar{k}} + \frac{(1-\delta)}{\bar{\gamma}}\right]\hat{k}_t + (1-\alpha)\frac{\bar{y}}{\bar{k}}\hat{H}_t - \frac{\bar{c}}{\bar{k}}\hat{c}_t - \frac{(1-\delta)}{\bar{\gamma}}\hat{\gamma}_t. \tag{7}$$

The labor-leisure choice determines aggregate labor supply as

$$\epsilon_H \hat{H}_t = -\hat{c}_t + \hat{w}_t \tag{8}$$

where ϵ_H is the inverse Frisch elasticity of labor supply. Factor prices are expressed as

$$\hat{w}_t = \alpha\hat{k}_t - \alpha\hat{H}_t \tag{9}$$

$$\hat{R}_t^K = \hat{\gamma}_t + (\alpha-1)\hat{k}_t + (1-\alpha)\hat{H}_t. \tag{10}$$

And the resource constraint provides

$$\alpha \hat{k}_t + (1-\alpha)\hat{H}_t = \frac{\bar{c}}{\bar{y}}\hat{c}_t + \frac{\bar{\imath}}{\bar{y}}\hat{\imath}_t. \tag{11}$$

Given our assumption about technological progress, equations (6) - (11) together with the expectations formation mechanism specified in the next section completely determine the aggregate dynamics of the economy.

3 Beliefs

Optimal decisions require households to forecast the evolution of future wages and returns to capital. They are assumed to use a simple econometric model, relating wages and the capital rental rate to the aggregate stock of capital. That is

$$\hat{R}^K_t = \omega^r_0 + \omega^r_1 \hat{k}_t + e^r_t, \tag{12}$$

$$\hat{w}_t = \omega^w_0 + \omega^w_1 \hat{k}_t + e^w_t \tag{13}$$

and

$$\hat{k}_{t+1} = \omega^k_0 + \omega^k_1 \hat{k}_t + e^k_t \tag{14}$$

where e^r_t, $e^r_t$and e^r_t are i.i.d. shocks. The beliefs contain the same variables that appear in the minimum state variable rational expectations solution to the model. And, while the rational expectations solution does not contain a constant, it has the natural interpretation under learning of capturing uncertainty about the steady state (equivalently the level of technology). It is assumed that wages and capital are forecast in normalized units as under rational expectations.

Rational Expectations. The model solution under rational expectations implies (to a first-order approximation) that labor and capital prices and the next-period capital stock are linearly related to aggregate capital, with *time-invariant* coefficients $\omega^r_0 = \omega^w_0 = \omega^k_0 = 0$ and $\omega^r_1 = \bar{\omega}^r_1$, $\omega^w_1 = \bar{\omega}^w_1$, $\omega^k_1 = \bar{\omega}^k_1$. Hence, the agents' forecasting model nests beliefs that would be observed in a rational expectations equilibrium. Furthermore, under rational expectations $e^r_t = \bar{\omega}^r_3 \hat{\gamma}_t$, $e^w_t = \bar{\omega}^w_3 \hat{\gamma}_t$ and $e^k_t = \bar{\omega}^k_3 \hat{\gamma}_t$.

Perpetual learning. Agents estimate equations (12) – (14), updating their coefficient estimates every period as new data become available. Following recent literature, households update their estimates using a discounted least-squares estimator, assigning lower weight to older observations to protect against structural change.[4] Letting $\omega' = (\omega_0, \omega_1)$, $z_t = \left(\hat{R}_t^K, \hat{w}_t, \hat{k}_{t+1}\right)$ and $q_{t-1} = \left(1, \hat{k}_t\right)$, the algorithm can be written in recursive terms as

$$\hat{\omega}_t = \hat{\omega}_{t-1} + gR_t^{-1}q_{t-1}\left(z_t - \hat{\omega}'_{t-1}q_{t-1}\right)' \tag{15}$$

$$R_t = R_{t-1} + g\left(q_{t-1}q'_{t-1} - R_{t-1}\right) \tag{16}$$

where $\hat{\omega}_t$ denotes the current period's coefficient estimate and $g \in (0,1)$ denotes the constant gain, determining the rate at which older observations are discounted. The constant gain assumption delivers perpetual learning, as market participants 'forget' the past. However, the model has the property that if beliefs were instead given by a recursive least-square algorithm, defined by $g = t^{-1}$, the learning process would converge to the rational expectations coefficients.[5] Under the constant gain algorithm, agents' estimates converge to a distribution. Evans and Honkapohja (2001) show that for a gain sufficiently close to zero the distribution is normal and centered around the time-invariant coefficients of the rational expectations coefficients. Put differently, the model naturally "nests" the rational expectations model with g arbitrarily small. Finally, the above constant gain algorithm can be interpreted as a Kalman filter of a random coefficients model, with specific priors on the coefficients' drift (see Appendix).

Timing and information. Agents update their estimates at the end of the period, after making consumption and labor supply decisions. This avoids simultaneous determination of the parameters defining agents' forecast functions and current prices and quantities. However, to compare the model under learning with the predictions under rational expectations, we assume that agents' expectations are determined simultaneously with consumption and labor supply decisions, so that agents observe all variables that are determined at time t, including

[4] Of course we consider an otherwise stationary model environment with a single shock so as to clearly isolate the role of expectations in generating business cycle fluctuations. Adding structural change would generate further volatility.

[5] In the temporal limit, agents have an infinite amount of data.

$\hat{k}_{t+1}$. For example, the one-period-ahead forecast of $\hat{R}_t^K$ is

$$\hat{E}_t \hat{R}_{t+1}^K = \hat{\omega}_{0,t-1}^r + \hat{\omega}_{1t-1}^r \hat{k}_{t+1}$$

where $\hat{\omega}_{0,t-1}^r$ and $\hat{\omega}_{1t-1}^r$ are the previous period's estimates of belief parameters that define the period t forecast function. Hence, they observe the same variables that a 'rational' agent would observe. The only difference is that they are attempting to learn the 'correct' coefficients that characterize optimal forecasts. An alternative approach would be to assume expectations are formed before taking decisions, but this would render comparison of the learning model to the benchmark real business cycle model difficult as rational expectations would not be a special case of the assumed belief structure.

That beliefs are updated a period after new data arrive is a key component of learning as a friction. Like models of sticky information — see, for example, Mankiw and Reis (2002) — where only some firms can update information about the state of the economy, we assume that all agents can revise their beliefs in response to new data, but only with a one-period lag subject to the constraint of the constant gain learning technology.[6]

It is assumed the innovation, $\hat{\gamma}_t$, is not used in equations (12) – (14). This does not imply $\hat{\gamma}_t$ is unobserved — indeed, (5) implies consumption decisions are in part determined by these innovations. The interpretation is that while individual households and firms observe these disturbances they do not know how they are mapped into market clearing prices in general equilibrium. This assumption is similar to, though arguably more appealing than, the imperfect common knowledge literature where it is often assumed that only certain kinds of aggregate data are public knowledge or only certain markets are available to trade state-contingent claims. Absent these assumptions prices would fully reveal information about which agents are assumed to have only imperfect understanding — there is no inference problem. In the present model, if the innovation was used in forecasting, agents would not face an inference problem and learn quickly given that the only disturbance in the model is the technology shock.[7]

[6] More generally, beliefs are state variables that are sluggish by assumption, much like habit formation, price indexation, investment adjustment costs and labor market search (the latter typically assuming that new matches are only productive in future periods).

[7] Formally, including the disturbance would generate a singularity in the regression if initial beliefs coincide

Finally in forecasting over the decision horizon agents do not take into account that they update their beliefs in subsequent periods. The model is therefore one of anticipated utility — see Sargent (1999).

True Data Generating Process. Using (12) – (14) to substitute for expectations in (6) and solving delivers the actual data generating process

$$
\begin{aligned}
z_t &= T_1(\hat{\omega}_{t-1}) q_{t-1} + T_2(\hat{\omega}_{t-1}) \hat{\gamma}_t && (17)\\
\hat{\omega}_t &= \hat{\omega}_{t-1} + g R_t^{-1} q_{t-1} \left(\left[\left(T_1(\hat{\omega}_{t-1}) - \hat{\omega}_{t-1}' \right) q_{t-1} + T_2(\hat{\omega}_{t-1}) \hat{\gamma}_t \right] \right)' && (18)\\
R_t &= R_{t-1} + g \left(q_{t-1} q_{t-1}' - R_{t-1} \right) && (19)
\end{aligned}
$$

and

$$
\begin{bmatrix} \hat{c}_t \\ \hat{\imath}_t \\ \hat{H}_t \end{bmatrix} = \Psi z_t, \qquad (20)
$$

where $T_1(\hat{\omega})$ and $T_2(\hat{\omega})$ are nonlinear functions of the previous period's estimates of beliefs and Ψ is a matrix comprised of composites of primitive model parameters. The actual evolution of z_t is determined by a time-varying coefficient equation in the state variables $\hat{k}_t$ and $\hat{\gamma}_t$, where the coefficients evolve according to (18) and (19). The evolution of z_t depends on $\hat{\omega}_{t-1}$, while at the same time $\hat{\omega}_t$ depends on z_t. Learning induces self-referential behavior. Agents use current prices and capital holdings to make inferences about future macroeconomic conditions, but in equilibrium prices depend on agents' beliefs, and prices in turn affect the evolution of beliefs. This dependence on z_t is related to the fact that outside the rational expectations equilibrium $T_1(\hat{\omega}_{t-1}) \neq \hat{\omega}_{t-1}'$ and similarly for T_2.

with the rational expectations equilibrium. When initial beliefs differ from the rational expectations equilibrium, the regression is well defined, but because there is no uncertainty about the forecasting model, beliefs quickly converge to the predictions of a rational expectations analysis (where the singularity would again emerge given infinite data. And with a small gain, as in our analysis, the regression's variance-covariance martix would still be close to singular). Finally, if the proposed interpretation is unappealing, this particular assumption could be relaxed by introducing an additional shock to the model. We refrain from doing this to ensure comparability with the standard real business cycle framework.

4 Calibration

The sample characteristics we seek to match are for U.S. data, 1955:Q1 to 2004:Q4. A short description of each series is contained in the Appendix. Concerning households' preferences we set the discount rate $\beta = 0.99$. We assume separable preferences between consumption and leisure with log-utility for consumption and close-to-linear utility of leisure.[8] Accordingly, the inverse Frisch elasticity of labor supply, ϵ_H, is set equal to 0.0025. Firms' technology is specified by a capital share $\alpha = 0.34$ and steady state growth rate of labor augmenting technical progress equal to $\bar{\gamma} = 1.0053$, consistent with the quarterly mean output growth over the sample.

Two parameters are left to calibrate: the standard deviation of the shock, σ_A, and the constant gain, g. We calibrate these two parameters by minimizing the sum-of-squared distances between the model implied volatility of HP-detrended output and the first autocorrelation coefficient of output growth and the corresponding data moments. To do this, at each iteration in the minimization problem the model is simulated for 20000 periods. The first 2000 periods of the simulation are discarded and required statistics are computed using the remaining observations. This insures that the model reaches its stationary distribution of belief parameters, implying that our calibration and subsequent results do not depend on the initial conditions on the belief parameters.

As illustration of the possible effects of initial beliefs on inference, consider the following example. Suppose that data are generated according to the process $x_t = \bar{x} + \varepsilon_t$, where $\bar{x} > 0$ is a constant and ε_t an i.i.d mean zero disturbance. Estimate the mean using all sample observations and assume that the initial condition on the expectation of this mean is $\bar{x}^e > \bar{x}$. Over time, beliefs about the mean of x_t will be revised down as realizations of this random variable fluctuate around the true mean. The resulting estimates exhibit positive autocorrelation. Yet the true model has zero serial correlation. Our simulation approach ensures inference is not driven by transitional dynamics of this kind.[9]

[8] This approximates the labor supply properties of a model of indivisable labor — see Hanson (1985) and Rogerson (1988).

[9] This is one way in which our work is distinguished from Milani (2006).

The procedure gives a gain of 0.0029. To interpret this magnitude, note the gain indexes the weight that agents assign to past data. This value of the gain implies that observations that are 50 years old receive a weight of $(1 - 0.0029)^{200} \simeq 0.5$, implying agents do not discount past data too heavily.[10] To gauge the relative magnitude across observations, the weight assigned to the most recent data observation is approximately one.

One concern about the analysis might be that this choice of gain provides excessive freedom to fit observed data. Several points are worth making. First, many deviations from benchmark theory involve increased parameterization. This is true when incorporating investment adjustment costs, variable capital utilization, financial market frictions and labor market search — see, for example, Andolfatto (1996), Burnside and Eichenbaum (1996) and Carlstrom and Fuerst (1997). All these model variants engender more highly parameterized models and all seek to match the kinds of properties discussed here. In the same spirit, learning is an example of alternative friction whose implications for model fit are being evaluated.

Second, our calibrated gain is considerably smaller than values found in the literature, which range from $0.01 - 0.05$ — see, for example, Milani (2006), which estimates the gain, and Orphanides and Williams (2005). Branch and Evans (2006) show that a simple VAR with constant gain performs well in forecasting output growth and inflation, with respect to alternative more sophisticated models. The constant gain model is also shown to approximate well the behavior of output growth forecasts in the Survey of Professional Forecasters. The choice of gain that maximizes the fit in their VAR is 0.007, which is also above our calibration. In a more sophisticated model, the gain parameter could be calibrated to replicate key properties of expectations data from surveys. We leave that to future research.

Finally note that under rational expectations we only have to choose σ_A to match the standard deviation of HP-filtered output.

[10] For this value of g, agents would give approximately zero weight to observations that are 500 years old.

5 Central Results

5.1 Statistical Properties

Tables 1 and 2 report summary statistics on the cyclical properties of various U.S. data series and the model under both rational expectations and learning dynamics. For each variable the relative standard deviation and correlation with output are reported. Table 1 reports these statistics for HP-filtered series (facilitating comparison to earlier studies based on filtered data), while Table 2 presents the corresponding statistics for the growth rates of each series (which is more natural given the assumed stochastic trend).

Table 1: HP filtered moments

	Statistic	Data	REE	Learning
Technology:	σ_A	-	1.22	0.98
Output:	σ_Y	1.54	1.54	1.52
Consumption:	σ_C/σ_Y	0.52	0.54	0.38
	$\rho_{Y,C}$	0.69	0.97	0.83
Investment:	σ_I/σ_Y	2.87	2.42	3.06
	$\rho_{Y,I}$	0.90	0.99	0.98
Hours:	σ_H/σ_Y	1.13	0.49	0.71
	$\rho_{Y,H}$	0.88	0.97	0.96
Wages:	σ_w/σ_Y	0.54	0.54	0.38
	$\rho_{Y,w}$	0.12	0.97	0.84
Labor Prod:	σ_{Pr}/σ_Y	0.68	0.54	0.38
	$\rho_{Y,Pr}$	0.52	0.97	0.84

The first row of Table 1 shows learning dynamics amplify the effects of technology shocks. To match the variance and serial correlation properties of output, the learning model requires a technology disturbance with a standard deviation that is about 20 percent smaller than required under rational expectations. Moreover, the relative volatility of hours and investment is 40 and 25 percent higher respectively, bearing closer resemblance to data implied moments than the rational expectations model. The former represents a significant success, being problematic for standard real business cycle theory — see Hansen (1985) and Rogerson (1988).[11] In regards to consumption, wages and labor productivity, the model performs less well. Given the high elasticity of the labor supply and the assumption of perfectly competitive markets the model predicts $\hat{C}_t \approx \hat{w}_t = \hat{Y}_t - \hat{H}_t$ and is therefore too stylized to capture the different dynamics of these variables. The source of the discrepancy between the model and the data resides in the behavior of consumption, discussed further below.

