THE EFFECTS OF NEWS SHOCKS AND BOUNDED RATIONALITY ON MACROECONOMIC VOLATILITY

by

BRIAN JOSEPH DOMBECK

A DISSERTATION

Presented to the Department of Economics
and the Graduate School of the University of Oregon
in partial fulfillment of the requirements
for the degree of
Doctor of Philosophy

June 2017

DISSERTATION APPROVAL PAGE

Student: Brian Joseph Dombeck

Title: The Effects of News Shocks and Bounded Rationality on Macroeconomic Volatility

This dissertation has been accepted and approved in partial fulfillment of the requirements for the Doctor of Philosophy degree in the Department of Economics by:

Bruce McGough	Co-Chair
George Evans	Co-chair
Jeremy Piger	Core Member
Arafaat Valiani	Institutional Representative

and

Scott L. Pratt	Dean of the Graduate School

Original approval signatures are on file with the University of Oregon Graduate School.

Degree awarded June 2017

DISSERTATION ABSTRACT

Brian Joseph Dombeck

Doctor of Philosophy

Department of Economics

June 2017

Title: The Effects of News Shocks and Bounded Rationality on Macroeconomic Volatility

This dissertation studies the impact embedding boundedly rational agents in real business cycle-type news-shock models may have on a variety of model predictions, from simulated moments to structural parameter estimates. In particular, I analyze the qualitative and quantitative effects of assuming agents are boundedly rational in a class of DSGE models which attempt to explain the observed volatility and comovements in key aggregate measures of U.S. economic performance as the result of endogenous responses to information in the form of "news shocks". The first chapter explores the theoretical feasibility of relaxing the rational expectations hypothesis in a three-sector real business cycle (RBC) model which generates boom-bust cycles as a result of periods of optimism and pessimism on the part of households. The second chapter determines whether agents forming linear forecasts of shadow prices in a nonlinear framework can lead to behavior approximately consistent with fully informed individuals in a one-sector real business cycle model. The third chapter analyzes whether empirical estimates of the relative importance of anticipated shocks may be biased by assuming rational expectations.

By merging the two hitherto separate but complementary strands of literature related to bounded rationality and news shocks I am able to conduct in-depth analysis of the importance of both the information agents have and what they choose to do with it. At its core, the study of news in macroeconomics is a study of the specific role alternative

information sets play in generating macroeconomic volatility. Adaptive learning on the other hand is concerned with the behavior of agents given an information set. Taken together, these fields jointly describe the input and the “black box” which produce model predictions from DSGE models. While previous research has been conducted on the effects of bounded rationality or news shocks in isolation, this dissertation marks the first set of research explicitly focused on the interaction of these two model features.

CURRICULUM VITAE

NAME OF AUTHOR: Brian Joseph Dombeck

GRADUATE AND UNDERGRADUATE SCHOOLS ATTENDED:

University of Oregon, Eugene, OR
Louisiana State University, Baton Rouge, LA

DEGREES AWARDED:

Doctor of Philosophy, Economics, 2017, University of Oregon
Master of Science, Economics, 2013, University of Oregon
Bachelor of Science, Economics, 20012, Louisiana State University

AREAS OF SPECIAL INTEREST:

Applied Structural Macroeconomics
Bounded Rationality
Information Choice

GRANTS, AWARDS AND HONORS:

Kleinsorge Summer Research Fellowship, University of Oregon, 2016
Kleinsorge Summer Research Fellowship, University of Oregon, 2014
Kleinsorge First Year Fellowship, University of Oregon, 2012
Graduate Teaching Fellowship, University of Oregon, 2012-2017

ACKNOWLEDGEMENTS

The completion of this dissertation would not have been possible without the support of many groups of people. I thank Professors Bruce McGough, George Evans, Jeremy Piger, Van Kolpin, Anne van den Nouweland , Kaj Gittings, Bulent Unel, Naci Mocan, Carter Hill, and Charles Roussel for all of their guidance and mentorship, and for teaching me to think like an economist. I thank Savannah Dombeck, Kellie Geldreich, and the rest of the University of Oregon Department of Economics staff for their assistance in navigating the myriad of requirements necessary for completing the Doctoral program. Finally, I thank my family for their love, patience, and understanding over the years.

For Savannah

TABLE OF CONTENTS

LIST OF FIGURES

Figure | Page

LIST OF TABLES

Table | Page

CHAPTER I

INTRODUCTION

A primary objective of applied structural macroeconomics is to identify the determinants of business cycles and create empirically relevant dynamic stochastic general equilibrium (DSGE) models. While most modern DSGE models assume agents have rational expectations and restrict attention to the effect of unanticipated shocks as drivers of the business cycle, my dissertation studies the theoretic and empirical effects of embedding boundedly rational agents in real business cycle-type news-shock models to determine whether and when these assumptions influence the predictions generated by our models.

News-shock models seek to generate positive comovement between consumption, investment, total labor supply, and output - which is consistent with postwar U.S. macroeconomic data - in response to news about the future state of the economy. In Chapter II I examine whether such models are compatible with a relaxation of the RE assumption in the context of the three-sector RBC model introduced in Beaudry and Portier (2004) by endowing agents with a perceived law of motion (PLM) for the evolution of the economy. Given their beliefs about the economy's transitional path, agents take actions which result in an actual law of motion (ALM) for the economy. Forecasting errors as captured by differences between agents' PLM and the ALM are then used to update their forecasting model via recursive least squares. This process of adaptive learning about the laws of motion for the economy plays out each period. If agents acting in this way eventually learn to behave in the same way as fully rational agents, the rational expectations equilibrium (REE) is said to be expectationally stable (E-stable). I find that the RE solution for this specific news shock model is E-Stable, and that this finding is robust to a variety of alternative parameterizations and timing assumptions. Furthermore I show that under certain assumptions regarding agents' information sets, if the REE is E-stable in any model without news shocks it will remain E-stable when news shocks are included. This implies the effects of news shocks on model predictions can be safely considered in a wide variety of existing models whose REE have been shown to be E-stable.

In Chapter III I explore the qualitative and quantitative effects of allowing agents to be boundedly optimal in the sense of Evans and McGough (2015) in the well-known news-shock model of Jaimovich and Rebelo (2009). Boundedly optimal agents are similar to their boundedly rational counterparts in that they are not endowed with knowledge of the conditional distribution of all variables in the model. However, instead of merely learning about the laws of motion for endogenous variables they are also endowed with forecasting models for the shadow prices of endogenous state variables. Forecasts of shadow prices are updated recursively which gives rise to a rich set of behavior as agents incorporate news shocks into their forecasts. By embedding this behavioral process in a nonlinear DSGE model, I am able to show through simulations that agents' behavior will converge close to that of fully rational agents, a remarkable feat considering how simplistic the process is. Furthermore I compare actual US business cycle statistics with those obtained by calibrating and simulating the model with shadow-price learning agents and find that bounded optimality has a significant effect on the properties of simulated data as compared to that generated under RE.

Finally, in Chapter IV I utilize Bayesian estimation techniques to determine whether the estimated relative importance of news in the news-shock model of Schmitt-Grohe and Uribe (2012) is robust to the inclusion of adaptive learning. News shocks and adaptive learning both act as expectations-based sources of business cycle activity in macroeconomic dynamic stochastic general equilibrium (DSGE) models. This chapter explores whether these sources act as complements or substitutes to each other by comparing estimates of the relative importance of news shocks obtained across expectation formation mechanisms. I find the relative importance of news shocks is amplified by up to 35% for key macroeconomic variables under learning as compared to rational expectations, implying that news shocks may be more important for driving business cycles than existing estimates would otherwise suggest. The chapters primary contributions are to clarify which assumptions the importance of news is robust to and to produce a new estimate of the constant-gain learning parameter from a novel joint news-and-learning estimation routine.

CHAPTER II

EXPECTATIONAL STABILITY AND THE "COMOVEMENT PROBLEM"

Introduction

Starting with Kydland and Prescott (1982) the economic literature seeking to explain stylized facts and comovements observed in postwar macroeconomic U.S. data through supply-side innovations has flourished. So-called real business cycle (RBC) theory maintains innovations to productivity are responsible for the booms and busts which comprise the business cycle. This implies a peculiar interpretation of cyclical activity: while expansions are driven by technological progress, recessions must be caused by technological regress.[1] In recent years a subset of RBC literature has begun exploring alternate mechanisms through which the economy might experience business cycles, focusing in particular on the potential role of information and expectations. For example in Beaudry and Portier (2004), the model of which is the focus for this chapter, news about the future generates expectationally driven business cycles of the sort described in Pigou (1927) in which the economy experiences boom-bust periods resulting from the actions of forward-looking agents responding to noisy signals of future technological innovation.

As noted in Krusell and McKay (2010), business cycles are characterized by positive comovement in aggregate consumption, investment, employment, and output. While standard RBC models generate positive pro-cyclical comovements in key variables in

[1] Stiglitz (2014) points out that "... by implication, the Great Depression was marked by an episode of acute amnesia, where in large parts of the world, people got less productive!"

response to contemporaneous productivity shocks, they are unable to generate these positive comovements in response to anticipated productivity shocks. Such models typically imply news about future total factor productivity (TFP) will cause consumption to move opposite investment, employment, and output. For example, the baseline calibration of Kydland and Prescott's original model suggests households will substitute investment for consumption in anticipation of higher future marginal factor productivity stemming from anticipated future technological growth. Because leisure is a normal good, households also reducing employment, causing a drop in the level of output. Thus, good news about the future state of the economy causes a recession today.

Several dynamic stochastic general equilibrium (DSGE) models have arisen in an effort to fix this "comovement problem".[2]Beaudry and Portier (2004) show that standard one and two sector RBC models are incapable of generating qualitatively realistic expectationally driven business cycles, that is, business cycles featuring positive comovement amongst key macroeconomic variables in response to news about the future. They propose a three-sector RBC model which utilizes a CES production function for consumption and short-term substitutability constraints on consumption and investment. Jaimovich and Rebelo (2009) augment the canonical RBC model to include variable capital utilization, investment adjustment costs, and a novel preference specification which allows the modeler to manipulate the strength of the wealth effect on labor supply. Lorenzoni (2009) studies a version of the Phelps-Lucas "island model" capable of generating empirically realistic comovements in response to anticipated demand shocks where firms receive a noisy public signal of future aggregate productivity and a private signal of their own productivity. Krusell and McKay (2010) modify a version of

[2]DSGE models are fully microfounded models of agent-level behavior in an economy. In addition to describing the decision making and expectation formation processes for households, firms, governments, central banks, and other economic agents.

the Diamond-Mortensen-Pissarides search-and-matching model to allow the number of firms, which is endogenous and interpreted as investment, to react to news about future productivity. Guo et al. (2015) show variable capital utilization and increasing returns to scale production are sufficient modifications to allow a standard RBC model to produce qualitatively and quantitatively realistic expectationally driven business cycles in response to news about future aggregate demand shocks.

Each of these approaches to fixing the comovement problem do so by changing the structure of the economy considered. But these departures from canon have not been accompanied by an analysis of whether the resulting rational expectations equilibrium (REE) are expectationally stable (E-stable); that is, whether the REE is "learnable". E-stability of an REE can be seen as a measure of the model's reliance on the rational expectations hypothesis. This paper proceeds by analyzing the expectational stability properties of the (unique) RE solution to the three-sector RBC model presented in Beaudry and Portier (2004). The goal is to discover whether news shock models can deliever REE which are simultaneously E-stable and capable of generating qualitatively realistic expectationally driven business cycles. My main results suggest that E-stability properties are robust to the inclusion of news shocks.

To this point the news-shock literature has exclusively relied on the assumption championed by Muth (1961) that agents have model consistent or rational expectations (RE). In other words, agents are perfectly informed as to the precise laws of motion governing the evolution of the economy. Critically, this implies their forecasts will be correct on average. This assumption elegantly addresses the critique of Lucas (1976) that changes in policy may naturally result in changes to the behavior of agents. While Lucas and Muth both carefully pointed out RE is not meant to be taken as a realistic description of the way in which individual agents behave, it is natural to ask whether weaker assumptions about the capabilities of agents could result in similar model predictions.[3] In particular, it is worth asking if an alternative expectation formation mechanism can be both a behavioral description and a useful tool for analyzing economic models.

One attractive alternative proposed in Marcet and Sargent (1989a,b,c) is to assume agents are engaged in "adaptive learning". The literature on boundedly rational agents[4] supposes agents are endowed with a perceived law of motion(PLM) about the true structure of the economy, but are unsure of the precise coefficients governing the laws of motion. Regarding their PLM as true the agents make optimal choices given their information set. Since the actions of agents today influence the state of the economy tomorrow, this gives rise to the actual law of motion (ALM) for the economy. Agents behave as econometricians and update their estimates of the laws of motion using a learning algorithm based on the observed forecast errors each period. A REE is said to be E-Stable if the PLM converges (asymptotically) to the REE, that is, if agents can eventually learn the true dynamics of the economy. Learning thus provides one way of checking the sensitivity of models to the standard rational expectations assumption: if a

---

[3]Specifically, Muth (1961) states

> "[Rational Expectations] *does not* assert that the scratch work of entrepreneurs resembles the system of equations in any way; nor does it state that predictions of entrepreneurs are perfect or that their expectations are all the same"

while Lucas (1978) states

> "...[Rational Expectations] does not describe the way agents think about their environment, how they learn, process information, and so forth. It is rather a property likely to be (approximately) possessed by the outcome of this unspecified process of learning and adapting."

[4]See Evans and Honkapohja (2001) for a comprehensive guide to the theory.

particular REE is not E-stable, then the predictions of the model may be perceived as less robust than a similar model featuring a learnable REE.

Given the rich set of behavior observed in models which contemplate changes to information sets and/or expectations formation mechanisms, it is natural to check whether the innovations necessary to generate expectationally driven business cycles are consistent with relaxing the assumption of rational expectations. This is particularly important in the context of new-shock models, as stability results have been shown to be quite sensitive to specific assumptions regarding the information sets of agents. For example, Bullard and Eusepi (2014) show that excluding the value of contemporaneous endogenous variables from agents' information sets breaks the link between determinacy and E-stability found in McCallum (2007). Since by construction news-shock models consider novel assumptions regarding agents' information sets while simultaneously altering modifying well-known DSGE models, an exploration of the E-stability properties in this class of models appears called for.

I begin by describing news shocks and how they are implemented into macroeconomic DSGE models. Next I introduce adaptive learning and demonstrate the procedure for determining whether a given REE is E-stable in the presence of news shocks. Finally, I apply the learning analysis to the 3-sector RBC model of Beaudry and Portier (2004). The chapter concludes with a brief discussion of the main results.

Consider an economy with temporary equilibria described by the stationary system of non-linear expectational difference equations

$$\hat{E}_t \left[f\left(X_{t+1}, X_t, X_{t-1}, \nu_t \right) \right] = 0$$

where X a vector of endogenous and exogenous variables and ν a vector of anticipated and unanticipated exogenous white-noise shocks. While non-linear solutions to these types of problems are increasingly being computed numerically, typically the system is linearized via a first-order Taylor-series expansion about the steady state $\bar{X}$. The model now takes the form

$$x_t = A + B\hat{E}_t x_{t+1} + Cx_{t-1} + D\nu_t$$

where x a vector of endogenous variables in terms of deviation from steady-state. Partitioning x_t into endogenous variables y_t and auxiliary state variables w_t, the model can be rewritten

$$y_t = \alpha + \beta y_{t-1} + \chi w_t + \delta \hat{E}_t y_{t+1} \tag{2.1}$$

$$w_t = \varphi w_{t-1} + M\nu_t \tag{2.2}$$

News shocks are exogenous stochastic shocks which are realized today but do not impact economic fundamentals until some time in the future. Unanticipated or “surprise” shocks thus can be thought of as a particular form of news shocks in which the realization and impact occur contemporaneously. In either case, the shocks are incorporated easily into the model by adding new auxiliary variables to the vector w_t, and appropriately defining the matrices φ and M.[5]

Adaptive Learning

E-Stability: Contemporaneous Expectations

I focus on the minimum state variable (MSV) solutions to the system given by (2.1) and (2.2), which corresponds to a solution with the same set of variables obtained by solving the model under rational expectations. Restricting attention to the MSV solutions helps facilitate a direct comparison of solutions obtained under different expectation formation assumptions.

Before continuing it is worth noting that under rational expectations it is irrelevant whether agents time t information sets include the values of the contemporaneous endogenous variables y_t or if they observe only lagged endogenous variables y_{t-1} and must instead forecast them: because the expectations are model consistent, the forecast will not be systematically incorrect. When considering bounded rationality, however, this is no longer the case; differences between the PLM and the ALM will drive a wedge between expected and realized values and these residuals may well be serially correlated. This

[5] For example, suppose agents receives news zero and one periods in advance. Then we can write equation (2.2) as

$$\begin{pmatrix} \varepsilon_t^0 \\ \varepsilon_t^1 \end{pmatrix} = \begin{pmatrix} 0 & 1 \\ 0 & 0 \end{pmatrix} \begin{pmatrix} \varepsilon_{t-1}^0 \\ \varepsilon_{t-1}^1 \end{pmatrix} + \begin{pmatrix} 1 & 0 \\ 0 & 1 \end{pmatrix} \begin{pmatrix} \nu_t^0 \\ \nu_t^1 \end{pmatrix}$$

where $w_t = \left(\varepsilon_t^0, \varepsilon_t^1\right)$ and $\nu_t = \left(\nu_t^0, \nu_t^1\right)$. In this case the total time t innovation to the exogenous variables in equation (2.1) is given by $\varepsilon_t^0 = \nu_t^0 + \varepsilon_{t-1}^1 = \nu_t^0 + \nu_{t-1}^1$. Agents receive the information about the partial innovation ν_{t-1}^1 in period $t-1$, but it does not impact the economy until period t; the partial innovation ν_t^0 is a surprise shock learned about in period t which affects fundamentals contemporaneously.

wedge can significantly impact the E-stability property of REE, and thus I consider both cases in what follows.

Regardless of the specific assumptions on the information set of agents, the MSV solutions of the system above have the form

$$y_t = a + by_{t-1} + cw_t \tag{2.3}$$

$$w_t = \varphi w_{t-1} + M\nu_t \tag{2.4}$$

Suppose agents are endowed with the form of the solution given in equation (2.3) as their PLM. Then given the information assumptions above, expectations are given by

$$\begin{aligned} \hat{E}_t y_t &= y_t \\ \hat{E}_t y_{t+1} &= a + b\hat{E}_t y_t + c\varphi w_t \\ &= a + by_t + c\varphi w_t \end{aligned}$$

which can be substituted into equation (2.1) to yield the ALM

$$(I - \delta b)\, y_t = (\alpha + \delta a) + \beta y_{t-1} + (\chi + \delta c\varphi)\, w_t \tag{2.5}$$

This provides a mapping from beliefs captured by the PLM to the ALM. This "T-map" is given by

$$T(a,b,c) = \left\{ (I - \delta b)^{-1} (\alpha + \delta a)\,, (I - \delta b)^{-1} \beta, (I - \delta b)^{-1} (\chi + \delta c\varphi) \right\}$$

E-stability of a solution is determined by the matrix differential equation

$$\frac{d}{d\tau}(a,b,c) = T(a,b,c) - (a,b,c) \tag{2.6}$$

It is worth noting that a REE $\left(\bar{a},\bar{b},\bar{c}\right)$ can be interpreted as the solution to the fixed point problem

$$0 = T\left(\bar{a},\bar{b},\bar{c}\right) - \left(\bar{a},\bar{b},\bar{c}\right)$$

i.e. a set of beliefs which are self-fulfilling and coincide on average with realizations such that

$$\left(I - \delta\left(I + \bar{b}\right)\right)\bar{a} = \alpha \tag{2.7a}$$

$$\delta\bar{b}^2 - \bar{b} + \beta = 0 \tag{2.7b}$$

$$\left(I - \delta\bar{b}\right)\bar{c} - \delta\bar{c}\varphi = \chi \tag{2.7c}$$

Inspection of equation (2.7b) shows the coefficient matrix b is independent of the assumed structure for anticipated and unanticipated shocks; that is, the coefficient matrix b is independent of φ and M from (2.2). Furthermore, for a given b equations (2.7a) and (2.7c) uniquely determine the coefficient matrices a and c. However, b is the solution of a matrix-quadratic which is not in general unique, and hence solution techniques such as the well-known methods of Blanchard and Kahn, Uhlig, Klein, or Sims, must be employed to obtain the REE.

The Jacobian of the vectorized matrix differential equation (2.6) evaluated at the REE $\bar{a}, \bar{b}, \bar{c}$ can be shown to be comprised of three blocks[6]

$$DT_a(\bar{a}, \bar{b}) = I \otimes \left(I - \delta\bar{b}\right)^{-1} \delta \tag{2.8a}$$

$$DT_b(b) = \left[\left(I - \delta\bar{b}\right)^{-1} \beta\right]' \otimes \left[\left(I - \delta\bar{b}\right)^{-1} \delta\right] \tag{2.8b}$$

$$DT_c(\bar{b}, \bar{c}) = \varphi' \otimes \left(I - \delta\bar{b}\right)^{-1} \delta \tag{2.8c}$$

Proposition 10.3 of Evans and Honkapohja (2001) suggests the REE is E-Stable if all eigenvalues of the matrices $DT_a(\bar{a}, \bar{b})$, $DT_b(\bar{b})$, and $DT_C(\bar{b}, \bar{c})$ have real parts less than 1. This can be easily established given a particular model, but the objective of this paper is to determine what effect - if any - the inclusion of anticipated shocks has on the E-stability of a given REE. This is established by Proposition 2.3.1

Proposition 2.3.1. *If agents' information sets include the value of contemporaneous endogenous variables and the REE corresponding to the core of a model featuring news shocks, defined as the system of equations which exists when news shocks are shut down, is E-stable, then the REE corresponding to the model with news shocks included is also E-stable.*

---

[6]For example, since the equation governing the evolution of b is independent of a and c we may start with the equation

$$T_b(b) = (I - \delta b)^{-1} \beta$$

the matrix differential of this equation with respect to b is obtained by applying the rule $dF^{-1} = -F^{-1}(dF)F^{-1}$, and hence

$$dT_b(b) = (I - \delta b)^{-1} \delta (db) (I - \delta b)^{-1} \beta$$

Since $dvecx = vecdx$ and $vec(ABC) = (C' \otimes A)vecB$ the vectorized Jacobian $DT_b = \partial vecT_b / \partial (vecb)'$- which determines local stability of this equation evaluated at a particular b- is given by

$$DT_b(b) = \left[(I - \delta b)^{-1} \beta\right]' \otimes \left[(I - \delta b)^{-1} \delta\right]$$

Similar operations can be employed to obtain DT_a and DT_c.

Proof. Suppose the REE is E-stable without anticipated shocks included in the model. Then since the value of $\bar{b}$ is independent of φ and M (which define the structure of exogenous shocks), and since a is uniquely determined for given b, it follows the structural modifications necessary to incorporate news shocks will have no effect on the eigenvalues of the matrices DT_a and DT_b. Furthermore, these modifications can always be done such that φ a block-upper triangular matrix with upper-left block corresponding to the no-news model and lower-right block a nilpotent lower-shift matrix. Since the eigenvalues of DT_c are the set of pairwise-products of eigenvalues of φ' and $\left(I - \delta\bar{b}\right)^{-1}\delta$, the inclusion of the nilpotent sub-matrix implies DT_c will acquire a set of zero eigenvalues in addition to exactly the same eigenvalues as in the no-news model, thereby preserving E-stability of the REE. □

Proposition 2.3.1 suggests that to check whether an REE for a particular model is E-stable it will be sufficient to consider the model *without* news shocks. If it is E-stable without news then it will continue to be E-stable in the presence of news. One immediate implication of this is that any models with well known unique and E-stable REE can be modified to include news. However, the conditions which imply a REE is E-stable depend critically on the information assumed to be known to agents at the time of making their forecast. In the next section I consider the effect of introducing an informational friction to the decision making process of agents.

E-Stability: Delayed Information Assumption

In contrast to the preceding subsection, I now assume the values of contemporaneous endogenous variables are unknown to agents at the time they must make their decisions. Instead, agents must form forecasts of these values using their PLM. This approach emphasizes the (possible) distinction between individual and aggregate-level endogenous variables in macroeconomic models; for example, individual households consumption/savings decisions are typically functions of future real interest rates, which are themselves determined by the aggregate consumption/savings decision of the economy as a whole. Thus each household must form a forecast of the future aggregate action.

I again consider the model given by equations (2.1) and (2.2). The PLM is again assumed to take the form of the MSV solution, which is still

$$y_t = a + by_{t-1} + cw_t \tag{2.9}$$

$$w_t = \varphi w_{t-1} + M\nu_t \tag{2.10}$$

However, the information assumptions now imply that expectations are given by

$$\begin{aligned} \hat{E}_t y_t &= a + by_{t-1} + cw_t \\ \text{and } \hat{E}_t y_{t+1} &= a + b\hat{E}_t y_t + c\varphi w_t \\ &= (I + b)a + b^2 y_{t-1} + (bc + c\varphi) w_t \end{aligned}$$

As before these can be inserted into equation (2.1) to obtain the ALM

$$y_t = [\alpha + \delta (I + b) a] + \left[\beta + \delta b^2\right] y_{t-1} + [\chi + \delta (bc + c\varphi)] w_t \tag{2.11}$$

Thus the mapping of beliefs from the PLM to the ALM is given by

$$T(a, b, c) = \left\{\alpha + \delta (I + b) a, \beta + \delta b^2, \chi + \delta (bc + c\varphi)\right\}$$

and expectational stability of a solution is determined by the matrix differential equation

$$\frac{d}{d\tau}(a, b, c) = T(a, b, c) - (a, b, c) \tag{2.12}$$

where again an REE can be seen as the solution to the fixed point problem

$$0 = T\left(\bar{a}, \bar{b}, \bar{c}\right) - \left(\bar{a}, \bar{b}, \bar{c}\right)$$

It is simple to verify the solution to this problem is identical to that implied by equations (2.7a), (2.7b), and (2.7c), establishing that the specific informational assumptions are equivalent under RE: when agents are perfectly informed to the true ALM for the economy it matters not whether they know the values of contemporaneous endogenous variables or forecast them, as their forecasts will by assumption be correct on average.

However, the additional complication for boundedly rational households can have non-trivial effects on the conditions necessary to ensure asymptotic convergence to the REE. In particular, Proposition 10.1 in Evans and Honkapohja (2001) states the RE solution $(\bar{a}, \bar{b}, \bar{c})$ is E-Stable if all eigenvalues of the Jacobian of the vectorized T-map, which again is comprised of three blocks, have real parts less than 1. Under this alternative timing regime, the blocks are given by[7]

$$DT_a(\bar{a}, \bar{b}) = I \otimes \delta + I \otimes \delta \bar{b}$$

$$DT_b(\bar{b}) = \bar{b}' \otimes \delta + I \otimes \delta \bar{b}$$

$$DT_c(\bar{b}, \bar{c}) = \varphi' \otimes \delta + I \otimes \delta \bar{b}$$

As with the contemporaneous timing assumption, the matrices DT_a and DT_b are unaffected by the modifications to φ and M necessary to include news shocks, and hence their eigenvalues are similarly unaffected. However, DT_c is clearly affected; furthermore, this matrix is the sum of two matrices. Very little can be said of the eigenvalues of the sum of matrices in general, and thus analysis of E-stability under this delayed information set must rely on numerical exercises. In what follows I apply the analytic E-stability results from this section to the model of Beaudry and Portier (2004), and find that the REE are E-stable under both informational assumptions for a wide range of parameter constellations and informational structures

Application to Beaudry and Portier (2004)

Any solution to the comovement problem must address two issues. First, good news about the future must lead to increased demand for investment when the information is obtained as opposed to when the shock actually materialize. And second, the increased investment must be financed by increasing employment as opposed to decreasing

[7]Since the Kronecker product is a bilinear map, we have

$$\begin{aligned} DT_a\left(\bar{a}, \bar{b}\right) &= I \otimes \delta \left(I + \bar{b}\right) \\ &= I \otimes \delta + I \otimes \delta \bar{b} \end{aligned}$$

consumption. There are a variety of methods available for achieving these desired results. In what follows I will apply the E-stability analysis of the previous section to the news-shock model of Beaudry and Portier (2004), in which good news about the future generates a boom in the current period: all major macroeconomic variables - consumption, investment, labor supply, and output - rise upon receipt of the positive information regarding future economic fundamentals.

The economy of Beaudry and Portier (2004) is comprised of three sectors: investment or "durable" goods, intermediate "nondurable" goods, and final composite consumption goods. Households make a consumption/savings decision using income earned by supplying labor to the durable and nondurable good markets and renting capital to the consumption good market; this decision is partially informed the arrival of information about a future technological innovation i.e. a news shock. All markets are competitive. The durable and nondurable goods are produced using CRTS technology from household labor and a fixed factor, and production is augmented by sector-specific technology. The consumption good is produced from the capital stock and the nondurable good using CES technology featuring complementarity between the inputs.

