Incorporating Diagnostic Expectations into the New Keynesian Framework*

Jean-Paul L'Huillier† Sanjay R. Singh‡ Donghoon Yoo§

August 2022

Abstract

Diagnostic expectations constitute a realistic behavioral model of inference. This paper shows that this approach to expectation formation can be productively integrated into the New Keynesian framework. To this end, we start by offering a technical treatment of diagnostic expectations in linear macroeconomic models. Diagnostic expectations generate *endogenous* extrapolation in general equilibrium. We show that diagnostic expectations generate extra amplification in the presence of nominal frictions; a fall in aggregate supply generates a Keynesian recession; fiscal policy is more effective at stimulating the economy; with imperfect information, diagnostic expectations generate *delayed* overreaction of aggregate variables. We perform Bayesian estimation of a rich medium-scale model that incorporates consensus forecast data. Our estimate of the diagnosticity parameter is in line with previous studies. Moreover, we find empirical evidence in favor of the diagnostic model. Diagnostic expectations offer new propagation mechanisms to explain fluctuations.

Keywords: Heuristics, representativeness, general equilibrium, shocks, volatility.

JEL codes: E12, E32, E71.

---

*This paper circulated previously under the title "Diagnostic Expectations and Macroeconomic Volatility". We thank Francesco Bianchi, Ryan Chahrour, Cosmin Ilut, Kiminori Matsuyama, Peter Maxted, Hikaru Saijo, Andrei Shleifer, Stephen Terry, Takayuki Tsuruga, and Harald Uhlig for comments and useful discussions. We have also benefited from comments at the AFES, NASMES, the CEF Conference, EEA-ESEM, and the SED Meeting, and seminar participants at Keio University, New School of Economics, VEAMS, and UC Davis. The views expressed herein are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Cleveland, or the Federal Reserve System. All errors are ours.

†Federal Reserve Bank of Cleveland and Department of Economics, Brandeis University (jplhuillier2010@gmail.com).

‡Department of Economics, University of California, Davis (sjrsingh@ucdavis.edu).

§Institute of Social and Economic Research, Osaka University (donghoonyoo@iser.osaka-u.ac.jp).

1 Introduction

Diagnostic expectations (DE) have emerged as an important departure from rational expectations in macroeconomics and finance. Among the host of possible deviations from rational expectations, there are three broad reasons that make diagnostic expectations a leading alternative to consider for macroeconomic modeling. First, diagnostic expectations constitute a microfounded deviation immune to the Lucas critique. Second, this approach lends itself to a great deal of tractability, as a number of recent efforts in macroeconomics and finance have demonstrated (see Bordalo, Gennaioli, and Shleifer 2018; Bordalo, Gennaioli, Ma, and Shleifer 2020; Bordalo, Gennaioli, Shleifer, and Terry 2021, among others). Third, based on the pathbreaking and influential work on the "representativeness heuristic" by Kahneman and Tversky (1972), one ought to consider this behavioral model as fundamentally realistic, and thereby portable across fields of economics.[1]

In this paper, we argue that diagnostic expectations can be productively incorporated into the New Keynesian (NK) framework. We demonstrate this claim in two parts, analytical and empirical. Analytically, using a three-equation NK model, we show how diagnostic expectations bring rich insights on four issues raised by the literature. Empirically, by integrating diagnostic expectations into a rich medium-scale DSGE model, we find that diagnostic expectations provide a superior fit of business cycle and consensus forecast data. Our analysis brings novel implications for the interpretation of fluctuations.

The first analytical issue we tackle is that of amplification and propagation in general equilibrium. As shown in previous work (Bordalo, Gennaioli, and Shleifer 2018, henceforth BGS), diagnostic expectations (DE) imply an extrapolation of current shocks into the future. Intuitively, this could generate extra volatility for endogenous variables. We show that this intuition is in fact not guaranteed. In the presence of nominal frictions (as in the NK model) DE generate extra volatility; in a frictionless representative agent real business cycle (RBC) model, general equilibrium channels shut down the effect of DE, and output is *less* volatile under DE than under rational expectations (RE).[2]

The second issue considered is whether a fall in aggregate supply can cause a demand shortage. Since the onset of the COVID-19 pandemic, there is a renewed interest

[1]Simply put, the representativeness heuristic is the general human tendency to over-estimate how representative a small sample is, a pattern documented in a large body of literature in psychology and behavioral economics. For a survey and more detailed discussion, see Kahneman, Slovic, and Tversky (1982).

[2]Bordalo, Gennaioli, Shleifer, and Terry (2021) consider financial frictions and how DE generate realistic credit cycles in a real economy. See Section 4 for a broader discussion.

on whether supply-side disruptions can ultimately generate shortfalls in aggregate demand (see Guerrieri, Lorenzoni, Straub, and Werning 2022; Fornaro and Wolf 2021; Caballero and Simsek 2021; Bilbiie and Melitz 2020, among others.) Whereas the rational expectations NK (RE-NK) model generates the opposite prediction, we show that adding DE into the NK framework (DE-NK) allows for the possibility of "Keynesian supply shocks": Following a negative supply shock, diagnostic agents extrapolate the shock into the future, and hence become excessively pessimistic. This pushes them to reduce consumption drastically, generating a Keynesian recession. Later, beliefs systematically revert, and the economy features a boom, as in the RE-NK model.

The third issue we tackle concerns government policy. We show how endogenous extrapolation arising from the evaluation of the inflation process by diagnostic agents can significantly raise the government spending multiplier. Current surprise inflation causes the diagnostic agent to expect future inflation thereby reducing the subjective real interest rate. When the diagnosticity parameter is higher than the coefficient governing the reaction of the monetary authority to inflation, the DE-NK model is able to generate a multiplier greater than 1 even with i.i.d. government spending shocks. We show how this analytical conclusion can be challenged by the degree of extrapolation of the exogenous shock process, which depends on the persistence of this shock. If the shock is persistent enough, the DE of future spending can completely crowd out current consumption and lead to a multiplier that is equal to 0, or even negative. Hence, the degree of diagnosticity allows the model to span a wide range of multipliers, highlighting the importance of the behavioral friction in this context.

With an eye to the large macroeconomics literature on information frictions, the fourth question we consider concerns under- and overreaction of expectations (Coibion and Gorodnichenko 2015a; Bordalo, Gennaioli, Ma, and Shleifer 2020). Based on previous work (Lorenzoni 2009; Blanchard, L'Huillier, and Lorenzoni 2013), we extend DE-NK model to a setting where the consumers receive noisy signals about the future path of their income. Beliefs about their long-run income determine aggregate consumption and output due to nominal rigidities. We show that a plausible calibration of the imperfect information DE-NK model can generate both short-run underreaction, and an overreaction over the medium-term. Combining diagnostic expectations with information frictions can deliver rich implications for the path of agents' beliefs in general equilibrium models.