Table 2 shows the same set of statistics in terms of growth rates, underscoring that the model under learning delivers a better fit. In particular, the model does not display the counterfactually large output growth volatility which occurs under rational expectations.[12]

Table 2: Growth rates

	Statistic	Data	REE	Learning
Output:	σ_{Δ_Y}	0.88	1.19	0.99
Consumption:	$\sigma_{\Delta_C}/\sigma_{\Delta_Y}$	0.60	0.52	0.54
	ρ_{Δ_Y,Δ_C}	0.51	0.98	0.80
Investment:	$\sigma_{\Delta_I}/\sigma_{\Delta_Y}$	2.54	2.45	2.82
	ρ_{Δ_Y,Δ_I}	0.71	0.99	0.94
Hours:	$\sigma_{\Delta_H}/\sigma_{\Delta_Y}$	0.93	0.50	0.65
	ρ_{Δ_Y,Δ_H}	0.70	0.98	0.87
Wages:	$\sigma_{\Delta_w}/\sigma_{\Delta_Y}$	0.60	0.52	0.54
	ρ_{Δ_Y,Δ_w}	0.08	0.98	0.80
Labor Prod:	$\sigma_{\Delta_{Pr}}/\sigma_{\Delta_Y}$	0.95	0.52	0.54
	$\rho_{\Delta_Y,\Delta_{Pr}}$	0.68	0.98	0.80

[11] Using different measures of hours and real activity, or different sample sizes, affects the specific values of σ_H/σ_Y, but does not alter our conclusions regarding the model's performance. For example, using non-farm business output as a measure of economic activity yields $\sigma_H/\sigma_Y = 0.83$ for the whole sample and $\sigma_H/\sigma_Y = 1.1$ for the sample 1982:Q3-2006:Q1. Using average weekly hours (from the BLS household survey data) and real GDP as a measure of output gives a relative standard deviation of 0.88, over the same period.

[12] The rational expectations model over-predicts the standard deviation of output growth by some 30 percent in contrast to 10 percent for the learning model.

Turning to the correlations between each series and output, all moments are closer to the data than are those under rational expectations. Of particular note are the weaker correlations of consumption, wages and labor productivity with output. To presage later discussion, these improvements in fit arise from learning endogenously generating dynamics that are qualitatively like those elicited by the presence of demand shocks. Revisions to beliefs shift the marginal utility of income. And for a given technology frontier these variations in marginal utility have qualitative similarities to demand shocks. As shown by Christiano and Eichenbaum (1992), the inclusion of such shocks in conjunction with technology disturbances can improve the fit of unconditional moments pertaining to labor market variables.

Since Cogley and Nason (1993, 1995), the internal propagation mechanisms of technology shocks have been a central preoccupation of real business cycle theory. These papers demonstrate that the impulse response functions of model variables are entirely determined by the assumed stochastic properties of technology shocks — the existence of capital as a state variable adds little propagation. Rotemberg and Woodford (1996), in related criticism of real business cycle theory, show that predictable variation in model simulated output, hours and consumption data is negligible, despite evidence of substantial forecastable variation of these series in observed data. Moreover, what predictable variation there is in the model is of the wrong kind.

Figure 1 plots the autocorrelation function for output growth together with model predictions under both rational expectations and learning. The rational expectations real business cycle model has virtually no propagation, having an autocorrelation function that is essentially equal to zero at all horizons — recall that the growth rate of technology is given by an i.i.d. process. In contrast, the learning model matches the first-order serial correlation properties, though generates little persistence beyond that. While matching this feature of the data was part of the calibration, we view it as a success given how well remaining model properties are captured.

Table 3 reports the autocorrelation properties of the growth rate of key model variables. Investment, output and hours growth are remarkably well matched relative to the predictions of the model under rational expectations. That wages, labor productivity and consumption are counterfactually predicted to have negative serial autocorrelation stems from the well known comovement problem in real business cycle theory (given that the dynamics of these series in the model are indistinguishable) emphasized by Barro and King (1984). While the impact effect of technology shocks does induce positive comovement, subsequent dynamics under learning are driven by revisions to beliefs. The next section shows that these revisions to beliefs are isomorphic to demand shocks in the sense that for a given production frontier shifts in expectations imply consumption and hours must be negatively correlated from the labor-leisure condition (3). This explains the observed positive serial correlation in hours and concomitant negative serial autocorrelation in consumption. The final section introduces an extension to the baseline model that resolves these counterfactual predictions.

Table 3: Autocorrelation in growth rates

	Statistic	Data	REE	Learning
Wages:	Δ_w	0.19	0.10	-0.14
Consumption:	Δ_C	0.25	0.11	-0.14
Investment:	Δ_I	0.34	-0.03	0.42
Output:	Δ_Y	0.28	0.00	0.28
Labor Prod:	Δ_{Pr}	0.05	0.10	-0.14
Hours:	Δ_H	0.58	-0.03	0.44

5.2 Impulse Response Functions

Further insight can be gleaned from impulse response functions to a unit technology shock. The effects of a disturbance depend on the precise beliefs maintained by households at the time of the shock. Impulse response functions for the learning model are therefore generated by simulating the model twice for $2000+T$ periods. The first 2000 periods guarantee convergence to the model stationary distribution and are discarded. The second simulation includes a unit shock in period 2001. The T-period impulse response to a unit technology shock is then given by the difference between these two trajectories. The simulation is repeated 3000 times.

For stationary variables, the impulse response functions are expressed in percentage deviations from steady state. For non-stationary series, the impulse responses are reported in percentage deviations from the trend growth rate.[13] For these later series, a unit technology shock results in a permanent increase in their level. In each plot the solid lines correspond to the median point-wise impulse response function, while the dotted lines provide a 75 percent band — that is, the 12.5 and 87.5 percentiles of the simulated impulse responses. The dashed line gives the corresponding impulse response predicted by a rational expectations analysis of the model.

Figures 2 - 6 report the impulse response functions for output, consumption, investment, hours and the rental rate of capital. For all series the impact effects of a technology shock are almost identical when comparing the median impulse response under learning and the impulse response under rational expectations. This is because agents' beliefs are distributed around the rational expectations prediction function, as shown in the next section. However, in the case of learning, there is variation in the impact effects. The observed amplification of technology shocks in the previous section is in part sourced to this variation. Depending on the precise beliefs of households and firms at the time of the shock, which along with the capital stock determine the state of the economy, the impact effect of the technology shock could be larger or smaller.

Output, hours and investment display a hump-shaped profile in response to a technology shock. This reflects earlier noted persistence properties induced by learning dynamics. At the time of the shock, belief coefficients are fixed so that the impact effects are on average the same. In subsequent periods, beliefs are revised in response to observed data with a one-period lag. This generates persistence in the actual data generating process for all series.

An interesting feature of the model concerns dynamics the period after the technology shock dissipates. In a rational expectations equilibrium, all model variables, appropriately normalized, are a linear function of the capital stock and the disturbance to the growth rate of technology. As the disturbance is assumed to be i.i.d., the observed dynamics one period after the shock are entirely determined by adjustment in the capital stock. Under learning, this is not the case. The technology shock leads to revisions in beliefs that commence the period after the disturbance. Subsequent dynamics are largely driven by revisions to beliefs.

[13] Equivalently, the dynamics are those observed in transition to the new steady state associated with the higher level of technology.

In the period after the disturbance, agents revise upwards their beliefs about the returns to investment and downwards their beliefs about wages — not just for the next period, but for all future periods in the household's decision horizon. Hence, the present discounted value of capital returns rises and the present discounted value of labor returns falls relative to the predictions of rational expectations. Figure 7 plots the time series of these sums under each belief structure. Recalling aggregate consumption dynamics given by equation (6), optimism about future returns — a steeper profile — tilts the consumption profile towards greater future consumption. This and the flatter expected wage path serve to increase the marginal utility of income relative to rational expectations, leading to larger labor supply and investment effects. Both predictions are, therefore, in part realized in equilibrium outcomes in the period after the shock: the return to capital rises and investment demand surges, while the real wage drops as aggregate labor supply increases.[14] Thus the model generates dynamics that are consistent with those described by Pigou (1927).

Learning amplifies the standard substitution and income effects that operate in real business cycle theory in response to a technology shock. In particular, the response in consumption and hours resemble the effects of an investment-specific shock affecting the expected future rate of return on investment. Agents forecast higher returns to capital which induce them to decrease current consumption and increase their labor supply through the familiar intertemporal substitution effect on leisure — see Greenwood, Hercowitz, and Huffman (1988) for a treatment of investment-specific shocks. This further amplifies the strong substitution effect already present under rational expectations. Shifts in expectations are endogenous to technology disturbances giving the model greater flexibility in fitting various second-order moments. Increased variation in marginal utility of income generates increased volatility in hours worked. Because these variations in hours are caused by variations in the supply of labor for a given production frontier, the model also better matches the various statistics

[14]The assumption that agents forecast normalized wages, which fall in response to a positive technology shock, is not important for the result. In a model with stationary but persistent techonology shocks, agents would forecast a rise in wages, but one that is smaller relative to rational expecations. Hence, it remains true that agents project a flatter profile for real wages. Only in the case of i.i.d. productivity shocks does learning induce higher wages and consumption than rational expectations. This is because under rational expectations the intertemporal substitution effects work in the opposite direction than under persistent shocks. Consumption and wages increase on impact and then converge to their steady state values.

relating to labor market variables — to wit, hours, wages and average labor productivity. Moreover, the tight correlation between consumption and output is broken.

5.3 Distributions of Beliefs

Because beliefs are central to our story it is useful to study their properties further. Consider the following thought experiment. An econometrician observes an economy with data generated according to the real business model under rational expectations. For each observed sample, the econometrician runs the exact regressions that comprise the beliefs in the learning model — recall equations (12) – (14) — calibrated with a gain equal to $g = 0.0029$. The coefficients are recorded for many simulations.[15]

The dashed line in Figure 8 plots a kernel estimate of the implied distribution of the resulting parameter estimates. Six distributions are reported corresponding to the intercept and slope coefficient in each of the three forecasting equations. Because the econometrician is outside the model — equivalently, the econometrician is small relative to the population of rational expectation agents — the distribution reflects pure sampling error: there is no feedback of this sampling error on the true data generating process. The distributions are centered on the rational expectations equilibrium, exhibit negligible bias, and have a fairly tight variance. This variance would go to zero as the gain parameter goes to zero, as this would imply that all data are given equal weight. But with the chosen positive gain it is evident that the econometrician has fairly accurate estimates of the parameters characterizing the true data generating process, and would therefore make comparably good forecasts of future returns as the rational agent.

Now imagine a world where all agents modeled by our real business cycle theory actually construct forecasts based on these estimated models. This is precisely the model discussed in this paper. The kernel estimate of the resulting ergodic distribution of beliefs is given by the solid lines. The distribution of the estimated coefficients on capital is not centered on the rational expectations parameters. The distributions are re-centered around the rational

[15] To compute the distribution of beliefs, the model is simulated 2250 periods and agents' estimates are recorded after discarding the first 2000 observations. The simulation is repeated 7000 times.

expectations coefficients to facilitate comparison with the non-feedback case.[16] However the median impact impulse responses shown in the previous section indicate that agents' median forecast is in line with rational expectations.

The variation in possible beliefs that can be held by agents is substantially more dispersed than in the previous thought experiment. This dispersion is what leads to the nonlinear impulse response functions and the associated uncertainty of their paths. This in turn generates the increased volatility in the learning model.

The figures show that the bulk of the dispersion in agents' beliefs is *endogenously* determined by the interaction between observed prices and updating of agents' beliefs. The dispersion in beliefs reflects that prices are less informative about future macroeconomic conditions. This model feature is further manifestation of shifting expectations as a source of business cycle fluctuations that is very much in the spirit of Pigou and Keynes. Shifting beliefs about the future returns to capital and wages, perhaps due to greater optimism about future investment opportunities, leads to changes in current market clearing prices for labor and capital. In turn, these prices reinforce beliefs.

These dynamics obviously relate to a number of recent papers on news shocks and business cycle dynamics — see for example Beaudry and Portier (2006) and Jaimovich and Rebelo (2008). The present analysis is distinct in the sense that there is only a single source of disturbance — technology shocks. The observed dynamics can be sourced to two kinds of variation: that due to the direct effects of the shock and that due to revisions in beliefs. Because the latter are endogenous to variations in technology they could arguably be termed "endogenous news shocks". Note, however, that the mechanism in each case is different. In our model, dynamics are generated by contemporaneous technology shocks and the endogenous pessimism and optimism reflected in revisions to beliefs. Endogenous news affects agents' decisions through intertemporal substitution effects on labor. In contrast, these other papers generate shifts in current equilibrium prices in response to signals about productivity at some future date that are exogenous to current technology: exogenous news affects agents' decisions through an income effect. Hence, in our model, negative comovement between consumption

[16] The "bias" in the estimates, a product of the nonlinearity of beliefs and linear regression methods, is about 6% for each coefficient. As the gain goes to zero this bias vanishes.

and hours implies higher hours and lower consumption; in models with exogenous news, negative comovement implies higher consumption and lower hours. Irrespective, learning clearly provides a mechanism through which expectations-driven business cycles emerge.

6 Expectations-Driven Business Cycles

Under learning dynamics, real business cycle theory still faces difficulty in matching two key characteristics of the data. The first is the relative volatility of hours and output — and labor market variables more generally. Without a high elasticity of labor supply, the model struggles to replicate the volatility of output. And while learning alleviates the magnitude of the discrepancy between data and model predictions, there remains the question of what other model features would better fit this dimension of the data.[17] The second data characterization regards the problem of comovement: hours and consumption are negatively correlated. Introducing an alternative belief structure can do little to resolve this model prediction. For a given production frontier, and under the assumption that consumption and leisure are normal goods, shifting beliefs, regardless of how they are modeled, cause variation in the marginal utility of income for which optimal decisions demand negative comovement in these variables.

An emerging literature under the rubric *expectations-driven business cycles* studies assumptions on preferences and technology that resolve this comovement problem. The motivating example is typically a news shock about the state of future technology. In the benchmark model under rational expectations it creates an increase in consumption and a decrease in hours and investment.

Beaudry and Portier (2006) explore primitive assumptions on production technology that, in a competitive environment, are consistent with positive comovement in these variables. They show that if production in a multi-sector model displays cost complementarities in intermediate goods inputs then an otherwise standard real business cycle model will produce expectations-driven business cycles: positive comovement between consumption, output, hours and investment in response to an expectations shock. A growing number of papers

[17] Introducing labor market search as in Andolfatto (1996) is one possible remedy, though this friction appears to have more success with persistence properties than as a source of amplification.

have proposed alternative resolutions to the comovement problem by considering more complex variants of the standard real business cycle model. Jaimovich and Rebelo (2008) propose modified preferences, variable capital utilization and adjustment costs to investment; Chen and Song (2007) introduce financial frictions; den Haan and Kaltenbrunner (2007) focus on labor market frictions; Floden (2006) considers a model with vintage capital; and Christiano, Motto, and Rostagno (2006) introduce monetary frictions.