The keys to generating positive comovement in this model are twofold. Because the nondurable good and capital are complements in production of the consumption good, news about the technology used in producing the nondurable good causes a contemporaneous change to the demand for investment. And because the consumption and investment decisions are essentially decoupled from each other - the value of investment is directly related only to the loss of leisure by increasing labor in the durable good sector, as opposed to the loss of utility from reducing consumption - the increased investment is purchased by households working more rather than less. Thus, the model produces qualitatively realistic expectationally driven business cycles in which good news about the future generates a boom in all key macroeconomic variables today.

The Model

Formally, the composite consumption good C_t is produced from the nondurable good X_t and the (predetermined) capital stock K_{t-1} according to

$$C_t = \left(aX_t^\nu + (1-a)K_{t-1}^\nu\right)^{\frac{1}{\nu}} \tag{2.13}$$

where $\nu \leq 0$ to ensure the inputs are complements in production. The nondurable good is produced from household labor $l_{x,t}$, a fixed factor F_x, and technology $\theta_{x,t}$ according to the constant returns to scale (CRTS) production function

$$X_t = \theta_{x,t} l_{x,t}^{\alpha_x} F_x^{1-\alpha_x} \tag{2.14}$$

where $0 \leq \alpha_x \leq 1$ captures labor's share of nondurable-good production. The law of motion for the aggregate capital stock is

$$K_t = (1-\delta)K_{t-1} + I_t \tag{2.15}$$

where $0 \leq \delta \leq 1$ the depreciation rate and I_t gross private investment produced by the durable goods sector from household labor $l_{k,t}$, a fixed factor F_k, and technology $\theta_{k,t}$ according to the CRTS production function

$$I_t = \theta_{k,t} l_{k,t}^{\alpha_k} F_k^{1-\alpha_k} \tag{2.16}$$

where $0 \leq \alpha_k \leq 1$ captures labor's share of durable-good production. The fixed factors are inelastically supplied by households and have the effect of introducing diminishing returns in labor supply while maintaining CRTS in overall production. These fixed factors can be thought of as any scarce resource that constrains production such as privately held land or managerial capital.

Households are infinitely lived and receive utility from consumption and disutility from supplying labor. The lifetime utility function is assumed separable in consumption and labor, and is given by

$$U = \hat{E}_0 \left[\sum_{t=0}^{\infty} \beta^t \left\{ \log(C_t) - v_0 \left(\bar{l} - l_{x,t} - l_{k,t} \right) \right\} \right] \tag{2.17}$$

where $0 < \beta < 1$ the discount factor, $v_0 > 0$ scales the disutility of supplying labor, and $\bar{l}$ total disposable time. $\hat{E}_t$ denotes the subjective expectations of the household given their time t information set. The flow budget constraint is

$$C_t + P_t I_t \leq W_{x,t} l_{x,t} + W_{k,t} l_{k,t} + R_t K_{t-1} + \Pi_{x,t} + \Pi_{k,t} \tag{2.18}$$

where P_t the price of the investment good in terms of the consumption good and $\Pi_{x,t}$ and $\Pi_{k,t}$ the returns to renting the fixed factor in the nondurable and durable goods sectors, respectively. The aggregate resource constraint is defined to be

$$Y_t = C_t + P_t I_t \tag{2.19}$$

where Y_t the total output of the economy. Technology of the durable good exhibits a deterministic trend such that

$$\theta_{k,t} = g_{0,k} e^{g_1 t} \tag{2.20}$$

while the technology of the nondurable grows stochastically according to

$$\theta_{x,t} = g_{0,x} e^{g_1 t} \hat{\theta}_{x,t} \tag{2.21}$$

$$\hat{\theta}_{x,t} = \hat{\theta}^{\lambda}_{x,t-1} e^{\varepsilon^0_t} \tag{2.22}$$

where $0 < \lambda < 1$. One may interpret innovations to nondurable good-specific technology as capturing the process of product differentiation e.g. the creation of higher quality or entirely new products. New goods will require a higher stock of infrastructure, and this complementarity between nondurable TFP and the capital stock is key to the model's ability to generate qualitatively realistic expectationally driven business cycles.

News itself is assumed to take the form of an anticipated shock to the growth of nondurable technology. In particular, $\varepsilon^0_t = \nu^0_t + \nu^n_{t-n}$ where n is the horizon for which the shock is anticipated. This is in contrast to the "noisy news" view, which models news shocks as noisy signals of future innovations. Both formulations in some sense support the notion of revisions and agents being surprised by realizations which differ from their expectations. But the anticipated shocks version has the advantage of being straightforward to incorporate into the systems of equations comprising modern macroeconomic models; furthermore, it is simple to consider alternative assumptions for the information flows e.g. the length and/or density of the anticipated shock structure.

Solving this model for a particular parameterization involves detrending the non-stationary variables, calculating the non-stochastic steady state, and then log-linearizing the system around this non-stochastic steady state; see Appendix A for details. Abusing notation, the (linear) system of expectational difference equations can be put in standard form

$$y_t = \alpha + \beta y_{t-1} + \chi w_t + \delta \hat{E}_t y_{t+1} \tag{2.23}$$

$$w_t = \varphi w_{t-1} + M\nu_t \tag{2.24}$$

where $y_t = \left(\tilde{C}_t, \tilde{l}_{x,t}, \tilde{l}_{k,t}, \tilde{I}_t, \tilde{K}_t, \tilde{X}_t, \tilde{Y}_t, \tilde{P}_t, \tilde{R}_t, \tilde{W}_{x,t}, \tilde{W}_{k,t}, \tilde{\Lambda}_t, \tilde{Q}_t, \tilde{\theta}_{k,t}, \tilde{\theta}_{x,t}, \tilde{\hat{\theta}}_{x,t}\right)'$ the endogenous variables, $w_t = (\varepsilon_t^0, \varepsilon_t^1, \cdots, \varepsilon_t^n)'$ a vector of auxiliary state variables designed to pass news shocks through time, and $\nu_t = (\nu_t^0, \nu_t^1, \cdots, \nu_t^n)'$ a vector of anticipated and unanticipated exogenous stochastic shocks.[8] The elements of the matrices $\alpha, \beta, \chi,$ and δ correspond to the log-linearized equations which describe a temporary equilibrium and are functions of the structural parameters; the non-zero elements of $\varphi,$ and M are selected to imply the desired structure for the arrival of anticipated and unanticipated shocks.

Beaudry and Portier (2004) calibrate the model parameters using results from previous studies or to achieve specific steady-state values. My analysis of the E-stability of their news-shock model begins with their baseline calibrations, which are summarized in Table 1.

[8]Technology for the durable good is assumed to be deterministic. Incorporating a stochastic element would simply require modifying w_t and ν_t to include the appropriate auxiliary and exogenous variables and making the corresponding revisions to the filtering matrices φ and M.

Baseline Parameterization	
Parameter	**Values**
β	0.98
δ	0.05
α_x	0.60
α_k	0.97
$\bar{l}$	2
v_0	1
F_x	1
F_k	1
λ	0.999
ν	-3.78

TABLE 1. Calibrated Parameters for Three-sector Model

Inclusion of News Shocks

Putting the model in standard form allows an immediate application of the E-stability results from the section which introduced adaptive learning. In particular, it is simple for any given parameterization to calculate the eigenvalues of the T-map's vectorized Jacobian which indicate whether or not the system is E-stable. I begin by setting $n = 0$, that is, the case where there is no news and all shocks are completely unanticipated. The largest real roots of $DT_a(\bar{a}, \bar{b}), DT_b(\bar{b})$, and $DT_c(\bar{b}, \bar{c})$ under the contemporaneous information assumption are 0.5199, 0.51947, and 0 respectively, while the corresponding numbers are 0.1547, 0.1538, and 0 under the delayed information assumption. Thus the RE solution for this model is E-Stable under both informational assumptions without news.

Proposition 2.3.1 implies that the eigenvalues should be unaffected by the inclusion of news shocks under the contemporaneous timing assumption. Indeed, setting $n = 1$ so that households receive information about the technological innovation 1-period ahead and keeping all other parameters fixed at their baseline levels results in no change to the eigenvalues for $DT_a(\bar{a}, \bar{b}), DT_b(\bar{b})$, or $DT_c(\bar{b}, \bar{c})$, as expected. Furthermore, while there is no corollary for the delayed timing assumption, numerical results suggest the E-stability of the REE is similarly unaffected by the inclusion of news under the delayed timing assumption.

The main contribution of this chapter is the finding that news shocks do not alter the E-stability properties for REE. This is important because E-stability results are, in general, sensitive to the assumptions regarding the timing and flow of information. While news shocks are essentially a modification of the exogenous driving forces for the economy which would not in general be expected to alter E-stability results, their impact is felt through the revision of expectations and subsequent behavioral changes on the part of households from the arrival of new information. This informational feedback effect is responsible for the fact that some REE which are E-stable under the contemporaneous timing assumption may fail to be E-stable under the delayed timing assumption. The main results of this chapter affirm the robustness of REE to news shocks and suggest modelers wishing to incorporate anticipated shocks into their models may do so without sacrificing on the robustness of their REE to alternate expectations formations assumptions.

Robustness

In what follows I seek to characterize what role, if any, the specific calibration choices for structural parameters and the new shock structure play in determining whether the REE is E-stable. I proceed by first considering alternate calibrations for the structural parameters and then exploring the effect of lengthier forecasting horizons and denser informational structures. I find that the E-stability of the REE is robust to all constellations consistent with the model being able to generate qualitatively realistic expectationally driven business cycles, with and without news shocks.[9]

To analyze whether alternative parameterizations may impact the E-stability results, I adopt the following strategy: I change the value of structural parameters one at a time while holding all others at their baseline calibration, searching only within the parameter space consistent with the model generating qualitatively realistic expectationally driven business cycles. The REE is E-stable for all parameterizations considered and under both informational timing assumptions. Table 2 displays the smallest and largest considered for each parameter.

Robustness Checks		
Parameter	**Smallest Value**	**Largest Value**
β	0.001	1
δ	0	1
α_x	0.001	0.999
α_k	0.001	0.999
ν	-100	-0.001
a	0.001	0.999

TABLE 2. Alternate Parameterizations

In general, the largest real parts of eigenvalues are larger when households discount the future less (smaller β), when the depreciation rate of capital is smaller (smaller δ), when the decreasing returns to producing the nondurable good are smaller (larger α_x), when the decreasing returns to producing the durable good are larger (smaller α_k), when the complementarity between capital and nondurable goods in producing the consumption good is weaker (larger ν), and as the relative importance of the nondurable good to capital in producing consumption decreases (smaller a). Even so, setting these parameters jointly

[9]The ability of the model to solve the comovement problem is dependent only the value of ν, which governs the substitutability between capital and the nondurable good in the production of the consumption good. As long as these factors are complements in production of the consumption good i.e. $\nu \leq 0$, the model will be capable of generating qualitatively realistic expectationally driven business cycles.

to values which on their own would tend to imply relatively large eigenvalues does not cause the REE to become expectationally unstable.

One might also be concerned about assumed structure for news shocks themselves, that is, the length and density of the forecasting horizon. To analyze what role this structure plays in the E-stability of the REE I proceed as follow. First, I extend the forecasting horizon while keeping the structural parameters at their baseline levels by considering $n = 1, 2, ..., 10$. Second, I allow for multiple anticipated shocks, that is, households receive multiple pieces of information for the eventual innovation. For example, households may receive news about the innovation four and eight periods ahead. Neither changes to the length of the forecasting horizon nor its density have any impact on the largest real parts of eigenvalues of $DT_a(\bar{a}, \bar{b}), DT_b(\bar{b})$, or $DT_c(\bar{b}, \bar{c})$ under either timing assumption.

Conclusion

That the three-sector news-shock model of Beaudry and Portier (2004) can generate qualitatively realistic expectationally driven business cycles while the RE solution is E-stable suggest there is nothing fundamental about news shocks *per se* to imply models incorporating them will be expectationally unstable. In fact, Proposition 2.3.1 shows that the REE of any model be E-stable when news shocks are included in the time t information set as long as it is E-stable when news shocks are not included. While no analytic result exists for alternative timing specifications, numerical experimentation suggests this may be true more generally, as alternate informational structures did not impact the largest eigenvalues of the associated T-map's vectorized Jacobian.

This is a particularly important finding given the well-known sensitivity of E-stability of REE to specific assumptions regarding agents information sets. Since news-shock models rely critically on novel changes to agent's information sets, and since no other studies have examined E-stability in the context of a news-shock model, this study suggests a potential justification for RE on the grounds that boundedly rational reduced-form learning agents can learn the RE coefficients over time. More work will need to be done to examine whether the REE of such news-shock models demonstrate similar learnability properties under alternative forms of bounded rationality, such as Euler-equation learning, infinite-horizon learning, or shadow-price learning.

CHAPTER III

SHADOW PRICE LEARNING IN A NEWS-SHOCK MODEL

Introduction

The relatively recent interest in using news to explain macroeconomic fluctuations began largely from the asset-price bubbles experienced in the late 1990s with the tech-boom. The "irrational exuberance" exhibited by the market was difficult to generate in most standard macroeconomic DSGE models, and it is this gap which the original news-shock model of Beaudry and Portier (2004) attempted to fill. As shown in Chapter II, by modifying the structure of a RBC model in a particular way the authors were able to generate boom-bust cycles fueled by households experiencing periods of optimism and pessimism which created liquidation cycles in the capital stock. Yet this is just one way of solving the "comovement problem" found in RBC models, and it comes at the expense of making significant changes to the standard RBC model. In particular, it requires three sectors and very strong complementarity of inputs to final goods production, both of which are somewhat unusual in the broader context of typical workhorse macroeconomic models.

Jaimovich and Rebelo (2009) on the other hand provide a news-shock model capable of solving the comovement problem while remaining approachable to researchers familiar with standard RBC models. The core of the model is essentially a discrete-time Ramsey model with elastic labor supply. Savings are converted to investment according to some technological process, and investment is used to create a depreciating capital stock. This textbook structure is augmented by assuming variable capacity utilization of capital, a cost to adjusting investment from its previous level, and a novel preference specification nesting the well known specifications of Greenwood et al. (1988) and King et al. (1988). While these features taken together represent a significant departure from standard RBC models, the individual elements are fairly standard.

The relative simplicity of the model makes it an excellent candidate for exploring the qualitative and quantitative effects of relaxing the RE assumption in favor of a weaker assumption. Furthermore, since it is quite similar to well-known RBC models, the results of such an experiment exist in an environment in which some context has already been provided. While many alternatives to RE exist, I ultimately choose to consider bounded optimality of the sort described in Evans and McGough (2015). To place this decision in a broader context, I begin by describing the two benchmark theories of bounded rationality: *Euler-equation learning* and *infinite-horizon learning*.

Under Euler-equation learning, found e.g. in Evans and Honkapohja (2006) and Honkapohja et al. (2012), agents behave in accordance with their Euler equations which relate control decisions today with the forecasted values of decisions tomorrow. This approach eliminates the reliance on RE in deriving the equilibrium dynamics of the economy found in the reduced-form learning approach of Chapter II. Rather than assuming agents are rational, solving their dynamic programming problem, and then relaxing the RE assumption, Euler-equation learning studies the evolution of agent's beliefs and the path of the economy when their expectations are boundedly rational even when solving their dynamic programming problem. One of the benefits of Euler-equation learning is its simplicity: agents need only make one-period ahead forecasts to determine their optimal control decisions.

However, a possible shortcoming is this approach requires agents to forecast the future value of control variables, a somewhat strange exercise in a representative agent model where these are completely under the control of the agents. The infinite-horizon approach to learning of Preston (2005) avoids this by formally expressing decisions today as being determined by agent's expectations of their future lifetime budget constraint and transversality conditions. Behavior is thus based on the relationship between future wealth and the implied control decisions. One of the main benefits of the infinite-horizon approach is the strict adherence to microfoundations which implies behavior is truly optimal given beliefs.

In Branch et al. (2012a) a "hybrid" of the infinite-horizon and Euler-equation approaches called *N-step optimal learning* is developed. N-step optimal learning assumes agents explicitly take current and future expected values of wealth over a finite range into account when making their control decisions, which are themselves anchored to the Euler equation as a behavioral primitive. This learning mechanism has the infinite-horizon approach as a limiting case, and may be viewed as its finite-horizon version. N-step optimal learning captures both the simplicity of Euler-equation learning and the rigorous adherence to fully optimal behavior imparted by infinite-horizon learning, but it still requires considerable sophistication on the part of agents in the model.

A recent alternative similar in spirit but requiring less agent-level sophistication is the *shadow-price learning* (SP-learning) approach of Evans and McGough (2015). Rather than taking the behavioral primitive to be the Euler equations, which often embody complicated nonlinear relations between contemporaneous control decisions and the evolution of the determinants of future income, SP-learning assumes agents base their decisions on the standard first-order necessary conditions (FONCs) derived from the intertemporal Lagrangian. These FONCs describe a set of conditions which must be satisfied by the agents control decisions in order to be considered optimal, and they are functions of the expected present value of shadow prices for endogenous state variables. SP-learners form forecasts of these shadow prices using a linear PLM and update their beliefs over time using recursive least squares; thus the behavior is by definition optimal *given beliefs*.

SP-learning is an attractive alternative to RE for many reasons. First, the informational assumptions are quite natural: agents need understand little more than their preferences and budget constraint and take as given many of the same variables that real households do e.g. wages and interest rates. Second, the behavior of agents is quite intuitive: they simply contemplate how changing their behavior today would impact key variables tomorrow based on their beliefs and take action accordingly. Third, the updating of beliefs occurs via recursive least squares, an exercise which any student of introductory econometrics could conduct. Finally, it lends itself to considering heterogeneity of agents along some dimension such as information or initial wealth, as these differences will cause different transition paths under learning for different households.

While much attention in the bounded rationality literature has focused on determining whether agents endowed with boundedly rational expectations may learn to behave rationally, there is a somewhat smaller literature which examines the effect of relaxing RE on the quantitative predictions generated by a given DSGE model. For example, Williams (2003a) compares data generated by simulations of an RBC and a NK model under RE and reduced-form learning. He finds that reduced-form learning has an extremely small effect on generated moments. He then considers an alternative learning approach where agents' decisions incorporate uncertainty about the structure of the economy into their decision rules. This "structural learning" approach is shown to substantially increase the volatility and persistence of key macroeconomic variables generated by an otherwise standard RBC model. Given the discussion above this is not too surprising: the "structural learning" approach embodies to varying degrees the same core idea behind Euler-equation and infinite-horizon learning, namely that agents should try to act optimally given their beliefs, and that these beliefs may be quite different from those prescribed by RE.

This chapter proceeds in the spirit of Williams (2003a) by trying to determine the quantitative effects on business cycle statistics generated by the Jaimovich and Rebelo (2009) news-shock model when agents are assumed to be SP-learners. The central goal of this chapter is to address what effect relaxing the assumption of RE in favor of SP-learning has on the empirical relevance of this particular news-shock model, and whether the behavior of SP-learning households will come to approximate that of their fully rational counterparts. I first describe the economic environment, including the calibration for the model, the equations governing a temporary equilibrium, and the way in which news shocks are modeled. I then present the results of simulation exercises to generate and compare data on business cycles within the model to those found in US data. The chapter concludes by reviewing the main results and discussing paths for future research.

The Model

The representative household chooses consumption C_t and hours worked h_t to maximize the lifetime utility function

$$\hat{E}_0 \sum_{t=0}^{\infty} \beta^t U(V_t) \tag{3.1}$$

where $0 < \beta < 1$ the household's discount factor and U the period utility function which takes the CRRA form

$$U(V_t) = \frac{V_t^{1-\sigma} - 1}{1 - \sigma} \tag{3.2}$$

where $\sigma > 0$ the inverse intertemporal elasticity of substitution and the argument V_t is given by

$$V_t = C_t - \frac{\psi h_t^{1+\frac{1}{\theta}} S_t}{1 + \frac{1}{\theta}} \tag{3.3}$$

where $\psi > 0$ scales the disutility of labor supply and $\theta > 0$ governs the Frisch elasticity of labor supply. S_t is a geometric average of current and past habit-adjusted consumption and takes the form

$$S_t = C_t^{\gamma} S_{t-1}^{1-\gamma} \tag{3.4}$$

where $0 \leq \gamma \leq 1$ governs the magnitude of the wealth elasticity of labor supply. This preference specification, often referred to as "JR preferences", allows the modeler to calibrate the wealth effect of labor supply to be "small" while permitting a balanced growth path. The comovement problem in typical RBC models is partially caused by the wealth effect of labor supply dominating the substitution effect, thereby causing consumption and hours worked to move in opposite directions upon receipt of good news, and hence this utility function makes it simple to develop qualitatively realistic expectationally driven business cycles.

Note this specification nests two well-known and important preference specifications. $\gamma = 0$ corresponds to the preferences of Greenwood et al. (1988) in which labor supply depends only on current real wages and is independent of the marginal utility of income, while $\gamma = 1$ corresponds to the preferences of King et al. (1988) which are compatible with a balanced growth path at the optimal steady state of the economy. Small values for γ allow the economy to be consistent with a balanced growth path while also implying a very weak wealth effect of labor supply, both of which are important features for any dynamic stochastic general equilibrium (DSGE) model hoping to explain movement in key macroeconomic variables through news shocks.[1]

Households are assumed to own physical capital K_t and rent it to firms in a competitive market. Each household's stock of capital evolves according to the law of motion

$$K_t = (1 - \delta(u_t)) K_{t-1} + I_t \left[1 - \Phi\left(\frac{I_t}{I_{t-1}}\right)\right] \tag{3.5}$$

where I_t gross private investment. Note the timing convention: subscripts correspond to the time period in which the variable is decided, hence in time t the household chooses K_t and takes as given the predetermined capital stock K_{t-1}, which was chosen in period $t-1$.

$\Phi\left(\frac{I_t}{I_{t-1}}\right)$ imposes a cost to adjusting investment from its previous level while $\delta(u_t)$ implies capital depreciation is a function of its utilization rate u_t. Both functions are convex in their arguments, and I follow SGU in assuming the quadratic functional forms

$$\delta(u_t) = \delta_0 + \delta_1(u_t - 1) + \frac{\delta_2}{2}(u_t - 1)^2 \tag{3.6}$$

[1] Note that the case of $\gamma = 0$ is not consistent with a balanced growth path: along a steady state the real interest rate must be growing at a constant rate, and thus the marginal utility of income will also be growing. A balanced growth path requires the wealth and substitution effects from the constantly growing real wage to cancel, but $\gamma = 0$ ensures that only the substitution effect exists, and hence labor supply will be changing at a rate inconsistent with the rest of the economy.

where $\delta_0 > 0$ the steady-state depreciation rate, $\delta_1 > 0$ determines the steady-state value of u_t, and $\delta_2 > 0$ captures the rental rate elasticity of capacity utilization and

$$\Phi\left(\mu_t^I\right) = \frac{\kappa}{2}\left(\mu_t^I - \mu^I\right)^2 \tag{3.7}$$

where $\mu_t^I \equiv \frac{I_t}{I_{t-1}}$ the growth rate of investment, $\kappa > 0$ a scaling parameter, and μ^I the steady-state growth rate of gross private investment. This specification implies $\Phi\left(\mu^I\right) = \Phi'\left(\mu^I\right) = 0$ and $\Phi''\left(\mu^I\right) > 0$, i.e. there are no investment-adjustment costs on a balanced growth path.

Each period the household receives labor income from working h_t hours at rate W_t, rental income from from renting $u_t K_{t-1}$ units of effective capital at gross rental rate R_t, and lump sum firm-profits of Π_t. The household uses this income to purchase consumption and investment goods. The flow budget constraint is given by

$$C_t + A_t I_t \leq W_t h_t + R_t\left(u_t K_{t-1}\right) + \Pi_t \tag{3.8}$$

where A_t is an exogenous process representing the current state of technology for producing investment goods from consumption goods with stationary growth rate $\mu_t^A \equiv \frac{A_t}{A_{t-1}}$ and steady-state value μ^A. In a decentralized equilibrium A_t may be interpreted as the relative price of investment goods in terms of consumption goods, that is a unit of the investment good may be exchanged for A_t units of the consumption good.

Households are assumed to make their choices to maximize their expected lifetime utility, and hence each period they solve a constrained optimization problem to maximize expected discounted utility. Formally, households choose a set of stochastic processes

$\{C_t, h_t, u_t, I_t, K_t, S_t\}_{t=0}^{\infty}$ to maximize 3.1 subject to the constraints given by equations 3.2-3.8 and initial conditions for the endogenous state variables I_{-1}, K_{-1}, and S_{-1}. This can be written as a standard dynamic constrained maximization problem:

$$\max_{C_t,u_t,h_t,I_t,K_t,S_t} \quad \hat{E}_t \sum_{t=0}^{\infty} \beta^t U(V_t)$$

subject to

$$U(V_t) = \frac{V_t^{1-\sigma} - 1}{1-\sigma}$$

$$V_t = C_t - \frac{\psi h_t^{1+\frac{1}{\theta}} S_t}{1+\frac{1}{\theta}}$$

$$S_t = C_t^{\gamma} S_{t-1}^{1-\gamma}$$

$$C_t + A_t I_t \leq W_t h_t + R_t (u_t K_{t-1}) + \Pi_t$$

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

The household's time t control decisions $\hat{u}_t = (C_t, u_t, h_t, I_t, K_t, S_t)'$ will depend on their expectations of the future shadow price of the endogenous state variables $x_t^1 = (I_{t-1}, K_{t-1}, S_{t-1})'$, which are variables the household can directly affect in the future through their actions today.[2] Under RE all agents within the economy know the conditional distributions of all variables, and hence the laws of motion for the endogenous and exogenous state variables are known and the expected future value of these shadow prices is straightforward to compute. Standard solution techniques under RE involve deriving Euler equations relating today's control decisions to those of tomorrow to describe the behavior of households and combining these and other optimality conditions, resource constraints, and laws of motion to arrive a system of expectational difference equations.

[2]While agents will also use their expectations of future realizations of the exogenous state variables in their decision making process, their actions do not affect the marginal values of these variables; hence, they need not concern themselves with the expected future shadow prices of these exogenous states.

Typically this system is nonlinear, and hence the solution to the first-order approximation is considered.

In what follows I will replace RE with a version of bounded rationality based on shadow prices called SP-learning. Under SP-learning, households do not know the conditional distribution of all variables, and in particular they do not know the precise way in which shadow prices depend on their behavior. Instead they estimate the value of these shadow prices using a linear forecasting rule which is a function of observables.

The solution technique under SP-learning stands in stark contrast to that of RE. A major difference is that while agents are endowed with linear forecasting rules, their behavior is embedded within the nonlinear specification of the model. This implies except for special cases, the relevant equilibrium notion is that of a *restricted-perceptions equilibrium* (RPE) as opposed to a REE. A RPE can be thought of as an equilibrium arising from agents' optimally misspecified beliefs which are consistent with the stochastic processes realized in the economy. Given their forecasting model agents are unable to detect a misspecification. The first task of this chapter is determining how similar this RPE is to the implied REE.

Denote the shadow prices of investment, capital, and habit-adjustment by λ_t^I, λ_t^K and λ_t^S, respectively. These have the simple interpretation as the time t value of an additional unit of their corresponding endogenous state in time t. For example, λ_t^K is the marginal value of an additional unit of preinstalled capital K_{t-1} in time t. With this interpretation the FONCs describing optimal household choices of C_t, h_t, and u_t given beliefs about the future values of these shadow prices can be obtained via simple variational arguments.[3]

[3]Note that given optimal decisions for consumption, labor supply, and capacity utilization the choices for the geometric average of habit-adjusted consumption S_t, gross investment I_t, and hence next period's capital stock K_{t+1} are pinned down by equations (3.4), (3.8), and (3.5) respectively.