On the empirical front, we let DE and RE compete within a standard medium-scale DSGE model. Using Bayesian methods, we evaluate the relative fitness of both approaches when applied to post-war U.S. data. We include both business cycle and forecast data in the estimation. In order to submit the behavioral expectational friction to a stringent empirical test, the model we consider contains a large number of benchmark frictions and shocks drawn from the seminal works by Christiano, Eichenbaum, and Evans (2005) and Smets and Wouters (2007). For the same reason, we include news shocks and information frictions in the form of noise shocks to expectations. We find empirical evidence in favor of DE. In comparison with the RE model, variance decomposition and parameter estimates indicate that the DE model relies significantly less on noise shocks when explaining expectations errors. Moreover, the DE model relies more on internal propagation mechanisms than on exogenous shocks to account for the dynamics of price and wage inflation in the data.

An organizing theme across applications is that diagnosticity generates extra volatility and systematic reversals in beliefs. In dynamic settings, diagnostic expectations (DE) introduce an extrapolative mechanism. Upon impact of a shock, beliefs overreact. Subsequently, beliefs systematically revert, which the diagnostic agents fail to anticipate. Our work shows that those two insights from BGS have implications for NK models.

A recurrent theme in our paper is that when agents have diagnostic beliefs about endogenous variables, instead of exogenous processes, new behavioral insights emerge. Endogenous extrapolation, as highlighted in our fiscal policy exercise, has remarkable economic implications. We provide two examples of models with endogenous extrapolation at the end of Section 3.

The paper provides a substantial technical contribution: We develop a solution method for a general class of linear DSGE models with diagnostic expectations. The key to our method is to formally establish the existence and uniqueness of a rational expectations representation of the diagnostic expectations model, a challenging task in the presence of endogenous states. This result allows us to compute the equilibrium diagnostic expectation of endogenous variables. We briefly make a few technical remarks about our solution method. First, we show that incorporating DE requires researchers to loglinearize the model from scratch rather than simply replacing the rational expectations operator with the corresponding diagnostic expectations operator in linear economies. For a given set of equilibrium conditions obtained from first principles, the presence of DE actually changes the loglinear equilibrium conditions that constitute a correct approximation.[3] We explain, in detail, how to obtain the correct approximation and provide a few examples. Loglinearization under DE brings forward

[3]This is different from many other departures from the full-information rational expectations case, as for example the introduction of imperfect information (Woodford 2002) or other behavioral models (Garcia-Schmidt and Woodford 2019), where the structure of equilibrium conditions of the loglinear model does not change.

novel economic insights in forward looking models. For instance, the Euler equation for the diagnostic consumer features a real rate that contains a new term pertaining to surprise in current inflation from previous forecasts (see Section 3.4.2). Second, we provide sharp results on the stability and the existence of a bounded solution with DE. While the stability conditions are same as in the corresponding RE model, we note that the solution under DE can be explosive for certain limiting values of the diagnosticity parameter. Researchers may need to exercise caution when applying DE to endogenous variables.

Related Literature. The paper is primarily related to the emerging literature on DE. See Gennaioli and Shleifer (2018) for a review. Most closely related are papers by Maxted (2022) and Bordalo, Gennaioli, Shleifer, and Terry (2021), who incorporate DE in macro-finance frameworks.[4] Maxted (2022) shows that incorporating DE into a macro-finance framework can reproduce several facts surrounding financial crises (see also Krishnamurthy and Li 2020). Bordalo, Gennaioli, Shleifer, and Terry (2021) show that DE can quantitatively generate countercyclical credit spreads in a heterogeneous firms business-cycle model. We complement these efforts by providing a general treatment of DE in linear macroeconomic models. In particular, we show how incorporating DE into NK models (Woodford 2003; Galí 2015) delivers rich new insights and significantly improves the fit to the data.

In parallel and complementary work, Bianchi, Ilut, and Saijo (2022) also investigate applications of DE in linear models. Although their work, like ours, is comprehensive, the main focus of their paper is distant memory, the notion that agents' reference distribution looks back more than one period. In such settings, the law of iterated expectations fails, and therefore the model with distant memory is time inconsistent. Our paper focuses exclusively on linear settings with time consistency, and shows that this baseline setup offers a number of insights useful for the NK literature. We outline, in detail, the steps from the exact equilibrium conditions to the loglinear approximation of medium-scale models. Our main empirical focus is evaluating the role of diagnosticity in a rich medium-scale DSGE model with news shocks and information frictions.

Our paper also speaks to the literature proposing deviations from the full-information rational expectations (FIRE) hypothesis. See, for example, Mankiw and Reis (2002), Coibion and Gorodnichenko (2015a), Angeletos, Huo, and Sastry (2020), Bordalo,

[4] D'Arienzo (2020) investigates the ability of DE to reconcile the overreaction of expectations of long rates relative to the expectations of short rates to news in bond markets. Ma, Ropele, Sraer, and Thesmar (2020) quantify the costs of managerial biases.

Gennaioli, Ma, and Shleifer (2020), Kohlhas and Walther (2021), among others.

Angeletos, Huo, and Sastry (2020) document delayed overreaction of beliefs in response to business cycle shocks. Bordalo, Gennaioli, Ma, and Shleifer (2020) propose a model of DE with dispersed information to study underreaction and overreaction in survey forecasts. See also Ma, Ropele, Sraer, and Thesmar (2020) and Afrouzi, Kwon, Landier, Ma, and Thesmar (2020). We complement these analyses by showing that one can obtain delayed overreaction in an imperfect information DE-NK model. With respect to earlier work, there are two innovations in our procedure. First, we use a microfounded behavioral friction. Second, we generate these patterns with expectations in general equilibrium models. In a related vein, our estimated DSGE model builds on work exploring business cycle models where agents receive advance information about future productivity that is subject to an information friction (Blanchard, L'Huillier, and Lorenzoni 2013; Chahrour and Jurado 2018).

Our paper fits into the macroeconomics literature that models departures from rational expectations with various behavioral assumptions. Some of the recent applications have focused on resolving puzzles in New Keynesian models by introducing behavioral assumptions. Angeletos and Lian (2018), Farhi and Werning (2019), Gabaix (2020), and Garcia-Schmidt and Woodford (2019) are some of the papers that propose departures from rational expectations to attenuate the strength of forward guidance. Iovino and Sergeyev (2020) study the effectiveness of central bank balance sheet policies with level-k thinking. Bianchi-Vimercati, Eichenbaum, and Guerreiro (2022) study the effectiveness of fiscal policy at the zero lower bound in a model with level-k thinking. Angeletos, Huo, and Sastry (2020, Sec. 6.4) argue that these leading departures from rational expectations exhibit a form of under-extrapolation. In contrast, DE allow beliefs to generate overreaction and systematic reversals as we demonstrate. Farhi and Werning (2020) study the role of monetary policy as a macro-prudential tool when agents form extrapolative expectations.