6.1 The Model

Motivated by this literature, the benchmark model is augmented as follows. First, following Beaudry and Portier (2004) and Eusepi (2008), a production technology with a small degree of increasing returns is introduced. Second, household preferences are assumed to be non-separable in consumption and leisure but consistent with a long-run balanced growth path. These model features resolve the comovement problem. Increasing returns tends to induce persistent positive comovement in investment, consumption and hours in periods after a technology shock as the production function shifts out over time due to the external economies. Moreover, the assumption limits the quantity of investment crowded out by consumption. If increasing returns are deemed unappealing a model of endogenous entry and exit delivers an isomorphic production structure — see, for example, Portier (1995) or Jaimovich (2007). Non-separable preferences raise the marginal utility of consumption when labor supply is high, delivering tighter comovement between these variables. This modeling choice is dictated by keeping the model as simple as possible — no state variable is added — and to provide a meaningful comparison with the benchmark real business cycle framework.

Households maximize

$$\hat{E}_t^j \sum_{T=t}^{\infty} \beta^{T-t} \frac{\left(C_t^j\right)^{1-\sigma} v\left(1-L_t^j\right)}{1-\sigma}$$

subject to

$$C_t^j + K_{t+1}^j = R_t^K \left(U_t K_t^j\right) + W_t H_t^j + \left(1-\delta\left(U_t\right)\right) K_t^j.$$

The notation remains as before, with the following additions. U_t is the utilization rate of capital in any period t. The function $\delta\left(\cdot\right)$ gives the associated capital depreciation costs attached to a given utilization rate of capital. We choose $\delta\left(U_t\right) = \theta^{-1} U_t^{\theta}$. It is included to address the potential criticism that the benchmark model is designed to minimize amplification and propagation under rational expectations. The results show that even in the presence of this friction learning amplifies volatility relative to rational expectations by a greater magnitude than in the benchmark analysis. The only other change in the household's problem is the more general utility function. The utility function is assumed to be consistent with constant hours on the balanced growth path: it displays a constant intertemporal elasticity of substitution and constant Frisch elasticity of labor supply.

Firms maximize

$$Y_T - W_T H_T - R_T^K \left(U_t K_t\right)$$

by choice of effective capital input, $U_t K_t$, and labor input, H_t, subject to the production technology

$$Y_t = \Psi_t \left(U_t K_t\right)^{\alpha} \left(X_t H_t\right)^{1-\alpha}$$

where

$$\Psi_t = \left[\left(U_t K_t\right)^{\alpha} \left(X_t H_t\right)^{1-\alpha}\right]^{\eta} \; X_t^{-\eta}.$$

The term, Ψ_t, denotes the external effects of aggregate capital, indexed by the constant $\eta \geq 0$. The term $X_t^{-\eta}$ guarantees that a balanced growth path exists in this model. The assumptions $\sigma = 1$, $\eta = 0$ and $U_t = 1$ for all t delivers our benchmark model. Details of the first-order conditions; log-linear approximation; and resulting model equations are found in the appendix.

6.2 Calibration

The inverse Frisch elasticity of labor supply is set at the same value as in the simple real business cycle model. There are two extra parameters with respect to the benchmark model. The first parameter, measuring the aggregate externality, is set as $\eta = 0.1$, consistent with the lowest estimate in Baxter and King (1991). This value implies a "small" degree of externality and a locally determinate equilibrium under rational expectations.[18] The second parameter is the household's intertemporal elasticity of substitution, σ, which is chosen to make the ratio

[18]The parameter implies a downward-sloping demand for labor. For the connection between externality and indeterminacy, see Benhabib and Farmer (1994).

of the standard deviations of consumption and output in the model and HP-filtered data as close as possible. This gives $\sigma = 1.5$.[19] The parameters σ_A and g are again calibrated to match the standard deviation of output in the filtered data and the first autocorrelation of output growth respectively. The gain is now $g = 0.0015$, half that in the benchmark model. This gain implies a 74% weight on observations that are 50 years old. The appendix shows the parameter indexing variable depreciation, θ, is pinned down by the steady state return on capital and the steady state depreciation rate.

6.3 Results

Table 4 reports a subset of earlier presented statistics for the generalized model. The model does well in most dimensions. Assuming an intertemporal elasticity of consumption equal to $\sigma = 1.5$ achieves a stronger correlation between consumption and hours, reflected in the positive autocorrelation of the former. This comes at the cost of slightly lower volatility of investment relative to the benchmark model. These results address some of the concerns regarding predictable movements laid out in Rotemberg and Woodford (1996). While no evidence is adduced here on the magnitude of predictable movements in model dynamics, what is true is that the movements will be of the right kind.

Table 4: Model with increasing returns and non-separable preferences

	Statistic									
	σ_Y	σ_C/σ_Y	σ_I/σ_Y	σ_H/σ_Y	$\rho_{Y,C}$	$\rho_{Y,H}$	Δ_C	Δ_Y	Δ_I	Δ_H
Data	1.54	0.52	2.87	1.13	0.69	0.88	0.25	0.28	0.34	0.58
Model:										
$\sigma = 1.5$	1.50	0.52	2.42	0.70	0.99	0.99	0.14	0.27	0.35	0.39
$\sigma = 1$	1.50	0.33	3.04	0.70	0.92	0.98	0.01	0.22	0.30	0.33

[19] Under these parameter choices, the preferences of the representative agent have the property that consumption is an inferior good. Florin Bilbiie is thanked for alerting the authors to this possibility. The appendix describes microfoundations with costly labor market participation in which individual household preferences have consumption and leisure being normal goods but in which aggregate preferences approximate those assumed above. Eusepi and Preston (2008) develop theoretical implications in detail.

The second row shows the performance of the model when σ is equal to 1. This weakens the autocorrelation properties of consumption, which, as before, is noticeably less volatile than output. One last result is that the extended model improves the overall fit with the data but also increases amplification considerably. The standard deviation of the shock that is required to match the volatility of output is more than 30% lower than the required value under rational expectations (not shown).

The impulse response functions in figures 9 - 13 confirm that the model can generate expectations-driven business cycles, as consumption, investment and hours rise also *after* the productivity shock has occurred. Interestingly, learning as an endogenous news shock generates greater amplification and propagation when compared to other recent models of news-driven business cycles. For example, in Jaimovich and Rebelo (2008), model implied statistics are fairly similar across models with and without the news shock — current technology largely determine time series properties in that paper. This is not the case in our model.

7 Robustness

Modeling learning dynamics introduces one free parameter. It might reasonably be asked how sensitive are our results to the choice of gain parameter. Furthermore, our approach might be criticized on the ground that it is well known that real business cycle models need to be augmented with additional frictions to replicate observed data. And that if we permitted the real business cycle model under rational expectations a one parameter deviation from the benchmark model it would provide a similarly good fit as the model under learning dynamics. Or that the presence of such frictions would mitigate the role of learning as an amplification and propagation mechanism. The following exercises allay such concerns, showing that:

- Large gain coefficients generate excess volatility and counterfactual autocorrelations in many variables and therefore inferior fit of observed data;
- Introducing other frictions to the benchmark model under rational expectations, such as variable capital utilization or adjustment costs in investment, are not as successful

in fitting the data as well as our one parameter deviation of learning dynamics; and

- Even when learning is introduced in conjunction with these frictions, it continues to provide significant amplification and propagation relative to the same model under rational expectations.

7.1 Alternative Parameter Assumptions

Table 5 reports a subset of statistics for a number of variants of the benchmark model. The calibration is held fixed at our benchmark values for the model under learning, so that the standard deviation of technology shocks remains unchanged across simulations. Models 1 and 2 show the benchmark results for the rational expectations and learning models. The latter reiterates earlier results for ease of comparison while the former gives the results under rational expectations assuming the same standard deviation of technology shocks as model 2. The improved amplification is again immediate. Models 3 and 4 show the cases of a lower elasticity of labor supply ($\epsilon_H = 0.25$). Under both rational expectations and learning, the volatility of output falls for a given standard deviation technology shock. Concomitantly, the relative volatility of investment and hours also decline, while the relative volatility of consumption increases. The serial correlation properties adjust accordingly. These results underscore the centrality of the elasticity of labor supply in generating plausible volatility in real business cycle models.

Model 5 shows the learning model under a higher gain, $g = 0.009$, which is three times as large as our benchmark case. It significantly increases volatility in all series, but tends to overshoot corresponding sample moments. This makes clear that the modeler is not unconstrained in choosing this parameter — increasingly larger gains do not translate into increasingly better correspondence with data.

The final row reports statistics for an alternative model of learning. Many recent papers have proposed analyses of learning dynamics in the context of models where agents solve infinite horizon decision problems, but without requiring that agents make forecasts more than one period into the future. In these papers, agents' decisions depend only on forecasts of future variables that appear in Euler equations used to characterize rational expectations equilibrium. Key contributions include Bullard and Mitra (2002) and Evans and Honkapohja (2003).

Table 5: Robustness

	Statistic								
	σ_Y	σ_C/σ_Y	σ_I/σ_Y	σ_H/σ_Y	$\rho_{Y,C}$	$\rho_{Y,H}$	Δ_C	Δ_Y	Δ_I
Data	1.54	0.52	2.87	1.13	0.69	0.88	0.25	0.28	0.34
Model:									
Baseline RE	1.24	0.54	2.42	0.49	0.97	0.97	0.11	0.00	−0.03
Baseline Learn	1.52	0.38	3.07	0.72	0.83	0.96	−0.14	−0.28	0.42
Low Elast. RE	1.13	0.56	2.35	0.39	0.98	0.97	0.10	0.01	−0.02
Low Elast. Learn	1.28	0.43	2.91	0.55	0.87	0.95	−0.17	0.22	0.41
High Gain	2.30	0.32	4.00	1.04	0.03	0.95	−0.35	0.44	0.26
Euler Equation	1.24	0.54	2.42	0.49	0.97	0.97	0.10	0.00	−0.03

Of particular relevance to the present study are the analyses of Williams (2003) and Carceles-Poveda and Giannitsarou (2007). The former studies precisely the question explored here: can learning be a source of business cycle fluctuations? The latter is similarly motivated, with specific focus on asset pricing implications of real business cycle theory. Both papers make use of models with learning dynamics in which only one-period-ahead expectations matter to expenditure and production plans of households and firms. Both conclude that learning of the kind considered here is unpromising in generating amplification and propagation.[20]

A final related paper is Huang, Liu, and Zha (2008). It considers the same model as Williams (2003) where only one-period-ahead expectations matter, but examines a belief structure that does not nest the rational expectations equilibrium of the model. In particular, a class of self-confirming equilibria is analyzed — see Sargent (1999). The resulting impulse

[20] Williams (2003) also considers learning about structural parameters rather than the equilibrium mapping between state variables and prices and concludes that such uncertainty gives rises to greater amplification. This is a conceptually distinct exercise to that pursued in the present paper.

response functions indicate that such beliefs help amplify technology shocks. However, no attempt is made to calibrate the model to fit observed data. Again, our paper is distinguished from that analysis by considering optimal decisions conditional on beliefs and by constraining the class of beliefs to nest the rational expectations equilibrium of the model. The analyses also differ insofar as we consider model properties implied by the ergodic distribution of beliefs to remove the effects of initial conditions. Huang, Liu, and Zha (2008), in contrast, consider one specific choice of initial beliefs.

The final row replicates this kind of analysis in the context of the model developed here. Williams (2003) proceeds assuming that the Euler equations predicted by a rational expectations analysis of the model represent decision rules of agents under learning. The only model equation to change is that for consumption demand. The Euler equation is

$$c_t = E_t c_{t+1} - E_t \left(\beta \bar{R} R^K_{t+1} + \hat{\gamma}_{t+1} \right). \tag{21}$$

The model under learning then assumes household consumption decisions are determined as

$$c_t = \hat{E}_t c_{t+1} - \hat{E}_t \left(\beta \bar{R} R^K_{t+1} + \hat{\gamma}_{t+1} \right) \tag{22}$$

This requires the further assumption that households directly forecast their own future consumption using regressions of the kind specified in section 2. Preston (2005) shows that this decision rule leads to suboptimal decisions — see also Marcet and Sargent (1989).[21] All remaining model equations are unchanged as they do not directly depend on the specification of beliefs.

Not modeling optimal decisions and assuming consumption decisions are made according to (22) leads to dramatically different conclusions. Learning dynamics fail to generate amplification and propagation. Model implied moments are essentially indistinguishable from a rational expectations analysis of the model. This negative finding has less to do with learning than it does with the assumed nature of economic decisions. In real business cycle theory the only intertemporal decision is the household's consumption and saving decision. To make

[21] That (21) describes optimal decisions under rational expectations and not learning reflects the property under rational expectations of equilibrium probability laws embedding information about all relevant constraints, including transversality conditions and intertemporal budget constraints. This is not true once beliefs are exogenously specified as in the learning model contemplated here.

this decision households must forecast the entire future sequence of wages and real interest rates. These beliefs about future prices determine current market clearing prices, which in turn determine beliefs. A consequence of the model of household behavior given by (22) is the connection between market prices that govern future consumption and investment opportunities and current allocations and prices is severed. The economic structure of the model is completely changed and revealed to matter greatly for implied model dynamics. Only by properly modeling the interactions of households' and firms' beliefs about the economy and the markets in which they operate can we fully understand the potential of near-rational beliefs to explain observed data.

7.2 Alternative Frictions

Table 6 presents two final exercises. First, under both belief structures, model implications under variable capital utilization are considered. Second, under rational expectations only, a model with investment adjustment costs is presented. This permits a comparison of learning dynamics with one popular friction employed in the real business cycle literature. A more exhaustive comparative exercise is beyond the scope of this paper. The data moments and benchmark results are again presented in the first three rows.

Including variable capital utilization serves to amplify technology shocks under both belief structures. However, learning still provides 23 percent greater volatility. Regardless of the nature of beliefs, the relative volatilities, covariances and autocorrelations are largely unchanged. Introducing investment adjustment costs of the form

$$K_{t+1} = I_t \left[1 - \phi\left(\frac{I_t}{I_{t-1}}\right)\right] + (1-\delta)K_t$$

with $\phi(\bar{\gamma}) = \phi'(\bar{\gamma}) = 0$ and $\phi''(\bar{\gamma}) > 0$ in the rational expectations model certainly improves correspondence with data on some dimensions — the first-order serial correlation properties of output and investment are much improved and output is more volatile.[22] But remaining moments are, if anything, further from the data. In particular, the relative volatility of investment is considerably dampened. Finally, while results for learning under investment

[22] In this experiment σ_A and $\phi''(\bar{\gamma})$ are chosen to match the volatility of HP-filtered output and the first order autocorrelation of output growth.

adjustment costs have not been presented, we conjecture that such frictions will only enhance the amplification and propagation of near-rational expectations. Frictions that introduce additional state variables make current quantities and prices more sensitive to households' and firms' beliefs about future economic conditions — and this is the heart of our theory of Pigou-type fluctuations.