The details of this derivation are presented in Appendix B. The end result is the following three FONCs in the controls

$$
\begin{aligned}
U_{C_t}(V_t) + \frac{\partial S_t}{\partial C_t}\beta \hat{E}_t \lambda^S_{t+1} &= \frac{\partial I_t}{\partial C_t}\left(\beta \hat{E}_t \lambda^I_{t+1} + \frac{\partial K_t}{\partial I_t}\beta \hat{E}_t \lambda^K_{t+1}\right) \\
-U_{h_t}(V_t) &= \frac{\partial I_t}{\partial h_t}\left(\beta \hat{E}_t \lambda^I_{t+1} + \frac{\partial K_t}{\partial I_t}\beta \hat{E}_t \lambda^K_{t+1}\right) \\
\frac{\partial K_t}{\partial u_t}\beta \hat{E}_t \lambda^K_{t+1} &= \frac{\partial I_t}{\partial u_t}\left(\beta \hat{E}_t \lambda^I_{t+1} + \frac{\partial K_t}{\partial I_t}\beta \hat{E}_t \lambda^K_{t+1}\right)
\end{aligned}
$$

and the following three FONCs in the endogenous state variables

$$
\begin{aligned}
\lambda^I_t &= \frac{\partial K_t}{\partial I_{t-1}}\beta \hat{E}_t \lambda^K_{t+1} \\
\lambda^K_t &= \frac{\partial I_t}{\partial K_{t-1}}\left(\beta \hat{E}_t \lambda^I_{t+1} + \frac{\partial K_t}{\partial I_t}\beta \hat{E}_t \lambda^K_{t+1}\right) + \frac{\partial K_t}{\partial K_{t-1}}\beta \hat{E}_t \lambda^K_{t+1} \\
\lambda^S_t &= U_{S_{t-1}}(V_t) + \frac{\partial S_t}{\partial S_{t-1}}\beta \hat{E}_t \lambda^S_{t+1}
\end{aligned}
$$

It can be shown that the endogenous shadow prices can be expressed in terms of the contemporaneous controls and states, and hence their actual value is determined as a function of the contemporaneous decisions of the household. However, the household makes these decisions based on their expectations of the value of future shadow prices without knowing the nonlinear way in which the actual values are determined. Given a set of beliefs and a linear PLM for the evolution of these shadow prices, SP-learners take optimal actions given their forecasts of future shadow prices. This behavior results in the ALM for the shadow prices. Beliefs are updated as weighted averages of current beliefs and forecast errors via the recursive least squares algorithm as is standard in the adaptive learning literature. A crucial question addressed by studying household behavior under this less stringent specification for expectation formation is whether the household's beliefs will (approximately) converge to those of a rational agent.

The model is closed by describing the production side of the economy. It is important to keep in mind that the only agents operating as SP-learners are the households; firms behave in a manner consistent with rational expectations. One possible justification for this discrepancy in assumptions about expectations formation is that these types of agents may be more likely to have the means to act in a highly sophisticated manner, whereas the average household likely does not.

The representative firm pays for h_t worker-hours and rents $u_t K_{t-1}$ units of effective capital to produce output Y_t using CRT technology according to the production function

$$Y_t = z_t \left(u_t K_{t-1}\right)^{1-\alpha} h_t^{\alpha} \tag{3.9}$$

where $\alpha \in (0,1)$ governs the labor share of output in steady state. As is typical in RBC-type models, the supply side of the economy is subjected to exogenous stochastic shocks, given here by a transitory shock z_t. Factor markets are competitive and hence the gross rental rate equals the value of the marginal product of effective capital

$$R_t = (1-\alpha)\frac{Y_t}{u_t K_{t-1}} \tag{3.10}$$

and the wage paid is equal to the value of the marginal product of labor

$$W_t = \alpha \frac{Y_t}{h_t} \tag{3.11}$$

Output is fungible and may be used for private consumption or gross investment. Total demand is thus given by

$$Y_t = C_t + A_t I_t \tag{3.12}$$

Equilibrium

A temporary equilibrium for all periods $t \geq 0$ is described collectively by the optimal behavior of households given beliefs:

$$U_{C_t}(V_t) + \frac{\partial S_t}{\partial C_t}\beta \hat{E}_t \lambda^S_{t+1} = \frac{\partial I_t}{\partial C_t}\left(\beta \hat{E}_t \lambda^I_{t+1} + \frac{\partial K_t}{\partial I_t}\beta \hat{E}_t \lambda^K_{t+1}\right) \tag{3.13a}$$

$$-U_{h_t}(V_t) = \frac{\partial I_t}{\partial h_t}\left(\beta \hat{E}_t \lambda^I_{t+1} + \frac{\partial K_t}{\partial I_t}\beta \hat{E}_t \lambda^K_{t+1}\right) \tag{3.13b}$$

$$\frac{\partial K_t}{\partial u_t}\beta \hat{E}_t \lambda^K_{t+1} = \frac{\partial I_t}{\partial u_t}\left(\beta \hat{E}_t \lambda^I_{t+1} + \frac{\partial K_t}{\partial I_t}\beta \hat{E}_t \lambda^K_{t+1}\right) \tag{3.13c}$$

$$K_t = (1 - \delta(u_t)) K_{t-1} + I_t \left[1 - \Phi\left(\frac{I_t}{I_{t-1}}\right)\right] \tag{3.13d}$$

$$C_t + A_t I_t \leq W_t h_t + R_t (u_t K_{t-1}) \tag{3.13e}$$

$$S_t = C_t^{\gamma} S_{t-1}^{1-\gamma} \tag{3.13f}$$

$$I_{-1}, K_{-1}, S_{-1} \text{ given} \tag{3.13g}$$

where the utility function, its argument, and the functional forms for $\delta(u_t)$ and $\Phi\left(\frac{I_t}{I_{t-1}}\right)$ are given by equations (3.2), (3.3), (3.6), and (3.7), respectively. Appendix B expresses these equations in terms of the underlying variables as opposed to derivatives, which I present here for the sake of expositional clarity. The market clearing conditions, aggregate resource constraint, production function, and factor-market prices are given by

$$Y_t = C_t + A_t I_t \tag{3.14a}$$

$$Y_t = z_t \left(u_t K_{t-1}\right)^{1-\alpha} \left(h_t\right)^{\alpha} \tag{3.14b}$$

$$W_t = \alpha \frac{Y_t}{h_t} \tag{3.14c}$$

$$R_t = (1-\alpha) \frac{Y_t}{u_t K_{t-1}} \tag{3.14d}$$

The laws of motion for the exogenous stochastic processes A_t and z_t along with the precise way in which news shocks are included in the model are described below.

News Shocks and Expectations Formation

It remains to describe the particular way in which shadow price learning agents form their expectations. The variables relevant to household decision making but outside of their control in time t - that is, the predetermined control variables, factor prices, and the exogenous state variables (plus a constant) - are collected into $x_t = (1, I_{t-1}, K_{t-1}, S_{t-1}, W_t, R_t, A_t, z_t)'$, and households are assumed to have a PLM for the endogenous-state shadow prices given by $\lambda_t = H_t x_t$, where $\lambda_t = \left(\lambda_t^I, \lambda_t^K, \lambda_t^S\right)'$ and H_t denotes the household's time t estimates of the coefficient-matrix relating the state variables to the value of the shadow price.

There is one issue that must be addressed: W_t and R_t are exact functions of h_t, $u_t K_{t-1}$, and the exogenous processes z_t and A_t, and hence including all of these variables in the regression will lead to perfect multicollinearity. To avoid this I drop W_t and R_t from the set of regressors utilized by the household and denote this subset of state variables as $\tilde{x}_t = (1, I_{t-1}, K_{t-1}, S_{t-1}, A_t, z_t)'$.[4]

The household's PLM for the shadow prices is thus given by the linear model

$$\lambda_t = H_t'\tilde{x}_t \tag{3.15}$$

where H_t is updated via recursive least squares according to the dynamic system

$$R_{H,t} = R^H_{t-1} + g_t\left(\tilde{x}_{t-1}\tilde{x}'_{t-1} - R^H_{t-1}\right) \tag{3.16}$$

$$H_t = H_{t-1} + g_t R^{-1}_{H,t}\tilde{x}_{t-1}\left(\lambda_{t-1} - H'_{t-1}\tilde{x}_{t-1}\right)' \tag{3.17}$$

$R_{H,t}$ is the household's time t estimate of the second-moment matrix for the regressors while the "gain" parameter g_t controls how much weight households put on new information. Much of the theoretical work exploring whether agents can learn the coefficient-matrix of an economy's REE, such as Evans and Honkapohja (2001) assumes a *decreasing gain* such as $g_t = t^{-1}$, which implies the household response to forecast errors vanishes asymptotically. Alternatively, studies focused on simulation or estimation of DSGE models under learning, such as Williams (2003b) and Milani (2007), employ a *constant gain* where $g_t = \bar{g}$, which implies the household is a lifelong learner and continually revises its coefficient estimates placing the most weight on the most recent observations. I will simulate the model under both assumptions. Under constant-gain

[4] One could also consider assuming the household does not observe some of the exogenous processes and instead uses wages and the gross rental rate in their forecasts. This would require specifying a PLM and dynamic system for updating beliefs for these price variables. It can also lead to serious issue with asymptotic multicollinearity as household's beliefs converge to those of rational agents, because as the residuals in the dynamic system for updating beliefs go to zero the actions of the household become perfectly collinear with market prices.

learning I calibrate the gain to be 0.0152, which is in line with recent estimates from the empirical adaptive learning literature.

The value of shadow prices in time t is not known *ex-ante* to the household, and is determined as a result of their control choices. Hence, time t beliefs are updated using information available through time $t-1$. This clarifies the central idea of shadow price learning: households make optimal decisions *given* their (misspecified) beliefs and the information available to them in the moment. That beliefs may be misspecified highlights a particularly salient feature of SP-learning in particular and bounded rationality in general by embracing the notion that agents in the economy may certainly fail to fully understand the dynamics governing its evolution and yet an equilibrium may still exist.

Taking expectations of equation (3.15) we have $\hat{E}_t\lambda_{t+1} = H_t\hat{E}_t\tilde{x}_{t+1}$, and hence households must forecast the future values of regressors.[5]. The values of the time $t+1$ endogenous state variables I_t, K_t, and S_t are pinned down by the time t flow budget constraint, capital accumulation equation, and geometric-average identity respectively, and it is natural to assume the household knows these values. In addition the transition equations for the exogenous processes z_t and A_t are assumed to be known to the household.[6]

[5]I have assumed certainty equivalence on the parts of households; that is, they believe their time t estimate of the coefficient matrix is the correct estimate now and in the future, and hence behave accordingly. Thus household's do not need to attempt to forecast the way in which their estimates will change in the future.

[6]Since there is no feedback between the household's decisions and the evolution of these processes this is a very simple estimation exercise from an econometric standpoint if the households actually observe all of this data. An alternative approach would carefully consider the variables a household seems likely to observe. Exogenous variables that directly influence the household, such as the technology available for converting consumption goods into investment goods A_t in the flow budget constraint seem quite natural; however it is less obvious that households would directly observe productivity z_t. As mentioned previously, one could drop these as regressors in favor of e.g. wages and/or the rental rate of capital, but this may cause the coefficient estimation routine to suffer from asymptotic multicollinearity and/or bias from sampling error present in using estimated data as an explanatory variable.

These expectations are augmented by the inclusion of news shocks to the information set. A news shock can be thought of as information which arrives exogenously to the household about the value of some future innovation to an exogenous process, but (importantly) does not impact any economic fundamentals contemporaneously. The action generated by news shocks is therefore entirely in the response of agents to this information about the future. News is modeled as an anticipated shock to the exogenous processes and hence has the interpretation of imparting incomplete (but accurate) information to the household about future economic fundamentals.

In particular let the law of motion for the exogenous processes, indexed by $w = (A, z)$, take the form

$$\begin{aligned}
\ln(w_t) &= \rho_w \ln(w_{t-1}) + \varepsilon^0_{w,t} \\
\varepsilon^0_{w,t} &= \varepsilon^1_{w,t-1} + \nu^0_{w,t} \left[\nu^0_{w,t} \in \mathcal{I}_t\right] \\
\varepsilon^1_{w,t} &= \varepsilon^2_{w,t-1} + \nu^1_{w,t} \left[\nu^1_{w,t-1} \in \mathcal{I}_t\right] \\
\vdots &= \qquad \vdots \\
\varepsilon^n_{w,t} &= \varepsilon^{n-1}_{w,t-1} + \nu^n_{w,t} \left[\nu^n_{w,t-n} \in \mathcal{I}_t\right]
\end{aligned}$$

where $\mathcal{I}_t$ the time t information set of the representative household, $\nu^k_{w,t-k}$ the k-period ahead anticipated shock, and $\left[\nu^k_{w,t-k} \in \mathcal{I}_t\right] = 1$ if $\nu^k_{w,t-k} \in \mathcal{I}_t$ and 0 otherwise. This specification allows the modeler to easily and parsimoniously consider a variety of assumptions regarding the precise details of informational acquisition by households.[7]

[7] In fact, this requires the creation of only as many auxiliary state variables as the lengthiest forecasting horizon. For example, if a household receives information four and eight periods in advance the longest forecasting horizon is eight periods, and thus eight auxiliary state variables must be generated. An alternative but equivalent specification would generate a set of k state variables for each signal $v^k_{w,t-k}$ so that e.g. if a household received information four and eight periods in advance one would need to generate 12 additional state variables.

The system above makes clear the sense in which news shocks here are being modeled as partial information about the total value of the innovation. This permits a compact autoregressive representation

$$\ln(w_t) = \rho_w \ln(w_{t-1}) + \varepsilon^0_{w,t} \tag{3.18}$$

$$\varepsilon_{w,t} = \varphi_w \varepsilon_{w,t-1} + M_w \nu_{w,t} \tag{3.19}$$

where $\nu_{w,t} = \left(\nu^0_{w,t}, \nu^1_{w,t}, ..., \nu^n_{w,t}\right)'$ distributes i.i.d normal with mean zero and variance-covariance matrix equal to the identity matrix. The matrix φ_w and vectors $\varepsilon_{w,t}$, $\nu_{w,t}$, and M_w are given in Appendix B.

To see how introducing news as anticipated shocks can have real effects, suppose households were to receive information four and eight periods in advance about the time $t+1$ shock, and that the exogenous process is subjected to an unanticipated shock which arrives in time $t+1$. Then the time t expectations about the evolution of process w in time $t+1$ would forecast its value to be

$$\hat{E}_t \ln(w_{t+1}) = \rho_w \ln(w_t) + \nu^4_{w,t+1-4} + \nu^8_{w,t+1-8}$$

while the actual value realized in time $t+1$ would be

$$\ln(w_{t+1}) = \rho_w \ln(w_t) + \nu^0_{w,t+1} + \nu^4_{w,t+1-4} + \nu^8_{w,t+1-8}$$

This enables the consideration of a variety of interesting behavior as households revise their expectations in response to new information. For example, the expected value for the innovation prior to time $t+1$ could fail to materialize, that is, $\nu^0_{w,t+1} = -\left(\nu^4_{w,t+1-4} + \nu^8_{w,t+1-8}\right)$, or could materialize exactly as expected i.e. $\nu^0_{w,t+1} = 0$. More generally one may think of the household continually adjusting expectations throughout time in response to the receipt of new information, and this information being descriptive of the innovation's final value to a varying degree in any given period.

Recall the household's PLM for shadow prices is a linear function of the time $t+1$ endogenous state variables and exogenous processes. Since households are assumed to know the law of motion for all exogenous processes they will incorporate the anticipated shocks directly into their decision rules for time t controls via their forecast of future shadow prices. That is, the time t expectation of future shadow prices based on the PLM $\lambda_t = H_t' \tilde{x}_t$ is

$$\hat{E}_t \lambda_{t+1} = H_t' \hat{E}_t \tilde{x}_{t+1}$$

$$\text{where } \hat{E}_t \tilde{x}_{t+1} = \begin{pmatrix} 1 \\ \hat{E}_t I_t \\ \hat{E}_t K_t \\ \hat{E}_t S_t \\ \hat{E}_t A_{t+1} \\ \hat{E}_t z_{t+1} \end{pmatrix} = \begin{pmatrix} 1 \\ I_t \\ K_t \\ S_t \\ \rho_A A_t + \hat{E}_t \varepsilon^0_{A,t+1} \\ \rho_z z_t + \hat{E}_t \varepsilon^0_{z,t+1} \end{pmatrix}$$

where the household's expectation of the future shock conditional on all information received up to time t, $\hat{E}_t \varepsilon^0_{w,t+1}$, is described above.

Calibration

The model is calibrated using a combination of commonly used values in the literature, estimates obtained from Schmitt-Grohe and Uribe (2012), and steady-state targets for some endogenous variables. $\sigma = 1$ which corresponds to logarithmic utility, $\theta =$

Parameter	Value	Description
σ	1	Intertemporal Elasticity of Substitution
θ	1.4	Frisch-labor Supply Elasticity
γ	0.001	Wealth Elasticity of Labor Supply
β	0.985	Subjective Discount Factor
α	0.64	Steady-state Labor Share
δ_0	0.025	Steady-state Depreciation Rate
u	1	Steady-state Capacity Utilization Rate
h	0.2	Steady-state Labor Supply
κ	1.3	Adjustment Cost Acceleration
ρ_z	0.5	Persistence of Investment-specific TFP Growth
ρ_A	0.9	Persistence of TFP Growth

TABLE 3. Calibrated Parameters for One-sector Model

1.4 so that the wage elasticity of labor supply is 2.5 when the wealth effect of labor supply is shut off, and γ is set to 0.001 which is simultaneously consistent with an extremely small wealth effect of labor supply and a balanced growth path. β is assumed to be 0.985 so that the steady-state gross real interest rate is 1.5 percent, and α is set to 0.64 so that labor's share of output in steady-state is 64 percent. Steady state quarterly depreciation δ_0 is set to 2.5 percent and δ_2 is chosen so that the elasticity of $\delta(u_t)$ is 0.15. δ_1 is calibrated to ensure steady-state capacity utilization equals 1, while the disutility-scale parameter ψ is chosen so that household's spent 20% of their time working. The second derivative of the investment adjustment cost function κ is set to 1.3, though this is subjected to robustness checks since the literature has little to say about this parameter. The autoregressive parameters for growth rates of z_t and A_t are set to 0.5 and 0.9, respectively, which are consistent with the estimated values obtained in Schmitt-Grohe and Uribe (2012). This is summarized in Table **??** The relative importance of anticipated vs surprise shocks for each exogenous process is set consistent with estimates from Schmitt-Grohe and Uribe (2012). In particular, the standard deviations for the surprise components of z_t and A_t are set to 0.21 and 0.65, respectively, while the corresponding standard deviations for the (cumulative) anticipated components are 0.32 and 0.2. This implies the majority (60%) of variation in the growth rate of investment-specific technology is anticipated, while just under 25% of variation in TFP is anticipated.

To develop a benchmark against which to compare the results from SP-learning, I linearize the temporary equilibrium around the non-stochastic steady state. The resulting system of first order expectational difference equations can be easily solved under rational expectations, and the resulting equilibrium is the REE.

The Response to News

I turn now to an exploration of the response to news shocks by key macroeconomic variables in the model. Figure 1 shows the responses by consumption, investment, hours worked, and output when a fully rational household learns at time $t = 0$ that there will be a 1 unit increase in the value of investment-specific or total factor productivity in time $t = 3$, and the expected innovation arrives as expected. In both cases the model generates positive comovement amongst all variables at the time the news is received: good news about the future causes all aggregate variables to rise.

The total response to news about a shock to investment-specific technology z_t is more subdued than that regarding a shock to TFP A_t, because the price of investment affects output only through its effect on capital accumulation while an increase in TFP directly increases output and factor prices everything else held constant. Interestingly this gives rise to a discrepancy in the relative impact of news on the overall response: most of the movement in key variables stemming from news about z_t occurs in the period the news is received, while most of the movement from news about A_t occurs in the period the shock occurs. Clearly including additional exogenous processes will lead to a richer set

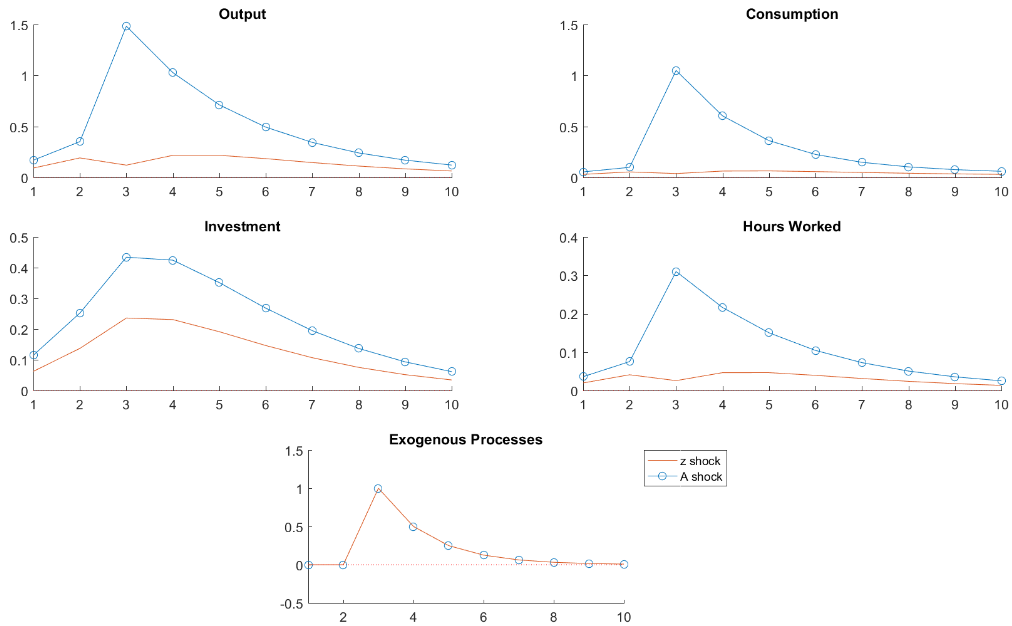

FIGURE 1. IRF for Accurate News

of possible reactions to news of each; indeed Schmitt-Grohe and Uribe (2012) include five additional exogenous disturbances and allow agents to receive information of their future values, in which case one should no longer speak generically of “news”.

Now suppose the news received by households turns out to be completely false. That is, at $t = 0$ households come to expect a 1 unit increase in the exogenous processes will occur in $t = 3$ which does not materialize. Figure 2 plots the IRF resulting from this thought experiment on rational households. Again, the receipt of news generates positive comovement in the key macroeconomic variables. However, once the news is shown to have been erroneously optimistic all variables tend back towards their initial steady state values. Contrary to the prediction of Beaudry and Portier (2004), in this model all variables remain above their steady-state levels for an extended period of time. Thus even in the face of incorrect news the positive comovement amongst variables remains is preserved.

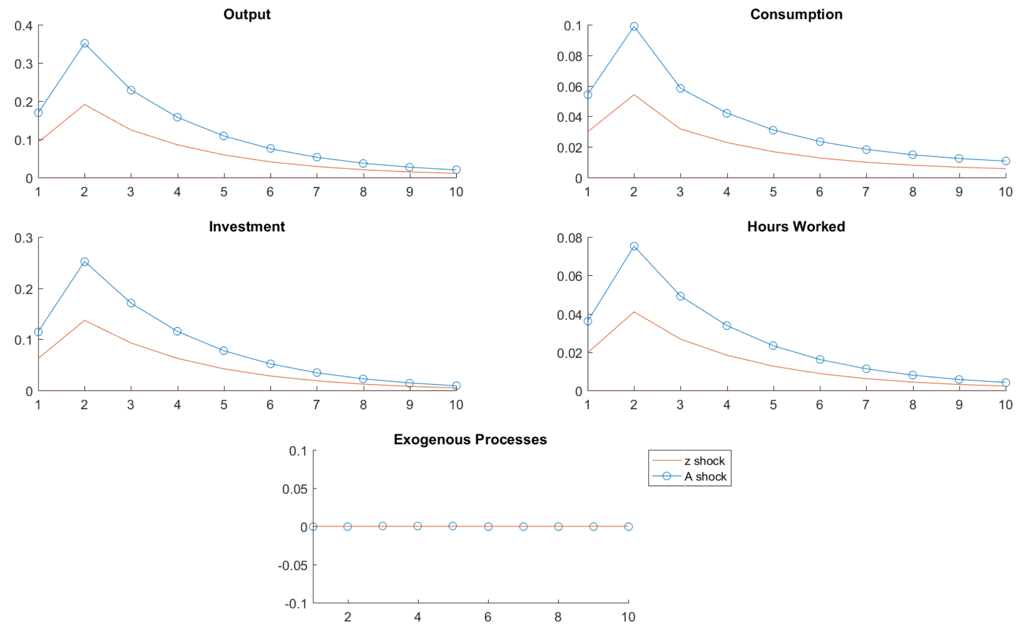

FIGURE 2. IRF for Inaccurate News

Simulation and Model Performance

Having demonstrated the ability of the model to generate qualitatively realistic expectationally driven business cycles in response to news about the future I now turn to an evaluation of the model's ability to generate quantitatively realistic empirical moments under SP-learning. I begin by determining whether the behavior of SP-learners will cause the economy to converge (approximately) to the REE. Figure 3 shows a simulation in which the initial beliefs of agents are perturbations of the linearized RE solution, the economy begins in steady state, and agents update their beliefs with a constant gain. For each variable the solid red line is the RE steady-state value, while the dashed line is the average of the actual variable realizations.

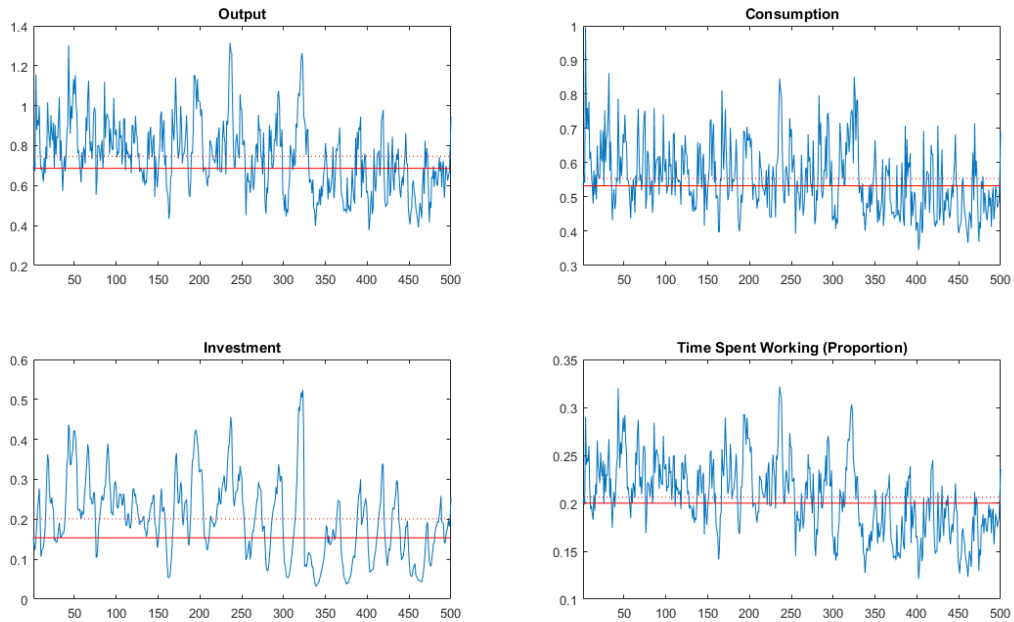

FIGURE 3. Model Simulation Under SP-learning

Since households employ a linear forecasting model in a nonlinear environment, convergence to the rational expectations equilibrium is not to be expected. Indeed, Figure 3 provides strong evidence that the economy converges to some RPE - the key macroeconomic variables all fluctuate around some stationary value - but the implied RPE seems quite far from the REE as judged by the respective central tendencies of the variables.

Tables 4 and 5 perform tests for equality of means for key model variables from simulating the model 1000 times each for 230 periods under varying assumptions regarding households expectations formation mechanism and their access to news. Table 4 displays the cross-sectional means of consumption, investment, hours worked, output, and the endogenous shadow prices under RE and SP-learning when agents receive news. Generally the cross-sectional means are higher under SP-learning than under RE, and the difference

	RE Mean	SPL Mean	pval	95% Low	95% High
Consumption	0.533	0.568	0.000	-0.036	-0.032
Labor Supply	0.200	0.212	0.000	-0.013	-0.012
Investment	0.154	0.172	0.000	-0.019	-0.018
Output	0.687	0.736	0.000	-0.052	-0.046
Investment SP	0.000	0.060	0.000	-0.061	-0.059
Capital SP	4.250	4.043	0.000	0.193	0.219
Habit-adjustment SP	-154.107	-155.120	0.000	0.704	1.321

TABLE 4. t-tests for Data Generating Process, News, 230 Periods

	RE Mean	SPL Mean	pval	95% Low	95% High
Consumption	0.534	0.563	0.000	-0.031	-0.027
Labor Supply	0.200	0.210	0.000	-0.011	-0.010
Investment	0.154	0.169	0.000	-0.016	-0.014
Output	0.688	0.728	0.000	-0.043	-0.038
Investment SP	-0.000	0.034	0.000	-0.035	-0.034
Capital SP	4.248	4.068	0.000	0.167	0.192
Habit-adjustment SP	-154.085	-154.949	0.000	0.558	1.170

TABLE 5. t-tests for Data Generating Process, No News, 230 Periods

is highly statistically significant: the p-value for the t-test of mean equality across expectations formation mechanisms imply rejection of the null hypothesis that the means are equal at any reasonable level of significance. Put another way, I am able to reject the null hypothesis that the data comes from the same data generating process, which of course is true. Similar findings are found for the case where households do not receive news in Table 5. Again the null hypothesis that the simulated data came from the same data generating process can be rejected at any meaningful level of significance.

While the discussion above suggests the rejection of mean equality should not be surprising, it is is worth noting each simulation is run for a relatively short amount of time. In Tables 6 and 7 I increase the simulation length to 1000 periods in order to allow the learning algorithm more time to converge. No significant changes arise: the RPE is statistically quite distinct from the REE.