Paper Organization. The paper is organized as follows. Section 2 starts with an example based on the classic demand and supply model by Muth (1961) to demonstrate the two key insights we get from diagnostic expectations: extra volatility and systematic reversals in beliefs. Section 3 presents our solution method, discusses stability, and provides examples illustrating endogenous propagation of diagnostic beliefs. Section 4 presents the analytical results from a 3-equation NK model. Section 5 presents the empirical evaluation of diagnostic expectations in a medium-scale DSGE model. Section 6 concludes. The Appendix provides supplementary materials and collects all the proofs.

2 Diagnostic Expectations on Endogenous Variables: A Demand and Supply Example

We begin by presenting the economic implications of diagnostic expectations on endogenous objects in a simple setting. To this end, we consider the classic model by Muth (1961), which we augment with the inference biases generated by diagnosticity (Kahneman and Tversky 1972). The benchmark for comparison is the case of rational expectations. The model by Muth is simple and transparent, allowing us to distill the behavioral insights offered by diagnostic expectations in the context of more complex models of forward-looking agents. This discussion will concentrate on two main features of diagnostic expectations: the excess volatility implied by belief overreaction, together with the systematic reversals that occur when uncertainty is realized. Appendix A provides details and proofs.

The model is as follows.[5] There is an isolated market for a commodity. The commodity demand at time t, Q_t^d, is a downward-sloping function of the price P_t (model variables are denoted in deviation from steady state):

$$Q_t^d = -\beta P_t, \quad \beta > 0 \tag{1}$$

The supply side is modeled with a time-to-build assumption. Suppliers invest, one period in advance, as a function of their expectations of the price next period:

$$I_t = \gamma \tilde{\mathbb{E}}_t[P_{t+1}], \quad \gamma > 0 \tag{2}$$

where I_t is the quantity invested and $\tilde{\mathbb{E}}_t$ is an expectation operator (rational or diagnostic). Supply at time t, Q_t^s, is equal to the quantity invested at time $t-1$ plus a persistent disturbance u_t:

$$Q_t^s = I_{t-1} + u_t \tag{3}$$

where $u_t = \rho_u u_{t-1} + \epsilon_t$, and ϵ_t is i.i.d. $N(0, \sigma_\epsilon^2)$. An absence of storage opportunities implies the market clearing condition: $Q_t^s = Q_t^d$.

We are interested in solving for the equilibrium behavior of suppliers, which depends on expectation formation, and on the implied equilibrium price dynamics. Under RE ($\tilde{\mathbb{E}}_t = \mathbb{E}_t$), suppliers' expectations are, on average, correct, with the exogenous

[5] See Muth (1961), Section 3, pp. 317-22.

disturbance as the only source of discrepancy between expectations of the price and actual realizations. In other words, deviations of the realized price from expectations are *unpredictable*. To see this, combine the equations above to get the equilibrium condition

$$P_t = -\frac{\gamma}{\beta}\tilde{\mathbb{E}}_{t-1}[P_t] - \frac{1}{\beta}u_t \tag{4}$$

Conjecture that the solution takes the form

$$I_t = A_1 u_t + A_2 \epsilon_t \tag{5}$$

$$P_t = B_1 u_{t-1} + B_2 \epsilon_t + B_3 \epsilon_{t-1} \tag{6}$$

where A_1, A_2, B_1, B_2 and B_3 are parameters to solve for. The RE solution is given by $A_1 = -\gamma\rho_u/(\beta+\gamma)$, $B_1 = A_1/\gamma$, and $B_2 = -1/\beta$ (with $A_2 = B_3 = 0$).

Taking the RE expectation on both sides of (4) shows that, in equilibrium, the expected price needs to satisfy $\mathbb{E}_{t-1}[P_t] = -(1/(\beta+\gamma))\mathbb{E}_{t-1}[u_t]$, which is guaranteed by the RE solution. Moreover, the only ex-post deviation from this forecast is $B_2\epsilon_t$, which is unpredictable.

Under DE (denoted as $\tilde{\mathbb{E}}_t = \mathbb{E}_t^\theta$, where $\theta > 0$ is a diagnosticity parameter), suppliers' expectations are, on average, excessively volatile. The BGS formula for the DE of an endogenous object, say the price, is $\mathbb{E}_t^\theta[P_{t+1}] = \mathbb{E}_t[P_{t+1}] + \theta\left(\mathbb{E}_t[P_{t+1}] - \mathbb{E}_{t-1}[P_{t+1}]\right)$.[6] This formula says that the DE is equal to the RE plus a distortion term capturing the extrapolative behavior of the agent. The solution under DE is given by $A_2 = -\beta\gamma\theta\rho_u/[(\beta+\gamma)(\beta+\gamma(1+\theta))]$, $B_3 = -A_2/\beta$ (with A_1, B_1 and B_2 identical to the RE case).

The implications of DE for the Muth model are as follows. First, investment is excessively volatile. By way of example, suppose that the economy is in steady state at $t-1$ and that a *negative* shock $\epsilon_t < 0$ hits. As under RE, suppliers anticipate a higher price to prevail at time $t+1$ due to the exogenous supply contraction at time t. Consequently, they increase their investment at time t ($A_1 < 0$). However, under DE, suppliers extrapolate the shock into the future. Their expectations about prices overreact. Thus, they increase investment *by more* than under rational expectations ($A_2 < 0$). Ex-post, at time $t+1$, there is a reversal. The market is glutted with an excessively high quantity of the commodity. The rosy expectations of diagnostic suppliers are disappointed, with the price rising by less than in the RE economy ($B_3 > 0$). Moreover, this discrepancy from expectations is systematic and predictable, since

[6] Section 3 will show that the same formula, obtained in BGS for exogenous objects, can indeed be applied to endogenous objects.

it depends purely on the past shock. These dynamics are neglected by agents. In sum, DE deliver excess investment volatility and a subsequent reversal. We will return to these two features in various applications of the New Keynesian model.

In his seminal article, Muth also was interested in the possibility of ad-hoc deviations from rational behavior. For a general MA(∞) specification for the exogenous process, Muth explored expected price dynamics that "over- or under-discount" the recent price history. A fascinating lesson that emerges is that the DE specification by BGS turns out to microfound this early idea. The diagnostic supplier over-reacts to the shock at t, resulting in an excessive discounting of information contained in prices up to time $t-1$.[7]

In variant of this simple model that includes inventory dynamics, a novel extrapolation mechanism emerges when DE is taken over endogenous variables. Muth introduced inventory speculation that gives rise to an inventory demand equation. Inventory demand, S_t, depends on the difference between the expected future price and the current price. The market clearing condition is now replaced with $Q_t^d+S_t = Q_t^s+S_{t-1}$, allowing for storage without depreciation.[8] Because now inventory stock holdings constitute an endogenous state variable, the DE and the RE solutions of the model do not coincide even when the shock is i.i.d. ($\rho_u = 0$). This is in contrast to the baseline case without inventory accumulation, where the two solutions coincide when shocks are i.i.d. This particular feature is attractive since endogenous state variables (such as capital, consumption habits, etc.) are ubiquitous in DSGE models. In the next section, we zoom into this point in the context of a tractable, univariate endogenous-state model. But first, we present a general solution method for linear DSGE models when diagnostic expectations are computed on endogenous variables.