Table 6: Alternative Frictions

	Statistic								
	σ_Y	σ_C/σ_Y	σ_I/σ_Y	σ_H/σ_Y	$\rho_{Y,C}$	$\rho_{Y,H}$	Δ_C	Δ_Y	Δ_I
Data	1.54	0.52	2.87	1.13	0.69	0.88	0.25	0.28	0.34
Model:									
Baseline RE	1.24	0.54	2.42	0.49	0.97	0.97	0.11	0.00	−0.03
Baseline Learn	1.52	0.38	3.07	0.72	0.83	0.96	−0.14	0.28	0.42
Var. Cap. Utl. RE	1.81	0.51	2.48	0.51	0.98	0.98	0.0	−0.02	−0.03
Var. Cap. Utl. Learn	2.23	0.38	2.91	0.66	0.93	0.98	−0.08	0.26	0.38
Inv. Adj. Costs RE	1.53	0.68	2.09	0.39	0.96	0.89	−0.04	0.16	0.56

8 Conclusion

This paper explores learning dynamics as a source of economic fluctuations, assessing its implications for the amplification and propagation of technology shocks in real business cycle models. In the spirit of Pigou (1927) a model is developed in which self-fulfilling expectations are possible in response to technology shocks. The benchmark model delivers volatility in output comparable to a rational expectations analysis with a standard deviation of technology shock that is 20 percent smaller, and has substantially more volatility in investment and hours. The model captures persistence in these series, unlike standard models. The improvement in fit stems from shifting beliefs having properties of demand shocks.

While introducing learning dynamics improves model fit relative to rational expectations, the benchmark model suffers a comovement problem between consumption, hours, output and investment. An augmented model that is consistent with expectations-driven business cycles, in the sense of Beaudry and Portier (2006), resolves these counterfactual predictions. This richer model produces additional amplification and propagation, requiring 30 percent smaller technology shocks than a rational expectations analysis, while providing a superior characterization of other second-order moments of observed data.

References

Andolfatto, D. (1996): "Business Cycles and Labor-Market Search," *American Economic Review*, 86(1), 112.

Beaudry, P., and F. Portier (2004): "When Can Changes in Expectations Cause Business Cycle Fluctuations in Neo-Classical Settings?," NBER Working Paper 10776.

——— (2006): "When Can Changes in Expectations Cause Business Cycle Fluctuations in Neo-Classical Settings?," *Journal of Economic Theory*, 135, 458–477.

Benhabib, J., and R. Farmer (1994): "Indeterminacy and Increasing Returns," *Journal of Economic Theory*, 63, 19–41.

Branch, W. A., and G. W. Evans (2006): "A Simple Recursive Forecasting Model," *Economics Letters*, (91), 158–166.

Bullard, J., and K. Mitra (2002): "Learning About Monetary Policy Rules," *Journal of Monetary Economics*, 49(6), 1105–1129.

Burnside, C., and M. Eichenbaum (1996): "Factor-Hoarding and the Propagation of Business-Cycle Shocks," *American Economic Review*, 86(5), 1154–1174.

Carceles-Poveda, E., and C. Giannitsarou (2007): "Asset Pricing with Adaptive Learning," forthcoming, Review of Economic Dynamics.

Carlstrom, C. T., and T. S. Fuerst (1997): "Agency Cost, Net Worth, and Business Cycle Fluctuations: A Computable General Equilibrium Analysis," *American Economic Review*, 87(5), 893–910.

Chen, K., and Z. Song (2007): "Financial Friction, Capital Re-allocation and News-driven Business Cycles," mimeo.

Christiano, L. J., and M. Eichenbaum (1992): "Current Real-Business-Cycle Theories and Aggregate Labor-Market Fluctuations," *American Economic Review*, 82, 430–450.

Christiano, L. J., R. Motto, and M. Rostagno (2006): "Monetary Policy and the Stock Market Boom-Burst Cycles," unpublished, Northwestern University.

Cogley, T., and J. M. Nason (1993): "Impulse Dynamics and Propagation Mechanisms in a Real Business Cycle Model," *Economics Letters*, 453, 77–81.

——— (1995): "Output Dynamics in Real-Business-Cycle Models," *American Economic Review*, 85(3), 492–511.

den Haan, W., and G. Kaltenbrunner (2007): "Anticipated Growth and Business Cycles in Matching Models," mimeo.

Edge, R., T. Laubach, and J. Williams (2007): "Learning and Shifts in Long-Run Productivity Growth," *Journal of Monetary Economics*, pp. 2421–2438.

Eusepi, S. (2008): "A Note on Expectations-Driven Business Cycles in Economies with Production Exernalities," forthcoming International Journal of Economic Theory.

Eusepi, S., and B. Preston (2008): "Preferences and Macroeconomic Comovement," unpublished, Columbia University.

Evans, G. W., and S. Honkapohja (2001): *Learning and Expectations in Economics*. Princeton, Princeton University Press.

——— (2003): "Expectations and the Stability Problem for Optimal Monetary Policies," *Review of Economic Studies*, 70(4), 807–824.

Floden, M. (2006): "Vintage Capital and Expectations-Driven Business Cycles," mimeo.

Greenwood, J., Z. Hercowitz, and G. W. Huffman (1988): "Investment, Capacity Utilization and the Real Business Cycle," *American Economic Review*, 78(3), 402–417.

Hansen, G. (1985): "Indivisible Labor and the Business Cycle," *Journal of Monetary Economics*, 16, 309–337.

HUANG, K., Z. LIU, AND T. ZHA (2008): "Learning, Adaptive Expectations, and Technology Shocks," unpublished, Federal Reserve Bank of Atlanta.

JAIMOVICH, N. (2007): "Firm Dynamics and Markup Variatopms: Implications for Sunspot Equilibria and Ednogenous Economic Fluctuations," forthcoming, Journal of Economic Theory.

JAIMOVICH, N., AND S. REBELO (2008): "Can News About the Future Drive the Business Cycle," unpublished, Northwestern University.

KING, R. G., C. I. PLOSSER, AND S. T. REBELO (1988): "Production, Growth and Business Cycles," *Journal of Monetary Economics*, 21, 195–232.

KING, R. G., AND S. T. REBELO (1999): "Resuscitating Real Business Cycles," in *Handbook of Macroeconomics*. Elsevier.

KREPS, D. (1998): "Anticipated Utility and Dynamic Choice," in *Frontiers of Research in Economic Theory*, ed. by D. Jacobs, E. Kalai, and M. Kamien, pp. 242–274. Cambridge: Cambridge University Press.

KYDLAND, F. E., AND E. C. PRESCOTT (1982): "Time to Build and Aggregate Fluctuations," *Econometrica*, 50(6), 1345–1370.

LORENZONI, G. (2008): "A Theory of Demand Shocks," unpublished, MIT.

MANKIW, N. G., AND R. REIS (2002): "Sticky Information versus Sticky Prices: A Proposal to Replace the New Keynesian Phillips Curve," *Quarterly Journal of Economics*, 4, 1295–1328.

MARCET, A., AND T. J. SARGENT (1989): "Convergence of Least-Squares Learning in Environments with Hidden State Variables and Private Information," *Journal of Political Economy*, pp. 1306–1322.

MILANI, F. (2006): "Expectations, Learning and Macroeconomic Persistence," *Journal of Monetary Economics*, 54, 2065–2082.

NIEUWERBURGH, S. V., AND L. VELDKAMP (2006): "Learning Asymmetries in Real Business Cycle Models," *Journal of Monetary Economics*, 53(4), 753–772.

ORPHANIDES, A., AND J. C. WILLIAMS (2005): "Imperfect Knowledge, Inflation Expectations, and Monetary Policy," in *The Inflation Targeting Debate*, ed. by B. S. Bernanke, and M. Woodford. University of Chicago Press.

PIGOU, A. C. (1927): *Industrial Fluctuations*. London: MacMillan.

PORTIER, F. (1995): "Business Formation and Cyclical Markups in French Business Cycles," *Annales D'Economie et de Statistique*, 37/38, 411–440.

PRESCOTT, E. (1986): "Theory Ahead of Business Cycles Measurement," *Carnegie-Rochester Conference Series on Public Policy*, 25, 11–66.

PRESTON, B. (2005): "Learning About Monetary Policy Rules when Long-Horizon Expectations Matter," *International Journal of Central Banking*, 1(2), 81–126.

ROGERSON, R. (1988): "Indivisible Labor, Lotteries and Equilibrium," *Journal of Monetary Economics*, 21, 3–16.

ROTEMBERG, J., AND M. WOODFORD (1996): "Real Business Cycle Models and the Forecastable Movements in Output, Hours and Consumption," *American Economic Review*, 86(1), 71–89.

SARGENT, T. J. (1999): *The Conquest of American Inflation*. Princeton University Press.

SCHMITT-GROHE, S. (2000): "Endogenous Business Cycles and the Dynamics of Output, Hours and Consumption," *American Economic Review*, 90, 1136–1159.

WILLIAMS, N. (2003): "Adaptive Learning and Business Cylces," unpublished, Princeton University.

A Appendix

A.1 Data

We use quarterly data for the US economy. The sample is 1955:Q1 to 2004:Q4. The variables are constructed as follows with DLX codes in parentheses. Output is Real Gross Domestic Product (GDPH); nominal consumption is computed as the sum of nondurable goods (CN), services (CS) and government expenditures (G); nominal investment is the sum of private non-residential investment structures (FNS), Equipment and software (FNE), private residential investment (FR) and consumption durable goods (CD). Consumption and investment are converted to real terms by using the GDP deflator (GDP/GDPH). Hours are measured by non-farm business hours (LXNFH). All variables a transformed to per capita terms by using the civilian non-institutional population above 16 years (LN16N). Productivity is measured as output per hour in the non-farm business sector (LXNFA). Finally, the hourly wage is measured by compensation per hour in the non-farm business sector (LXNFC). Real wage is obtained by using the non-farm output price deflator (LXNFI). We also document the volatility of hours by using (as an alternative measure) the average hours of all persons at work from the household survey (LENCLWHN). For this series, we use the sample 1982:Q3-2006:Q1.

A.2 Model

This section delineates the general model that includes capacity utilization, non-separability between consumption and leisure and externalities of production.

A.2.1 Households

Consumers choose consumption, leisure and capital to maximize

$$\hat{E}_t \sum_{T=t}^{\infty} \beta^{T-t} u\left(C_T, L_T\right)$$

subject to

$$\begin{aligned} C_t + K_{t+1} &= R_t^K \left(u_t K_t\right) + W_t H_t + \left(1 - \delta\left(U_t\right)\right) K_t \\ L_t &= 1 - H_t. \end{aligned}$$

The first order conditions are

$$\begin{aligned}
C_t &: u_c(C_t, L_t) = \Lambda_t \\
K_{t+1} &: \beta \hat{E}_t \Lambda_{t+1} R^K_{t+1} u_{t+1} - \Lambda_t + \beta \hat{E}_t \left[\Lambda_{t+1} \left(1 - \delta \left(U_{t+1}\right)\right)\right] = 0 \\
L_t &: u_L(C_t, L_t) = -\Lambda_t W_t \\
U_t &: R^K_t = \delta' \left(U_t\right).
\end{aligned}$$

In the sequel we assume

$$u \left(C_t, L_t\right) = \frac{C_t^{1-\sigma} v \left(1 - L_t\right)}{1 - \sigma}$$

where $\nu' \left(\bar{H}\right) > 0$, $\bar{H}$ is steady state hours worked, and $\epsilon_\nu = \nu'' \bar{H} / \nu > 0$. Also, we assume

$$\delta \left(U_t\right) = \frac{1}{\theta} U_t^\theta$$

with $\delta \left(\bar{U}\right) = \delta$ in steady state.

Normalized non-stationary variables are denoted by lower case letters. Stationary variables are left unchanged. Hence, for any trending variable G_t define $g_t = G_t / X_t$ as the corresponding normalized variable. The model is then studied in log deviation from a non-stochastic steady state in these normalized variables so that $\hat{g}_t = \ln \left(g_t / \bar{g}\right)$, with $\bar{g}$ denoting the steady state value of g_t.

In terms of normalized variables the first-order conditions are as follows. For consumption:

$$\lambda_t \equiv X_t^\sigma \Lambda_t = X_t^\sigma u_c(C_t, L_t) = X_t^\sigma C_t^{-\sigma} v \left(H_t\right) = c_t^{-\sigma} v \left(H_t\right).$$

For capital:

$$1 = \beta \hat{E}_t \left[\frac{\lambda_{t+1}}{\lambda_t} \frac{1}{\gamma_{t+1}^\sigma} \left(R^K_{t+1} U_{t+1} + \left(1 - \delta \left(U_{t+1}\right)\right)\right)\right]$$

For Leisure:

$$\lambda_t w_t = X_t^{\sigma - 1} \Lambda_t W_t = -\frac{c_t^{1-\sigma}}{1 - \sigma} v' \left(H_t\right).$$

A log-linear approximation to these relations around a balanced growth path provides:

1. Marginal utility of consumption:

$$\hat{\lambda}_t = -\sigma \hat{c}_t - \psi \left(1 - \sigma\right) \hat{H}_t$$

where in steady state

$$\psi \equiv \frac{\bar{H}v'\left(\bar{H}\right)}{v\left(\bar{H}\right)}\left(\sigma-1\right)^{-1}=\frac{\bar{w}\bar{H}}{\bar{c}}.$$

2. Euler equation:

$$\beta\hat{E}_t\left[\beta^{-1}\left(\hat{\lambda}_{t+1}-\hat{\lambda}_t-\sigma\hat{\gamma}_{t+1}\right)+\left(\beta^{-1}-\frac{(1-\delta)}{\bar{\gamma}^{\sigma}}\right)\left(\hat{R}^K_{t+1}+\hat{U}_{t+1}\right)-\frac{\delta}{\bar{\gamma}^{\sigma}}\theta\hat{U}_{t+1}\right]=0$$

which on using the steady state relation

$$\frac{\bar{R}^K\bar{U}}{\bar{\gamma}^{\sigma}}=\left(\beta^{-1}-\frac{(1-\delta)}{\bar{\gamma}^{\sigma}}\right)=\frac{\theta\delta}{\bar{\gamma}^{\sigma}}$$

becomes

$$\hat{E}_t\left[\left(\hat{\lambda}_{t+1}-\hat{\lambda}_t-\sigma\hat{\gamma}_{t+1}\right)+\beta\left(\beta^{-1}-\frac{(1-\delta)}{\gamma^{\sigma}}\right)\hat{R}^K_{t+1}\right]=0.$$

3. Labor-leisure choice:

$$\left(1-\sigma\right)\hat{c}_t+\epsilon_{\nu}\hat{H}_t=\hat{\lambda}_t+\hat{w}_t,$$

which, combined with the expression for marginal utility, gives:

$$\sigma^{-1}\hat{\lambda}_t+\hat{w}_t=\epsilon_H\hat{H}_t$$

where

$$\epsilon_H=\epsilon_{\nu}-\frac{\left(\sigma-1\right)^2}{\sigma}\psi>0$$

is the inverse Frisch elasticity of labor supply.[23]