A deeper inspection of the exact mechanisms driving this wedge between the RPE and REE may proceed from several points of observation. First, the data-generating process under SP-learning is inherently nonlinear. This contrasts with the REE which is obtained via a first-order approximation to the equilibrium dynamics of the model. Exploring higher-order approximations may shed light on precisely how important the nonlinearities are to the data-generating processes. Second, comparing tables 4 and 5 suggest that including anticipated shocks in agents' information sets increases the size of the wedge between the RPE and REE. While the cross-sectional mean values under REE are the same whether news is included or not, those obtained under SP-learning are closer to that of REE when there is no news. This suggest something about the news pushes SP-learning household behavior away from that of their ration counterparts, which is especially interesting given the results of the previous chapter showing news shocks should not matter to the learnability of an REE.

Finally, it appears that SP-learning agents systematically behave in such a way as to cause the shadow price of investment to be positive. This will occur if household behavior is such that the gross level of investment is increasing ($I_t > I_{t-1}$) and capital is being utilized at a rate above the rational expectations steady state ($u_t > 1$), or if the opposite cases are true. It is telling that this is exactly the behavior prescribed by the model in the event that the household receives news about future productivity, and likely explains why news drives a wedge between RE and NRE: under NRE households may go through periods in which they respond too strongly (relative to RE) to news of the future.

In contrast to Williams (2003a), where reduced-form learning converged to the REE and simulated data did not differ meaningfully whether agents were rational or boundedly rational, these results suggest that the convergence to the RPE should cause simulated business cycle statistics to differ from those obtained under RE. To explore this further,

	RE Mean	SPL Mean	pval	95% Low	95% High
Consumption	0.533	0.635	0.000	-0.103	-0.101
Labor Supply	0.200	0.238	0.000	-0.038	-0.037
Investment	0.154	0.192	0.000	-0.039	-0.039
Output	0.687	0.824	0.000	-0.138	-0.136
Investment SP	-0.000	0.117	0.000	-0.118	-0.116
Capital SP	4.251	3.620	0.000	0.625	0.638
Habit-adjustment SP	-154.125	-162.404	0.000	8.017	8.542

TABLE 6. t-tests for Data Generating Process, News, 1000 Periods

	RE Mean	SPL Mean	pval	95% Low	95% High
Consumption	0.533	0.600	0.000	-0.068	-0.066
Labor Supply	0.200	0.223	0.000	-0.024	-0.023
Investment	0.153	0.177	0.000	-0.023	-0.023
Output	0.686	0.774	0.000	-0.088	-0.087
Investment SP	-0.000	0.038	0.000	-0.038	-0.038
Capital SP	4.251	3.797	0.000	0.449	0.460
Habit-adjustment SP	-154.070	-158.385	0.000	4.056	4.573

TABLE 7. t-tests for Data Generating Process, No News, 1000 Periods

Figure 8 compares simulated data from the model with and without news under the assumption of RE and SP-learning against that of actual quarterly U.S. data. The data is from Jaimovich and Rebelo (2009) and covers the range 1944:Q1-2004:Q4. The logarithm of the simulated data is detrended by the Hodrick-Prescott (HP) filter using a smoothing parameter of 1600, consistent with quarterly data. I conducted 1000 simulations of 230 periods each, which is the length of the US data sample. I report the statistics for the cross-sectional average standard deviations of hours worked, investment, and consumption relative to that of output across all simulations along with the correlations between output and hours worked, investment, and consumption.

US data suggests investment is the most volatile, while hours worked tends to be roughly as volatile as output and consumption is much smoother. Significant differences exist in the volatility of simulated data between RE and SP-learning, whether news

	US Data	RE (News)	SPL (News)	RE (No News)	SPL (No News)
σ_h/σ_Y	0.968	0.714	0.716	0.714	0.715
σ_I/σ_Y	3.103	2.386	2.167	2.282	2.023
σ_C/σ_Y	0.712	0.737	0.706	0.746	0.725
$corr(Y,h)$	0.860	1.000	1.000	1.000	1.000
$corr(Y,I)$	0.890	0.850	0.879	0.922	0.912
$corr(Y,C)$	0.770	0.969	0.969	0.977	0.986

TABLE 8. Predicted Business Cycle Statistics

shocks are included or not. All model simulations except SP-learning with news shocks suggest hours worked are much too smooth, and relative volatility of consumption under SP-learning is extremely close to the data. Only the specification with SP-learning and news shocks matches the stylized facts pertaining to the ordering of relative volatilities of investment, hours worked, and consumption. The simulated correlation between hours worked and output is unit, which does not match the data, but is expected: output and hours worked are perfectly correlated by assumption in the RBC model. The correlation between output and investment under SP-learning is closer to that of the data than under RE regardless of whether news is included or not, while both expectations formations mechanisms do a poor job capturing the correlation between output and consumption.

Conclusion

The analysis of this paper has shown the behavior implied by SP-learning households is quite distinct from that of their rational counterparts. While agents in this news-shock model can not converge to the exact REE, the process of making decisions based on linear forecasting rules of shadow prices and updating these forecasts via adaptive least squares leads agents to converge to an RPE. The difference between the implied RPE and the REE seem to be exacerbated by the inclusion of anticipated shocks in agents' information sets. Furthermore, this behavior causes the model to generate simulated moments for key business cycle statistics which are quite distinct from those obtained under RE. In some cases the model under SP-learning appears to better approximate US data, while in others the RE approach is more accurate.

One interesting extension would be to consider informational heterogeneity in this environment. Angeletos and La'O (2010) considers the qualitative and quantitative effects of informational heterogeneity in a Lucas-Phelps "island" model where households on a continuum of islands receive public and private signals about the productivity of aggregate and island-specific productivity. They are somewhat surprised to find that information heterogeneity does not have any serious effect on welfare or quantitative features of the model. But Hellwig (2010) observes this goes back to an insight going back to Hayek (1945) suggesting that markets parsimoniously convey all relevant information through prices, and thus heterogeneity by itself is not likely to represent a significant source of amplification or persistence in simulated data. However, SP-learning introduces a behavioral friction in the way agents utilize information which is likely to be exacerbated by informational frictions stemming from heterogeneity. While it is purely speculative, I suspect SP-learning (and potentially other forms of bounded rationality) may provide a simple means of making informational heterogeneity matter in DSGE models.

CHAPTER IV

LEARNING VS NEWS: WHAT DRIVES BUSINESS CYCLES?

Introduction

Having established in Chapter II that news shocks do not necessarily impinge upon the E-stability properties of a REE, and having conducted a calibrate and simulate exercise in Chapter III to understand the interaction between information flows, bounded rationality, and restricted perceptions equilibrium, I turn now to an empirical exploration of the relationship between news shocks and expectation formation. The central goal of this chapter is to determine whether the estimated importance of news - measured as the relative contribution of anticipated shocks versus surprise shocks in generating macroeconomic volatility - is affected by the way in which agents are assumed to form their expectations.

In what is widely regarded to be the seminal work in the modern empirical news shock literature, Beaudry and Portier (2006) explores the relationship between stock prices, the growth rate of technology, and business cycles using several specifications of VARs. They find evidence that innovations to total factor productivity (TFP) are often anticipated, and further that responses to this anticipation are responsible for a large fraction of business cycle fluctuations. This contrasts with the typical RBC view of technological innovation by surprise which had begun to be questioned as early as Basu et al. (2004) in which improvements to technology are shown to typically have no contemporaneous impact on output.

Beaudry and Lucke (2010) allows news shocks to compete with more traditional sources of volatility in a VAR identified using short-run and long-run timing assumptions and find that surprise shocks play little role in economic volatility compared to news shocks. Barsky and Sims (2012) allow traditional news shocks to compete with "animal spirit shocks" *a la* Lorenzoni (2009). The authors conclude that much of the innovation in measured consumer confidence is due to traditional news shocks and that most of the relation between changes in consumer confidence and economic activity are due to news. In Beaudry et al. (2011) the focus is on what the authors refer to as "optimism shocks" which are essentially sunspots that exogenously shift expectations. These shocks are identified using sign restrictions and are shown to permanently affect current and future economic activity in a way strongly resembling news shocks. The narrative approaches of Alexopoulos (2011) and Ramey (2011) have similarly found support for the notion that news is an important determinant of volatility in aggregate economic variables.

Schmitt-Grohe and Uribe (2012) modifies the well-known RBC-type news-shock model of Jaimovich and Rebelo (2009) to include seven exogenous variables which are subjected to both anticipated and unanticipated shocks. Using likelihood-based classical and Bayesian estimation techniques they find anticipated shocks account for about half of observed volatility in output and consumption. Conversely, Khan and Tsoukalas (2012) allow news shocks to compete against traditional sources of macroeconomic volatility in a small-scale monetary RBC model similar to that considered in Smets and Wouters (2007). Their results imply anticipated shocks are responsible for only 15% of volatility in output. These strongly contrasting results highlight the preeminent role of model specification in DSGE estimation.

While much of the empirical news shock literature has approached the topic from a reduced form or structural VAR perspective, Blanchard et al. (2013) and Beaudry and

Portier (2014) point out estimation via VARs in news-rich environments may suffer from a nonfundamentalness problem. Alternatively, one may estimate the relative importance of anticipated shocks in a news-rich dynamic stochastic general equilibrium (DSGE) model. However, the construction of such a model requires the modeler to make a number of assumptions. One of the most important is the way in which agents' expectations are formed, and all previous authors on the subject have utilized the rational expectations hypothesis.

While much early work on adaptive learning focuses on whether particular *rational expectations equilibrium* (REE) are *expectationally stable* (E-stable), Williams (2003a) and Adam (2005) each explores the impact of adaptive learning on macroeconomic volatility and persistence in calibrated and simulated DSGE models.[1] Their results suggest that relaxing rational expectations (RE) in favor of an adaptive learning approach not only increases persistence and volatility in key macroeconomic variables, but may also provide a better fit to the data than models based on RE.

More recently, adaptive learning has been used in estimated DSGE models. Milani (2007) estimates a monetary-NK DSGE model in which agents are assumed to have what he refers to as "near-rational expectations" (NRE), that is, they form expectations using a correctly specified forecasting model but are unsure of the precise coefficients governing the laws of motion of the economy.[2] In stark contrast to the results obtained under RE, the model under NRE simultaneously fits the data better and ascribes little importance to mechanical sources of volatility and persistence such as habit formation

[1] A REE is said to be E-stable if the beliefs of boundedly rational agents converge (asymptotically) to those of fully rational agents. Besides reflecting the sensitivity of models to the specific assumptions placed on agents' expectation formation mechanisms and serving as an asymptotic justification for the use of RE, E-stability has been proposed as a selection criterion for indeterminate models.

[2] This is also referred to as "Euler equation learning" because agents are learning about endogenous variables which appear in the Euler equations of otherwise rational agents.

and inflation indexation. Similar results obtain in e.g. Milani and Rajbhandari (2012) and Slobodyan and Wouters (2012) in various DSGE models and under various forms of bounded rationality.

Relatedly, Eusepi and Preston (2011) show that periods of optimism and pessimism arising from the erroneous forecasts (and subsequent updating of beliefs) of boundedly rational agents in an otherwise standard stochastic growth model generate the same type of expectationally driven business cycles as those found in news-shock models. This observation presents a natural question: to what extent are estimates of the importance of anticipated shocks impacted by the specific assumptions governing agents' expectation formation mechanism? Do news shocks serve as proxies for an underlying learning process in the data? Or could it be that the combination of news shocks and adaptive learning could actually enhance the relative importance of anticipated shocks by exacerbating the waves of optimism and pessimism?

My main results from this chapter suggest the importance of news is robust to the inclusion of adaptive learning. In particular, the relative importance of news is estimated to be up to thirty-five percent greater for key macroeconomic variables such as output, consumption, investment, and hours worked under NRE than under RE. However, these differences arise from movement in endogenous variables caused by the arrival of the anticipated event as opposed to movement caused by the anticipation itself, what Sims (2016) refers to as the impact of "realized news" and "pure news", respectively.

These findings speak to current differences in structural estimates of the importance of news, which appear to be sensitive to the structure of the model considered. Since parameter estimates obtained from DSGE models under learning often attenuate the importance of specific mechanical sources of volatility and persistence relative to RE, this study will clarify whether the importance of news is also sensitive to specific assumptions governing expectation formations.

Furthermore, it provides additional evidence regarding the merits of using adaptive learning in empirical applications. Despite a preponderance of evidence suggesting its implausibility as a description of individual or aggregate behavior, RE is still the paradigm for describing expectations in dynamic macroeconomic models.[3] It is noteworthy that while the present study suggests the model under learning and RE produce similar model predictions for volatility and persistence, the models produce dissimilar estimates regarding the relative importance of news, implying forecasts generated from the underlying models may be quite different.

The chapter proceeds as follows. I first describe the news-shock RBC-type model which is to be estimated. Next, I describe the specific assumptions made for agent-level expectation formation. This is followed by a detailed description of the estimation methodology and the presentation of the main results for the chapter. The chapter concludes with a discussion of the main results which attempts to place them in the context of recent empirical macroeconomic literature.

The Model

The economy considered is that of Schmitt-Grohe and Uribe (2012) which itself augments the news-shock RBC model of Jaimovich and Rebelo (2009) considered in the previous chapter. Aggregate demand consists of a government which levies lump-sum taxes to finance expenditure and a representative household making the usual

[3]See e.g. Assenza et al. (2014) for a summary of the state of learning-to-forecast and learning-to-optimize laboratory experiments, which have largely failed to find support for RE as a description of aggregate behavior.

consumption/savings and labor supply decisions while also choosing the rate at which to utilize their existing capital stock. Adjustments to gross investment are assumed to be costly. In addition, household preferences incorporate direct and indirect internal habit-formation.

Aggregate supply consists of competitive firms which produce a fungible final good from labor supplied through monopolistically-competitive labor unions and physical capital rented from the household. The model features seven exogenous variables which have been shown in the literature to be empirically important mechanical sources of volatility and persistence. Finally, each of the exogenous variables are subjected to disturbances in the form of anticipated and unanticipated exogenous stochastic white-noise shocks, that is, news and surprise shocks.

Households

The representative household maximizes lifetime utility which is given by

$$E_0 \sum_{t=0}^{\infty} \beta^t \zeta_t U(V_t) \tag{4.1}$$

where $\beta \in (0, 1)$ the household's discount factor, ζ_t a stationary exogenous stochastic time-preference shock, and U the period utility function which takes the CRRA form

$$U(V_t) = \frac{V_t^{1-\sigma} - 1}{1 - \sigma} \tag{4.2}$$

where $\sigma > 0$ the inverse intertemporal elasticity of substitution. The argument of the utility function is given by

$$V_t = C_t - bC_{t-1} - \psi h_t^{\theta} S_t \tag{4.3}$$

where C_t and h_t are private consumption and hours worked, respectively. $b \in [0, 1)$ controls the degree of internal habit formation, $\psi > 0$ scales the disutility of labor supply, and $\theta > 1$ governs the Frisch elasticity of labor supply.[4] S_t is a geometric average of current and past habit-adjusted consumption and evolves according to

$$S_t = (C_t - bC_{t-1})^{\gamma} S_{t-1}^{1-\gamma} \tag{4.4}$$

where $\gamma \in (0, 1]$ governs the magnitude of the wealth elasticity of labor supply.

This preference specification was first used in Jaimovich and Rebelo (2009). The comovement problem in RBC models is partially caused by the wealth effect of labor supply dominating the substitution effect, thereby causing consumption and hours worked to move in opposite directions upon receipt of good news about the future. The JR specification for utility allows the modeler to change the strength of the wealth effect of labor supply by calibrating γ.

Households are assumed to own and rent physical capital to firms in a competitive market. The capital stock accumulates according to

$$K_t = (1 - \delta(u_t)) K_{t-1} + z_t^I I_t \left[1 - \Phi\left(\frac{I_t}{I_{t-1}}\right)\right] \tag{4.5}$$

[4] In the special case where there is no internal or external habit formation i.e. $b = \gamma = 0$, the Frisch labor supply is $\eta^{\lambda} = \frac{1}{\theta - 1}$, hence the restriction that $\theta > 1$

where K_{t-1} the predetermined capital stock determined in period $t - 1$, I_t gross private investment, and z_t^I a stationary exogenous stochastic shock to the technology for converting investment into installed capital.

The function $\Phi(\cdot)$ imposes an investment-adjustment cost on households while $\delta(\cdot)$ implies capital depreciation is a function of the current-period capital utilization rate u_t. Both functions are increasing, convex and take quadratic functional forms. In particular we have

$$\delta(u_t) = \delta_0 + \delta_1(u_t - 1) + \frac{\delta_2}{2}(u_t - 1)^2 \tag{4.6}$$

where $\delta_0 > 0$ the steady-state depreciation rate, $\delta_1 > 0$ determines the steady-state value of u_t, and $\delta_2 > 0$ captures the rental rate elasticity of capacity utilization. The time-varying investment-adjustment cost function is given by

$$\Phi\left(\mu_t^I\right) = \frac{\kappa}{2}\left(\mu_t^I - \mu^I\right)^2 \tag{4.7}$$

consistent with Christiano et al. (2005), where $\mu_t^I \equiv \frac{I_t}{I_{t-1}}$ the growth rate of investment with steady state μ^I and $\kappa > 0$ a scaling parameter which implies the growth rates of investment and output are autocorrelated. This specification implies $\Phi\left(\mu^I\right) = \Phi'\left(\mu^I\right) = 0$, i.e. there are no investment-adjustment costs on a balanced growth path, and $\kappa = \Phi''\left(\mu^I\right)$.

The household receives labor income from supplying h_t hours at rate W_t^*, rental income from from renting $u_t K_{t-1}$ units of effective capital at gross rental rate R_t, and lump-sum dividends from labor-union membership and firm-profits of Π_t. The household uses this income to pay lump-sum taxes T_t and purchase private consumption and gross private investment. The flow budget constraint is thus given by

$$C_t + A_t I_t + T_t \leq W_t^* h_t + R_t \left(u_t K_{t-1}\right) + \Pi_t \tag{4.8}$$

where A_t a difference-stationary exogenous stochastic process representing the current state of technology for producing investment goods from consumption goods with growth rate $\mu_t^A \equiv \frac{A_t}{A_{t-1}}$. In a decentralized equilibrium A_t may be interpreted as the relative price of investment goods in terms of consumption goods; that is, a unit of the investment good may be exchanged for A_t units of the consumption good.

Labor Market

Households are assumed to have market power, albeit in an indirect way. There is a continuum of monopolistically competitive labor unions which households supply labor to, and these unions negotiate wages with final goods producers. The market power of these unions is captured by a stationary exogenous stochastic process ω_t, which can be interpreted simply as a time-varying markup in wages. The solutions to the final-good producers' cost-minimization and unions' profit-maximization problems together imply all unions charge the same price for labor services and pay the same wage to member-households, and hence the wage received by households is given by $W_t^* = W_t/\omega_t$ where W_t the price charged by unions to firms for labor services and $\mu_t^W \geq 1$ the exogenously time-varying wage markup which controls the level of market power the labor unions have with steady-state value μ^W. This friction drives a wedge between the wage paid by firms and that received by workers. Profits of the labor unions are remitted to all member households as a lump-sum dividend. Details of the solution to the optimization problems of firms and unions are presented in Appendix C.

The representative firm contracts h_t worker-hours and rents $u_t K_{t-1}$ units of effective capital to produce a final good Y_t^{out} using CRTS technology according to

$$Y_t^{out} = z_t \left(u_t K_{t-1}\right)^{\alpha_k} \left(X_t h_t\right)^{\alpha_h} \left(X_t F\right)^{1-\alpha_k-\alpha_h} \tag{4.9}$$

where the parameters $\alpha_k, \alpha_h \in (0,1)$ govern the steady-state output shares of effective capital and effective labor, respectively, and satisfy $\alpha_k + \alpha_h \leq 1$. F is a fixed factor (e.g land or managerial capital) which introduces production exhibits diminishing returns to scale in effective capital and labor while ensuring production exhibits overall CRTS. Production is augmented by a stationary exogenous stochastic Hicks-neutral productivity shock z_t and a difference-stationary exogenous stochastic Harrod-neutral (i.e. labor-augmenting) productivity shock X_t with growth rate $\mu_t^X = \frac{X_t}{X_{t-1}}$ and steady-state growth μ^X.

The market for effective capital is competitive and hence the gross rental rate equals the value of the marginal product of effective capital

$$R_t = \alpha_k \frac{Y_t^{out}}{u_t K_{t-1}} \tag{4.10}$$

and the wage paid by firms to labor unions is similarly equal to the value of the marginal product of labor

$$W_t = \alpha_h \frac{Y_t^{out}}{h_t} \tag{4.11}$$

The final good is fungible and may be used for private consumption, private investment, or government spending G_t, and hence the aggregate resource constraint is

$$Y_t^{out} = C_t + A_t I_t + G_t \tag{4.12}$$

where Y_t^{out} is total output. Because A_t and X_t are non-stationary, output displays a stochastic trend which can be shown to be $X_t^Y = X_t A_t^{\alpha_k/(\alpha_k-1)}$. The government maintains a balanced budget such that $G_t = T_t$. Furthermore, government spending exhibits a stochastic trend $X_t^G = \left(X_{t-1}^G\right)^{\rho_{xg}} \left(X_{t-1}^Y\right)^{1-\rho_{xg}}$ which co-integrated with the trend in output so that the share of government spending in output is stationary; however $\rho_{xg} \in [0, 1)$ allows the trend to be smoother than that of output. Detrended government spending $g_t \equiv G_t/X_t^G$ is a stationary exogenous stochastic process with steady state g. Since X_t^G is predetermined, and since detrended government spending is stationary, the current level of government spending is independent of contemporaneous changes to exogenous variables; however, these may affect government spending with a lag via corresponding changes to the trend path of output.

Temporary Equilibrium

The temporary equilibrium for this economy obtains where all households maximize utility given expectations, all firms maximize profits, and all markets clear. This can be described the nonlinear system of expectational difference equations

$$K_t = (1-\delta(u_t))\,K_{t-1} + z_t^I I_t \left[1-\Phi\left(\frac{I_t}{I_{t-1}}\right)\right] \tag{4.13a}$$

$$Y_t^{out} = C_t + A_t I_t + G_t \tag{4.13b}$$

$$Y_t^{out} = z_t\,(u_t K_{t-1})^{\alpha_k}\,(X_t h_t)^{\alpha_h}\,(X_t F)^{1-\alpha_k-\alpha_h} \tag{4.13c}$$

$$S_t = (C_t - bC_{t-1})^{\gamma}\,S_{t-1}^{1-\gamma} \tag{4.13d}$$

$$V_t = C_t - bC_{t-1} - \psi h_t^{\theta} S_t \tag{4.13e}$$

$$\Lambda_t = \left(\zeta_t V_t^{-\sigma} - \frac{\gamma S_t}{C_t - bC_{t-1}}\Pi_t\right) - b\beta\hat{E}_t\left[\left(\zeta_{t+1}V_{t+1}^{-\sigma} - \frac{\gamma S_{t+1}}{C_{t+1}-bC_t}\Pi_{t+1}\right)\right] \tag{4.13f}$$

$$\Lambda_t = \frac{\theta\zeta_t V_t^{-\sigma}\psi h_t^{\theta-1} S_t}{W_t/\omega_t} \tag{4.13g}$$

$$\Pi_t = \zeta_t V_t^{-\sigma}\psi h_t^{\theta} + \beta\hat{E}_t\left[(1-\gamma)\frac{S_{t+1}}{S_t}\Pi_{t+1}\right] \tag{4.13h}$$

$$Q_t\Lambda_t = \beta\hat{E}_t\Lambda_{t+1}\left[(R_{t+1}u_{t+1} + Q_{t+1}(1-\delta(u_{t+1})))\right] \tag{4.13i}$$

$$R_t = \delta'(u_t)Q_t \tag{4.13j}$$

$$A_t\Lambda_t = Q_t\Lambda_t z_t^I\left(1-\Phi\left(\frac{I_t}{I_{t-1}}\right) - \Phi'\left(\frac{I_t}{I_{t-1}}\right)\left(\frac{I_t}{I_{t-1}}\right)\right) \tag{4.13k}$$

$$+\,\beta\hat{E}_t\left[Q_{t+1}\Lambda_{t+1}z_{t+1}^I\Phi'\left(\frac{I_{t+1}}{I_t}\right)\left(\frac{I_{t+1}}{I_t}\right)^2\right] \tag{4.13l}$$

$$W_t = \alpha_h\frac{Y_t^{out}}{h_t} \tag{4.13m}$$

$$R_t = \alpha_h\frac{Y_t^{out}}{u_t K_{t-1}} \tag{4.13n}$$

where Λ_t the marginal value of income, Π_t the marginal value of household's stock of current and past habit-adjusted consumption (i.e. the Lagrange multiplier from relaxing the constraint on S_t), and $\Lambda_t Q_t$ the marginal value of pre-installed capital. Q_t can thus be interpreted as marginal Tobin's Q. $\hat{E}_t$ denotes the (possibly non-rational) expectations of the household given their time t beliefs and information set. The temporary equilibrium can be concisely described by

$$\hat{E}_t \left[f \left(Y_{t+1}, Y_t, Y_{t-1}, \nu_t \right) \right] = 0 \tag{4.14}$$

where Y_t and ν_t are collections of endogenous and exogenous variables and exogenous stochastic white-noise shocks, respectively.

Stationarity, Information Flows, and the Log-linearized System

Because A_t and X_t are non-stationary the system (4.14) must be transformed to be in terms of stationary variables. The details of this transformation and the resulting stationary nonlinear system of expectational difference are presented in Appendix C. Letting $\hat{Y}_t$ be a collection of (detrended) endogenous and exogenous variables, the resulting stationary system can be concisely describe as

$$\hat{E}_t \left[f \left(\hat{Y}_{t+1}, \hat{Y}_t, \hat{Y}_{t-1}, \nu_t \right) \right] = 0 \tag{4.15}$$

$\hat{Y}_t$ includes the seven stationary exogenous variables $x = \left\{ z, \zeta, z^I, g, \omega, \mu^A, \mu^X \right\}$ which are assumed to follow AR(1) processes such that

$$\ln x_t = \rho_x \ln x_{t-1} + \epsilon^0_{x,t}$$

where $\rho_x \in (0, 1)$. In addition to an unanticipated or "surprise" shock, households receive information about each of the exogenous variables 4 and 8 periods in advance. These news shocks can be parsimoniously modeled by introducing auxiliary variables $\varepsilon^n_{x,t}$ for

$n = 0, 1, ..., q$ where q the longest forecasting horizon and

$$\varepsilon_{x,t}^{0} = \varepsilon_{x,t-1}^{1} + \sigma_x^0 \nu_{x,t}^0$$
$$\varepsilon_{x,t}^{1} = \varepsilon_{x,t-1}^{2}$$
$$\varepsilon_{x,t}^{2} = \varepsilon_{x,t-1}^{3}$$
$$\varepsilon_{x,t}^{3} = \varepsilon_{x,t-1}^{4}$$
$$\varepsilon_{x,t}^{4} = \varepsilon_{x,t-1}^{5} + \sigma_x^4 \nu_{x,t}^4$$
$$\varepsilon_{x,t}^{5} = \varepsilon_{x,t-1}^{6}$$
$$\varepsilon_{x,t}^{6} = \varepsilon_{x,t-1}^{7}$$
$$\varepsilon_{x,t}^{7} = \varepsilon_{x,t-1}^{8}$$
$$\varepsilon_{x,t}^{8} = \sigma_x^8 \nu_{x,t}^8$$

where $\nu_{x,t}^k$ for $k = 0, 4, 8$ are assumed to be Gaussian white-noise shocks with unit variance and effective standard deviation σ_x^k. Repeated substitution of the auxiliary variables implies $\varepsilon_{x,t}^0 = \sigma_x^0 \nu_{x,t}^0 + \sigma_x^4 \nu_{x,t-4}^4 + \sigma_x^8 \nu_{x,t-8}^8$, and hence household expectations are given by

$$\hat{E}_t \varepsilon_{x,t+n}^0 = \begin{cases} 0 & \text{if } n > 8 \\ \sigma_x^8 \nu_{x,t+n-8}^8 & \text{if } 8 \geq n > 4 \\ \sigma_x^4 \nu_{x,t+n-4}^4 + \sigma_x^8 \nu_{x,t+n-8}^8 & \text{if } 4 \geq n > 1 \end{cases}$$

i.e. household's expectations of future exogenous variables are constantly updated as news shocks flow into their information sets.

While the anticipated components do not affect contemporaneous economic fundamentals, the information reveals something about the future state of the economy. Households incorporate this information into their decision rules, and it is precisely the fact that news shocks affect various endogenous variables in different ways at different times which allows the econometrician to identify the effect of a particular news shock.[5] Note that the inclusion of such news shocks implies no departure from rational expectations: The anticipated shocks are fundamental to the model, and hence rational households must incorporate them into their forecasts of the future.