3 Solution Method

We present a solution method for a general class of linear models. Agents use diagnostic expectations to form beliefs about the evolution of all variables, exogenous and endogenous. Our strategy consists in obtaining a rational expectations (RE) representation of the diagnostic expectations (DE) model. Based on this step, the model can be solved using standard techniques. We also establish results for the existence and stability of the solution under DE.

[7]Specifically, as the appendix shows, solving the Muth model under the general process (3.6) (p. 319) and DE, microfounds the Muth factor $f_1 = 1+\theta$ (where, following the author's terminology, the case $f_1 > 1$ corresponds to over-discounting).

[8]See Muth (1961), Section 4, pp. 322-330.

3.1 General Formulation and Rational Expectations Representation

3.1.1 Exogenous Processes

We start by specifying the exogenous drivers of the economy. Exogenous variables are stacked in a $(n \times 1)$ vector $\mathbf{x}_t$ that is assumed to follow the multivariate AR(1) stochastic process

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{v}_t \tag{7}$$

where $\mathbf{v}_t$ is a $(k \times 1)$ vector of Normal and orthogonal exogenous shocks, $\mathbf{v}_t \sim N(0, \Sigma_{\mathbf{v}})$, and $\mathbf{A}$ is a diagonal matrix of persistence parameters. Since vector $\mathbf{x}_{t+1}$ follows a multivariate normal distribution, we can write its true (or non-distorted) pdf as $f(\mathbf{x}_{t+1}|\mathbf{x}_t) \propto \varphi((\mathbf{x}_{t+1} - \mathbf{A}\mathbf{x}_t)'\Sigma_{\mathbf{v}}^{-1}(\mathbf{x}_{t+1} - \mathbf{A}\mathbf{x}_t))$, where $\varphi(x)$ is the density of a standard normal distribution, $\varphi(x) = \frac{1}{\sqrt{2\pi}}e^{-\frac{1}{2}x^2}$.

3.1.2 Diagnostic Expectations

Extending the approach by Bordalo, Gennaioli, and Shleifer (2018) (henceforth BGS), the multivariate diagnostic distribution of x_{t+1} is defined as

$$f_t^{\theta}(\mathbf{x}_{t+1}) = f(\mathbf{x}_{t+1}|G_t) \cdot \left[\frac{f(\mathbf{x}_{t+1}|G_t)}{f(\mathbf{x}_{t+1}|-G_t)}\right]^{\theta} \cdot C \tag{8}$$

where G_t and $-G_t$ are conditioning events. G_t encodes current conditions: $G_t \equiv \{\mathbf{x}_t = \check{\mathbf{x}}_t\}$, where $\check{\mathbf{x}}_t$ denotes the realization of $\mathbf{x}_t$.[9] $-G_t$ encodes a reference group (i.e. a reference event), that is used to compute the reference distribution $f(\mathbf{x}_{t+1}|-G_t)$. Due to the representativeness heuristic, agents overweight the last realization of $\mathbf{x}_t$ (relative to the reference group) when forming beliefs about the future realization of $\mathbf{x}_{t+1}$. The likelihood ratio $f(\mathbf{x}_{t+1}|G_t)/f(\mathbf{x}_{t+1}|-G_t)$ distorts beliefs to a degree governed by the diagnosticity parameter $\theta \geq 0$. C is a constant ensuring that $f_t^{\theta}(\mathbf{x}_{t+1})$ integrates to 1.

Following BGS, we impose that, in the presence of uncertainty about $\mathbf{x}_{t+1}$, the reference event $-G_t$ carries "no news" at time t (henceforth no-news assumption or NNA).

Assumption 1 (Multivariate No-News Assumption)

$$f(\mathbf{x}_{t+1}|-G_t) = f(\mathbf{x}_{t+1}|\mathbf{x}_t = \mathbf{A}\check{\mathbf{x}}_{t-1}) \tag{9}$$

[9] We do not use the same notation $\hat{\mathbf{x}}_t$ for realizations as BGS, since we have reserved hats over variables for loglinear deviations below.

We make Assumption 1 throughout the paper. To understand the meaning of this assumption, consider an agent forming beliefs about future $\mathbf{x}_{t+1}$. Under the NNA, these beliefs are formed conditional on the event that the random variable $\mathbf{x}_t$, conditional on the past realization $\check{\mathbf{x}}_{t-1}$, is what it was expected to be, so $\mathbf{v}_t = \mathbb{E}[\mathbf{v}_t] = 0$, which is equivalent to $\mathbf{x}_t = \mathbf{A}\check{\mathbf{x}}_{t-1}$. The diagnostic distribution is thus written as

$$f_t^{\theta}(\mathbf{x}_{t+1}) = f(\mathbf{x}_{t+1}|\mathbf{x}_t = \check{\mathbf{x}}_t) \cdot \left[\frac{f(\mathbf{x}_{t+1}|\mathbf{x}_t = \check{\mathbf{x}}_t)}{f(\mathbf{x}_{t+1}|\mathbf{x}_t = \mathbf{A}\check{\mathbf{x}}_{t-1})}\right]^{\theta} \cdot C \tag{10}$$

Notice that the distribution (10) is conditional on two elements: first, it is conditional on the current realization of $\mathbf{x}_t$, written $\check{\mathbf{x}}_t$, because this enters the true distribution of $\mathbf{x}_{t+1}$; second, it is conditional on the reference event $-G_t \equiv \{\mathbf{x}_t = \mathbf{A}\check{\mathbf{x}}_{t-1}\}$, which depends on the realization at $t-1$, $\check{\mathbf{x}}_{t-1}$.