4. Capacity utilization:

$$\hat{U}_t=\frac{1}{(\theta-1)}\hat{R}^K_t.$$

A.2.2 Firms

The firms's problem is

$$\max_{U_tK_t,H_t} Y_T-W_TH_T-R^K_T\left(U_tK_t\right)$$

subject to the production technology

$$Y_t=\Psi_t\left(U_tK_t\right)^{\alpha}\left(X_tH_t\right)^{1-\alpha}\;\text{ where }\Psi_t=\left[\left(U_tK_t\right)^{\alpha}\left(X_tH_t\right)^{1-\alpha}\right]^{\eta}\;X_t^{-\eta}$$

[23]The restriction $\epsilon_H>0$ guarantees the concavity of the utility function.

denotes the external effects of aggregate capital. The term $X_t^{-\eta}$ guarantees that a balanced growth path with exogenous growth exists in this model. Output is made stationary by the following transformation

$$\Psi_t \gamma_t^{-\alpha} k_t^{\alpha} U_t^{\alpha} H_t^{1-\alpha} = y_t$$

which is log-linearized to

$$\hat{\Psi}_t - \alpha\hat{\gamma}_t + \alpha\hat{k}_t + \alpha\hat{U}_t + (1-\alpha)\hat{H}_t = \hat{y}_t. \tag{23}$$

The external effects can be expressed in terms of stationary variables

$$\Psi_t = \left[\left(\frac{U_t}{\gamma_t} k_t\right)^{\alpha} H_t^{1-\alpha}\right]^{\eta},$$

and after log-linearizing

$$\hat{\Psi}_t = \eta\alpha\left(\hat{U}_t + \hat{k}_t - \hat{\gamma}_t\right) + (1-\alpha)\eta\hat{H}_t. \tag{24}$$

The first order condition with respect to hours is

$$-W_t + (1-\alpha)\Psi_t (U_t K_t)^{\alpha} (X_t)^{1-\alpha} H_t^{-\alpha} = 0$$

which becomes

$$(1-\alpha)\Psi_t X_{t-1}^{\alpha}\left(U_t \frac{K_t}{X_{t-1}}\right)^{\alpha} \frac{(X_t)^{1-\alpha}}{X_t} H_t^{-\alpha} = \frac{W_t}{X_t}$$

and hence

$$(1-\alpha)\Psi_t \gamma_t^{-\alpha} (U_t k_t)^{\alpha} H_t^{-\alpha} = w_t.$$

Combined with the definition of output gives

$$w_t = (1-\alpha)\frac{y_t}{H_t}$$

which in log-linear form becomes

$$\hat{w}_t = \hat{y}_t - \hat{H}_t. \tag{25}$$

The capital input decision gives:

$$0 = -R_t^K + \alpha\Psi_t\left(\frac{U_t}{\gamma_t} k_t\right)^{\alpha-1} H_t^{1-\alpha}.$$

Using the definition of output yields

$$R_t^K = \alpha\gamma_t \frac{y_t}{U_t k_t}$$

which in log-linear form is

$$\hat{R}_t^K = \hat{\gamma}_t + \hat{y}_t - \hat{U}_t - \hat{k}_t. \tag{26}$$

Finally, the evolution of capital in log-linear terms is described by:

$$\hat{k}_{t+1} = \frac{\bar{\imath}}{\bar{k}}\hat{\imath}_t + \frac{(1-\delta)}{\bar{\gamma}}\left(\hat{k}_t - \hat{\gamma}_t\right) - \frac{\delta\theta}{\bar{\gamma}}\hat{U}_t. \tag{27}$$

A.3 Consumption decision rule

The final task is to derive the optimal consumption decision rule under arbitrary expectations. Households choose a path for consumption, taking as given their initial capital holdings, capital and labor prices and their expectations about future prices. The flow budget constraint can be expressed in terms of stationary variables

$$c_t + k_{t+1} = (\gamma_t)^{-1} R_t^K (U_t k_t) + w_t H_t + (1 - \delta(U_t))(\gamma_t)^{-1} k_t.$$

Log-linearization gives

$$\bar{\gamma}^{\sigma-1}\beta^{-1}\hat{k}_t = \left\{ \begin{array}{c} \frac{\bar{c}}{\bar{k}}\hat{c}_t + \hat{k}_{t+1} - \\ \bar{R}\bar{\gamma}^{\sigma-1}\left(\hat{R}_t^K + \hat{U}_t + \frac{1-\alpha}{\alpha}\hat{w}_t + \frac{1-\alpha}{\alpha}\hat{H}_t - \hat{\gamma}_t\right) - \\ \frac{(1-\delta)}{\bar{\gamma}}\left[-\hat{\gamma}_t - \frac{\delta}{(1-\delta)}\theta\hat{U}_t\right] \end{array} \right\}$$

employing the relations

$$\frac{\bar{w}\bar{H}}{\bar{k}} = \frac{1-\alpha}{\alpha}\frac{\bar{U}\bar{R}^k}{\bar{\gamma}} = \frac{1-\alpha}{\alpha}\bar{R}\bar{\gamma}^{\sigma-1} \text{ and } \left[\frac{\bar{R}^k\bar{U}}{\bar{\gamma}} + \frac{1-\delta}{\bar{\gamma}}\right] = \bar{\gamma}^{\sigma-1}\beta^{-1} = \tilde{\beta}^{-1}.$$

Defining

$$\begin{aligned} \bar{R}\bar{\gamma}^{\sigma-1} &= \bar{\gamma}^{\sigma-1}\left(\beta^{-1} - \frac{(1-\delta)}{\bar{\gamma}^{\sigma}}\right) \\ &= \tilde{\beta}^{-1} - \frac{(1-\delta)}{\bar{\gamma}} \\ &= \tilde{R} \end{aligned}$$

and $\frac{\bar{R}^k \bar{U}}{\bar{\gamma}} = \frac{\delta\theta}{\bar{\gamma}}$, the expression above can be further simplified to

$$\hat{k}_t = \tilde{\beta}\left[\frac{\bar{c}}{\bar{k}}\hat{c}_t + \hat{k}_{t+1} + \tilde{\beta}^{-1}\hat{\gamma}_t - \tilde{R}\left(\hat{R}_t^K + \frac{1-\alpha}{\alpha}\hat{w}_t + \frac{1-\alpha}{\alpha}\hat{H}_t\right)\right].$$

Using the labor supply condition

$$\sigma^{-1}\hat{\lambda}_t + \hat{w}_t = \epsilon_H \hat{H}_t$$

and the definition of marginal utility gives the following constant-consumption labor supply

$$\left[\epsilon_H - \frac{\sigma-1}{\sigma}\psi\right]\hat{H}_t = -\hat{c}_t + \hat{w}_t \tag{28}$$

Substituting for labor supply decision $\hat{H}_t$ using the household's first-order condition gives

$$\hat{k}_t = \tilde{\beta}\left(\epsilon_c \hat{c}_t + \tilde{\beta}^{-1}\hat{\gamma}_t + \hat{k}_{t+1} - \epsilon_w \hat{w}_t - \tilde{R}\hat{R}_t^K\right)$$

where

$$\begin{aligned}
\epsilon_c &= \frac{\bar{c}}{\bar{k}} + \left[\epsilon_H - \psi\frac{(\sigma-1)}{\sigma}\right]^{-1}\tilde{R}\frac{1-\alpha}{\alpha} \\
\epsilon_w &= \left(1 + \left[\epsilon_H - \psi\frac{(\sigma-1)}{\sigma}\right]^{-1}\right)\tilde{R}\frac{1-\alpha}{\alpha}.
\end{aligned}$$

Solving forward and taking expectations yields the intertemporal budget constraint

$$\epsilon_c \hat{E}_t \sum_{T=t}^{\infty}\tilde{\beta}^{T-t}\hat{c}_T = \tilde{\beta}^{-1}\hat{k}_t + \hat{E}_t\sum_{T=t}^{\infty}\tilde{\beta}^{T-t}\left(\epsilon_w \hat{w}_T + \tilde{R}\hat{R}_T^K - \tilde{\beta}^{-1}\hat{\gamma}_T\right)$$

as of time t. Solving the Euler equation backward from time T and taking expectations at t yields

$$\hat{E}_t\left(\sigma\hat{c}_T + \psi(1-\sigma)\hat{H}_T\right) = \sigma\hat{c}_t + \psi(1-\sigma)\hat{H}_t + \hat{E}_t\left[\sum_{T=t}^{T-1}\left(\tilde{\beta}\tilde{R}\hat{R}_{T+1}^K - \sigma\hat{\gamma}_{T+1}\right)\right].$$

Substituting for the constant-consumption labor supply we obtain

$$\hat{E}_t\left(\left[(1-\chi)\sigma\hat{c}_T + \chi\sigma\hat{w}_T\right]\right) = \sigma\hat{c}_t + \psi(1-\sigma)\hat{H}_t + \hat{E}_t\left[\sum_{T=t}^{T-1}\left(\tilde{\beta}\tilde{R}\hat{R}_{T+1}^K - \sigma\hat{\gamma}_{T+1}\right)\right]$$

where

$$\chi = \frac{\psi(1-\sigma)}{\sigma\epsilon_H + \psi(1-\sigma)}.$$

Rearranging in terms of expected consumption and substituting into the intertemporal budget constraint we get

$$\hat{c}_t + \sigma^{-1}\psi\,(1-\sigma)\,\hat{H}_t =$$

$$\frac{(1-\chi)\left(1-\tilde{\beta}\right)}{\epsilon_c\tilde{\beta}}\hat{k}_t +$$

$$\hat{E}_t\sum_{T=t}^{\infty}\tilde{\beta}^{T-t}\left[\frac{(1-\chi)\left(1-\tilde{\beta}\right)}{\epsilon_c}\left((\epsilon_w+\tilde{\epsilon}_c\chi)\,\hat{w}_T+\tilde{R}\hat{R}_T^K-\tilde{\beta}^{-1}\hat{\gamma}_T\right)-\tilde{\beta}\left(\sigma^{-1}\tilde{\beta}\tilde{R}\hat{R}_{T+1}^K-\hat{\gamma}_{T+1}\right)\right].$$

Finally, we obtain the consumption decision rule, depending only on forecast of prices that are beyond the control of the household

$$\begin{aligned}
\hat{c}_t + \sigma^{-1}\psi\,(1-\sigma)\,\hat{H}_t \;=\;& \frac{(1-\chi)\left(1-\tilde{\beta}\right)}{\epsilon_c}\left[\tilde{\beta}^{-1}\hat{k}_t+\tilde{R}\hat{R}_t^K-\tilde{\beta}^{-1}\hat{\gamma}_t+(\epsilon_w+\epsilon_c\chi)\,\hat{w}_t\right]\\
&+\hat{E}_t\sum_{T=t}^{\infty}\tilde{\beta}^{T-t}\left[\tilde{\beta}-\frac{(1-\chi)\left(1-\tilde{\beta}\right)}{\epsilon_c}\right]\hat{\gamma}_{T+1}\\
&+\hat{E}_t\sum_{T=t}^{\infty}\tilde{\beta}^{T-t}\left[\frac{(1-\chi)\left(1-\tilde{\beta}\right)}{\epsilon_c}-\tilde{\beta}\sigma^{-1}\right]\tilde{\beta}\tilde{R}\hat{R}_{T+1}^K\\
&+\hat{E}_t\sum_{T=t}^{\infty}\tilde{\beta}^{T-t}\frac{(1-\chi)\left(1-\tilde{\beta}\right)}{\epsilon_c}\tilde{\beta}\,(\epsilon_w+\epsilon_c\chi)\,\hat{w}_{T+1}.
\end{aligned}$$

Setting $\sigma = 1$ and $\chi = 0$ we get back to the simple real busisiness cycle model.

A.4 Steady State

From the Euler equation we get

$$\begin{aligned}
\frac{\bar{R}^k\bar{U}}{\bar{\gamma}} &= \bar{\gamma}^{\sigma-1}\beta^{-1}-\frac{1-\delta}{\bar{\gamma}}\\
&= \tilde{\beta}^{-1}-\frac{1-\delta}{\bar{\gamma}}
\end{aligned}$$

and from the capacity utilization first-order condition

$$\frac{\bar{R}^k\bar{U}}{\bar{\gamma}} = \frac{\delta\theta}{\bar{\gamma}} \Longrightarrow \theta = \frac{\bar{R}^k\bar{U}}{\delta}$$

which defines θ, allowing to determine U and therefore R^k. The ratios

$$\frac{\bar{y}}{\bar{k}} = (\alpha)^{-1} \frac{\bar{R}^k \bar{U}}{\bar{\gamma}}; \frac{\bar{\imath}}{\bar{k}} = 1 - \frac{1-\delta}{\bar{\gamma}}; \frac{\bar{c}}{\bar{k}} = \frac{\bar{y}}{\bar{k}} - \frac{\bar{\imath}}{\bar{k}} \text{ and } \frac{\bar{c}}{\bar{y}} = \frac{\bar{c}}{\bar{k}} / \frac{\bar{y}}{\bar{k}}.$$

Finally the steady state level ψ, for a given choice of $\bar{H}$, is determined by

$$\begin{aligned} \psi &= \frac{\bar{H} v'\left(\bar{H}\right)}{v\left(\bar{H}\right)} (\sigma - 1)^{-1} = \frac{\bar{w}\bar{H}}{\bar{k}} \frac{\bar{k}}{\bar{c}} \\ &= \frac{1-\alpha}{\alpha} \bar{R} \bar{\gamma}^{\sigma-1} \frac{\bar{k}}{\bar{c}} (\sigma - 1)^{-1} \\ &= \frac{1-\alpha}{\alpha} \tilde{R} \frac{\bar{k}}{\bar{c}} (\sigma - 1)^{-1}. \end{aligned}$$

A.5 The model with costly participation

The preferences described above suffer from the problem that, for a given σ, if the Frisch elasticity of labor supply increases beyond some threshold level, consumption becomes an inferior good. In this section we show how a simple model of costly labor market participation gives a similar labor supply and consumption decision rule. We assume that each 'household' is composed of a continuum of family members. Labor is indivisible: each member of the household decides whether to work a fixed amount of hours or to not participate in the labor market. Participating in the labor market entails a cost. We assume perfect risk sharing within the household.