The system describing the informational flow structure can be represented by collecting auxiliary variables into vectors $\varepsilon_t^x = \left(\varepsilon_{x,t}^0, \varepsilon_{x,t}^1, ..., \varepsilon_{x,t}^8\right)'$, the Gaussian white-noise shocks into vectors $\nu_t^x = \left(\nu_{x,t}^0, \nu_{x,t}^1, ..., \nu_{x,t}^8\right)'$, and writing

$$w_t = \varphi w_{t-1} + M\nu_t \tag{4.16}$$

where $w_t = \left(\varepsilon_t^z, \varepsilon_t^\zeta, \varepsilon_t^{z^I}, \varepsilon_t^g, \varepsilon_t^\omega, \varepsilon_t^{\mu^A}, \varepsilon_t^{\mu^X}\right)'$ and $\nu_t = \left(\nu_t^z, \nu_t^\zeta, \nu_t^{z^I}, \nu_t^g, \nu_t^\omega, \nu_t^{\mu^A}, \nu_t^{\mu^X}\right)'$. Given the assumed structure of information flows, φ is an upper-shift matrix with 1's on the super-diagonal and zeros elsewhere while M is simply a sparse matrix containing the effective standard deviations of the anticipated and unanticipated shocks.[6]

Following the majority of applied structural macroeconomic literature, I log-linearize the stationary system (4.15) around its non-stochastic steady state. Denote by $\tilde{y}_t$ the collection of endogenous and exogenous variables in terms of their percent-deviation from

[5] See Schmitt-Grohe and Uribe (2012) or Beaudry and Portier (2014) for more on the issue of identification in news-shock DSGE models.

[6] In what follows I assume the exogenous stochastic shocks are pure white-noise and hence are uncorrelated across forecasting horizons and across exogenous variables. It would be simply to modify these assumptions, however, and would amount to appropriately changing φ and M to capture the new relationships.

non-stochastic steady state. Then the log-linearized system of equations, including those defining the information flow structure, can be written compactly as

$$\tilde{y}_t = \alpha + \beta \tilde{y}_{t-1} + \chi w_t + \delta \hat{E}_t \tilde{y}_{t+1} \tag{4.17a}$$

$$w_t = \varphi w_{t-1} + M \nu_t \tag{4.17b}$$

A variety of well-known solution techniques exist for obtaining the solution to a linear system of first-order expectational difference equations like that defined by equations (4.17a) and (4.17b) under the assumption of rational expectations e.g. Blanchard and Kahn (1980), Uhlig et al. (1995), Klein (2000), or Sims (2002). However, the objective of this chapter is to compare and contrast the estimated importance of anticipated shocks when agents are assumed to be rational versus when the rational expectations hypothesis is slightly relaxed. Thus, the next section describes the assumed expectation formation mechanism and adaptive learning process for the boundedly rational agents.

Adaptive Learning

As is clear from the temporary equilibrium described by (4.17a), household actions are conditional on their expectations of the future state of the economy. I depart from RE by assuming households behave as econometricians equipped with a forecasting model. Their beliefs, referred to as the perceived law of motion (PLM), take the form of the minimum-state variable (MSV) solution obtained under rational expectations; that is, agents form forecasts according to

$$\begin{aligned} \tilde{y}_t &= a_{t-1} + b_{t-1} \tilde{y}_{t-1} + c_{t-1} w_t \\ w_t &= \varphi w_{t-1} + M \nu_t \end{aligned} \tag{4.18}$$

This model is "correct" in the sense that its structure is consistent with the REE; there are no omitted or included irrelevant explanatory variables.[7] However, the exact values may be different from those implied by the REE. The model is closed by assuming households update their beliefs over time according to the constant-gain least squares (CGLS) algorithm. Denoting household beliefs as $\xi_t' = (a_t, b_t, c_t)$ we have

$$\xi_t = \xi_{t-1} + \mathbf{\bar{g}}\hat{R}_t^{-1} Z_{t-1} \left(\tilde{y}_t - \xi_{t-1}' Z_{t-1}\right)' \tag{4.19}$$

$$\hat{R}_t = \hat{R}_{t-1} + \mathbf{\bar{g}} \left(Z_{t-1} Z_{t-1}' - \hat{R}_{t-1}\right) \tag{4.20}$$

where the data used to forecast is captured by $Z_{t-1}' = \left(1, \tilde{y}_{t-1}', w_t'\right)$ and $\mathbf{\bar{g}}$ the "constant-gain" parameter which describes the relative importance of recent forecast errors. The system (4.19)-(4.20) defines a household's time t estimate of the coefficients ξ_t and the matrix of second-moments $\hat{R}_t$ as a weighted average of their previous estimates and the new information contained in the forecast error $\epsilon_t = \tilde{y}_t - \xi_{t-1}' Z_{t-1}$. Note that setting $\mathbf{\bar{g}} = t^{-1}$ would result in agents using the well-known recursive least squares (RLS) algorithm for updating beliefs; furthermore, if the REE is *expectationally stable* (E-stable) then it obtains as a special limiting case where $t \to \infty$ and $\mathbf{\bar{g}} \to 0$.

[7] I am implicitly assuming household's understand the informational flow structure; that is, they understand which news is relevant for which exogenous variables and that they receive news 0, 4, and 8 periods in advance. This is an intuitively appealing assumption, and also a weak one, given that the lack of a feedback-loop implies agents would very quickly learn the true laws of motion governing the evolution of this exogenous system.

Given the PLM described by equation (4.18) and assuming agents do not observe the contemporaneous value of endogenous variables i.e. $\hat{E}_t\tilde{y}_t \neq \tilde{y}_t$[8], expectations are given by

$$\hat{E}_t\tilde{y}_t = a_{t-1} + b_{t-1}\tilde{y}_{t-1} + c_{t-1}w_t \tag{4.21}$$

$$\begin{aligned} \hat{E}_t\tilde{y}_{t+1} &= a_{t-1} + b_{t-1}\hat{E}_t\tilde{y}_t + c_{t-1}\varphi w_t \\ &= \left[\left(I + b_{t-1}\right)a_{t-1}\right] + b_{t-1}^2\tilde{y}_{t-1} + \left[b_{t-1}c_{t-1} + c_{t-1}\varphi\right]w_t \end{aligned} \tag{4.22}$$

which can be substituted into the system (4.17a) to yield the actual law of motion (ALM) for the economy

$$\begin{aligned} \tilde{y}_t &= \left[\alpha + \delta\left(I + b_{t-1}\right)a_{t-1}\right] + \left[\beta + \delta b_{t-1}^2\right]\tilde{y}_{t-1} + \left[\chi + \delta\left(b_{t-1}c_{t-1} + c_{t-1}\varphi\right)\right]w_t \\ w_t &= \varphi w_{t-1} + M\nu_t \end{aligned} \tag{4.23}$$

Equation (4.23) emphasizes the feedback mechanism implied by adaptive learning: household perceptions of the economy - captured by $a_{t-1}, b_{t-1},$ and c_{t-1} - directly impact the actual state of the economy through their influence over household decision making. Collecting the endogenous, exogenous, and auxiliary variables into a vector $\mathbb{S}_t = \left(\tilde{y}_t, w_t\right)'$, the ALM for the economy can be written

$$\mathbb{S}_t = A_t + F_t\mathbb{S}_{t-1} + G_t\nu_t \tag{4.24}$$

where $\nu_t \overset{i.i.d}{\sim} N\left(0, I\right)$ and $A_t, F_t,$ and G_t are time-varying coefficient matrices formed from deep structural parameters of the economy and household beliefs. The central goal of this chapter is to compare estimates of the deep parameters obtained under RE against those

[8]This distinction does not matter under RE as all agents are endowed with model-consistent expectations, and at first blush may appear strange in the context of a representative agent model. However, it is a natural assumption to make in general equilibrium settings where individual agents must forecast the aggregate behavior of all other agents to determine their individual behavior.

obtained under near-rational expectations, specifically those corresponding to the effective standard deviation of the anticipated and unanticipated shocks.

Estimation Methodology

Following the recent literature in estimating DSGE models with learning e.g. Milani (2007), Slobodyan and Wouters (2012), and Milani and Rajbhandari (2012), I choose to estimate the parameters of (4.27) using Bayesian Markov Chain Monte Carlo (MCMC) methods. In the present context there are three main advantages to Bayesian versus classical maximum likelihood estimation.

First, as described in An and Schorfheide (2007), maximum likelihood procedures do not allow an econometrician to avoid estimates which are known to be unlikely given previously acquired institutional knowledge. Bayesian estimation allows for the augmentation of the data through a joint prior distribution which re-weights the likelihood. In this way the econometrician can easily and transparently allow information which is not contained in the data to have an impact on the estimated results.

Second, Bayesian methods result in an entire posterior distribution, as opposed to point estimates obtained under maximum likelihood. This emphasizes the fact that our models merely approximate the economy by allowing the econometrician to report results in terms of probabilities rather than a single "true" parameter value. This facilitates an honest communication of results in line with the norms in a variety of disciplines ranging from political science to meteorology.

Finally, the likelihood functions associated with DSGE models can present major problems for optimization routines. Because these models are approximations of a data generating process, the corresponding likelihood functions often feature many peaks and cliffs. MCMC approaches such as the Random-Walk Metropolis Hastings (RWMH) algorithm are robust to these topographic challenges as their success derives precisely because they explore the entire parameter space, thereby eliminating the well-known problem of Newtonian optimizers getting stuck in inappropriate spaces due to initial starting values.

The object of interest is the posterior distribution $p(\Theta|Y^T)$ which describes the probability of a particular parameter constellation Θ given data Y^T. The posterior distribution is obtained through Bayes' Law as

$$p(\Theta|Y^T) \propto p(Y^T|\Theta)p(\Theta) \tag{4.25}$$

where the likelihood function $p(Y^T|\Theta)$ captures the probability of observing the data in Y^T given Θ and the prior distribution $p(\Theta)$ summarizes the econometrician's *a priori* knowledge of the parameters Θ. In general the posterior distribution does not take any known form, and hence I utilize the RWMH algorithm to sample from the target distribution.

Because the resulting sequence of random samples is a Markov chain, the proposal distribution should be chosen to ensure proposals are accepted neither too often nor too infrequently. If samples are accepted too frequently then the sequence will be highly serially correlated; if samples are accepted too infrequently than the algorithm will fail to explore the entirety of parameter space. But choosing a proposal distribution to yield an acceptance rate which satisfies this criteria can be extremely difficult, especially in high-dimensional settings, because the various elements may behave in strikingly different ways.

To overcome this challenge I utilize the Log Adaptive Metropolis (LAP) procedure of Shaby and Wells (2010). In addition to recursively updating the variance-covariance matrix of the proposal distribution vis-a-vis the adaptive Metropolis algorithm of Haario et al. (2001), LAP multiplicatively updates the tuning parameter proportionally to the deviation of the actual acceptance rate from its target. This results in a proposal distribution which quickly establishes an acceptance rate in close proximity to the desired one, while also respecting the relationships between estimated parameters revealed by the history of draws.

The Observables

To estimate the DSGE model given by (4.24) I must specify a link between the unobserved state-variables of the model and their observed empirical counterparts. I utilize quarterly U.S. data provided by Schmitt-Grohe and Uribe (2012) on the demeaned growth rates of real per-capita output, consumption, investment, government expenditure, and hours worked along with the growth rates of total factor productivity (TFP) and the relative price of investment denoted g_t^Y, g_t^C, g_t^I, g_t^G, g_t^h, g_t^{TFP}, and $g_t^{P_I}$, respectively. I use this data to facilitate comparisons between my results and previous estimates of the relative importance of anticipated shocks.

The data runs from 1955:Q2 to 2006:Q4 which corresponds to 207 observations.[9] The growth rate of real per-capita output is assumed to be measured with error, which is required by the fact that the RBC model implies a linear restriction between these seven

[9]Because the model does not feature structural change I am implicitly assuming the nature of "news" has remained the same over the entire period. This may not be an entirely satisfying assumption, and future work will examine the impact of splitting the sample *a la* Beaudry et al. (2011) in which optimism shocks are found to be much more important in their post-1983 subsample than in their pre-1983 subsample.

observables which is not satisfied by the data.[10] Formally, the observable variables are linked to the non-stationary variables from (4.13a)-(4.13n) by

$$\begin{pmatrix} g_t^Y \\ g_t^C \\ g_t^I \\ g_t^g \\ g_t^h \\ g_t^{TFP} \\ g_t^{P_I} \end{pmatrix} = \begin{pmatrix} \Delta \log (Y_t) \\ \Delta \log (C_t) \\ \Delta \log (A_t I_t) \\ \Delta \log (G_t) \\ \Delta \log (h_t) \\ \Delta \log \left(z_t X_t^{1-\alpha_k} \right) \\ \Delta \log (A_t) \end{pmatrix} + \begin{pmatrix} \sigma_{me}^Y & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix} \begin{pmatrix} e_t^Y \\ e_t^C \\ e_t^I \\ e_t^g \\ e_t^h \\ e_t^{TFP} \\ e_t^{P_I} \end{pmatrix}$$

or concisely as

$$\mathbb{Y}_t = H\mathbb{S}_t + De_t \tag{4.26}$$

where $e_t \overset{i.i.d}{\sim} N(0, I)$ a vector of white-noise measurement error shocks and σ_{me}^Y the effective standard deviation of measurement error in output growth. H is a selection matrix of which defines the observed variables in terms of the unobserved states. The system of transition equations (4.24) together with the system of measurement equations (4.26) imply estimation can be conducted by analyzing the (potentially time-varying) state-space model

$$\begin{aligned} \mathbb{S}_t &= A_t + F_t \mathbb{S}_{t-1} + G_t \nu_t \\ \mathbb{Y}_t &= H\mathbb{S}_t + De_t \end{aligned} \tag{4.27}$$

[10]This is the well-known *stochastic singularity* problem in estimating linearized DSGE models is called. See e.g. Ruge-Murcia (2007) for more on this and its impact on estimation.

which is a linear-Gaussian system in the econometric-state variables $\mathbb{S}_t$ and variables with directly observed counterparts $\mathbb{Y}_t$.

The Likelihood Function

Given the state-space model in (4.27), the value of the likelihood function evaluated at a given parameter constellation Θ can be easily calculated using the Kalman Filter. The introduction of adaptive learning requires the coefficient-matrices of the transition equations to be updated using estimates of the state variables based on the CGLS algorithm. I follow Slobodyan and Wouters (2008) and use the simple filtered estimates for the states rather than the smoothed estimates.

The additional updating step presents two potential computational difficulties. First, agents updated estimates may imply the system is non-stationary; thus I incorporate a *projection facility* as in Marcet and Sargent (1989a) which causes agents to skip their updating step when they realize their new estimates are non-nonsensical. Second, MSV learning in linearized-DSGE models can produce estimated second-moment matrices with very small minimum eigenvalues due to perfect-multicollinearity from the models' implied linear restrictions; hence, I incorporate a *ridge correction mechanism* as in Slobodyan and Wouters (2012) which conditionally adds an arbitrarily small constant to the diagonal of the estimated second-moment matrix.[11]

Provided initial values for the states $\mathbb{S}_{0|0}$, the estimated mean square error (MSE) matrix $\mathbb{P}_{0|0}$, beliefs for the second moment matrix $\hat{R}_0$, and beliefs for the coefficient matrices $x_0 = (a_0, b_0, c_0)$, the Kalman Filtering routine proceeds by repeating 5 steps

[11] I set the value of this small ridge to be $1e-6$ times a conformable identity matrix.

1. Calculate the *a priori* estimates of the time t states $\mathbb{S}_{t|t-1}$ and MSE matrix $\mathbb{P}_{t|t-1}$

2. Use these estimates to produce *a priori* estimates of the time t observables $\mathbb{Y}_{t|t-1}$ and corresponding variance-covariance matrix $\Omega_{t|t-1}$

3. Calculate the *a posteriori* time t value of the (log) likelihood function

4. Update the estimated coefficient matrices A_t, F_t, and G_t implied by agents updating their beliefs $\hat{R}_t$ and ξ_t according to the CGLS algorithm

5. Calculate the *a posteriori* estimates of the time t states $\mathbb{S}_{t|t}$ and MSE matrix $\mathbb{P}_{t|t}$

for $t = 1, 2, ...T$. The projection facility and ridge-correction mechanism are conditionally applied during step 4, which does not exist under RE. Initial values for the states and estimated MSE matrix are set to the unconditional mean and variance-covariance matrix of the states, respectively, while initial beliefs are assumed to be equal to the REE. Since the REE is E-stable, this implies agents' estimates will fluctuate around the REE, thereby facilitating a simple comparison between the estimated posterior densities.

The Prior

The parameter vector Θ contains all of the parameters of the model, some of which are calibrated. In particular, the discount factor β is set to 0.99 implying a steady-state quarterly real-interest rate of 1% if $\delta_1 = \delta_2 = 0$. σ is set to 1 to yield logarithmic utility while α_k and α_h are set to 0.225 and 0.675, respectively, implying mild decreasing returns to scale. δ_0 is the steady-state depreciation rate and is set to 2.5 percent. The fixed factor F and steady-state values for the Hicks-neutral technology shock z, preference shock ζ, and investment-specific technology shock z^I are normalized to 1. The steady-state value

Calibrated Parameters

Parameter	Values	Description
β	0.99	Discount factor
σ	1	Logarithmic Utility
α_k	0.225	Effective Capital Share of Output
α_h	0.675	Effective Labor Share of Output
δ_0	0.025	Steady-state Depreciation Rate
F	1	Fixed Factor
z	1	Hicks-neutral technology shock
ζ	1	Preference shock
z^I	1	Investment-specific technology shock
ω	1.15	Wage-markup shock
μ^A	0.9972	Growth rate, Relative Price of Investment shock
μ^X	1.0032	Growth rate, Harrod-neutral technology shock
h	0.2	Steady-state Hours Worked
u	1	Steady-state Capacity Utilization Rate

TABLE 9. Calibrated Parameters for Estimated One-sector Model

for the wage-markup shock ω is set to 1.15 and g is set such that government's share of output is 0.20 in steady state. The steady-state values of the growth of Harrod-neutral technology μ^X and the relative price of investment μ^A are set to 1.0032 and 0.9957, respectively, to match the long-run averages implied by the data. Finally, δ_1 and ψ are calibrated to ensure capacity utilization and hours worked in steady state are 1 and 0.2, respectively. This information is summarized in Table 9.

Each of the remaining parameters which are to be estimated require an prior distribution. These priors may reflect knowledge about the parameters obtained from previous studies, and often help ensure the estimation routine focuses on "sensible" values e.g. ensuring standard deviations are non-negative. I choose my priors to be consistent with those in Schmitt-Grohe and Uribe (2012). In particular, the parameter θ governing the Frisch elasticity of labor supply is distributed Gamma with mean 5 and standard deviation 2, the curvature of the depreciation function δ_2 is distributed inverse-Gamma with mean 0.01 and standard deviation 0.05, and the curvature of the investment adjustment cost function κ is distributed Gamma with mean 10 and standard deviation 2. These correspond to rather loose priors, and the inverse-Gamma on κ ensures it will be strictly positive. The standard deviations for the structural shocks are assumed to follow Gamma distributions to allow for the possibility that some shocks are completely unimportant i.e. have standard deviations equal to zero. The surprise shocks and output's measurement error have mean and standard deviation equal to 0.1, and the anticipated shocks have mean and standard deviation such that surprise shocks account for 50% of total variance in each exogenous variables at the prior.

The internal habit adjustment parameter b as well as the autoregressive parameters for each shock ρ_x are assumed to follow Beta distributions with mean 0.5 and standard deviation 0.2 which correspond to hump-shaped prior distributions over the unit interval. The habit-adjustment term γ from the JR preferences is uniformly distributed over the unit interval. Finally, the constant-gain parameter $\bar{\mathbf{g}}$ is distributed Gamma with mean and standard deviation 0.035, which is consistent with priors used in recent learning-estimation literature while placing non-trivial probability at near-zero values.

Main Results

Table 10 reports the median and 95% highest posterior distribution interval (HPDI) implied by the priors and 500,000 draws from the estimated posterior distribution under RE and NRE. The most notable differences between models occur with respect to the estimated standard deviations for the wage-markup shock (σ_ω^0,σ_ω^4,σ_ω^8) and the growth rate of the relative price of investment ($\sigma_{\mu^A}^0$,$\sigma_{\mu^A}^4$,$\sigma_{\mu^A}^8$). Under RE, surprise shocks are responsible for about half of the variance in the wage-markup shock at the posterior median, while under NRE they are responsible for almost none. A similar shift occurs

Prior and Posterior Distribution for Estimated Parameters

	Prior				**Posterior - RE**			**Posterior - Learning**		
Param.	Distr.	95% Low	Med.	95% High	95% Low	Med.	95% High	95% Low	Med.	95% High
σ_z^0	Gamma	0.000	0.069	0.303	0.649	0.730	0.814	0.639	0.733	0.824
σ_z^4	Gamma	0.000	0.049	0.214	0.000	0.057	0.252	0.000	0.062	0.283
σ_z^8	Gamma	0.000	0.049	0.212	0.000	0.051	0.194	0.000	0.051	0.249
σ_ζ^0	Gamma	0.000	0.069	0.300	2.795	3.398	4.081	2.741	3.352	4.043
σ_ζ^4	Gamma	0.000	0.049	0.214	0.000	0.051	0.215	0.000	0.049	0.219
σ_ζ^8	Gamma	0.000	0.049	0.215	0.000	0.050	0.219	0.000	0.051	0.205
$\sigma_{z^I}^0$	Gamma	0.000	0.069	0.302	4.701	5.599	6.563	4.634	5.602	6.560
$\sigma_{z^I}^4$	Gamma	0.000	0.049	0.214	0.000	0.051	0.230	0.000	0.047	0.215
$\sigma_{z^I}^8$	Gamma	0.000	0.049	0.214	0.000	0.051	0.211	0.000	0.047	0.207
σ_ω^0	Gamma	0.000	0.069	0.302	0.000	0.206	1.115	0.000	0.079	0.410
σ_ω^4	Gamma	0.000	0.049	0.212	0.000	0.153	1.103	0.000	0.054	0.245
σ_ω^8	Gamma	0.000	0.049	0.215	0.000	0.133	1.091	0.917	1.038	1.169
σ_g^0	Gamma	0.000	0.069	0.302	0.000	0.073	0.309	0.000	0.067	0.279
σ_g^4	Gamma	0.000	0.049	0.214	3.686	4.236	4.816	3.718	4.289	4.900
σ_g^8	Gamma	0.000	0.049	0.212	0.000	0.051	0.223	0.000	0.053	0.246
$\sigma_{\mu^A}^0$	Gamma	0.000	0.069	0.303	0.015	0.277	0.372	0.001	0.214	0.352
$\sigma_{\mu^A}^4$	Gamma	0.000	0.049	0.212	0.000	0.115	0.341	0.000	0.180	0.352
$\sigma_{\mu^A}^8$	Gamma	0.000	0.049	0.213	0.000	0.083	0.306	0.000	0.096	0.337
$\sigma_{\mu^X}^0$	Gamma	0.000	0.069	0.304	0.032	0.078	0.149	0.024	0.076	0.148
$\sigma_{\mu^X}^4$	Gamma	0.000	0.049	0.214	0.000	0.021	0.065	0.000	0.024	0.074
$\sigma_{\mu^X}^8$	Gamma	0.000	0.049	0.215	0.000	0.044	0.124	0.000	0.044	0.125
σ_{me}^Y	Gamma	0.000	0.069	0.305	0.541	0.594	0.651	0.542	0.594	0.652
$\rho_{x,g}$	Beta	0.125	0.500	0.871	0.117	0.420	0.733	0.131	0.439	0.738
ρ_z	Beta	0.131	0.501	0.873	0.797	0.858	0.914	0.798	0.861	0.920
ρ_ζ	Beta	0.129	0.502	0.869	0.015	0.109	0.231	0.013	0.103	0.225
ρ_{z^I}	Beta	0.127	0.499	0.868	0.729	0.798	0.868	0.724	0.796	0.864
ρ_ω	Beta	0.115	0.500	0.858	0.917	0.950	0.980	0.923	0.955	0.985
ρ_g	Beta	0.132	0.499	0.873	0.903	0.942	0.979	0.906	0.944	0.979
ρ_{μ^A}	Beta	0.128	0.500	0.869	0.351	0.465	0.575	0.359	0.469	0.580
ρ_{μ^X}	Beta	0.132	0.500	0.873	0.889	0.951	0.990	0.877	0.950	0.988
θ	Gamma	1.503	4.760	8.961	2.936	3.696	4.490	3.057	3.803	4.645
κ	Gamma	3.171	9.479	18.048	4.709	5.889	7.203	4.625	5.713	7.017
δ_2	InvGamma	0.000	0.000	0.002	0.030	0.035	0.042	0.029	0.035	0.041
γ	Unif	0.001	0.499	0.951	0.000	0.001	0.008	0.000	0.001	0.006
b	Beta	0.503	0.707	0.885	0.813	0.841	0.867	0.812	0.840	0.865
$\bar{\mathbf{g}}$	Gamma	0.000	0.024	0.106	0	0	0	0.000	0.007	0.025

TABLE 10. Estimated Parameters

in the implied variance for the growth rate of the relative price of investment: under RE the surprise shock is responsible for about 80% of total variance, while it is responsible for only about 50% under NRE. Thus, the learning model ascribes greater importance to anticipated shocks for driving exogenous variables than the model with RE.

That the values for mechanical sources of volatility and persistence such as the habit formation parameter b and the rental rate elasticity of capacity utilization δ_2 are so similar across models is somewhat surprising. Previous estimation-with-learning exercises as in Milani and Rajbhandari (2012) and Slobodyan and Wouters (2012) found that mechanical sources of volatility and persistence failed to maintain their importance when agents were assumed to be boundedly rational. However, my results are consistent with the earlier finding from Slobodyan and Wouters (2007) that differences in estimates between RE and models with learning are to a large extent driven by differences in the information sets of agents. Because agents' initial beliefs correspond to the REE the information set of agents under NRE is quite similar to that of their rational counterparts.

To assess the quantitative importance of anticipated shocks versus surprise shocks, I conduct a forecast-error variance decomposition (FEVD) exercise on the implied growth rates of output, consumption, investment, hours worked, and government spending. For surprise shocks typical FEVDs have a simple interpretation: because the shock is realized in the same period it is announced, the induced volatility can be directly attributed to the change in fundamentals.

Anticipated shocks, however, have two separate impacts on the system. Sims (2016) refers to these impacts as "pure news' and "realized news", respectively. Pure news captures volatility resulting from the change in agents' expectations of future fundamentals before those fundamentals have actually changed. Realized news captures volatility resulting from the actual arrival of the anticipated event. For instance, if news arrives four periods in advance, horizons up to four periods will represent pure news, while horizons after four periods will represent both pure and realized news.

Since an unconditional FEVD is an asymptotic result, it will mix these two effects. If we are interested in the relative importance of anticipated versus surprise shocks then we should look at FEVDs conducted on shorter horizons. Figure 4 shows the 95% HPDI of the estimated variance for the growth rates of output, consumption, investment, and hours worked explained by anticipated shocks across models and FEVD horizons from 7500 draws from the posterior distributions, while Figure 5 displays the same information for the growth rate of government expenditure.[12] The vertical axis gives the median percent of volatility explained by anticipated shocks, while the horizontal axis gives the FEVD horizon. The light-shaded (dark-shaded) region corresponds to the 95% HPDI from the model under NRE (RE).

In all cases the importance of news is greater under the learning model. Furthermore the HPDIs are larger under NRE than under RE. This suggests that news is not merely a proxy for the endogenous learning process, but that in fact the combination of news and learning leads to additional volatility than would otherwise be implied. However it is worth noting that the entirety of this difference occurs from the realized news component: consistent with the results from Sims (2016), the relative importance of pure news is similarly low under both models.

Given the observed differences in the relative importance of anticipated vs unanticipated shocks across expectation formation mechanisms, it is natural to wonder

[12]I calculate FEVDs for horizons $h = \{0, 1, 2, 3, 4, 5, 6, 7, 8, 16\}$. Technically an unconditional FEVD is taken in the limit as the forecasting horizon goes to infinity. In practice this convergence occurs relatively quickly. Horizons longer than 16 result in variance shares which are nearly identical to those generated at 16.