Extending the definition of BGS to the multivariate normal vector $\mathbf{x}_{t+1}$, the DE is the expectation, element by element, under the density (10). We write this expectation as $\mathbb{E}_t^{\theta}[\mathbf{x}_{t+1}]$.[10] Using a multivariate version of Proposition 1 in BGS, we obtain the formula[11]

$$\mathbb{E}_t^{\theta}[\mathbf{x}_{t+1}] = \mathbb{E}_t[\mathbf{x}_{t+1}] + \theta(\mathbb{E}_t[\mathbf{x}_{t+1}] - \mathbb{E}_{t-1}[\mathbf{x}_{t+1}]) \tag{11}$$

3.1.3 Stochastic Difference Equation

The class of forward-looking models we analyze is written as a stochastic difference equation. Uncertainty is modeled under the diagnostic distribution (10). Let $\mathbf{y}_t$ denote a $(m \times 1)$ vector of endogenous variables (including jump variables and states) and $\mathbf{x}_t$, as above, denote the $(n \times 1)$ vector of exogenous states. The model is:

$$\mathbb{E}_t^{\theta}[\mathbf{F}\mathbf{y}_{t+1} + \mathbf{G_1}\mathbf{y}_t + \mathbf{M}\mathbf{x}_{t+1} + \mathbf{N_1}\mathbf{x}_t] + \mathbf{G_2}\mathbf{y}_t + \mathbf{H}\mathbf{y}_{t-1} + \mathbf{N_2}\mathbf{x}_t = 0 \tag{12}$$

where $\mathbf{F}$, $\mathbf{G_1}$, $\mathbf{G_2}$, $\mathbf{M}$, $\mathbf{N_1}$, $\mathbf{N_2}$, and $\mathbf{H}$, are matrices of parameters. $\mathbf{F}$, $\mathbf{G_1}$, $\mathbf{G_2}$, and $\mathbf{H}$ are $(m \times m)$ matrices, $\mathbf{N_1}$ and $\mathbf{N_2}$ are $(m \times n)$ matrices. This diagnostic expectation is taken over the diagnostic density of $\mathbf{F}\mathbf{y}_{t+1} + \mathbf{G_1}\mathbf{y}_t + \mathbf{M}\mathbf{x}_{t+1} + \mathbf{N_1}\mathbf{x}_t$. For generality, in equation (12), we specify current variables both inside the diagnostic expectations operator in linear combination with future variables (e.g. $\mathbf{N_1}\mathbf{x}_t$) as well as outside the

[10] The diagnostic distribution depends on two separate information sets, G_t and G_{-t}, drawing information available at dates t and $t-1$. So, one could denote it by $\mathbb{E}_{t,t-1}^{\theta}$. However, to avoid confusion, we prefer to stick to the notation used in BGS and the surrounding literature. Similarly, when denoting the RE operator $\mathbb{E}_t$, the subindex indicates the date at which the expectation is taken (in which case it coincides with the information set's date.)

[11] See Lemma 2 in the appendix.

expectations operator (e.g. $\mathbf{N_2 x}_t$).[12]

3.1.4 Solution Procedure

The remaining steps are the following. First, postulate a form for the solution. Second, determine how to handle the diagnostic expectation $\mathbb{E}_t^\theta[\mathbf{F}\mathbf{y}_{t+1}+\mathbf{G_1}\mathbf{y}_t+\mathbf{M}\mathbf{x}_{t+1}+\mathbf{N_1}\mathbf{x}_t]$, which is a linear combination of endogenous and exogenous variables, some of which are future, and some of which are current (known at time t). Third, obtain a rational expectations representation of the model. Fourth, solve for the model expressed in terms of rational expectations using standard tools (as the method of undetermined coefficients, for instance).

Form of the Solution. We look for a solution of the form

$$\mathbf{y}_t = \mathbf{P}\mathbf{y}_{t-1} + \mathbf{Q}\mathbf{x}_t + \mathbf{R}\mathbf{v}_t \tag{13}$$

We make this guess based on the behavioral properties afforded by DE. In the context of RE models, the correct conjecture is of the form $\mathbf{y}_t = \mathbf{P}\mathbf{y}_{t-1} + \mathbf{Q}\mathbf{x}_t$. As shown by BGS, DE generate overreaction in the context of exogenous processes. We allow for this possibility in the context of the endogenous dynamics of $\mathbf{y}_t$ using the extra term $\mathbf{R}\mathbf{v}_t$.

Diagnostic Expectation of Linear Combinations of Endogenous and Exogenous Variables. Under (13), $\mathbf{y}_{t+1}$ follows a multivariate normal distribution. Since the vector of exogenous drivers $\mathbf{x}_{t+1}$ also follows a multivariate normal distribution, we know that the linear combination $\mathbf{F}\mathbf{y}_{t+1}+\mathbf{G_1}\mathbf{y}_t+\mathbf{M}\mathbf{x}_{t+1}+\mathbf{N_1}\mathbf{x}_t$ is *also* distributed following a multivariate normal density. This Gaussian property is the key to the solution to the model. Using (11), it allows us to express the diagnostic expectation $\mathbb{E}_t^\theta[\mathbf{F}\mathbf{y}_{t+1}+\mathbf{G_1}\mathbf{y}_t+\mathbf{M}\mathbf{x}_{t+1}+\mathbf{N_1}\mathbf{x}_t]$ in terms of the RE operator $\mathbb{E}_t$. Indeed, the expression for the DE present in model (12) can be expressed as:

$$\begin{aligned}\mathbb{E}_t^\theta[\mathbf{F}\mathbf{y}_{t+1}+\mathbf{G_1}\mathbf{y}_t+\mathbf{M}\mathbf{x}_{t+1}+\mathbf{N_1}\mathbf{x}_t] = {}& \mathbb{E}_t[\mathbf{F}\mathbf{y}_{t+1}+\mathbf{G_1}\mathbf{y}_t+\mathbf{M}\mathbf{x}_{t+1}+\mathbf{N_1}\mathbf{x}_t] \\ & +\theta\big(\mathbb{E}_t[\mathbf{F}\mathbf{y}_{t+1}+\mathbf{G_1}\mathbf{y}_t+\mathbf{M}\mathbf{x}_{t+1}+\mathbf{N_1}\mathbf{x}_t] \\ & -\mathbb{E}_{t-1}[\mathbf{F}\mathbf{y}_{t+1}+\mathbf{G_1}\mathbf{y}_t+\mathbf{M}\mathbf{x}_{t+1}+\mathbf{N_1}\mathbf{x}_t]\big)\end{aligned} \tag{14}$$

[12] After loglinearization we will encounter expressions of this form. Throughout the paper we present a few examples to make this point concrete.

Notice that even the current vectors $\mathbf{x}_t$ and $\mathbf{y}_t$ undergo a diagnostic transformation in equation (14), since expectations are taken over a linear combination involving future variables. This does not however mean that if an agent were to be asked about their expectation of $\mathbf{x}_t$ in the absence of uncertainty, they would respond something different than $\mathbf{x}_t$. We return to this discussion in Section 3.3.

We are now in a position to obtain the representation of the model in terms of rational expectations.