The maximization problem for the household is

$$E_t \sum_{T=t}^{\infty} \beta^{T-t} u\left(C_T, L_T\right)$$

subject to

$$\begin{aligned} C_t + K_{t+1} + X_t \Phi\left(e_t\right) &= R_t^K \left(u_t K_t\right) + W_t H_t + \left(1 - \delta\left(U_t\right)\right) K_t \\ L_t &= 1 - H_t, \end{aligned}$$

where e_t denotes the fraction of household members that are working. Household consumption is defined as

$$C_t = e_t C_t^e + \left(1 - e_t\right) C_t^u$$

where C_t^e denotes consumption of employed members and C_t^u is consumption of the unemployed. The utility function is defined as

$$u(C_t, L_t) = e_t \frac{(C_t^e)^{1-\sigma} \nu(h)}{1-\sigma} + (1-e_t) \frac{(C_t^u)^{1-\sigma} \nu(0)}{1-\sigma},$$

where

$$L_t = 1 - H_t = 1 - e_t h.$$

Finally, the function $\Phi(e_t)$ denotes the cost attached to labor market participation and has the following properties

$$\Phi_e(e_t) > 0, \quad \Phi_{ee}(\bar{e}) > 0.$$

Total household consumption satisfies a standard Euler equation, while the first order conditions for the employed and unemployed are

$$(C_t^e)^{-\sigma} \nu(h) = \Lambda_t \tag{29}$$

$$(C_t^u)^{-\sigma} \nu(0) = \Lambda_t \tag{30}$$

implying the risk-sharing condition

$$\frac{C_t^e}{C_t^u} = \left[\frac{\nu(h)}{\nu(0)}\right]^{\frac{1}{\sigma}}.$$

Employed household members enjoy greater consumption in compensation of work effort.

The first order condition with respect to participation gives

$$\frac{1}{1-\sigma}\left[-(C_t^e)^{1-\sigma} \nu(h) + (C_t^u)^{1-\sigma} \nu(0)\right] = \Lambda_t \left[W_t h - C_t^e + C_t^u - \Phi_{e,t}\right]$$

which, rearranging, becomes

$$\frac{\sigma}{\sigma-1}(C_t^e - C_t^u) = W_t h - \Phi_{e,t}.$$

Expressing the variables in stationary levels and log-linearizating provides

$$\begin{aligned} \hat{c}_t^e &= \hat{c}_t^u \\ \hat{c}_t &= \frac{\bar{e}(\bar{c}^e - \bar{c}^u)}{\bar{c}} \hat{H}_t + \hat{c}_t^e \\ \frac{\sigma}{\sigma-1} \frac{e(\bar{c}^e - \bar{c}^u)}{\bar{c}} \hat{c}_t^e &= \psi \hat{w}_t - \epsilon_\phi \bar{\phi} \hat{H}_t, \end{aligned} \tag{31}$$

where

$$\psi = \frac{\bar{H}\bar{w}}{\bar{c}}, \quad \epsilon_\phi = \frac{\Phi_{ee}\bar{e}}{\Phi_e}, \quad \bar{\phi} = \frac{\Phi_e \bar{e}}{\bar{c}}.$$

In steady state the following holds

$$\frac{\bar{e}\left(\bar{c}^e - \bar{c}^u\right)}{\bar{c}} = \frac{\sigma - 1}{\sigma}\left(\psi - \bar{\phi}\right) > 0.$$

Substituting for $\hat{c}_t^e$ and using the above steady state relation we get

$$\left(\psi - \bar{\phi}\right)\left[\hat{c}_t - \frac{\sigma - 1}{\sigma}\left(\psi - \bar{\phi}\right)\hat{H}_t\right] = \psi\hat{w}_t - \epsilon_\phi\bar{\phi}\hat{H}_t,$$

which gives the following constant-consumption labor supply

$$\left[\frac{\epsilon_\phi\bar{\phi}}{\left(\psi - \bar{\phi}\right)} - \frac{\sigma - 1}{\sigma}\left(\psi - \bar{\phi}\right)\right]\hat{H}_t = -\hat{c}_t + \frac{\psi}{\left(\psi - \bar{\phi}\right)}\hat{w}_t.$$

To derive the optimal decision rule, the flow budget constraint can be expressed in terms of stationary variables as

$$c_t + k_{t+1} + \Phi\left(e_t\right) = \left(\gamma_t\right)^{-1} R_t^K\left(U_t k_t\right) + w_t H_t + \left(1 - \delta\left(U_t\right)\right)\left(\gamma_t\right)^{-1} k_t.$$

Log-linearization gives

$$\bar{\gamma}^{\sigma-1}\beta^{-1}\hat{k}_t = \left\{\begin{array}{c} \frac{\Phi_e\bar{e}}{\bar{c}}\frac{\bar{c}}{\bar{k}}\hat{H}_t + \frac{\bar{c}}{\bar{k}}\hat{c}_t + \hat{k}_{t+1} - \\ \bar{R}\bar{\gamma}^{\sigma-1}\left(\hat{R}_t^K + \hat{U}_t + \frac{1-\alpha}{\alpha}\hat{w}_t + \frac{1-\alpha}{\alpha}\hat{H}_t - \hat{\gamma}_t\right) - \\ \frac{(1-\delta)}{\bar{\gamma}}\left[-\hat{\gamma}_t - \frac{\delta}{(1-\delta)}\theta\hat{U}_t\right] \end{array}\right\},$$

which becomes

$$\hat{k}_t = \tilde{\beta}\left[\frac{\bar{c}}{\bar{k}}\hat{c}_t + \bar{\phi}\frac{\bar{c}}{\bar{k}}\hat{H}_t + \hat{k}_{t+1} + \tilde{\beta}^{-1}\hat{\gamma}_t - \tilde{R}\left(\hat{R}_t^K + \frac{1-\alpha}{\alpha}\hat{w}_t + \frac{1-\alpha}{\alpha}\hat{H}_t\right)\right].$$

Using the for the constant-consumption labor supply

$$\hat{k}_t = \tilde{\beta}\left(\epsilon_c\hat{c}_t + \tilde{\beta}^{-1}\hat{\gamma}_t + \hat{k}_{t+1} - \epsilon_w\hat{w}_t - \tilde{R}\hat{R}_t^K\right)$$

where

$$\begin{aligned} \epsilon_c &= \frac{\bar{c}}{\bar{k}} + \left[\frac{\epsilon_\phi\bar{\phi}}{\left(\psi - \bar{\phi}\right)} - \frac{\sigma - 1}{\sigma}\left(\psi - \bar{\phi}\right)\right]^{-1}\left[\tilde{R}\frac{1-\alpha}{\alpha} - \bar{\phi}\frac{\bar{c}}{\bar{k}}\right] \\ \epsilon_w &= \left(1 + \frac{\psi}{\left(\psi - \bar{\phi}\right)}\left[\frac{\epsilon_\phi\bar{\phi}}{\left(\psi - \bar{\phi}\right)} - \frac{\sigma - 1}{\sigma}\left(\psi - \bar{\phi}\right)\right]^{-1}\right)\tilde{R}\frac{1-\alpha}{\alpha}. \end{aligned}$$

Iterating forward and taking expectations provides

$$\epsilon_c \hat{E}_t \sum_{T=t}^{\infty} \tilde{\beta}^{T-t} \hat{c}_T = \tilde{\beta}^{-1} \hat{k}_t + \hat{E}_t \sum_{T=t}^{\infty} \tilde{\beta}^{T-t} \left(\epsilon_w \hat{w}_T + \tilde{R} \hat{R}_T^K - \tilde{\beta}^{-1} \hat{\gamma}_T \right)$$

the intertemporal budget constraint as of time t. Combining (29) and (31) we obtain

$$\hat{c}_t = \frac{\sigma - 1}{\sigma} \left(\psi - \bar{\phi} \right) \hat{H}_t - \sigma^{-1} \hat{\lambda}_t.$$

Solving the Euler equation backward from time T and taking expectations

$$\begin{aligned} \hat{E}_t \hat{\lambda}_T &= \hat{\lambda}_t + \hat{E}_t \left[\sum_{T=t}^{T-1} \left(\tilde{\beta} \tilde{R} \hat{R}_{T+1}^K - \sigma \hat{\gamma}_{T+1} \right) \right] \\ \hat{E}_t \left(\sigma \hat{c}_T + \left(\psi - \bar{\phi} \right) (1 - \sigma) \hat{H}_T \right) &= \sigma \hat{c}_t + \left(\psi - \bar{\phi} \right) (1 - \sigma) \hat{H}_t + \hat{E}_t \left[\sum_{T=t}^{T-1} \left(\tilde{\beta} \tilde{R} \hat{R}_{T+1}^K - \sigma \hat{\gamma}_{T+1} \right) \right]. \end{aligned}$$

Substituting for the constant-consumption labor supply yields

$$\hat{E}_t \left(\left[(1 - \chi) \sigma \hat{c}_T + \chi \sigma \hat{w}_T \right] \right) = \sigma \hat{c}_t + \psi (1 - \sigma) \hat{H}_t + \hat{E}_t \left[\sum_{T=t}^{T-1} \left(\tilde{\beta} \tilde{R} \hat{R}_{T+1}^K - \sigma \hat{\gamma}_{T+1} \right) \right]$$

by using

$$\chi = \left[\frac{\left(\psi - \bar{\phi} \right)^2 (1 - \sigma)}{\sigma \epsilon_\phi \bar{\phi} + \left(\psi - \bar{\phi} \right)^2 (1 - \sigma)} \right].$$

Rearranging in terms of expected consumption and substituting into the intertemporal budget constraint we get

$$\begin{aligned} & \epsilon_c \hat{E}_t \sum_{T=t}^{\infty} \tilde{\beta}^{T-t} \left[\frac{1}{1-\chi} \left\{ \hat{c}_t + \sigma^{-1} \left(\psi - \bar{\phi} \right) (1 - \sigma) \hat{H}_t + \hat{E}_t \left[\sum_{T=t}^{T-1} \left(\sigma^{-1} \tilde{\beta} \tilde{R} \hat{R}_{T+1}^K - \hat{\gamma}_{T+1} \right) \right] - \chi \hat{w}_T \right\} \right] \\ = & \ \tilde{\beta}^{-1} \hat{k}_t + \hat{E}_t \sum_{T=t}^{\infty} \tilde{\beta}^{T-t} \left(\epsilon_w \hat{w}_T + \tilde{R} \hat{R}_T^K - \tilde{\beta}^{-1} \hat{\gamma}_T \right). \end{aligned}$$

Further simplification leads to

$$\hat{c}_t + \sigma^{-1} \left(\psi - \bar{\phi} \right) (1 - \sigma) \hat{H}_t =$$

$$\frac{(1 - \chi) \left(1 - \tilde{\beta} \right)}{\epsilon_c \tilde{\beta}} \hat{k}_t +$$

$$\hat{E}_t \sum_{T=t}^{\infty} \tilde{\beta}^{T-t} \left[\frac{(1 - \chi) \left(1 - \tilde{\beta} \right)}{\epsilon_c} \left((\epsilon_w + \tilde{\epsilon}_c \chi) \hat{w}_T + \tilde{R} \hat{R}_T^K - \tilde{\beta}^{-1} \hat{\gamma}_T \right) - \tilde{\beta} \left(\sigma^{-1} \tilde{\beta} \tilde{R} \hat{R}_{T+1}^K - \hat{\gamma}_{T+1} \right) \right].$$

Finally, we obtain the consumption decision rule, depending only on forecast of prices that are beyond the control of the household,

$$\begin{aligned}
\hat{c}_t + \sigma^{-1}\left(\psi - \bar{\phi}\right)(1-\sigma)\hat{H}_t \quad = \quad & \frac{(1-\chi)\left(1-\tilde{\beta}\right)}{\epsilon_c}\left[\tilde{\beta}^{-1}\hat{k}_t + \tilde{R}\hat{R}_t^K - \tilde{\beta}^{-1}\hat{\gamma}_t + (\epsilon_w + \epsilon_c\chi)\hat{w}_t\right] \\
& + \hat{E}_t \sum_{T=t}^{\infty} \tilde{\beta}^{T-t}\left[\tilde{\beta} - \frac{(1-\chi)\left(1-\tilde{\beta}\right)}{\epsilon_c}\right]\hat{\gamma}_{T+1} \\
& + \hat{E}_t \sum_{T=t}^{\infty} \tilde{\beta}^{T-t}\left[\frac{(1-\chi)\left(1-\tilde{\beta}\right)}{\epsilon_c} - \tilde{\beta}\sigma^{-1}\right]\tilde{\beta}\tilde{R}\hat{R}_{T+1}^K \\
& + \hat{E}_t \sum_{T=t}^{\infty} \tilde{\beta}^{T-t}\frac{(1-\chi)\left(1-\tilde{\beta}\right)}{\epsilon_c}\tilde{\beta}(\epsilon_w + \epsilon_c\chi)\hat{w}_{T+1}.
\end{aligned}$$

Setting $\sigma = 1$ and $\chi = 0$ ($\sigma = 1$) we get back to the simple RBC model. For low values of $\bar{\phi}$ (the cost of participating) the decision rule approximates our representative agent model. To give the intuition of the equivalence between the representative agent and the model with costly participation consider an increase in the representative agent's income that leaves unchanged the price of capital and the price of labor. Under the current calibration the agent's preferences imply that consumption is an inferior good. Therefore with higher income the agent decreases consumption and increases leisure. In the case of a "family" with costly labor market participation, the positive income transfer induces a higher fraction of family members to exit the labor market and consume leisure, while the family members that are still working do not decrease their consumption. However aggregate consumption decreases because of a composition effect. As shown above, family members that do not work are allocated lower consumption. Since their number increases, aggregate consumption also decreases. Also notice that assuming $\sigma > 1$, by letting $\Phi_{ee}, \Phi_e \rightarrow 0$ the model becomes Rogerson's lottery model with non-separable preferences described in King and Rebelo (1999).

A.6 Constant gain learning and the Kalman filter

Agents update their beliefs using the following constant-gain algorithm

$$\hat{\omega}_t = \hat{\omega}_{t-1} + gR_t^{-1}q_{t-1}\left(\tilde{z}_t - \hat{\omega}_{t-1}'q_{t-1}\right) \tag{32}$$

$$R_t = R_{t-1} + g\left(q_{t-1}q'_{t-1} - R_{t-1}\right)$$

where we now assume for simplicity that $\tilde{z}_t = T(\hat{\omega}_{t-1})q_{t-1}$ is one-dimensional (for example k_{t+1}) and q_t is a *two*-dimensional vector. Following Evans and Honkapohja (2001) and Sargent and Williams (2005), the limiting behavior of the estimates are approximated by the following system of ordinary differential equations[24]

$$\begin{aligned}\dot{\hat{\omega}} &= R^{-1}M_q\left(\hat{\omega}\right)\left[T\left(\hat{\omega}\right) - \hat{\omega}\right] \\ \dot{R} &= M_q\left(\hat{\omega}\right) - R\end{aligned}$$

where $M_q\left(\hat{\omega}\right) = E\left(q_{t-1}q'_{t-1}\right)$.[25] Asymptotically R converges to $M_q\left(\hat{\omega}\right)$.