Share of Volatility Due to Anticipated Shocks, RE vs NRE

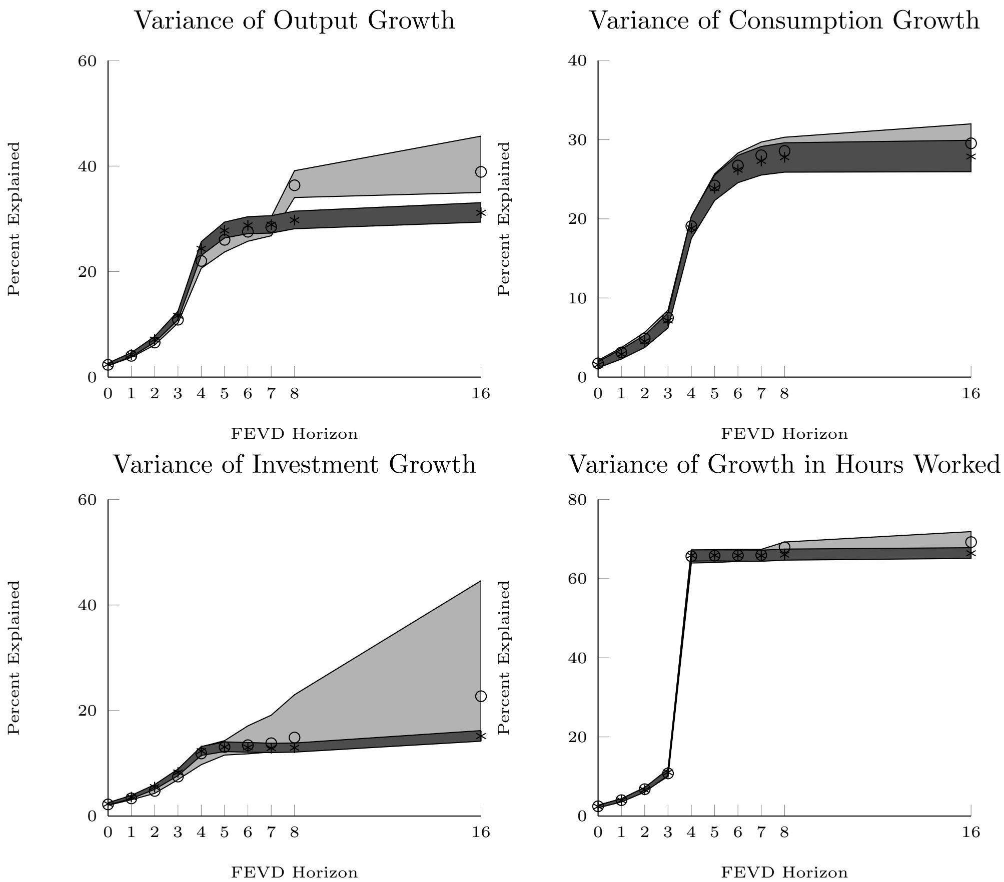

FIGURE 4. Relative Importance of News: Key Macroeconomic Variables, RE vs NRE

whether there are differences in each model's ability to fit the data. Table 11 compares the standard deviation (relative to output), correlation with output, and first-order autocorrelation coefficients for the growth rates of output, consumption, investment, hours worked, government expenditure, TFP, and the relative price of investment obtained from the actual data to that obtained from 7500 simulations of each model. The deep structural parameters for each simulation were generated via a random draw of the respective model's posterior distribution. The summary statistics are the median value across these simulations. The variables Y, C, I, h, g, TFP and A refer to the growth rates of output, consumption, investment, hours worked, government expenditure, TFP, and the relative

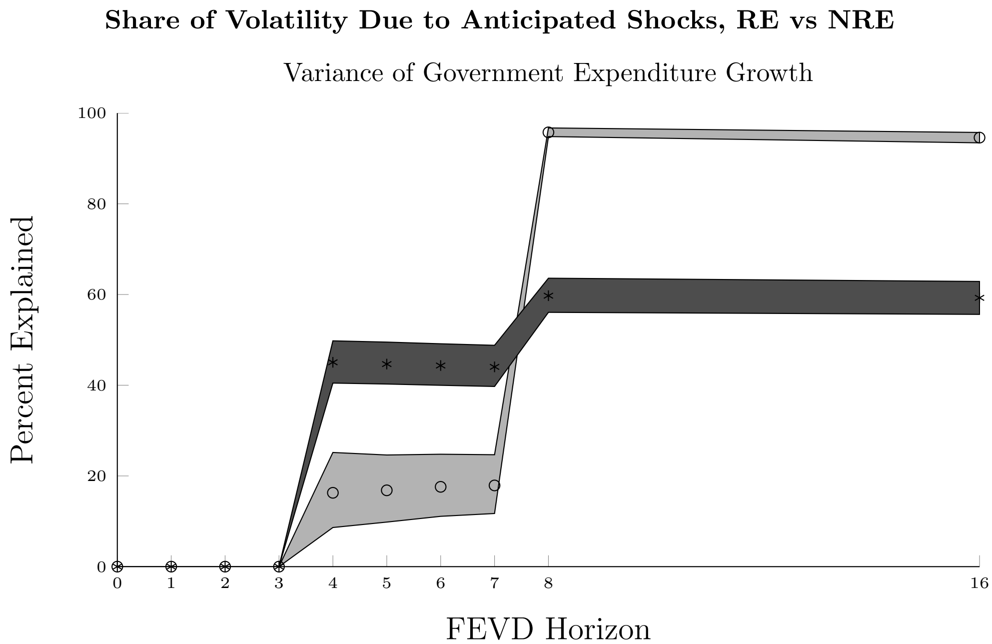

FIGURE 5. Relative Importance of News: Government Spending, RE vs NRE

price of investment, respectively, and relative standard deviations are the ratio of the standard deviation of the variable to that of the growth rate of output.

The summary statistics produced by both models are extremely similar to each other and are broadly consistent with the data. The growth rates of consumption and investment are estimated to be overly volatile, while the growth rate of the relative price of investment is somewhat less volatile; however, the underlying standard deviation of output growth is almost identical to that of the data. Furthermore the relative standard deviations of growth in hours worked, government spending, and TFP are all quite accurate compared to the data. The estimated correlations with the data are weaker than those observed in the data for the growth rates of investment, hours worked, and TFP, but match quite well for consumption and government spending. Finally, the estimated

Model Predictions: Volatility and Persistence, Data vs NRE vs RE

	Y	C	I	h	g	TFP	A
	Relative Standard Deviation						
Data	1.00	0.56	2.52	0.92	1.25	0.83	0.45
Model - NRE	1.00	0.70	3.06	0.97	1.15	0.85	0.35
Model -RE	1.00	0.70	3.05	0.98	1.18	0.88	0.37
	Correlation with Output						
Data	1.00	0.50	0.69	0.72	0.25	0.40	-0.12
Model - NRE	1.00	0.51	0.57	0.40	0.26	0.27	0.01
Model -RE	1.00	0.49	0.55	0.37	0.27	0.29	0.01
	Autocorrelation						
Data	0.28	0.20	0.53	0.60	0.05	-0.01	0.49
Model - NRE	0.32	0.46	0.63	0.18	0.03	-0.01	0.46
Model -RE	0.34	0.48	0.65	0.19	0.03	-0.00	0.46

TABLE 11. Model Predictions: Volatility and Persistence, Data vs NRE vs RE

persistence of consumption is too high, while the estimated persistence of hours worked is too low.[13]

Conclusion

This chapter has demonstrated that the estimated relative importance of anticipated news versus unanticipated news, as interpreted by the previous literature, is robust to the inclusion of learning. In particular it appears that news and learning serve to augment each other: in general, the estimated macroeconomic volatility in key variables is more a function of news under NRE than under RE. This is especially interesting given that both models produce quantitatively similar predictions for overall volatility and persistence.

[13]My estimates are broadly similar but generally closer to the data than those produced in Schmitt-Grohe and Uribe (2012). For example, the sum of squared deviations of data vs estimated relative standard deviation is around 0.33 under both NRE and RE under my estimates compared to 1.6 under theirs. This is driven to a large extent to the relative standard deviation of investment to output growth which is much greater under their estimates due to a much lower estimated standard deviation of output (0.91 in the data and my estimates, 0.73 under their estimates).

However, pure news is responsible for almost none of the estimated volatility in either model; it is only once the anticipated shock materializes that news begins to be important. It is possible that this finding is sensitive to the particular form of learning assumed. The NRE approach undertaken in this chapter represents a relatively small departure from RE.

Future research will look at the relative impact on model predictions and estimates across a variety of bounded rationality assumptions. Departures from RE such as the infinite-horizon learning of Preston (2005), finite-horizon learning of Branch et al. (2012a), and shadow-price learning of Evans and McGough (2015) all reflect larger departures from RE, and provide a natural avenue for future projects studying the empirical importance of both news shocks and adaptive learning in DSGE model estimation. It will be particularly interesting to observe whether differences emerge between between bounded rationality of the "learning-to-forecast" and "learning-to-optimize" types emerge.

APPENDIX A

APPENDIX FOR CHAPTER II

Competitive Equilibrium

Firms

The representative firm in the competitive durable goods sector solves

$$\max_{l_{k,t}} \quad P_t I_t - W_{k,t} l_{k,t} - R_{k,t} \tilde{l}_x$$
$$\text{s.t.} \quad I_t = \exp(\theta_{k,t}) l_{k,t}^{\alpha_k} \tilde{l}_k^{1-\alpha_k}$$

where P_t the price of the investment good in terms of the consumption good, $W_{k,t}$ and $R_{k,t}$ the rents paid to labor $l_{k,t}$ and the fixed factor $\tilde{l}_k$, respectively, and $\theta_{k,t}$ is a measure of technology.[1] The FONCs are given by

$$W_{k,t} = \alpha_k P_t \frac{I_t}{l_{k,t}} \tag{A.1}$$
$$R_{k,t} = (1 - \alpha_k) p_t \frac{I_t}{\tilde{l}_k} \tag{A.2}$$

Thus labor and the fixed factor in this sector are paid their marginal products in terms of the numeraire good.

---

[1] Technology in the durable goods sector evolves deterministically according to

$$\log \theta_{k,t} = g_{0,k} + g_1 * t$$

Consumption Good

The representative firm solves

$$\max_{X_t, K_{t-1}} \quad C_t - P_{x,t}X_t - R_t K_{t-1}$$
$$\text{s.t} \quad C_t = (aX_t^v + (1-a)K_{t-1}^v)^{\frac{1}{v}}$$

where $P_{x,t}$ the price of the nondurable good X_t in terms of the consumption good and R_t the gross real interest rate. The FONCs are given by

$$R_t = (1-a)C_t^{1-v}K_{t-1}^{v-1} \tag{A.3}$$

$$P_{x,t} = aC_t^{1-v}X_t^{v-1} \tag{A.4}$$

This will be used below to determine the value of labor and the fixed factor in the non-durable goods sector X_t.

The nondurable good is an intermediate good in production, and hence is only consumed by final goods producers. The representative firm in the competitive nondurable goods sector solves

$$\max_{l_{x,t}, \tilde{l}_x} \quad P_{x,t}X_t - W_{x,t}l_{x,t} - R_{x,t}\tilde{l}_x$$
$$\text{s.t} \quad X_t = \exp(\theta_{x,t}) l_{x,t}^{\alpha_x} \tilde{l}_x^{1-\alpha_x}$$

where $W_{x,t}$ and $R_{x,t}$ the rents paid to labor $l_{x,t}$ and the fixed factor $\tilde{l}_x$, respectively, and $\theta_{x,t}$ is a measure of technology.[2]

The FONCs are given by

$$W_{x,t} = \alpha_x X_t / l_{x,t} \tag{A.5}$$

$$R_{x,t} = \frac{r_{x,t}}{(1-\alpha_x) X_t / \tilde{l}_x} \tag{A.6}$$

$$\tag{A.7}$$

We can use the demand from the representative final good producer to determine the wage and rental rates. Equations (A.5) and (A.4) together imply

$$W_{x,t} = aC_t^{1-v} \alpha_x \frac{X_t^v}{l_{x,t}} \tag{A.8}$$

and equations (A.6) and (A.4) together imply

$$r_{x,t} = aC_t^{1-v} (1-\alpha_x) \frac{X_t^v}{\tilde{l}_x} \tag{A.9}$$

[2] Technology in the nondurable goods sector evolves according to

$$\log \theta_{x,t} = g_{0,x} + g_1 * t + \log \hat{\theta}_{x,t}$$
$$\log \hat{\theta}_{x,t} = \lambda \log \hat{\theta}_{x,t-1} + \epsilon_t$$

where ϵ_t an i.i.d mean zero shock which is known to the firms when making their production decision and v_{t-n} is the news shock revealed to households.

Households

The household problem is given by

$$\max_{C_t,l_{x,t},l_{k,t},I_t,K_t} \log C_t + v_0(\bar{l} - l_{x,t} - l_{k,t}) \tag{A.10}$$

subject to the law of motion for capital and the flow budget constraint

$$K_t = (1-\delta)K_{t-1} + I_t \tag{A.11}$$

$$C_t + P_t I_t = W_{x,t} l_{x,t} + W_{k,t} l_{k,t} + R_t K_{t-1} + \Pi_t \tag{A.12}$$

The LaGrangian is thus

$$\mathcal{L} = \hat{E}_0 \sum_{t=0}^{\infty} \beta^t \left\{ \log C_t + v_0(\bar{l} - l_{x,t} - l_{k,t}) \right\} + \Lambda_t \left\{ W_{x,t} l_{x,t} + W_{k,t} l_{k,t} + R_t K_{t-1} + \Pi_t - C_t - P_t I_t \right\}$$
$$+ \Lambda_t Q_t \left\{ (1-\delta) K_{t-1} + I_t - K_t \right\}$$

where Λ_t the marginal value of income, Q_t Tobin's marginal Q, and hence $\Lambda_t Q_t$ the marginal value of capital in terms of consumption. The first-order necessary conditions for an interior solution to the household's problem are given by

$$\frac{\partial \mathcal{L}}{\partial C_t} = 0 \iff \Lambda_t = \frac{1}{C_t} \tag{A.13a}$$

$$\frac{\partial \mathcal{L}}{\partial l_{x,t}} = 0 \iff v_0 = \Lambda_t W_{x,t} \tag{A.13b}$$

$$\frac{\partial \mathcal{L}}{\partial l_{k,t}} = 0 \iff v_0 = \Lambda_t W_{k,t} \tag{A.13c}$$

$$\frac{\partial \mathcal{L}}{\partial I_t} = 0 \iff P_t = Q_t \tag{A.13d}$$

$$\frac{\partial \mathcal{L}}{\partial K_t} = 0 \iff \Lambda_t Q_t = \beta \hat{E}_t \left[\Lambda_{t+1} \left(R_{t+1} + Q_{t+1}(1-\delta)\right)\right] \tag{A.13e}$$

Equilibrium

A temporary equilibrium is a set $\left\{C_t, l_{x,t}, l_{k,t}, I_t, K_t, X_t, Y_t, P_t, R_t, W_{x,t}, W_{k,t}, \Lambda_t, Q_t, \theta_{k,t}, \theta_{x,t}, \tilde{\theta}_{x,t}\right\}$ such that consumers and firms are utility maximizing and markets clear:

$$Y_t = C_t + P_t I_t \tag{A.14a}$$

$$C_t = \left(aX_t^\nu + (1-a)K_{t-1}^\nu\right)^{\frac{1}{\nu}} \tag{A.14b}$$

$$X_t = \theta_{x,t} l_{x,t}^{\alpha_x} F_x^{1-\alpha_x} \tag{A.14c}$$

$$I_t = \theta_{k,t} l_{k,t}^{\alpha_k} F_k^{1-\alpha_k} \tag{A.14d}$$

$$K_t = (1-\delta)K_{t-1} + I_t \tag{A.14e}$$

$$\Lambda_t = \frac{1}{C_t} \tag{A.14f}$$

$$v_0 = \Lambda_t W_{x,t} \tag{A.14g}$$

$$v_0 = \Lambda_t W_{k,t} \tag{A.14h}$$

$$P_t = Q_t \tag{A.14i}$$

$$\Lambda_t Q_t = \beta \hat{E}_t \left[\Lambda_{t+1} R_{t+1} + \Lambda_{t+1} Q_{t+1}(1-\delta)\right] \tag{A.14j}$$

$$R_t = (1-a)\left(\frac{C_t}{K_{t-1}}\right)^{1-v} \tag{A.14k}$$

$$W_{x,t} = a\alpha_x C_t^{1-v} \frac{X_t^v}{l_{x,t}} \tag{A.14l}$$

$$W_{k,t} = P_t \alpha_k \frac{I_t}{l_{k,t}} \tag{A.14m}$$

$$\theta_{k,t} = g_{0,k} e^{g_1 t} \tag{A.14n}$$

$$\theta_{x,t} = g_{0,x} e^{g_1 t} \tilde{\theta}_{x,t} \tag{A.14o}$$

$$\tilde{\theta}_{x,t} = \tilde{\theta}_{x,t-1}^{\lambda} e^{\epsilon_t} \tag{A.14p}$$

$$\tag{A.14q}$$

where Λ_t and $\Lambda_t Q_t$ the marginal value (or shadow price) of income and installed capital, respectively. Q_t is the ratio of these marginal values and can be interpreted as the relative price of installed capital in terms of the consumption good, i.e. *marginal* Tobin's Q.

Equation (A.14a) is the economy's aggregate resource constraint and equations (A.14b)-(A.14d) are the production functions for the consumption, nondurable, and durable goods respectively. Equation (A.14e) is the capital accumulation equation. Optimal household decision making for consumption, nondurable good labor supply, durable good labor supply, investment, and capital yields equations (A.14f)-(A.14j), respectively. Markets are competitive and hence factors of production are paid the value of their marginal products, as specified by equations (A.14k)-(A.14m). Finally, technology for producing the durable and nondurable goods is given by equations (A.14n)-(A.14p).

Define the for any non-stationary variable x the stationary counterpart $x_t = \hat{x}_t e^{g_1 t}$. The temporary equilibrium in terms of detrended variables is

$$\hat{Y}_t = \hat{C}_t + P_t \hat{I}_t \tag{A.15a}$$

$$\hat{C}_t = \left(a\hat{X}_t^{\nu} + (1-a)\left(\frac{\hat{K}_{t-1}}{e^{g_1}}\right)^{\nu}\right)^{\frac{1}{\nu}} \tag{A.15b}$$

$$\hat{X}_t = \hat{\theta}_{x,t} l_{x,t}^{\alpha_x} F_x^{1-\alpha_x} \tag{A.15c}$$

$$\hat{I}_t = \hat{\theta}_{k,t} l_{k,t}^{\alpha_k} F_k^{1-\alpha_k} \tag{A.15d}$$

$$\hat{K}_t = (1-\delta)\frac{\hat{K}_{t-1}}{e^{g_1}} + \hat{I}_t \tag{A.15e}$$

$$\hat{\Lambda}_t = \frac{1}{\hat{C}_t} \tag{A.15f}$$

$$v_0 = \hat{\Lambda}_t \hat{W}_{x,t} \tag{A.15g}$$

$$v_0 = \hat{\Lambda}_t \hat{W}_{k,t} \tag{A.15h}$$

$$P_t = Q_t \tag{A.15i}$$

$$e^{g_1}\hat{\Lambda}_t Q_t = \beta \hat{E}_t \left[\hat{\Lambda}_{t+1} R_{t+1} + \hat{\Lambda}_{t+1} Q_{t+1}(1-\delta)\right] \tag{A.15j}$$

$$R_t = (1-a)\left(\frac{\hat{C}_t}{\hat{K}_{t-1}} e^{g_1}\right)^{1-v} \tag{A.15k}$$

$$\hat{W}_{x,t} = a\alpha_x \hat{C}_t^{1-v} \frac{\hat{X}_t^{v}}{l_{x,t}} \tag{A.15l}$$

$$\hat{W}_{k,t} = P_t \alpha_k \frac{\hat{I}_t}{l_{k,t}} \tag{A.15m}$$

$$\hat{\theta}_{k,t} = g_{0,k} \tag{A.15n}$$

$$\hat{\theta}_{x,t} = g_{0,x}\tilde{\theta}_{x,t} \tag{A.15o}$$

$$\check{\theta}_{x,t} = \check{\theta}_{x,t-1}^{\lambda} e^{\epsilon_t} \tag{A.15p}$$

The entire detrended, log-linearized system of equilibrium conditions is given by

$$\tilde{Y}_t - \left(\frac{C}{Y}\right)\tilde{C}_t - \left(\frac{PI}{Y}\right)\tilde{P}_t - \left(\frac{PI}{Y}\right)\tilde{I}_t = 0 \tag{A.16}$$

$$\tilde{C}_t - a\left(\frac{X}{C}\right)^v \tilde{X}_t = (1-a)\left(\frac{K}{C}\right)^v \tilde{K}_{t-1} \tag{A.17}$$

$$\tilde{X}_t - \tilde{\theta}_{x,t} - \alpha_x \tilde{l}_{x,t} = 0 \tag{A.18}$$

$$\tilde{I}_t - \tilde{\theta}_{k,t} - \alpha_k \tilde{l}_{k,t} = 0 \tag{A.19}$$

$$\tilde{K}_t - \left(\frac{I}{K}\right)\tilde{I}_t = \left(\frac{1-\delta}{e^{g_1}}\right)\tilde{K}_{t-1} \tag{A.20}$$

$$\tilde{\Lambda}_t + \tilde{C}_t = 0 \tag{A.21}$$

$$\tilde{\Lambda}_t + \tilde{W}_{x,t} = 0 \tag{A.22}$$

$$\tilde{\Lambda}_t + \tilde{W}_{k,t} = 0 \tag{A.23}$$

$$\tilde{Q}_t - \tilde{P}_t = 0 \tag{A.24}$$

$$\tilde{\Lambda}_t + \tilde{Q}_t = \hat{E}_t\tilde{\Lambda}_{t+1} + \left(\frac{1}{R+Q(1-\delta)}\right)\hat{E}_t\tilde{R}_{t+1} + \left(\frac{Q(1-\delta)}{R+Q(1-\delta)}\right)\hat{E}_t\tilde{Q}_{t+1} \tag{A.25}$$

$$\tilde{R}_t + (R(v-1))\,\tilde{C}_t = (R(v-1))\,\tilde{K}_{t-1} \tag{A.26}$$

$$\tilde{l}_{x,t} + \tilde{W}_{x,t} - (1-v)\,\tilde{C}_t - v\tilde{X}_t = 0 \tag{A.27}$$

$$\tilde{l}_{k,t} + \tilde{W}_{k,t} - \tilde{P}_t - \tilde{I}_t = 0 \tag{A.28}$$

$$\tilde{\theta}_{k,t} = 0 \tag{A.29}$$

$$\tilde{\theta}_{x,t} - \tilde{\tilde{\theta}}_{x,t} = 0 \tag{A.30}$$

$$\tilde{\tilde{\theta}}_{x,t} = \lambda\tilde{\tilde{\theta}}_{x,t-1} + \epsilon_t \tag{A.31}$$

where a tilde over a variable indicates the percent deviation from steady state and variables without a time subscript indicate their steady state value.[3]

[3]Note that the gross interest rate is expressed in "deviation from steady state" rather than "percent deviation from steady state", since the variable is already expressed in percentage terms.

APPENDIX B

APPENDIX FOR CHAPTER III

Solving the Model

Suppose the household was acting optimally and contemplated reallocating a small amount of consumption for investment. The marginal cost of the reduction in consumption is the direct loss of utility today as well as the expected discounted value of changing the geometric average of habit-adjusted consumption. The marginal benefit is the expected discounted value of the additional investment, which affects future investment adjustment costs and increases tomorrow's capital stock. Hence the household's first-order necessary condition for C_t is

$$U_{C_t}(V_t) + \frac{\partial S_t}{\partial C_t}\beta \hat{E}_t \lambda^S_{t+1} = \frac{\partial I_t}{\partial C_t}\left(\beta \hat{E}_t \lambda^I_{t+1} + \frac{\partial K_t}{\partial I_t}\beta \hat{E}_t \lambda^K_{t+1}\right) \tag{B.1}$$

Now suppose the household considered a small increase in their labor supply. The additional labor income could be used to increase investment, and hence the marginal benefit is the expected discounted value of the additional investment they can afford, which again affects future investment adjustment costs and increases tomorrow's capital stock. The marginal cost is the disutility associated with higher labor supply, and hence the household's first-order necessary condition for h_t is

$$-U_{h_t}(V_t) = \frac{\partial I_t}{\partial h_t}\left(\beta \hat{E}_t \lambda^I_{t+1} + \frac{\partial K_t}{\partial I_t}\beta \hat{E}_t \lambda^K_{t+1}\right) \tag{B.2}$$

If the household were to utilize more of their predetermined capital stock they would receive rental income which could be used to finance additional investment, again affecting tomorrow's capital stock and investment adjustment costs. However working capital more intensely also increase the rate of depreciation, and thus the first-order necessary condition for u_t is

$$\frac{\partial K_t}{\partial u_t}\beta \hat{E}_t \lambda_{t+1}^K = \frac{\partial I_t}{\partial u_t}\left(\beta \hat{E}_t \lambda_{t+1}^I + \frac{\partial K_t}{\partial I_t}\beta \hat{E}_t \lambda_{t+1}^K\right) \tag{B.3}$$

The optimal choices of C_t, h_t, and u_t pin down the value for I_t through the flow budget constraint, which pins down the value for K_t through the capital accumulation equation. Likewise, the optimal choice of C_t generically determines the value of S_t. Thus equations (B.1), (B.2), and (B.3) fully describe the optimal behavior of the household as a function of their beliefs about the future shadow prices.