Proposition 1 (Multivariate Rational Expectations Representation) *Under the multivariate NNA, model (12) admits the following RE representation:*

$$\begin{aligned} \mathbf{F}\mathbb{E}_t[\mathbf{y}_{t+1}] + \mathbf{G}\mathbf{y}_t + \mathbf{H}\mathbf{y}_{t-1} + \mathbf{M}\mathbb{E}_t[\mathbf{x}_{t+1}] + \mathbf{N}\mathbf{x}_t & \\ +\mathbf{F}\theta\big(\mathbb{E}_t[\mathbf{y}_{t+1}] - \mathbb{E}_{t-1}[\mathbf{y}_{t+1}]\big) & \\ +\mathbf{M}\theta\big(\mathbb{E}_t[\mathbf{x}_{t+1}] - \mathbb{E}_{t-1}[\mathbf{x}_{t+1}]\big) & \\ +\mathbf{G}_1\theta\big(\mathbf{y}_t - \mathbb{E}_{t-1}[\mathbf{y}_t]\big) & \\ +\mathbf{N}_1\theta\big(\mathbf{x}_t - \mathbb{E}_{t-1}[\mathbf{x}_t]\big) = 0 & \end{aligned} \tag{15}$$

where $\mathbf{G} = \mathbf{G}_1 + \mathbf{G}_2$ *and* $\mathbf{N} = \mathbf{N}_1 + \mathbf{N}_2$*. Moreover, this representation is unique.*

The proof of this result is based on equality (14), together with the additivity property of the RE expectations operator.

Solution. Armed with this representation, we verify that equation (13) indeed constitutes a solution. Appendix C presents the detailed steps to solve for the relevant matrices following Uhlig (1995).

3.2 Stability

It turns out that the model under DE is subject to the same stability conditions as the model under RE. More precisely, consider the same model above, but under rational expectations ($\theta = 0$):

$$\mathbf{F}\mathbb{E}_t[\mathbf{y}_{t+1}] + \mathbf{G}\mathbf{y}_t + \mathbf{H}\mathbf{y}_{t-1} + \mathbf{M}\mathbb{E}_t[\mathbf{x}_{t+1}] + \mathbf{N}\mathbf{x}_t = 0 \tag{16}$$

where the matrices $\mathbf{F}$, $\mathbf{G}$, $\mathbf{H}$, $\mathbf{M}$ and $\mathbf{N}$ are defined above. The following result holds.

Proposition 2 (Stability) *Assume a bounded solution exists for the DE model given by equations* (7) *and* (12)*. The stability conditions for this DE model are identical to the stability conditions for the RE model given by* (7) *and* (16)*.*

While the stability conditions are exactly same as under the RE model, we note that the existence of a bounded solution under DE requires an additional assumption. We formalize this requirement in the following proposition.

Proposition 3 (Existence of a Bounded Solution) *Assume a bounded solution exists for the RE model given by equations* (7) *and* (12) *with* $\theta = 0$. *Then a bounded solution for the DE model exists if* $(1+\theta)\mathbf{FP} + \mathbf{G} + \theta\mathbf{G}_1$ *is full-rank.*

Example 1 below will illustrate how DE can affect the existence of a bounded solution, even when RE models have a bounded and stable solution.

3.3 Technical Remark: Predetermined Variables and the Diagnostic Expectations Operator

A novelty in DSGE models, when compared to earlier DE models in the literature, is the presence of *predetermined* variables. This fact raises the question of how to compute the DE of such variables. Thus far, this issue was not explicit since in models of class (12), agents do not compute the DE of predetermined variables in isolation, but in combination with future variables. Nevertheless, from a technical perspective, some readers may wonder about the right way to think about this question. For completeness, here we provide a brief discussion.

The main aspect to recognize is that, because in these models predetermined variables are in linear combination of normally distributed *future* variables, the linear combination is also normally distributed. By implication, uncertainty is present, and diagnosticity is active in the mind of the agent. This carries implications for predetermined variables. This can be seen from (14), since even the vectors $\mathbf{x}_t$ and $\mathbf{y}_t$ undergo a diagnostic transformation when expectations are taken over the linear combination.

This observation offers the following approach to define the DE of a predetermined variable, say x_t.[13] Suppose that x_t follows an univariate AR(1) process, $x_t = \rho_x x_{t-1} + \varepsilon_t$, where $\varepsilon_t \sim i.i.d.\ N(0, \sigma_\varepsilon^2)$. Introduce an arbitrarily small amount of uncertainty by adding white noise δ_{t+1}, with variance σ_δ^2. We are interested in beliefs about the linear combination $x_t + \delta_{t+1}$. Since this object is also normally distributed, we obtain, using the BGS formula:

$$\mathbb{E}_t^\theta[x_t + \delta_{t+1}] = (1+\theta)\mathbb{E}_t[x_t + \delta_{t+1}] - \theta\mathbb{E}_{t-1}[x_t + \delta_{t+1}] \tag{17}$$

Taking the limit when σ_δ approaches 0 from above, we obtain the following expression

[13] We thank an anonymous referee for suggesting this approach.

for the DE of the predetermined variable:

$$\mathbb{E}_t^\theta[x_t] \equiv \lim_{\sigma_\delta \to 0^+} \mathbb{E}_t^\theta[x_t + \delta_{t+1}] = (1+\theta)x_t - \theta\mathbb{E}_{t-1}[x_t] \tag{18}$$

These steps show that, in these models, inference on predetermined variables is also affected by the representativeness heuristic.[14] In the appendix, we provide a formal derivation of (18) recurring to the Dirac delta distribution.

What is the behavioral basis for equation (18)? For the purpose of providing intuition, consider a recursive equation for a price, say $p_{t+1} = a + bp_t + \nu_{t+1}$, with $b > 0$ and a white noise ν_{t+1}. A higher current price p_t signals a higher future price p_{t+1}. Memory is overly influenced by the statistical positive association of current and future prices. Indeed, the presence of the future shock ν_{t+1} introduces uncertainty and activates the representativeness heuristic. Seeing a higher p_t causes the agent to over-sample past histories in which *next period's* price was high. In sum, p_t cues the agent to extrapolate into the future period's price, p_{t+1}. In a way, one might interpret this behavior as one where beliefs overshoot already for date-t variables, as (18) shows.[15]

Note, however, that the validity of this limit argument rests on the presence of uncertainty surrounding predetermined variables in these models. This does not mean, of course, that if an agent were to be asked, hypothetically, about their expectation of x_t in the absence of uncertainty ($\sigma_\delta = 0$), they would respond something different than x_t. Consistent with BGS and Gennaioli and Shleifer (2010), a complete absence of uncertainty would *deactivate* diagnosticity (equivalently, dropping the NNA) and eliminate any inference bias:[16]

$$\mathbb{E}_t^\theta\left[x_t | \sigma_\delta = 0\right] = x_t \tag{19}$$

As we explain in Example 1 below, expression (18) carries interesting behavioral insights into the amplification mechanism operating in DSGE models. Moreover, the limit argument offers an approach to operationalize the definition of the DE of predetermined variables, in a way preserves consistency within the DSGE. Also, it delivers an additivity result for the DSGE model, which is however relegated to the appendix.

We conclude this technical remark by emphasizing that adhering to this limit ar-

[14]Note that this bias appears even for i.i.d. shocks in the presence of surrounding uncertainty: take a white noise process $x_t = \epsilon_t$. We have that $\lim_{\sigma_\delta \to 0^+} \mathbb{E}_t^\theta[\epsilon_t + \delta_{t+1}] = (1+\theta)\epsilon_t$, despite $\mathbb{E}_t^\theta[\epsilon_{t+1}] = 0$.