As a way to justify and interpret the use of constant-gain algorithms it is common to relate them to the Kalman filter. Assume agents believe that the data generating process is the following random walk model of coefficient variation

$$\tilde{z}_t = \omega'_{t-1}q_{t-1} + \tilde{e}^z_t$$

$$\omega_t = \omega_{t-1} + \tilde{e}^\omega_t$$

where for simplicity we assume that $\tilde{z}_t$ is one-dimensional and q_t is a n-dimensional vector. The shock $\tilde{e}^z_t$ has standard deviation σ_A and variance-covariance matrix of $\tilde{e}^\omega_t$ is assumed to be $\Sigma^\omega << \sigma^2_z I$. The matrix Σ^ω defines agents' prior about the variance in the coefficients' drift. Let $\hat{\omega}_{t|t-1}$ the optimal estimate of ω_t conditional on information up to date $t-1$. This is obtained from the following Kalman filtering equations

$$\hat{\omega}_{t+1|t} = \hat{\omega}_{t|t-1} + \frac{P_t q_{t-1}}{1 + q'_{t-1}P_t q_{t-1}}\left(\tilde{z}_t - \hat{\omega}'_{t|t-1}q_{t-1}\right) \tag{33}$$

$$P_{t+1} = P_t - \frac{P_t q_{t-1}q'_{t-1}P_t}{1 + q'_{t-1}P_t q_{t-1}} + \frac{1}{\sigma^2_z}\Sigma^\omega \tag{34}$$

where we use

$$E\left[\left(\omega_t - \hat{\omega}_{t|t-1}\right)\left(\omega_t - \hat{\omega}_{t|t-1}\right)'\right] = \sigma^2_z P_t.$$

[24] This ODE system is called the mean dynamics of the estimates. Sargent and Williams (2005) investigate a second ODE system which describe the escape dynamics, which are not the focus of this paper. In the simulations conducted with the calibrated model we did not observe escape dynamics.

[25] The unconditional expectation has finite value if the system is E-Stable. See also Evans and Honkapohja (2001).

Sargent and Williams (2005) propose the following approximation to the filtering equations. For large t, (34) can be approximated by

$$P_{t+1} = P_t - P_t M_q\left(\hat{\omega}\right) P_t + \frac{1}{\sigma_z^2}\Sigma^{\omega}.$$

Further assuming that $1/\left(1 + q_{t-1}' P_t q_{t-1}\right) \approx 1$, the filtering equations can be re-written as

$$\begin{aligned} \hat{\omega}_{t+1|t} &= \hat{\omega}_{t|t-1} + P_t q_{t-1}\left(\tilde{z}_t - \hat{\omega}_{t|t-1}' q_{t-1}\right) \\ P_{t+1} &= P_t - P_t M_q\left(\hat{\omega}\right) P_t + \frac{1}{\sigma_z^2}\Sigma^{\omega}. \end{aligned} \tag{35}$$

The asymptotic behavior of (35) can be shown to be equivalent to the asymptotic behavior of constant-gain least squares, provided agents' priors on Σ^{ω} satisfy

$$\Sigma^{\omega} = g^2 \sigma_z^2 M_q\left(\hat{\omega}\right)^{-1}. \tag{36}$$

To show this, the matrix P converges asymptotically to a unique positive definite matrix which solves the Riccati equation

$$P M_q\left(\hat{\omega}\right) P = \frac{1}{\sigma_z^2}\Sigma^{\omega}.$$

Using (36) the solution becomes $P = g M_q\left(\hat{\omega}\right)^{-1}$. Hence, in large samples, P_t converges to P and R_t converges to $M_q\left(\hat{\omega}\right)$, implying that the constant gain algorithm and the Kalman filter have the same asymptotic behavior. As shown in Sargent and Williams (2005), the two algorithms share the same asymptotic behavior in large samples but their transitional dynamics display differences in small samples. In this paper, we analyze the dynamics of agents' beliefs at their stationary distribution, and therefore evaluate the learning algorithm in large samples.

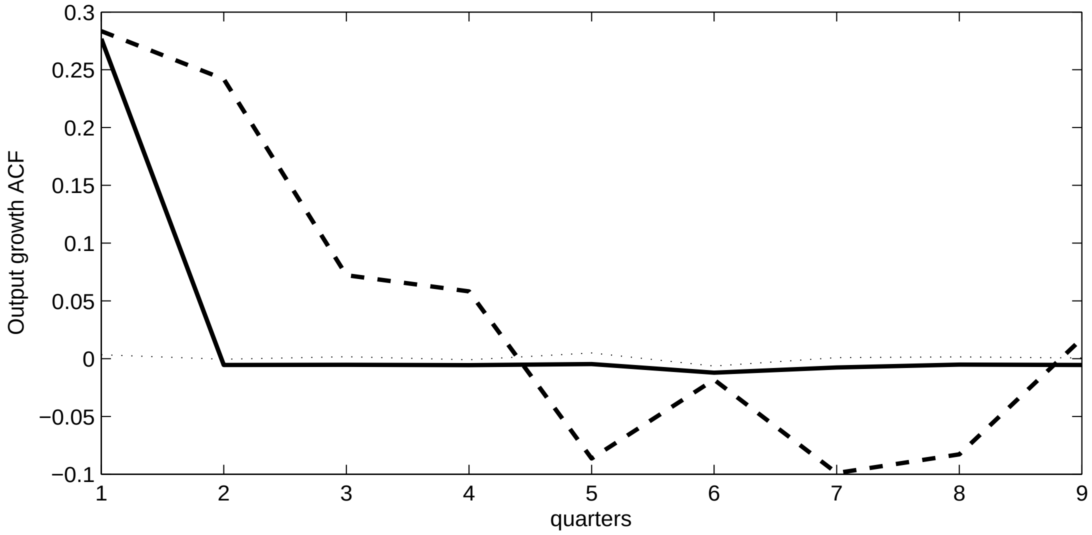

Figure 1: Output autocorrelation function. The thick dashed line denotes US data, the thick solid line denoted the model with learning, while the dotted line denotes the model under rational expectations.

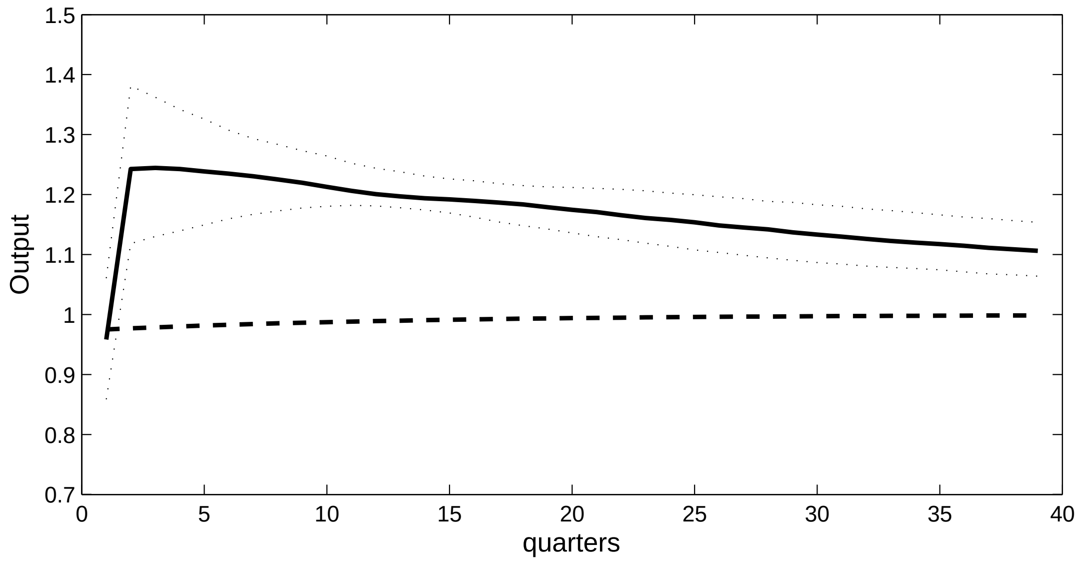

Figure 2: Dotted lines denote the 75% bands, solid line denotes the median impulse response under learning. The dashed line denotes the impulse response under rational expectations.

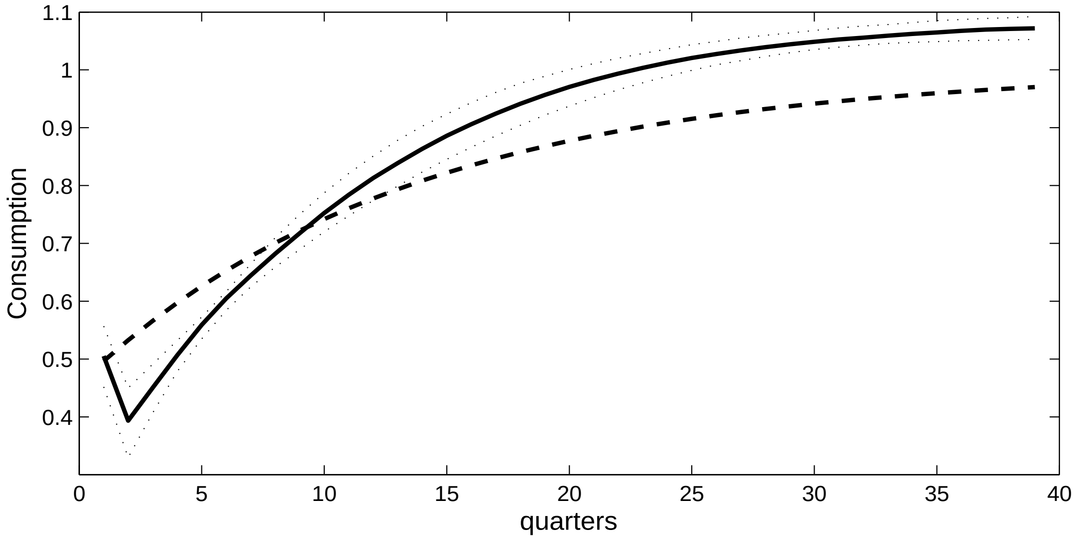

Figure 3: Dotted lines denote the 75% bands, solid line denotes the median impulse response under learning. The dashed line denotes the impulse response under rational expectations.

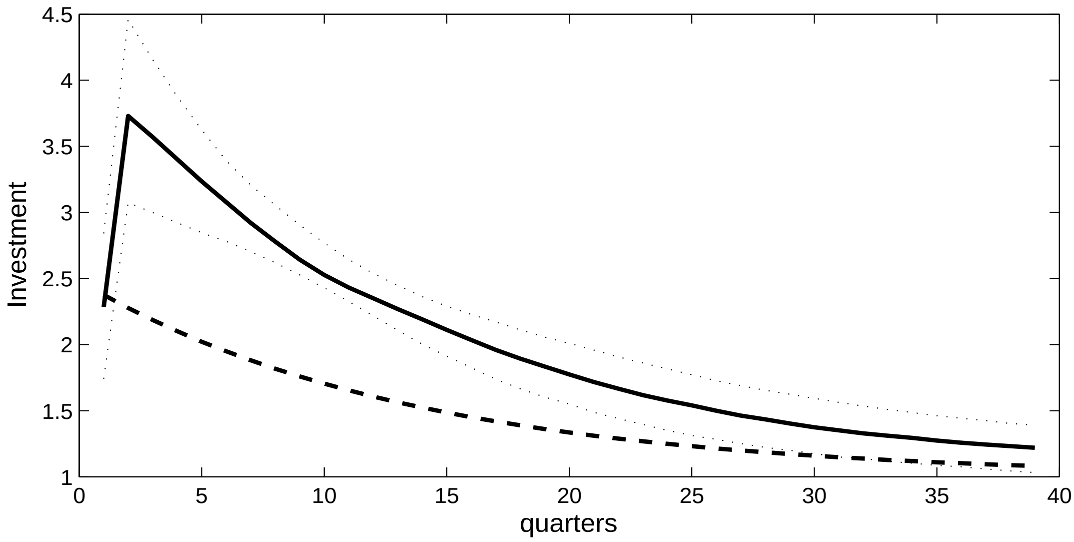

Figure 4: Dotted lines denote the 75% bands, solid line denotes the median impulse response under learning. The dashed line denotes the impulse response under rational expectations.

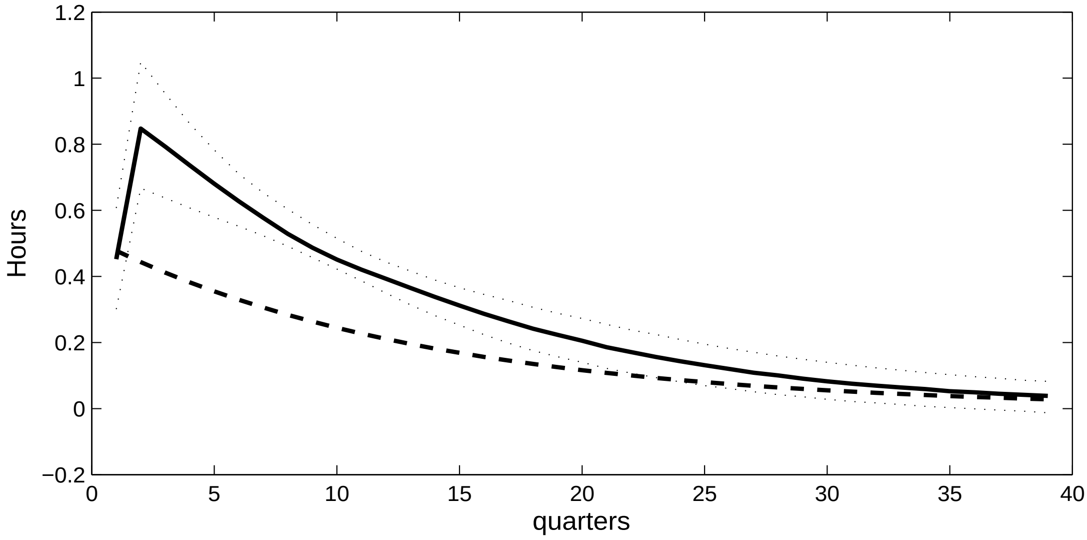

Figure 5: Dotted lines denote the 75% bands, solid line denotes the median impulse response under learning. The dashed line denotes the impulse response under rational expectations.

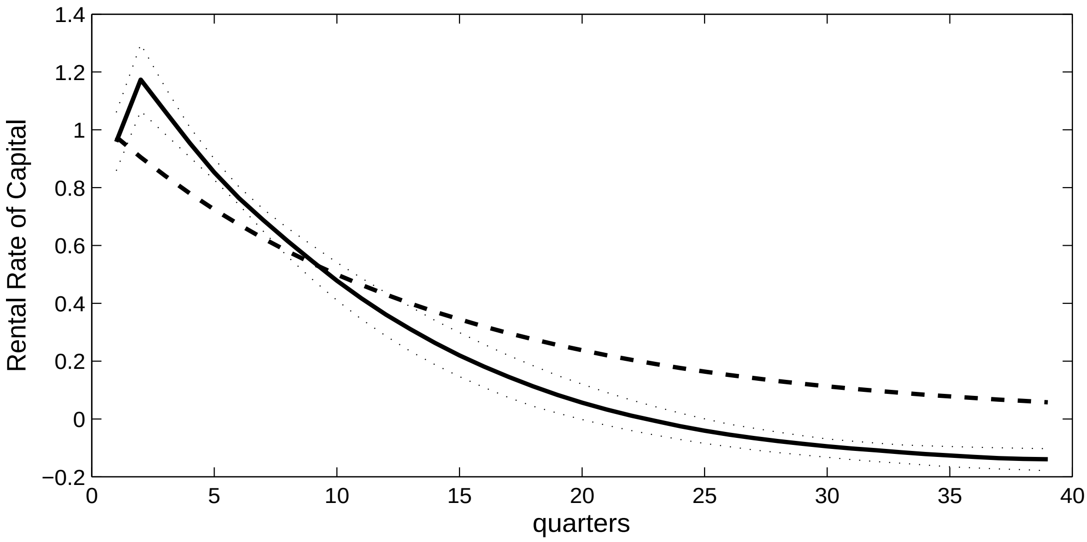

Figure 6: Dotted lines denote the 75% bands, solid line denotes the median impulse response under learning. The dashed line denotes the impulse response under rational expectations.