The expected value of these shadow prices may also be derived using variational arguments. The time t value of additional I_{t-1} in time t holding everything else constant is the resulting change to investment adjustment costs which changes the size of tomorrow's capital stock, that is

$$\lambda_t^I = \frac{\partial K_t}{\partial I_{t-1}}\beta \hat{E}_t \lambda_{t+1}^K \tag{B.4}$$

The time t value of additional preinstalled capital in time t is the expected discounted value of the additional rental income (which could be invested) plus the value of directly increasing tomorrow's capital stock, that is

$$\lambda_t^K = \frac{\partial I_t}{\partial K_{t-1}}\left(\beta \hat{E}_t \lambda_{t+1}^I + \frac{\partial K_t}{\partial I_t}\beta \hat{E}_t \lambda_{t+1}^K\right) + \frac{\partial K_t}{\partial K_{t-1}}\beta \hat{E}_t \lambda_{t+1}^K \tag{B.5}$$

Finally, the time t value of a small change to the predetermined level of habit-adjusted consumption in time t is the direct change to utility today, as well as the expected discounted value of the resulting change in the level of habit-adjusted consumption, that is

$$\lambda_t^S = U_{S_{t-1}}(V_t) + \frac{\partial S_t}{\partial S_{t-1}} \beta \hat{E}_t \lambda_{t+1}^S \tag{B.6}$$

The solution to the household's optimization problem satisfies the FONCs for the controls given by

$$\left(C_t - \psi h_t^\theta S_t\right)^{-\sigma} \left(1 - \frac{\gamma \psi h_t^\theta S_t}{C_t}\right) + \frac{\gamma S_t}{C_t} \beta \hat{E}_t \lambda_{t+1}^S = \frac{1}{A_t} \beta \hat{E}_t \left(\lambda_{t+1}^I + \left(1 - \Phi\left(\frac{I_t}{I_{t-1}}\right) - \Phi'\left(\frac{I_t}{I_{t-1}}\right)\left(\frac{I_t}{I_{t-1}}\right)\right) \lambda_{t+1}^K\right) \tag{B.7a}$$

$$\left(C_t - \psi h_t^\theta S_t\right)^{-\sigma} \left(\frac{\theta \psi h_t^{\theta-1} S_t}{W_t}\right) = \frac{1}{A_t} \beta \hat{E}_t \left(\lambda_{t+1}^I + \left(1 - \Phi\left(\frac{I_t}{I_{t-1}}\right) - \Phi'\left(\frac{I_t}{I_{t-1}}\right)\left(\frac{I_t}{I_{t-1}}\right)\right) \lambda_{t+1}^K\right) \tag{B.7b}$$

$$\frac{\delta'(u_t)}{R_t} \beta \hat{E}_t \lambda_{t+1}^K = \frac{1}{A_t} \beta \hat{E}_t \left(\lambda_{t+1}^I + \left(1 - \Phi\left(\frac{I_t}{I_{t-1}}\right) - \Phi'\left(\frac{I_t}{I_{t-1}}\right)\left(\frac{I_t}{I_{t-1}}\right)\right) \lambda_{t+1}^K\right) \tag{B.7c}$$

$$\lambda_t^I = \Phi'\left(\frac{I_t}{I_{t-1}}\right)\left(\frac{I_t}{I_{t-1}}\right)^2 \beta \hat{E}_t \lambda_{t+1}^K \tag{B.7d}$$

$$\lambda_t^K = (R_t u_t) \frac{1}{A_t} \beta \hat{E}_t \left(\lambda_{t+1}^I + \left(1 - \Phi\left(\frac{I_t}{I_{t-1}}\right) - \Phi'\left(\frac{I_t}{I_{t-1}}\right)\left(\frac{I_t}{I_{t-1}}\right)\right) \lambda_{t+1}^K\right) + (1 - \delta(u_t)) \beta \hat{E}_t \lambda_{t+1}^K \tag{B.7e}$$

$$\lambda_t^S = \left(C_t - \psi h_t^\theta S_t\right)^{-\sigma} \left(-\frac{(1-\gamma)\psi h_t^\theta S_t}{S_{t-1}}\right) + \beta \frac{(1-\gamma) S_t}{S_{t-1}} \hat{E}_t \lambda_{t+1}^S \tag{B.7f}$$

It can be shown that the shadow prices can be calculated as

$$\lambda_t^I = \Phi'\left(\frac{I_t}{I_{t-1}}\right)\left(\frac{I_t}{I_{t-1}}\right)^2 \left(\frac{R_t}{\delta'(u_t)}\right) \left(C_t - \psi h_t^\theta S_t\right)^{-\sigma} \left(\frac{\theta \psi h_t^{\theta-1} S_t}{W_t}\right) \tag{B.8a}$$

$$\lambda_t^K = (1 + u_t \delta'(u_t) - \delta(u_t)) \left(\frac{R_t}{\delta'(u_t)}\right) \left(C_t - \psi h_t^\theta S_t\right)^{-\sigma} \left(\frac{\theta \psi h_t^{\theta-1} S_t}{W_t}\right) \tag{B.8b}$$

$$\lambda_t^S = \left[\left(C_t - \psi h_t^\theta S_t\right)^{-\sigma} \left(\frac{(1-\gamma) S_t}{S_{t-1}}\right)\right] \left[\left(\frac{-C_t}{\gamma S_t}\right)\left(1 - \frac{\gamma \psi h_t^\theta S_t}{C_t} - \frac{\theta \psi h_t^{\theta-1} S_t}{W_t}\right) - \psi h_t^\theta\right] \tag{B.8c}$$

The full dynamic system under SP-learning is described by the following equations, comprised of the optimality conditions for households and firms, resource constraints, and identities:

$$
\begin{aligned}
\left(C_t - \psi h_t^\theta S_t\right)^{-\sigma}\left(1 - \frac{\gamma\psi h_t^\theta S_t}{C_t}\right) + \frac{\gamma S_t}{C_t}\beta \hat{E}_t \lambda_{t+1}^S &= \frac{1}{A_t}\beta \hat{E}_t\left(\lambda_{t+1}^I + \left(1 - \Phi\left(\frac{I_t}{I_{t-1}}\right) - \Phi'\left(\frac{I_t}{I_{t-1}}\right)\left(\frac{I_t}{I_{t-1}}\right)\right)\lambda_{t+1}^K\right) \\
\left(C_t - \psi h_t^\theta S_t\right)^{-\sigma}\left(-\frac{\theta\psi h_t^{\theta-1} S_t}{W_t}\right) &= \frac{1}{A_t}\beta \hat{E}_t\left(\lambda_{t+1}^I + \left(1 - \Phi\left(\frac{I_t}{I_{t-1}}\right) - \Phi'\left(\frac{I_t}{I_{t-1}}\right)\left(\frac{I_t}{I_{t-1}}\right)\right)\lambda_{t+1}^K\right) \\
\frac{\delta'(u_t)}{R_t}\beta \hat{E}_t \lambda_{t+1}^K &= \frac{1}{A_t}\beta \hat{E}_t\left(\lambda_{t+1}^I + \left(1 - \Phi\left(\frac{I_t}{I_{t-1}}\right) - \Phi'\left(\frac{I_t}{I_{t-1}}\right)\left(\frac{I_t}{I_{t-1}}\right)\right)\lambda_{t+1}^K\right) \\
\lambda_t^I &= \Phi'\left(\frac{I_t}{I_{t-1}}\right)\left(\frac{I_t}{I_{t-1}}\right)^2\left(\frac{R_t}{\delta'(u_t)}\right)\left(C_t - \psi h_t^\theta S_t\right)^{-\sigma}\left(\frac{\theta\psi h_t^{\theta-1} S_t}{W_t}\right) \\
\lambda_t^K &= \left(1 + u_t\delta'(u_t) - \delta(u_t)\right)\left(\frac{R_t}{\delta'(u_t)}\right)\left(C_t - \psi h_t^\theta S_t\right)^{-\sigma}\left(\frac{\theta\psi h_t^{\theta-1} S_t}{W_t}\right) \\
\lambda_t^S &= \left[\left(C_t - \psi h_t^\theta S_t\right)^{-\sigma}\left(\frac{(1-\gamma)S_t}{S_{t-1}}\right)\right] \times \\
&\quad \left[\left(\frac{-C_t}{\gamma S_t}\right)\left(1 - \frac{\gamma\psi h_t^\theta S_t}{C_t} - \frac{\theta\psi h_t^{\theta-1} S_t}{W_t}\right) - \psi h_t^\theta\right] \\
K_t &= \left(1 - \delta\left(u_t\right)\right) K_{t-1} + I_t\left[1 - \Phi\left(\frac{I_t}{I_{t-1}}\right)\right] \\
C_t + A_t I_t &\leq W_t h_t + R_t\left(u_t K_{t-1}\right) \\
S_t &= C_t^\gamma S_{t-1}^{1-\gamma} \\
Y_t &= C_t + A_t I_t \\
Y_t &= z_t\left(u_t K_{t-1}\right)^{1-\alpha}\left(h_t\right)^\alpha \\
W_t &= \alpha\frac{Y_t}{h_t} \\
R_t &= (1-\alpha)\frac{Y_t}{u_t K_{t-1}} \\
\hat{E}_t \lambda_t^I &= H_t^I \tilde{x}_t \\
\hat{E}_t \lambda_t^K &= H_t^K \tilde{x}_t \\
\hat{E}_t \lambda_t^S &= H_t^S \tilde{x}_t \\
R_{H,t} &= R_{H,t-1} + g_t\left(\tilde{x}_{t-1}\tilde{x}_{t-1}' - R_{H,t-1}\right) \\
H_t &= H_{t-1} + g_t R_{H,t}^{-1}\tilde{x}_{t-1}\left(\lambda_{t-1} - H_{t-1}'\tilde{x}_{t-1}\right)'
\end{aligned}
$$

Recursive Representation of News-Shocks

We can write the evolution of any exogenous variable indexed by w as

$$\ln(w_t) = \rho_w \ln(w_{t-1}) + \varepsilon^0_{w,t} \tag{B.9}$$

$$\varepsilon_{w,t} = \varphi_w \varepsilon_{w,t-1} + M_w \nu_{w,t} \tag{B.10}$$

where $\varepsilon_{w,t} = \left(\varepsilon^0_{w,t}, \varepsilon^1_{w,t}, .., \varepsilon^n_{w,t}\right)$ a vector of auxiliary state variables which carry the news shocks through time, φ a lower-shift matrix with 1's on the super-diagonal and zeros elsewhere, $\nu_{w,t} = \left(\nu^0_{w,t}, \nu^1_{w,t}, ..., \nu^n_{w,t}\right)'$ a vector of (anticipated and unanticipated) shocks, and $M_w = \left(\left[\nu^0_{w,t} \in \mathcal{I}_t\right], \left[\nu^1_{w,t-1} \in \mathcal{I}_t\right], ..., \left[\nu^n_{w,t-n} \in \mathcal{I}_t\right]\right)$ a row-vector of 1's and 0's respecting the specific assumptions regarding the information obtained by households.

APPENDIX C

APPENDIX FOR CHAPTER IV

The Labor Market

Final-goods producers demand a composite labor good $h_t^c = \left[\int_0^1 h_t(j)^{\frac{1}{1+\mu_t}} dj\right]^{1+\mu_t}$ where $h_t(j)$ denotes labor of type $j \in [0,1]$ and $\mu_t \geq 0$ is the exogenous stochastic wage markup. The cost minimization problem for a final-goods producer is

$$\min_{h_t(j)} \quad \int_0^1 W_t(j) h_t(j) dj$$
$$\text{subject to} \quad h_t^c \leq \left[\int_0^1 h_t(j)^{\frac{1}{1+\mu_t}} dj\right]^{1+\mu_t}$$

the solution to which yields the conditional factor demand functions $h_t^*(j) = h_t^c \left(\frac{W_t(j)}{W_t^c}\right)^{-\frac{(1+\mu_t)}{\mu_t}}$ where $W_t^c = \left[\int_0^1 W_t(j)^{-\frac{1}{\mu_t}} dj\right]^{-\mu_t}$ the cost of a single unit of the composite labor input.

Labor $h_t(j)$ is provided by monopolistically competitive unions. Since labor is freely mobile all unions pay their members the same wage rate W_t and charge firms $W_t(j)$. The profit maximization problem for these unions is

$$\max_{W_t(j)} \quad (W_t(j) - W_t)\, h_t(j)$$
$$\text{subject to} \quad h_t(j) = h_t^*(j)$$

the solution to which is $W_t^*(j) = (1 + \mu_t) W_t$. It is clear from this expression all unions will charge firms the same price denoted W_t^*, and that this wage is marked up in accordance with the value of market power captured by μ_t. This implies firms will demand identical quantities of each type of labor, that is $h_t(j) = h_t^c$ for all j.

The profits of each union j are thus $\Pi_t(j) = \mu_t W_t h_t^c$, which is the same for all unions. These profits are rebated to member-households as a lump sum. Since the unions determine the hours of labor allocated, firms choose their labor demand, and households choose how much labor to provide, an equilibrium requires $\int_0^1 h_t(j)dj = h_t^c = \int h_t(\omega)d\omega$.

Inducing Stationarity

Many of the endogenous variables inherit the non-stationarity of A_t and X_t. In particular, it can be shown that

- C_t, S_t, Y_t^{out}, and W_t grow at rate $\mu_t^Y = \frac{X_t^Y}{X_{t-1}^Y}$

- K_t and I_t grow at rate $\mu_t^K = \frac{X_t^K}{X_{t-1}^K}$ where $X_t^K = \frac{A_t}{X_t^Y}$

- G_t grows at rate $\mu_t^G = \frac{X_t^G}{X_{t-1}^G}$

- R_t and Q_t grow at rate μ_t^A

- Π_t and Λ_t grow at rate $\left(\mu_t^Y\right)^{-\sigma}$

Thus we define the stationary variables $\hat{Y}_t^{out} = \frac{Y_t^{out}}{X_t^Y}$, $\hat{C}_t = \frac{C_t}{X_t^Y}$, $\hat{W}_t = \frac{W_t}{X_t^Y}$, $\hat{I}_t = \frac{A_t I_t}{X_t^Y}$, $\hat{K}_t = \frac{A_t K_t}{X_t^Y}$, $g_t = \frac{G_t}{X_t^G}$, $\hat{R}_t = \frac{R_t}{A_t}$, $\hat{Q}_t = \frac{Q_t}{A_t}$, $\hat{\Pi}_t = \left(X_t^Y\right)^\sigma \Pi_t$, and $\hat{\Lambda}_t = \left(X_t^Y\right)^\sigma \Lambda_t$. u_t and h_t are already stationary variables. By manipulating the equations defining a non-stationary temporary equilibrium one can express the system in terms of the stationary variables. In particular, one can show the stationarized dynamic system of equilibrium conditions to be given by

$$\hat{K}_t = (1-\delta(u_t))\frac{\hat{K}_{t-1}}{\mu_t^K} + z_t^I \hat{I}_t\left[1-\Phi\left(\frac{\hat{I}_t}{\hat{I}_{t-1}}\mu_t^K\right)\right]$$

$$\hat{Y}_t = \hat{C}_t + \hat{I}_t + g_t X_t^{G,Y}$$

$$\hat{Y}_t = z_t\left(u_t\frac{\hat{K}_{t-1}}{\mu_t^K}\right)^{\alpha_k}(h_t)^{\alpha_h}(F)^{1-\alpha_k-\alpha_h}$$

$$\hat{V}_t = \hat{C}_t - b\frac{\hat{C}_{t-1}}{\mu_t^Y} - \psi h_t^\theta \hat{S}_t$$

$$\hat{S}_t = \left(\hat{C}_t - b\frac{\hat{C}_{t-1}}{\mu_t^Y}\right)^\gamma\left(\frac{\hat{S}_{t-1}}{\mu_t^Y}\right)^{1-\gamma}$$

$$\hat{\Lambda}_t = \left(\zeta_t\hat{V}_t^{-\sigma} - \frac{\gamma\hat{S}_t}{\hat{C}_t - b\frac{\hat{C}_{t-1}}{\mu_t^Y}}\hat{\Pi}_t\right) - b\beta\hat{E}_t\left[\left(\mu_{t+1}^Y\right)^{-\sigma}\left(\zeta_{t+1}\hat{V}_{t+1}^{-\sigma} - \frac{\gamma\hat{S}_{t+1}}{\hat{C}_{t+1} - b\frac{\hat{C}_t}{\mu_{t+1}^Y}}\hat{\Pi}_{t+1}\right)\right]$$

$$\hat{\Lambda}_t = \frac{\theta\zeta_t\hat{V}_t^{-\sigma}\psi h_t^{\theta-1}\hat{S}_t}{\hat{W}_t/\omega_t}$$

$$\hat{\Pi}_t = \zeta_t\hat{V}_t^{-\sigma}\psi h_t^\theta + \beta(1-\gamma)\hat{E}_t\left[\frac{\hat{S}_{t+1}}{\hat{S}_t}\hat{\Pi}_{t+1}\left(\mu_{t+1}^Y\right)^{1-\sigma}\right]$$

$$\hat{Q}_t\hat{\Lambda}_t = \beta\hat{E}_t\left[\mu_{t+1}^A\left(\mu_{t+1}^Y\right)^{-\sigma}\hat{\Lambda}_{t+1}\left(\hat{R}_{t+1}u_{t+1} + \hat{Q}_{t+1}(1-\delta(u_{t+1}))\right)\right]$$

$$\hat{R}_t = \delta'(u_t)\hat{Q}_t$$

$$\begin{aligned}\hat{\Lambda}_t = {} & \hat{Q}_t\hat{\Lambda}_t z_t^I\left(1-\Phi\left(\frac{\hat{I}_t}{\hat{I}_{t-1}}\mu_t^K\right) - \Phi'\left(\frac{\hat{I}_t}{\hat{I}_{t-1}}\mu_t^K\right)\left(\frac{\hat{I}_t}{\hat{I}_{t-1}}\mu_t^K\right)\right) \\ & + \beta\hat{E}_t\left[\hat{Q}_{t+1}\hat{\Lambda}_{t+1}z_{t+1}^I\mu_{t+1}^A\left(\mu_{t+1}^Y\right)^{-\sigma}\Phi'\left(\frac{\hat{I}_{t+1}}{\hat{I}_t}\mu_{t+1}^K\right)\left(\frac{\hat{I}_{t+1}}{\hat{I}_t}\mu_{t+1}^K\right)^2\right]\end{aligned}$$

$$\hat{W}_t = \alpha_h\frac{\hat{Y}_t}{h_t}$$

$$\hat{R}_t = \alpha_k\frac{\hat{Y}_t}{u_t\frac{\hat{K}_{t-1}}{\mu_t^K}}$$

where the functional forms for the depreciation rate and investment-adjustment costs are given by

$$\delta\left(u_t\right)=\delta_0+\delta_1\left(u_t-1\right)+\frac{\delta_2}{2}\left(u_t-1\right)^2$$
$$\Phi\left(\frac{\hat{I}_t}{\hat{I}_{t-1}}\mu_t^K\right)=\frac{\kappa}{2}\left(\frac{\hat{I}_t}{\hat{I}_{t-1}}\mu_t^K-\mu^I\right)^2$$

The laws of motion for the exogenous stochastic processes can be written

$$\begin{aligned}
\ln\left(z_t/z\right) &= \rho_z \ln\left(z_{t-1}/z\right)+\epsilon^0_{z,t}\\
\ln\left(\zeta_t/\zeta\right) &= \rho_\zeta \ln\left(\zeta_{t-1}/\zeta\right)+\epsilon^0_{\zeta,t}\\
\ln\left(z^I_t/z^I\right) &= \rho_{z^I} \ln\left(z^I_{t-1}/z^I\right)+\epsilon^0_{z^I,t}\\
\ln\left(g_t/g\right) &= +\rho_g \ln\left(g_{t-1}/g\right)+\epsilon^0_{g,t}\\
\ln\left(\omega_t/\omega\right) &= \rho_\omega \ln\left(\omega_{t-1}/\omega\right)+\epsilon^0_{\omega,t}\\
\ln\left(\mu^A_t/\mu^A\right) &= \rho_{\mu^A} \ln\left(\mu^A_{t-1}/\mu^A\right)+\epsilon^0_{\mu^A,t}\\
\ln\left(\mu^X_t/\mu^X\right) &= \rho_{\mu^X} \ln\left(\mu^X_{t-1}/\mu^X\right)+\epsilon^0_{\mu^X,t}\\
\mu^Y_t &= \mu^X_t\left(\mu^A_t\right)^{\alpha_k/(\alpha_k-1)}\\
\mu^K_t &= \mu^X_t\left(\mu^A_t\right)^{1/(\alpha_k-1)}\\
X^{G,Y}_t &= \frac{\left(X^{G,Y}_{t-1}\right)^{\rho_{x,g}}}{\mu^Y_t}=\frac{X^G_t}{X^Y_t}
\end{aligned}$$

where the shock structure for any stochastic process $x = \left\{z, \zeta, z^I, g, \omega, \mu^A, \mu^X\right\}$ is given by

$$\epsilon^0_{x,t} = \epsilon^1_{x,t-1} + \nu^0_{x,t}$$
$$\epsilon^1_{x,t} = \epsilon^2_{x,t-1}$$
$$\epsilon^2_{x,t} = \epsilon^3_{x,t-1}$$
$$\epsilon^3_{x,t} = \epsilon^4_{x,t-1}$$
$$\epsilon^4_{x,t} = \epsilon^5_{x,t-1} + \nu^4_{x,t}$$
$$\epsilon^5_{x,t} = \epsilon^6_{x,t-1}$$
$$\epsilon^6_{x,t} = \epsilon^7_{x,t-1}$$
$$\epsilon^7_{x,t} = \epsilon^8_{x,t-1}$$
$$\epsilon^8_{x,t} = \nu^8_{x,t}$$

Finally, the growth rates of the seven key macroeconomic variables considered in terms of the stationary variables is

$$g^Y_t = \frac{Y_t}{Y_{t-1}} = \frac{\hat{Y}_t}{\hat{Y}_{t-1}}\frac{X^Y_t}{X^Y_{t-1}} = \frac{\hat{Y}_t}{\hat{Y}_{t-1}}\mu^Y_t$$
$$g^C_t = \frac{C_t}{C_{t-1}} = \frac{\hat{C}_t}{\hat{C}_{t-1}}\frac{X^Y_t}{X^Y_{t-1}} = \frac{\hat{C}_t}{\hat{C}_{t-1}}\mu^Y_t$$
$$g^I_t = \frac{A_t I_t}{A_{t-1} I_{t-1}} = \frac{\hat{I}_t}{\hat{I}_{t-1}}\frac{X^Y_t}{X^Y_{t-1}} = \frac{\hat{I}_t}{\hat{I}_{t-1}}\mu^Y_t$$
$$g^h_t = \frac{h_t}{h_{t-1}}$$
$$g^G_t = \frac{G_t}{G_{t-1}} = \frac{g_t}{g_{t-1}}\left(X^{G,Y}_{t-1}\right)^{\rho_{xg}-1}$$
$$g^{TFP}_t = \frac{z_t}{z_{t-1}}\left(\mu^X_t\right)^{1-\alpha_k}$$
$$g^A_t = \mu^A_t$$

Steady State Computation and Calibration

Let x be the steady-state value of the stationary random variable $\hat{x}_t$. The steady states for z, ζ, and z^I are normalized to 1, g is set such that the output share of government spending is 0.20, ω is set to 1.15, and μ^Y and μ^A are set to the mean gross growth rate of per capita output and relative price of investment observed in the data which, given the stochastic trends in output and government spending, implies $\mu^X = \mu^Y \, {\mu^A}^{\frac{\alpha_k}{1-\alpha_k}}$ and $X^{G,Y} = \left(\mu^Y\right)^{1/(\rho_{x,g}-1)}$. The parameter μ^I is set to μ^K so that investment-adjustment costs (and their derivative) are zero in steady state, while δ_1 is set to $\frac{1}{\beta \mu^A {\mu^Y}^{-\sigma}} - (1-\delta_0)$ to ensure the steady-state capacity utilization rate u is 1. The scale parameter ψ governing the disutility of labor in the utility function is calibrated to ensure the steady-state value of h is 0.20. With these calibrations for the steady-state values of exogenous processes and free parameters, the non-stochastic steady state can be easily computed from the equilibrium conditions and constraints given above. In particular, the steady state can be shown to be

$$R = \delta_1$$

$$h = 0.2$$

$$u = 1$$

$$Q = 1$$

$$K = \left(\frac{\alpha_k z \left(\frac{\mu^K}{u}\right)^{1-\alpha_k} h^{\alpha_h}}{R} \right)^{\frac{1}{1-\alpha_k}}$$

$$I = \left(1 - \frac{1-\delta_0}{\mu^K}\right) K$$

$$Y = \frac{RuK}{\alpha_k \mu^K}$$

$$W = \alpha_h \frac{Y}{h}$$

$$g = 0.2$$

$$C = Y - I - gX^{GY}$$

$$S = C(1 - b/\mu^Y) \left(\mu^Y\right)^{\frac{\gamma-1}{\gamma}}$$

$$\psi = \left[\left(\frac{\theta h^{\theta-1} S}{W/\omega} \right) \left(1 - b\beta \left(\mu^Y\right)^{-\sigma} \right)^{-1} + \left(\frac{\gamma S}{(C - bC/\mu^Y)} \right) \left(\frac{h^\theta}{1 - \beta(1-\gamma) \left(\mu^Y\right)^{1-\sigma}} \right) \right]^{-1}$$

$$V = \left(C - \frac{bC}{\mu^Y} \right) - \psi h^\theta S$$

$$\Pi = \left(\zeta V^{-\sigma} \psi h^\theta\right) \left(1 - \beta(1-\gamma) \left(\mu^Y\right)^{1-\sigma} \right)^{-1}$$

$$\Lambda = \frac{\theta \zeta V^{-\sigma} \psi h^{\theta-1} S}{W/\omega}$$

Log-Linearization

The stationarized system can be log-linearized around its non-stochastic steady state. Denote by a $\widetilde{\text{tilde}}$ the deviation from steady state, that is, for any variable x we have $\tilde{x}_t = \log\left(\frac{x_t}{x}\right)$.

1. The capital accumulation equation is

$$\hat{K}_t = (1 - \delta(u_t))\frac{\hat{K}_{t-1}}{\mu_t^K} + z_t^I \hat{I}_t \left[1 - \Phi\left(\frac{\hat{I}_t}{\hat{I}_{t-1}}\mu_t^K\right)\right]$$

Taking logs of both sides we have

$$\log \hat{K}_t = \log\left((1 - \delta(u_t))\frac{\hat{K}_{t-1}}{\mu_t^K} + z_t^I \hat{I}_t \left[1 - \Phi\left(\frac{\hat{I}_t}{\hat{I}_{t-1}}\mu_t^K\right)\right]\right)$$

Noting that $\log K = \log\left((1-\delta(u))\frac{K}{\mu^K} + z^I I\left(1 - \Phi\left(\frac{I}{I}\mu^K\right)\right)\right)$, the first-order Taylor series expansion around the steady state can be shown to be

$$\tilde{K}_t = \left(\frac{1-\delta(u)}{\mu^K}\right)\tilde{K}_{t-1} - \left(\frac{\delta'(u)u}{\mu^K}\right)\tilde{u}_t - \left(\frac{1-\delta(u)}{\mu^K}\right)\tilde{\mu}_t^K + \left(\frac{z^I I}{K}\right)\tilde{z}_t^I + \left(\frac{z^I I}{K}\right)\tilde{I}_t$$

2. The aggregate resource constraint is

$$\hat{Y}_t = \hat{C}_t + \hat{I}_t + g_t X_t^{G,Y}$$

Taking logs of both sides we have

$$\log \hat{Y}_t = \log\left(\hat{C}_t + \hat{I}_t + g_t X_t^{G,Y}\right)$$

Noting that $\log \hat{Y} = \log\left(\hat{C} + \hat{I} + gX^{G,Y}\right)$, the first-order Taylor series expansion around the steady state can be shown to be

$$\tilde{Y}_t = \left(\frac{C}{Y}\right)\tilde{C}_t + \left(\frac{I}{Y}\right)\tilde{I}_t + \left(\frac{gX^{G,Y}}{Y}\right)\tilde{X}_t^{G,Y} + \left(\frac{gX^{G,Y}}{Y}\right)\tilde{g}_t$$

3. The production function is

$$\hat{Y}_t = z_t\left(u_t\frac{\hat{K}_{t-1}}{\mu_t^K}\right)^{\alpha_k}(h_t)^{\alpha_h}(F)^{1-\alpha_k-\alpha_h}$$

Taking logs on both sides yields

$$\log \hat{Y}_t = \log z_t + \alpha_k \log u_t + \alpha_k \log \hat{K}_{t-1} - \alpha_k \log \mu_t^K + \alpha_h \log h_t + (1-\alpha_k-\alpha_h)\log F$$

Since the equation is log-linear, subtracting each side by $\log Y = \log z + \alpha_k \log u + \alpha_k \log \hat{K} - \alpha_k \log \mu^K + \alpha_h \log h + (1-\alpha_k-\alpha_h)\log F$, the log-linearized production function can be shown to be

$$\tilde{Y}_t = \tilde{z}_t + \alpha_k\tilde{u}_t + \alpha_k\tilde{K}_{t-1} - \alpha_k\tilde{\mu}_t^K + \alpha_h\tilde{h}_t$$

4. The utility function's argument is

$$\hat{V}_t = \hat{C}_t - b\frac{\hat{C}_{t-1}}{\mu_t^Y} - \psi h_t^\theta \hat{S}_t$$

Taking logs of both sides yields

$$\log \hat{V}_t = \log\left(\hat{C}_t - b\frac{\hat{C}_{t-1}}{\mu_t^Y} - \psi h_t^\theta \hat{S}_t\right)$$

Noting that $\log \hat{V} = \log\left(\hat{C} - b\frac{\hat{C}}{\mu^Y} - \psi h^\theta \hat{S}\right)$, the first-order Taylor series expansion around the steady state can be shown to be

$$\tilde{V}_t = \left(\frac{C}{V}\right)\tilde{C}_t - \left(\frac{bC/\mu^Y}{V}\right)\tilde{C}_{t-1} + \left(\frac{bC/\mu^Y}{V}\right)\tilde{\mu}_t^Y - \left(\frac{\psi h^\theta S\theta}{V}\right)\tilde{h}_t - \left(\frac{\psi h^\theta S}{V}\right)\tilde{S}_t$$

5. The identity for the geometric average of current and past habit-adjusted consumption is

$$\hat{S}_t = \left(\hat{C}_t - b\frac{\hat{C}_{t-1}}{\mu_t^Y}\right)^\gamma \left(\frac{\hat{S}_{t-1}}{\mu_t^Y}\right)^{1-\gamma}$$

Taking logs of both sides yields

$$\log \hat{S}_t = \gamma \log\left(\hat{C}_t - b\frac{\hat{C}_{t-1}}{\mu_t^Y}\right) + (1-\gamma)\log\left(\frac{\hat{S}_{t-1}}{\mu_t^Y}\right)$$

Noting that $\log \hat{S} = \gamma \log\left(\hat{C} - b\frac{\hat{C}}{\mu^Y}\right) + (1-\gamma)\log\left(\frac{\hat{S}}{\mu^Y}\right)$, the first-order Taylor series expansion around the steady state can be shown to be

$$\tilde{S}_t = \left(\frac{\gamma}{1-b/\mu^Y}\right)\tilde{C}_t - \left(\frac{\gamma b/\mu^Y}{1-b/\mu^Y}\right)\tilde{C}_{t-1} + (1-\gamma)\tilde{S}_{t-1} + \left(\frac{\gamma b/\mu^Y}{1-b/\mu^Y} - (1-\gamma)\right)\tilde{\mu}_t^Y$$

6. The FONC for consumption is

$$\hat{\Lambda}_t = \left(\zeta_t \hat{V}_t^{-\sigma} - \frac{\gamma \hat{S}_t}{\hat{C}_t - b\frac{\hat{C}_{t-1}}{\mu_t^Y}}\hat{\Pi}_t\right) - b\beta \hat{E}_t\left[\left(\zeta_{t+1}\hat{V}_{t+1}^{-\sigma} - \frac{\gamma \hat{S}_{t+1}}{\hat{C}_{t+1} - b\frac{\hat{C}_t}{\mu_{t+1}^Y}}\hat{\Pi}_{t+1}\right)\left(\mu_{t+1}^Y\right)^{-\sigma}\right]$$

Taking logs of both sides yields

$$\log \hat{\Lambda}_t = \log \left(\left(\zeta_t \hat{V}_t^{-\sigma} - \frac{\gamma \hat{S}_t}{\hat{C}_t - b\frac{\hat{C}_{t-1}}{\mu_t^Y}} \hat{\Pi}_t \right) - b\beta \hat{E}_t \left[\left(\zeta_{t+1} \hat{V}_{t+1}^{-\sigma} - \frac{\gamma \hat{S}_{t+1}}{\hat{C}_{t+1} - b\frac{\hat{C}_t}{\mu_{t+1}^Y}} \hat{\Pi}_{t+1} \right) \left(\mu_{t+1}^Y\right)^{-\sigma} \right] \right)$$

Noting that $\log \hat{\Lambda} = \log \left(\left(\zeta \hat{V}^{-\sigma} - \frac{\gamma \hat{S}}{\hat{C} - b\frac{\hat{C}}{\mu^Y}} \hat{\Pi} \right) - b\beta \hat{E} \left[\left(\zeta \hat{V}^{-\sigma} - \frac{\gamma \hat{S}}{\hat{C} - b\frac{\hat{C}}{\mu^Y}} \hat{\Pi} \right) \left(\mu^Y\right)^{-\sigma} \right] \right)$, the first-order Taylor series expansion around the steady state can be shown to be

$$\begin{aligned}
\tilde{\Lambda}_t = {} & \phi_1 \left(\tilde{\zeta}_t - \sigma \tilde{V}_t - \phi_3 \left(\hat{E}_t \tilde{\zeta}_{t+1} - \sigma \hat{E}_t \tilde{V}_{t+1} \right) \right) \\
& - \phi_2 \Bigg\{ \tilde{S}_t + \tilde{\Pi}_t + \left(\frac{b/\mu^Y}{1 - b/\mu^Y} \right) \tilde{C}_{t-1} - \left(\frac{b/\mu^Y}{1 - b/\mu^Y} \right) \tilde{\mu}_t^Y - \left(\frac{1 + b^2 \beta \left(\mu^Y\right)^{-\sigma-1}}{1 - b/\mu^Y} \right) \tilde{C}_t \\
& \qquad - \phi_3 \left(\hat{E}_t \tilde{S}_{t+1} + \hat{E}_t \tilde{\Pi}_{t+1} - \left(\frac{1}{1 - b/\mu^Y} \right) \hat{E}_t \tilde{C}_{t+1} \right) \Bigg\} \\
& + \left[\phi_3 \left(\sigma \left(\phi_1 - \phi_2 \right) - \frac{\phi_2 b/\mu^Y}{1 - b/\mu^Y} \right) \right] \hat{E}_{t+1} \mu_{t+1}^Y
\end{aligned}$$

where $\phi_1 = \frac{\zeta V^{-\sigma}}{\Lambda}$, $\phi_2 = \frac{\gamma S \Pi}{\Lambda C (1 - b/\mu^Y)}$, and $\phi_3 = b\beta \left(\mu^Y\right)^{-\sigma}$.