[15]To see how the date-t variable undergoes a transformation, notice that $\mathbb{E}_t^\theta[p_{t+1}] = a + \mathbb{E}_t^\theta[bp_t + \nu_{t+1}] = a + b\left(p_t + \theta\left(p_t - \mathbb{E}_{t-1}[p_t]\right)\right)$.

[16]In technical terms, beliefs feature a discontinuity at the point of no uncertainty: no bias at exactly zero uncertainty, together with a discrete jump for strictly positive uncertainty about the future.

gument is *not* crucial for the purposes of solving these models. One can directly apply (14) to the linear combination of future and contemporaneous variables present in model (12) by acknowledging that the combination is normally distributed. See Bianchi, Ilut, and Saijo (2022), who also take this approach.[17] Both approaches are equivalent.

3.4 Examples: Endogenous Extrapolation

These two examples illustrate how DE generate endogenous extrapolation in dynamic models. (Example 1 also discusses unbounded solutions; Example 2 also discusses the loglinearization of equations with non-stationary variables.)

3.4.1 Example 1: Univariate Endogenous State Variable Model

The main purpose of this example is to illustrate the following point. When DE are taken over exogenous variables, there is no extrapolation if shocks are i.i.d. There is in fact an equivalence between RE and DE. To see this, consider the AR(1) process x_t, assume $\rho_x = 0$ and compute $\mathbb{E}_t^\theta[x_{t+1}]$. A simple calculation using formula (11) shows that RE and DE are equivalent in this case. Instead, in the context of DE over endogenous variables, state variables can activate extrapolation *even* when shocks are i.i.d. We label this property 'endogenous extrapolation'.[18] Modeling diagnostic expectations on endogenous variables provides a novel, internal propagation mechanism for DSGE models.

Consider the following model:

$$y_t = a\mathbb{E}_t^\theta[y_{t+1}] + cy_{t-1} + \varepsilon_t \tag{20}$$

where $|a + c| < 1$ and ε_t is white noise.

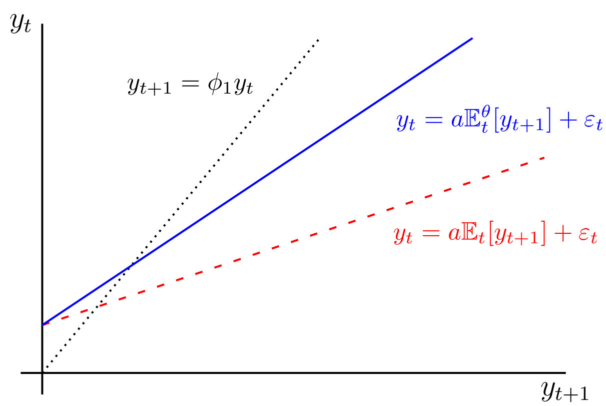
The solution of the RE model ($\theta = 0$) can be derived analytically using the minimum state variable solution method:

$$y_t = \phi_1 y_{t-1} + \frac{1}{1 - a\phi_1}\varepsilon_t \tag{21}$$

[17] In this vein, notice that all of the results of subsections 3.1 and 3.2 did not use the limit argument.

[18] We briefly discussed this endogenous extrapolation property in the context of the Muth model with inventory dynamics in Section 2.

Figure 1: Amplification in the Univariate Model



where $\phi_1 \equiv \frac{1-\sqrt{1-4ac}}{2a}$.[19] Under DE, the minimum state variable solution is given by

$$y_t = \phi_1 y_{t-1} + \frac{1}{1-(1+\theta)a\phi_1}\varepsilon_t \tag{22}$$

Notice from equation (22) that computing the DE over the endogenous variable y_{t+1} delivers extrapolation and amplification, even though the exogenous process is i.i.d. To see this, notice that since $1 - a\phi_1 > 0$, for small enough θ (more on this below), a positive shock ε_t generates an overreaction of y_t.

To get intuition, consider Figure 1. On the (y_{t+1}, y_t) plane, we plot the form of the solution, $y_{t+1} = \phi_1 y_t$ (dotted line) and the forward-looking reaction functions $y_t = a\mathbb{E}_t^\theta[y_{t+1}] + \varepsilon_t$ (full line) and $y_t = a\mathbb{E}_t[y_{t+1}] + \varepsilon_t$ (dashed line). We assume the economy is in steady state before the shock, and thus $y_{t-1} = 0$. Under RE, the reaction function collapses to $y_t = ay_{t+1} + \varepsilon_t$. Under DE, the reaction function is, instead, $y_t = a(1+\theta)y_{t+1} + \varepsilon_t$. The intersection of the dotted line with either reaction function (RE or DE) gives the solution. Because of extrapolation, the reaction function is steeper under DE, signifying the higher expectation of y_{t+1} in the mind of the agent. This extrapolation is the source of amplification at date t.[20]

With this example, we can also illustrate the result obtained in Proposition 3: When $\theta \to \frac{1}{a\phi_1} - 1$ or $\theta \to \infty$, then the DE solution explodes even though there exists

[19]Specifically, using the method of undetermined coefficients, we get the following requirement: $\phi_1 = a\phi_1^2 + c$. Imposing that $\phi_1 \to 0$ as $c \to 0$, we arrive at the solution. $|a + c| < 1$ ensures that the model is stable in the sense of Proposition 2 and that the RE solution is bounded.

[20]In other words: Given the solution $y_{t+1} = \phi_1 y_t + 1/(1-(1+\theta)a\phi_1)\varepsilon_{t+1}$, the uncertainty about ε_{t+1} activates diagnosticity in the mind of the agent when computing $\mathbb{E}_t^\theta[y_{t+1}]$, which becomes $\phi_1 (y_t + \theta(y_t - \mathbb{E}_{t-1}[y_t]))$. Therefore, there is a transformation of the DE of y_t, which generates amplification.

a unique bounded RE solution. The lesson of this example is therefore that in practice the researcher may need to be mindful of bifurcation points. In particular, bifurcation values could affect search over the parameter space in the context of structural estimation. In our application to NK models, we compute the conditions such that the DE solution explodes, and verify that the associated limit values for θ are very large. Therefore, this does not materially affect our results.