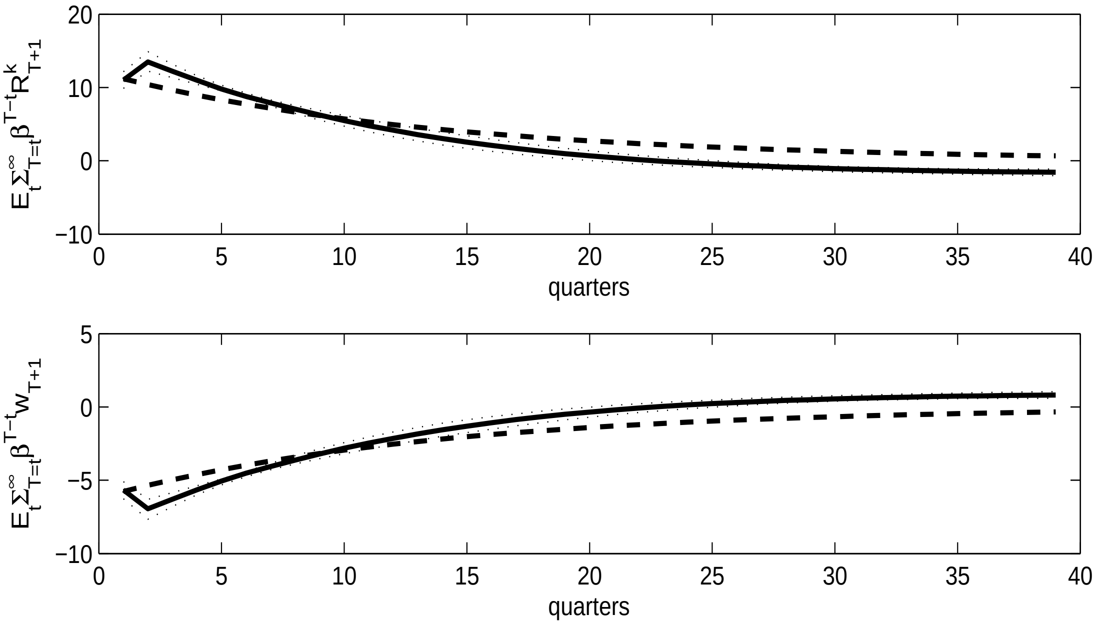

Figure 7: Dotted lines denote the 75% bands, solid line denotes the median impulse response under learning. The dashed line denotes the impulse response under rational expectations. The top panel is the present discounted value of returns to capital and the bottom panel the corresponding value for wages.

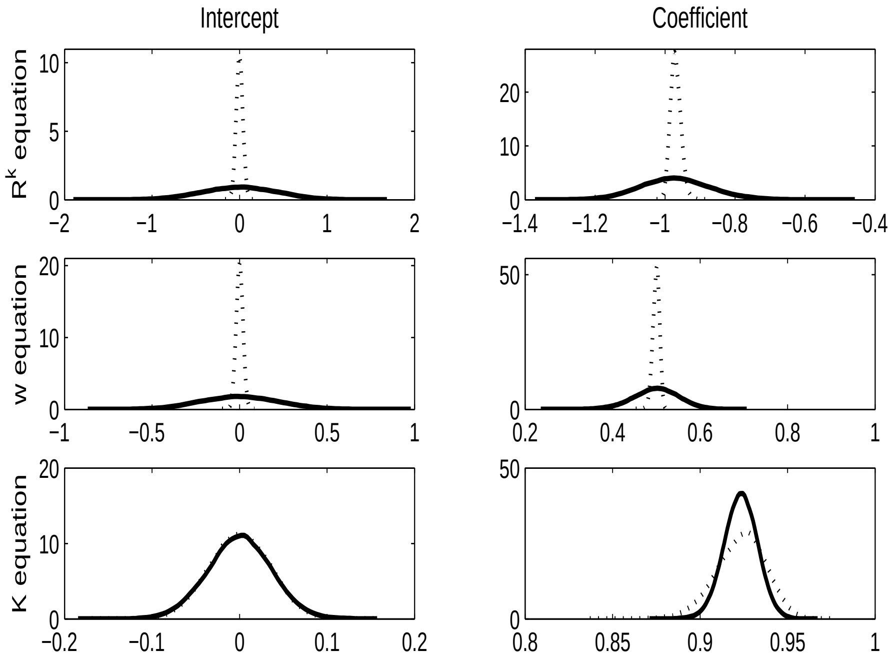

Figure 8: Solid line: model with feedback. Dotted line, model without feedback.

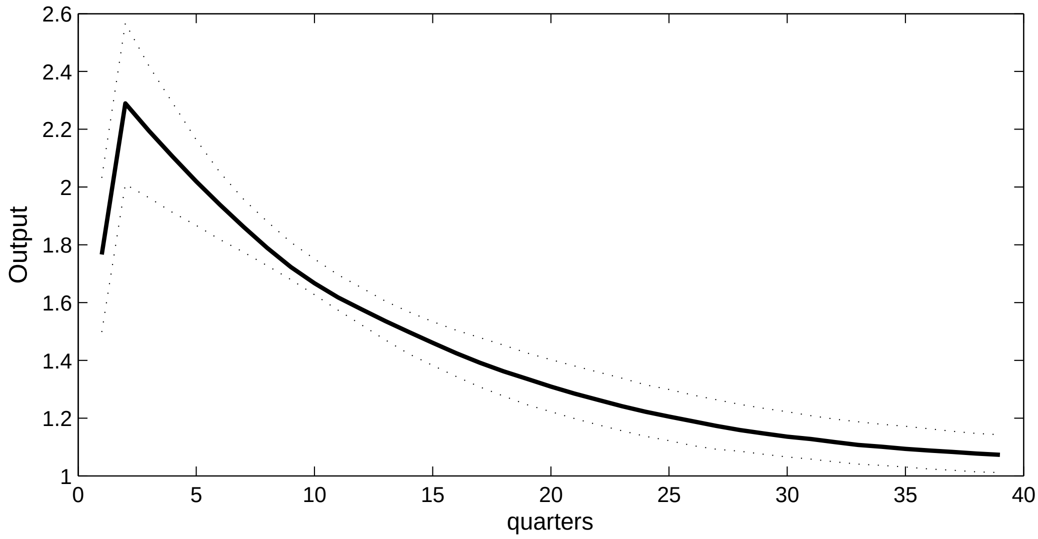

Figure 9: Output dynamics in response to a technology shock with increasing returns and non-separable preferences. The thick solid line denotes the model with learning, with the dotted lines showing the 12.5 and 87.5 percentiles.

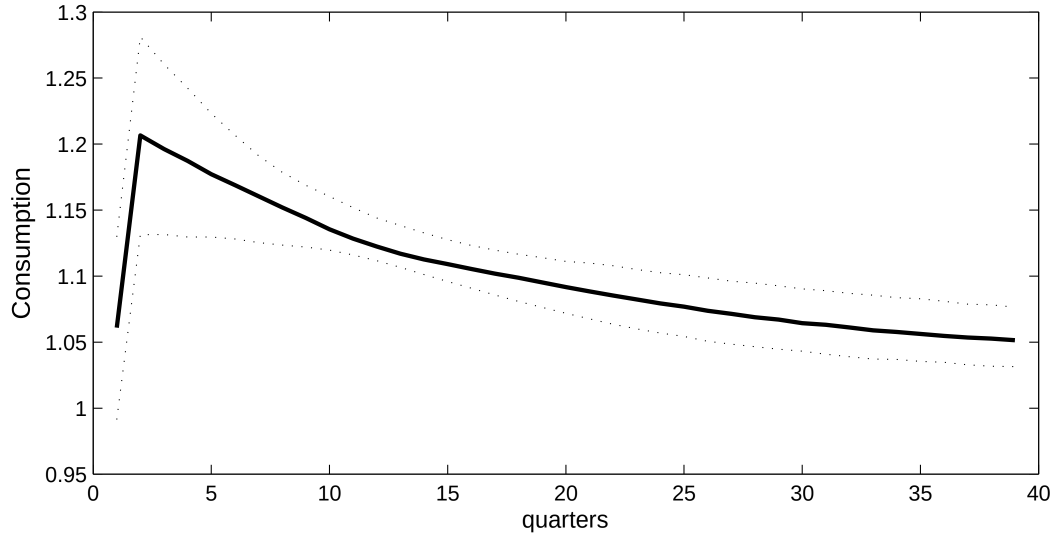

Figure 10: Consumption dynamics in response to a technology shock with increasing returns and non-separable preferences. The thick solid line denotes the model with learning, with the dotted lines showing the 12.5 and 87.5 percentiles.

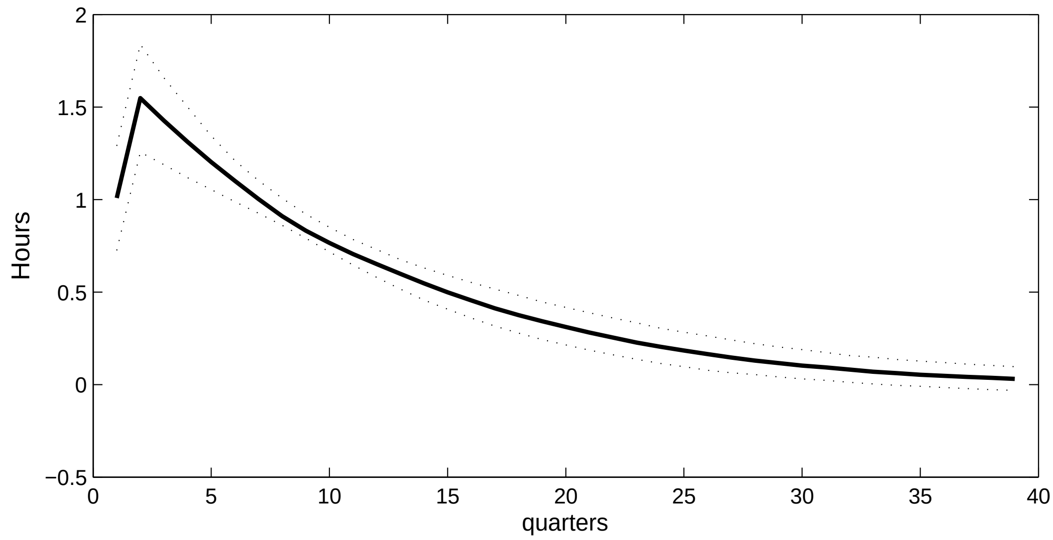

Figure 11: Hours dynamics in response to a technology shock with increasing returns and non-separable preferences. The thick solid line denotes the model with learning, with the dotted lines showing the 12.5 and 87.5 percentiles.

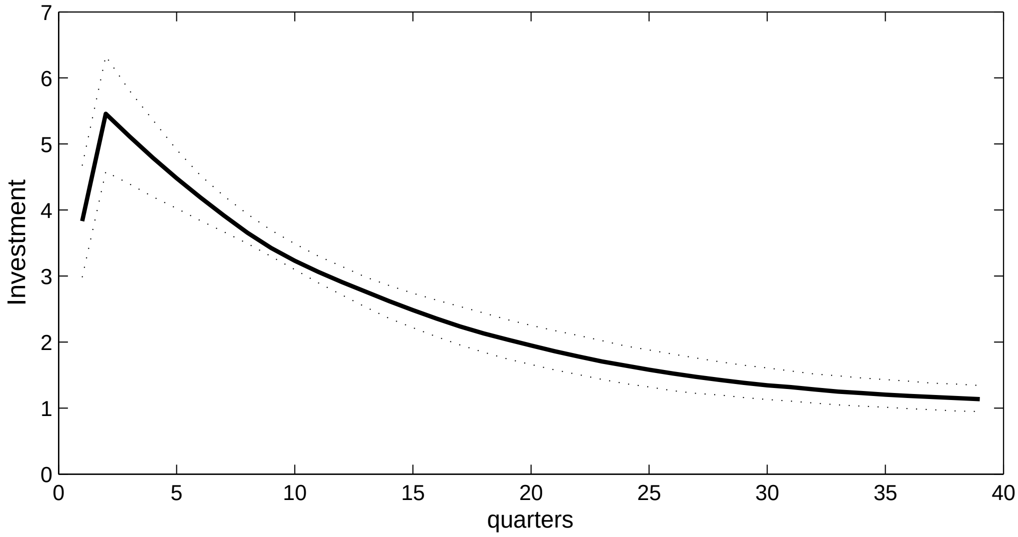

Figure 12: Investment dynamics in response to a technology shock with increasing returns and non-separable preferences. The thick solid line denotes the model with learning, with the dotted lines showing the 12.5 and 87.5 percentiles.

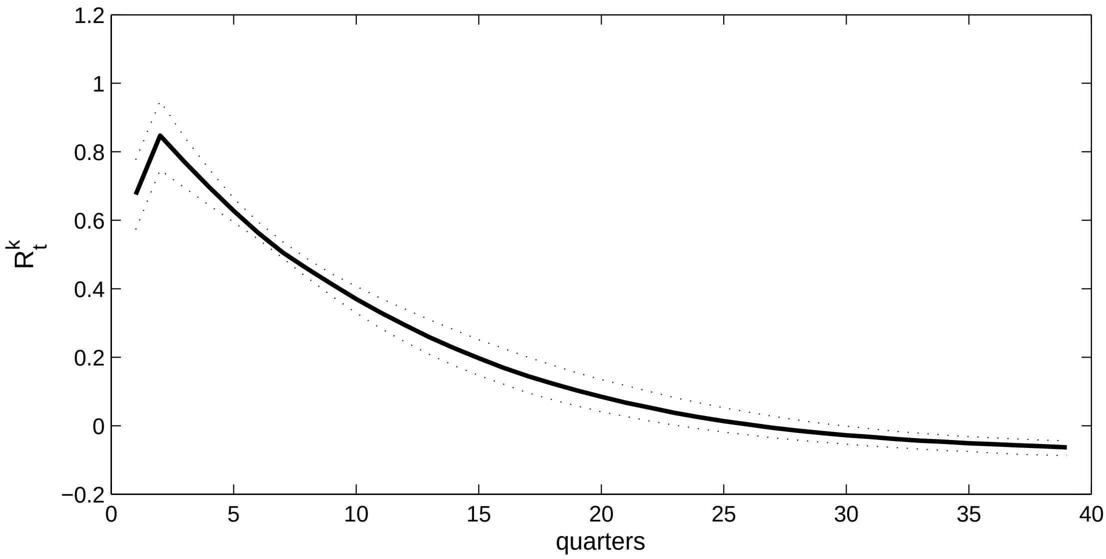

Figure 13: Rental rate dynamics in response to a technology shock with increasing returns and non-separable preferences. The thick solid line denotes the model with learning, with the dotted lines showing the 12.5 and 87.5 percentiles.