7. The FONC for labor supply is

$$\hat{\Lambda}_t = \frac{\theta \zeta_t \hat{V}_t^{-\sigma} \psi h_t^{\theta-1} \hat{S}_t}{\hat{W}_t / \omega_t}$$

Taking logs of both sides yields

$$\log \hat{\Lambda}_t = \log \left(\frac{\theta \zeta_t \hat{V}_t^{-\sigma} \psi h_t^{\theta-1} \hat{S}_t}{\hat{W}_t / \omega_t} \right)$$

This equation is linear in logs; subtracting $\log \hat{\Lambda} = \log \left(\frac{\theta \zeta \hat{V}^{-\sigma} \psi h^{\theta-1} \hat{S}}{\hat{W} / \omega} \right)$ directly yields

$$\tilde{\Lambda}_t = \tilde{\zeta}_t - \sigma \tilde{V}_t + (\theta - 1) \tilde{h}_t + \tilde{S}_t - \tilde{W}_t + \tilde{\omega}_t$$

8. The FONC for the geometric average of current and past habit-adjusted consumption is

$$\hat{\Pi}_t = \zeta_t \hat{V}_t^{-\sigma} \psi h_t^{\theta} + \beta \hat{E}_t \left[(1-\gamma) \frac{\hat{S}_{t+1}}{\hat{S}_t} \hat{\Pi}_{t+1} \left(\mu_{t+1}^Y\right)^{1-\sigma} \right]$$

Taking logs of both sides yields

$$\log \hat{\Pi}_t = \log \left(\zeta_t \hat{V}_t^{-\sigma} \psi h_t^{\theta} + \beta \hat{E}_t \left[(1-\gamma) \frac{\hat{S}_{t+1}}{\hat{S}_t} \hat{\Pi}_{t+1} \left(\mu_{t+1}^Y\right)^{1-\sigma} \right] \right)$$

Noting that $\log \hat{\Pi} = \log \left(\zeta \hat{V}^{-\sigma} \psi h^{\theta} + \beta \left[(1-\gamma) \frac{\hat{S}}{\hat{S}} \hat{\Pi} \left(\mu^Y\right)^{1-\sigma} \right] \right)$, the first-order Taylor series expansion around the steady state can be shown to be

$$\Pi \tilde{\Pi}_t = \left(\zeta V^{-\sigma} \psi h^{\theta} \right) \left(\tilde{\zeta}_t - \sigma \tilde{V}_t + \theta \tilde{h}_t \right) + \left(\beta (1-\gamma) \Pi \left(\mu^Y\right)^{1-\sigma} \right) \left(\hat{E}_t \tilde{S}_{t+1} - \tilde{S}_t + \hat{E}_t \tilde{\Pi}_{t+1} + (1-\sigma) \hat{E}_t \tilde{\mu}_{t+1}^Y \right.$$

9. The FONC for capital is

$$\hat{Q}_t \hat{\Lambda}_t = \beta \hat{E}_t \left[\hat{\Lambda}_{t+1} \left(\hat{R}_{t+1} u_{t+1} + \hat{Q}_{t+1} (1 - \delta(u_{t+1})) \right) \left(\mu_{t+1}^A \left(\mu_{t+1}^Y\right)^{-\sigma} \right) \right]$$

Taking logs of both sides yields

$$\log \hat{Q}_t + \log \hat{\Lambda}_t = \log \left(\beta \hat{E}_t \left[\hat{\Lambda}_{t+1} \left(\hat{R}_{t+1} u_{t+1} + \hat{Q}_{t+1} (1 - \delta(u_{t+1})) \right) \left(\mu_{t+1}^A \left(\mu_{t+1}^Y\right)^{-\sigma} \right) \right] \right)$$

Noting that $\log \hat{Q} + \log \hat{\Lambda} = \log \left(\beta \left[\hat{\Lambda} \left(\hat{R}u + \hat{Q}(1-\delta(u)) \right) \left(\mu^A \left(\mu^Y \right)^{-\sigma} \right) \right] \right)$, the first-order Taylor series expansion around the steady state can be shown to be

$$\tilde{Q}_t + \tilde{\Lambda}_t = \left(\frac{\beta \mu^A \left(\mu^Y \right)^{-\sigma}}{Q} \right) \Big\{ (Ru + Q(1-\delta(u))) \left(\hat{E}_t \tilde{\Lambda}_{t+1} + \hat{E}_t \tilde{\mu}^A_{t+1} - \sigma \hat{E}_t \tilde{\mu}^Y_{t+1} \right) + (Ru) \hat{E}_t \tilde{R}_{t+1}$$
$$+ \left((R - Q\delta'(u)) \, u \right) \hat{E}_t \tilde{u}_{t+1} + (Q(1-\delta(u))) \, \hat{E}_t \tilde{Q}_{t+1} \Big\}$$

10. The FONC for the capacity utilization rate of capital is

$$\hat{R}_t = \delta'(u_t) \hat{Q}_t$$

Taking logs of both sides yields

$$\log \hat{R}_t = \log \left(\delta'(u_t) \right) + \log \hat{Q}_t$$

Thus the first-order Taylor series expansion around the steady state can be shown to be

$$\tilde{R}_t = \frac{\delta''(u)u}{\delta'(u)} \tilde{u}_t + \tilde{Q}_t$$

11. The FONC for investment is

$$\hat{\Lambda}_t = z^I_t \left(1 - \Phi \left(\frac{\hat{I}_t}{\hat{I}_{t-1}} \mu^K_t \right) - \Phi' \left(\frac{\hat{I}_t}{\hat{I}_{t-1}} \mu^K_t \right) \left(\frac{\hat{I}_t}{\hat{I}_{t-1}} \mu^K_t \right) \right) \hat{Q}_t \hat{\Lambda}_t$$
$$+ \beta \hat{E}_t \left[\left(z^I_{t+1} \Phi' \left(\frac{I_{t+1}}{I_t} \mu^K_{t+1} \right) \left(\frac{I_{t+1}}{I_t} \mu^K_{t+1} \right)^2 \hat{Q}_{t+1} \hat{\Lambda}_{t+1} \right) \left(\mu^A_{t+1} \left(\mu^Y_{t+1} \right)^{-\sigma} \right) \right]$$

Taking logs of both sides yields

$$\log\hat{\Lambda}_t = \log\left(z_t^I\left(1-\Phi\left(\frac{\hat{I}_t}{\hat{I}_{t-1}}\mu_t^K\right)-\Phi'\left(\frac{\hat{I}_t}{\hat{I}_{t-1}}\mu_t^K\right)\left(\frac{\hat{I}_t}{\hat{I}_{t-1}}\mu_t^K\right)\right)\hat{Q}_t\hat{\Lambda}_t\right.$$
$$\left.+\beta\hat{E}_t\left[\left(z_{t+1}^I\Phi'\left(\frac{I_{t+1}}{I_t}\mu_{t+1}^K\right)\left(\frac{I_{t+1}}{I_t}\mu_{t+1}^K\right)^2\hat{Q}_{t+1}\hat{\Lambda}_{t+1}\right)\left(\mu_{t+1}^A\left(\mu_{t+1}^Y\right)^{-\sigma}\right)\right]\right)$$

Noting that at steady state

$$\log\hat{\Lambda} = \log\left(z^I\left(1-\Phi\left(\frac{\hat{I}}{\hat{I}}\mu^K\right)-\Phi'\left(\frac{\hat{I}}{\hat{I}}\mu^K\right)\left(\frac{\hat{I}}{\hat{I}}\mu^K\right)\right)\hat{Q}\hat{\Lambda}\right.$$
$$\left.+\beta\hat{E}\left[\left(z^I\Phi'\left(\frac{I}{I}\mu^K\right)\left(\frac{I}{I}\mu^K\right)^2\hat{Q}\hat{\Lambda}\right)\left(\mu^A\left(\mu^Y\right)^{-\sigma}\right)\right]\right)$$

the first-order Taylor series expansion around the steady state can be shown to be

$$\begin{aligned}\tilde{\Lambda}_t = z^I Q\Big\{&\tilde{z}_t^I+\tilde{Q}_t+\tilde{\Lambda}_t\\ &+\Phi''\left(\mu^K\right)\left(\mu^K\right)^2\Big[\tilde{I}_{t-1}-\tilde{\mu}_t^K-\tilde{I}_t\\ &+\beta\mu^A\left(\mu^Y\right)^{-\sigma}\mu^K\left(\hat{E}_t\tilde{I}_{t+1}+\hat{E}_t\tilde{\mu}_{t+1}^K-\tilde{I}_t\right)\Big]\Big\}\end{aligned}$$

12. The wage valuation equation is

$$\hat{W}_t = \alpha_h\frac{\hat{Y}_t}{h_t}$$

Taking logs of both sides yields

$$\log\hat{W}_t = \log\alpha_h + \log\hat{Y}_t - \log h_t$$

This is linear in logs. Subtracting both sides by $\log \hat{W} = \log \alpha_h + \log \hat{Y} + \log h$ the log-linearized equation can be shown to be

$$\tilde{W}_t = \tilde{Y}_t - \tilde{h}_t$$

13. The effective-capital valuation equation is

$$\hat{R}_t = \alpha_k \frac{\hat{Y}_t}{u_t \frac{\hat{K}_{t-1}}{\mu_t^K}}$$

Taking logs of both sides yields

$$\log \hat{R}_t = \log \alpha_k + \log \hat{Y}_t - \log u_t - \log \hat{K}_{t-1} + \log \mu_t^K$$

This is linear in logs. Subtracting both sides by $\log \hat{R} = \log \alpha_k + \log \hat{Y} - \log u - \log \hat{K}_{t-1} + \log \mu^K$ the log-linearized equation can be shown to be

$$\tilde{R}_t = \tilde{Y}_t - \tilde{u}_t - \tilde{K}_{t-1} + \tilde{\mu}_t^K$$

14. Output growth is given by

$$\mu_t^Y = \mu_t^X \left(\mu_t^A\right)^{\alpha_k/(\alpha_k - 1)}$$

Taking logs of both sides yields

$$\log \mu_t^Y = \log \mu_t^X - \left(\frac{\alpha_k}{1 - \alpha_k}\right) \log \mu_t^A$$

This is linear in logs. Subtracting both sides by $\log \mu^Y = \log \mu^X - \left(\frac{\alpha_k}{1-\alpha_k}\right) \log \mu^A$ the log-linearized equation can be shown to be

$$\tilde{\mu}_t^Y = \tilde{\mu}_t^X - \left(\frac{\alpha_k}{1-\alpha_k}\right) \tilde{\mu}_t^A$$

15. Capital (and investment) growth is given by

$$\mu_t^K = \frac{\mu_t^Y}{\mu_t^A}$$

Taking logs of both sides yields

$$\log \mu_t^K = \log \mu_t^Y - \log \mu_t^A$$

This is linear in logs. Subtracting both sides by $\log \mu^K = \log \mu^Y - \log \mu^A$ the log-linearized equation can be shown to be

$$\tilde{\mu}_t^K = \tilde{\mu}_t^Y - \tilde{\mu}_t^A$$

16. The stochastic trend for government spending is

$$X_t^{G,Y} = \frac{\left(X_{t-1}^{G,Y}\right)^{\rho_{x,g}}}{\mu_t^Y}$$

Taking logs of both sides yields

$$\log X_t^{G,Y} = \rho_{x,g} \log X_{t-1}^{G,Y} - \log \mu_t^Y$$

This is linear in logs. Subtracting both sides by $\log X^{G,Y} = \rho_{x,g} \log X^{G,Y} - \log \mu^{Y}$ the log-linearized equation can be shown to be

$$\tilde{X}_t^{G,Y} = \rho_{x,g} \tilde{X}_{t-1}^{G,Y} - \tilde{\mu}_t^{Y}$$

17. The stationary exogenous stochastic processes $x = \left\{z, \zeta, z^I, g, \omega, \mu^A, \mu^X\right\}$ all evolve according to

$$x_t = x^{1-\rho_x} x_{t-1}^{\rho_x} \epsilon_{x,t}^0$$

Taking logs of both sides yields

$$\ln x_t = (1 - \rho_x)\, x + \rho_x \ln x_{t-1} + \varepsilon_{x,t}^0$$

where $\varepsilon_{x,t}^0 = \ln \epsilon_{x,t}^0$. This is linear in logs. Subtracting both sides by $\ln x = (1 - \rho_x)\, x + \rho_x \ln x$ the log-linearized equation can be shown to be

$$\tilde{x}_t = \rho_x \tilde{x}_{t-1} + \varepsilon_{x,t}^0$$

REFERENCES CITED

Adam, K. (2005). Learning to Forecast and Cyclical Behavior of Output and Inflation. *Macroeconomic Dynamics*, 9(01):1–27.

Aiyagari, S. R. (1994). "Uninsured Idiosyncratic Risk and Aggregate Saving". *The Quarterly Journal of Economics*, pages 659–684.

Alexopoulos, M. (2011). "Read All about It!! What Happens Following a Technology Shock?". *American Economic Review*, 101(4):1144–79.

Allen, S. G. (1985). "Why Construction Industry Productivity is Declining". *The Review of Economics and Statistics*, 67(4):661–69.

An, S. and Schorfheide, F. (2007). Bayesian Analysis of DSGE Models. *Econometric reviews*, 26(2-4):113–172.

Angeletos, G.-M., Hellwig, C., and Pavan, A. (2007). "Dynamic Global Games of Regime Change: Learning, Multiplicity, and the Timing of Attacks". *Econometrica*, 75(3):711–756.

Angeletos, G.-M. and La'O, J. (2010). "Noisy Business Cycles". In *NBER Macroeconomics Annual 2009, Volume 24*, pages 319–378. University of Chicago Press.

Angeletos, G.-M. and Pavan, A. (2004). "Transparency of Information and Coordination in Economies with Investment Complementarities". Technical report, National Bureau of Economic Research.

Assenza, T., Bao, T., Hommes, C., Massaro, D., et al. (2014). Experiments on expectations in macroeconomics and finance. *Experiments in macroeconomics*, 17:11À70.

Barsky, R. B. and Sims, E. R. (2012). "Information, Animal Spirits, and the Meaning of Innovations in Consumer Confidence". *American Economic Review*, 102(4):1343–77.

Basu, S., Fernald, J., and Kimball, M. (2004). Are Technology Improvements Contractionary? Technical report, National Bureau of Economic Research.

Basu, S., Fernald, J. G., and Kimball, M. S. (2006). Are Technology Improvements Contractionary? *The American Economic Review*, 96(5):1418–1448.

Beaudry, P. and Lucke, B. (2010). Letting Different Views about Business Cycles Compete. In *NBER Macroeconomics Annual 2009, Volume 24*, pages 413–455. University of Chicago Press.

Beaudry, P., Nam, D., and Wang, J. (2011). Do Mood Swings Drive Business Cycles and Is It Rational? Technical report, National Bureau of Economic Research.

Beaudry, P. and Portier, F. (2004). "An Exploration into Pigou's Theory of Cycles". *Journal of Monetary Economics*, 51(6):1183–1216.

Beaudry, P. and Portier, F. (2006). "Stock Prices, News, and Economic Fluctuations". *American Economic Review*, 96(4):1293–1307.

Beaudry, P. and Portier, F. (2014). "News-Driven Business Cycles: Insights and Challenges". *Journal of Economic Literature*, 52(4):993–1074.

Becker, R. A. (1980). "On the Long-run Steady State in a Simple Dynamic Model of Equilibrium with Heterogeneous Households". *The Quarterly Journal of Economics*, 95(2):375–382.

Benhabib, J. and Farmer, R. E. (1994). "Indeterminacy and Increasing Returns". *Journal of Economic Theory*, 63(1):19–41.

Bernanke, B. S. and Mishkin, F. S. (1997). "Inflation Targeting: A New Framework for Monetary Policy?". Technical report, National Bureau of Economic Research.

Blanchard, O. J. and Kahn, C. M. (1980). The Solution of Linear Difference Models under Rational Expectations. *Econometrica: Journal of the Econometric Society*, pages 1305–1311.

Blanchard, O. J., L'Huillier, J.-P., and Lorenzoni, G. (2013). "News, Noise, and Fluctuations: An Empirical Exploration". *American Economic Review*, 103(7):3045–70.

Blinder, A. S. (1999). *"Central Banking in Theory and Practice"*. Mit press.

Blinder, A. S. (2013). *"After the Music Stopped: The Financial Crisis, the Response, and the Work Ahead"*. Penguin.

Branch, W., Evans, G. W., and McGough, B. (2012a). Finite-horizon learning. In *CDMA Working Paper*.

Branch, W., Evans, G. W., and McGough, B. (2012b). "Finite Horizon Learning". CDMA Working Paper Series 201204, Centre for Dynamic Macroeconomic Analysis.

Bullard, J. and Eusepi, S. (2014). ”When Does Determinacy Imply Expectational Stability?”. *International Economic Review*, 55(1):1–22.

Burnside, C., Eichenbaum, M., and Rebelo, S. (1995). “Capital Utilization and Returns to Scale”. In *NBER Macroeconomics Annual 1995, Volume 10*, pages 67–124. MIT Press.

Cho, S. (2015). “Determinacy and E-stability Under Reduced-form Learning”. *Economics Letters*.

Christiano, L. J., Eichenbaum, M., and Evans, C. L. (2005). Nominal Rigidities and the Dynamic Effects of a Shock to Monetary Policy. *Journal of political Economy*, 113(1):1–45.

Dombeck, B. (2016). “Implications of Solving the Co-movement Problem vis-a-vis Expectational Stability”.

Ellison, M. and Pearlman, J. (2011). “Saddlepath Learning”. *Journal of Economic Theory*, 146(4):1500–1519.

Eusepi, S. and Preston, B. (2010). “Central Bank Communication and Expectations Stabilization”. *American Economic Journal: Macroeconomics*, 2(3):235–71.

Eusepi, S. and Preston, B. (2011). “Expectations, Learning, and Business Cycle Fluctuations”. *American Economic Review*, 101(6):2844–72.

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

Evans, G. W. and Honkapohja, S. (2006). “Monetary Policy, Expectations and Commitment”. *The Scandinavian Journal of Economics*, 108(1):15–38.

Evans, G. W. and McGough, B. (2015). ”Learning to Optimize”. Working paper.

Farmer, R. E. and Guo, J.-T. (1994). “Real Business Cycles and the Animal Spirits Hypothesis”. *Journal of Economic Theory*, 63(1):42–72.

Greenwood, J., Hercowitz, Z., and Huffman, G. W. (1988). “Investment, Capacity Utilization, and the Real Business Cycle”. *The American Economic Review*, pages 402–417.

Guerrón-Quintana, P. and Nason, J. M. (2012). ”Bayesian Estimation of DSGE Models.

Guo, J.-T., Sirbu, A.-I., and Weder, M. (2015). “News About Aggregate Demand and the Business Cycle”. *Journal of Monetary Economics*, (0).

Haario, H., Saksman, E., and Tamminen, J. (2001). An Adaptive Metropolis Algorithm. *Bernoulli*, pages 223–242.

Hellwig, C. (2005). “Heterogeneous Information and the Benefits of Transparency”. Technical report, UCLA mimeo.

Hellwig, C. (2010). ”Comment on ”Noisy Business Cycles”. Technical report, National Bureau of Economic Research.

Hommes, C., Sonnemans, J., Tuinstra, J., and Van De Velden, H. (2007). “Learning in Cobweb Experiments”. *Macroeconomic Dynamics*, 11(S1):8–33.

Honkapohja, S., Mitra, K., and Evans, G. W. (2012). ”Notes on Agents Behavioral Rules Under Adaptive Learning and Studies of Monetary Policy”.

Jaimovich, N. and Rebelo, S. (2009). “Can News about the Future Drive the Business Cycle?”. *American Economic Review*, 99(4):1097–1118.

Karnizova, L. (2010). ”News versus Sunspot Shocks in a New Keynesian Model”. *Economics: The Open-Access, Open-Assessment E-Journal*, 4.

Karnizova, L. V. (2007). *”News versus Sunspot Shocks in Linear Rational Expectations Models”*. Department of Economics, University of Ottawa= Dép. de Science économique, Université d’Ottawa.

Khan, H. and Tsoukalas, J. (2012). “The Quantitative Importance of News Shocks in Estimated DSGE Models”. *Journal of Money, Credit and Banking*, 44(8):1535–1561.

King, R. G., Plosser, C. I., and Rebelo, S. T. (1988). “Production, Growth and Business Cycles: I. The Basic Neoclassical Model”. *Journal of monetary Economics*, 21(2):195–232.

Klein, P. (2000). Using the Generalized Schur Form to Solve a Multivariate Linear Rational Expectations Model. *Journal of Economic Dynamics and Control*, 24(10):1405–1423.

Krusell, P. and McKay, A. (2010). “News Shocks and Business Cycles”. *FRB Richmond Economic Quarterly*, 96(4):373–397.

Kydland, F. E. and Prescott, E. C. (1982). “Time to Build and Aggregate Fluctuations”. *Econometrica: Journal of the Econometric Society*, pages 1345–1370.

Leeper, E. M., Richter, A. W., and Walker, T. B. (2012). “Quantitative Effects of Fiscal Foresight”. *American Economic Journal: Economic Policy*, 4(2):115–44.

Lorenzoni, G. (2009). "A Theory of Demand Shocks". *American Economic Review*, 99(5):2050–84.

Lucas, R. E. (1976). "Econometric Policy Evaluation: A Critique". In *Carnegie-Rochester conference series on public policy*, volume 1, pages 19–46. Elsevier.

Lucas, R. E. (1978). "Asset Prices in an Exchange Economy". *Econometrica: Journal of the Econometric Society*, pages 1429–1445.

Marcet, A. and Sargent, T. J. (1989a). "Convergence of Least-Squares Learning in Environments with Hidden State Variables and Private Information". *The Journal of Political Economy*, pages 1306–1322.

Marcet, A. and Sargent, T. J. (1989b). "Convergence of Least Squares Learning Mechanisms in Self-Referential Linear Stochastic Models". *Journal of Economic theory*, 48(2):337–368.

Marcet, A. and Sargent, T. J. (1989c). "Least-Squares Learning and the Dynamics of Hyperinflation". In *International Symposia in Economic Theory and Econometrics, edited by William Barnett, John Geweke, and Karl Shell*, pages 119–137.

McCallum, B. T. (2007). "E-stability vis-a-vis Determinacy Results for a Broad Class of Linear Rational Expectations Models". *Journal of Economic Dynamics and Control*, 31(4):1376–1391.

McFadden, D. (1989). "A Method of Simulated Moments for Estimation of Discrete Response Models Without Numerical Integration". *Econometrica: Journal of the Econometric Society*, pages 995–1026.

McGough, B. (2006). "Shocking Escapes". *Economic Journal*, 116(511):507–528.

Milani, F. (2007). "Expectations, Learning and Macroeconomic Persistence". *Journal of Monetary Economics*, 54(7):2065–2082.

Milani, F. and Rajbhandari, A. (2012). Expectation Formation and Monetary DSGE Models: Beyond the Rational Expectations Paradigm. *Advances in Econometrics*, 28:253.

Mitra, K., Evans, G. W., and Honkapohja, S. (2013). "Policy Change and Learning in the RBC Model". *Journal of Economic Dynamics and Control*, 37(10):1947 – 1971.

Morris, S. and Shin, H. S. (2002). "Social Value of Public Information". *The American Economic Review*, 92(5):1521–1534.

Morris, S., Shin, H. S., and Tong, H. (2006). “Social Value of Public Information: Morris and Shin (2002) Is Actually Pro-Transparency, Not Con: Reply”. *The American Economic Review*, pages 453–455.

Muth, J. F. (1961). “Rational Expectations and the Theory of Price Movements”. *Econometrica: Journal of the Econometric Society*, pages 315–335.

Pigou, A. C. (1927). *“Industrial Fluctuations”*. Macmillan.

Preston, B. J. (2005). ”Learning About Monetary Policy Rules When Long-horizon Expectations Matter”. 1.

Ramey, V. A. (2009). “Identifying Government Spending Shocks: It’s All in the Timing”. Working Paper 15464, National Bureau of Economic Research.

Ramey, V. A. (2011). “Identifying Government Spending Shocks: It’s all in the Timing”. *The Quarterly Journal of Economics*, 126(1):1–50.

Ruge-Murcia, F. J. (2007). Methods to Estimate Dynamic Stochastic General Equilibrium Models. *Journal of Economic Dynamics and Control*, 31(8):2599–2636.

Schmitt-Grohe, S. and Uribe, M. (2012). “What’s News in Business Cycles”. *Econometrica*, 80(6):2733–2764.

Shaby, B. and Wells, M. T. (2010). Exploring an adaptive metropolis algorithm. Technical report.

Sims, C. A. (2002). Solving Linear Rational Expectations Models. *Computational economics*, 20(1):1–20.

Sims, E. (2016). What’s news in News? A cautionary note on using a variance decomposition to assess the quantitative importance of news shocks. *Journal of Economic Dynamics and Control*, 73(1):41–60.

Slobodyan, S. and Wouters, R. (2007). Adaptive Learning in an Estimated Medium–Size DSGE Model. Technical report, mimeo.

Slobodyan, S. and Wouters, R. (2008). Estimating a Medium–scale DSGE Model with Expectations Based on Small Forecasting Models. Technical report, mimeo.

Slobodyan, S. and Wouters, R. (2012). Learning in an Estimated Medium-scale DSGE Model. *Journal of Economic Dynamics and control*, 36(1):26–46.

Smets, F. and Wouters, R. (2007). “Shocks and Frictions in US Business Cycles: A Bayesian DSGE Approach”. *American Economic Review*, 97(3):586–606.

Stiglitz, J. E. (2014). “Reconstructing Macroeconomic Theory to Manage Economic Policy”. Working Paper 20517, National Bureau of Economic Research.

Svensson, L. E. O. (2006). “Social Value of Public Information: Comment: Morris and Shin (2002) Is Actually Pro-Transparency, Not Con”. *American Economic Review*, 96(1):448–452.

Tong, H. (2004). “Do Transparency Standards Improve Macroeconomic Forecasting?”. *Unpublished Paper*, 455.

Uhlig, H. et al. (1995). A Toolkit for Analyzing Nonlinear Dynamic Stochastic Models Easily.

Veldkamp, L. L. (2011). *“Information Choice in Macroeconomics and Finance”*. Princeton University Press.

Williams, N. (2003a). ”Adaptive Learning and Business Cycles”.

Williams, N. (2003b). ”Adaptive Learning and Business Cycles. *Manuscript, Princeton University.*

Woodford, M. (2007). “The Case for Forecast Targeting as a Monetary Policy Strategy”. *Journal of Economic Perspectives*, 21(4):3–24.

Woodford, M. (2013). “Fedspeak: Does it Matter How Central Bankers Explain Themselves?”.

Yang, S.-C. S. and Traum, N. (2011). “When Does Government Debt Crowd Out Investment?”. In *2011 Meeting Papers*, number 479. Society for Economic Dynamics.