3.4.2 Example 2: Nominal Euler Equation

Consider the following Euler equation of a nominal economy:

$$\frac{u'(C_t)}{P_t} = \beta(1+i_t)\mathbb{E}_t^\theta\left[\frac{u'(C_{t+1})}{P_{t+1}}\right] \tag{23}$$

where C_t is consumption, P_t is the price level, i_t is the nominal rate, $u(\cdot) = \log(\cdot)$ is period utility, and β is the discount factor.[21]

In order to loglinearize (23) one needs to take the path dependence implied by DE into consideration. Because of the reference distribution, previous beliefs held at date $t-1$ constitute a state variable. One way to appreciate this fact is, for instance, by looking at the BGS formula (11) and notice that the DE involves past held beliefs $\mathbb{E}_{t-1}$. Therefore, different from the RE case, one cannot multiply by P_t on both sides of the equation and introduce P_t inside the DE operator. Instead, loglinearizing (23) directly, we obtain:

$$\hat{c}_t = \mathbb{E}_t^\theta\left[\hat{c}_{t+1}\right] - (\hat{i}_t - (\mathbb{E}_t^\theta[\hat{p}_{t+1}] - \hat{p}_t)) \tag{24}$$

where $\{\hat{c}_t, \hat{i}_t, \hat{p}_t\}$ denote loglinear deviations of consumption and the interest rate from their respective steady states, and of the price level from an initial price level, respectively. Using the BGS formula (11) and algebraic manipulation delivers the loglinear diagnostic Euler equation[22]

$$\hat{c}_t = \mathbb{E}_t^\theta[\hat{c}_{t+1}] - (\hat{i}_t - \mathbb{E}_t^\theta[\hat{\pi}_{t+1}]) + \theta(\hat{\pi}_t - \mathbb{E}_{t-1}[\hat{\pi}_t]) \tag{25}$$

According to the last term, current surprise inflation induces an *expansionary* channel by reducing the subjective real rate computed by diagnostic agents. The reason is as follows. Due to path dependence, computation of a real rate of interest involves

[21]Section 4 derives this equation from first principles. Following BGS, the diagnostic distribution for non-linear processes is also defined as a distorted likelihood that over-weights states representative of recent news. We provide a formal definition in Appendix D.

[22]See Appendix D for the derivation.

the price level at $t-1$.[23] Since the agent is extrapolating from yesterday $(t-1)$ into tomorrow $(t+1)$, today's inflation innovation $\hat{\pi}_t - \mathbb{E}_{t-1}[\hat{\pi}_t]$ is also extrapolated into tomorrow when making a forecast for price level p_{t+1}: Current surprise inflation causes the diagnostic agent to expect future inflation, to a degree θ, thereby reducing the subjective real interest rate. The intuition for why this is the case is same as discussed above. Uncertainty about future variables, p_{t+1} in this instance, entails a transformation of the current variables when they enter in linear combination with future variables. Furthermore, this effect is present even in the case of *i.i.d. shocks*, once again underscoring the novelty of computing DE on endogenous variables. We will exploit this channel in Section 4 by emphasizing its implications for fiscal policy.

3.5 A Practical Guide to the Implementation of Diagnostic Expectations in DSGE Models

We conclude this section with the following summary. A researcher interested in using diagnostic expectations within a DSGE model can take the following steps.

1. Obtain the exact equilibrium conditions of the model. (Section 4 provides an example in the context of a 3-equation NK model, and Section 5 in the context of a medium-scale DSGE model.)
2. Loglinearize the model, being careful not to introduce contemporaneous variables in-and-out of the DE operator. (See the appendix for examples.)
3. Obtain the RE representation of the model (Proposition 1).
4. Solve the RE model based on a software package that can handle expectations conditional on previous period's information set $(\mathbb{E}_{t-1})$.
5. Check that the parameter values considered does not cover bifurcation values (Proposition 3 and Example 1).

4 Analysis Using a New Keynesian Model

In this section, we derive a three-equation New Keynesian model augmented by diagnostic expectations. Our goal is to revisit a number of prominent themes in this

[23]To see this, multiply on both sides of (23) by P_{t-1} and use P_t inside the DE to obtain:

$$u'(C_t)\frac{P_{t-1}}{P_t} = \beta(1+i_t)\mathbb{E}_t^{\theta}\left[u'(C_{t+1})\frac{P_{t-1}}{P_t}\frac{P_t}{P_{t+1}}\right] \tag{26}$$

which can then be loglinearized to arrive at (25).

context.

4.1 Diagnostic New Keynesian Model

We set up the model from first principles. There are three sets of agents in the economy: households, firms and the government.

4.1.1 Households

Households maximize the following lifetime utility

$$\log C_t - \frac{\omega}{1+\nu} L_t^{1+\nu} + \mathbb{E}_t^\theta \left[\Sigma_{s=t+1}^{\infty} \beta^{s-t} \left(\log(C_s) - \frac{\omega}{1+\nu} L_s^{1+\nu} \right) \right] \tag{27}$$

where L_t is labor supply, $\nu > 0$ is the inverse of the Frisch elasticity of labor supply, β is the discount factor β, satisfying $0 < \beta < 1$, $\omega > 0$ is a parameter that pins down the steady-state level of hours.[24] Maximization is subject to a budget constraint:

$$P_t C_t + \frac{B_{t+1}}{(1+i_t)} = B_t + W_t L_t + D_t + T_t \tag{28}$$

where P_t is the price level, B_{t+1} is the demand of nominal bonds that pay off $1 + i_t$ interest rate in the following period, W_t is the wage, D_t and T_t are dividends from firm-ownership and lump-sum government transfers, respectively.

Notice that we write dynamic maximization problems, as this one, by explicitly separating time t choice variables from the expectation of future choice variables. This separation is crucial for solving the model with diagnostic expectations, and is a consequence of the DE path dependence discussed in Section 3.[25]

4.1.2 Firms

Monopolistically competitive firms, indexed by $j \in [0, 1]$, produce a differentiated good, $Y_t(j)$. We assume a Dixit-Stiglitz aggregator that aggregates intermediate goods into a final good, Y_t. Intermediate goods' demand is given by $Y_t(j) = \left(\frac{P_t(j)}{P_t}\right)^{-\epsilon_p} Y_t$, where $\epsilon_p > 1$ is the elasticity of substitution, $P_t(j)$ is the price of intermediate good j, and P_t is the price of final good Y_t. Each intermediate good is produced using the technology

[24] Following BGS, the diagnostic distribution for non-linear processes is also defined as a distorted likelihood that overweights states representative of recent news. We provide a formal definition in Appendix D.

[25] The reader may wonder whether DE introduces time inconsistency in agents' choices. It turns out that this is not the case in the loglinear approximation when the reference distribution is based on $t-1$. By the law of iterated expectations (which then holds for the diagnostic expectation), time $t+1$ policy functions are in fact consistent with agents' expectations (about their time $t+1$ policy functions).

$Y_t(j) = A_t L_t(j)$, where $\hat{a}_t \equiv \log(A_t)$ is an aggregate TFP process that follows an AR(1) process with persistence coefficient ρ_a:

$$\hat{a}_t = \rho_a \hat{a}_{t-1} + \varepsilon_{a,t} \tag{29}$$

and $\varepsilon_{a,t} \sim iid\, N(0, \sigma_a^2)$. The firm pays a quadratic adjustment cost $\frac{\psi_p}{2}\left(\frac{P_t(j)}{P_{t-1}(j)} - 1\right)^2 P_t Y_t$, in units of the final good (Rotemberg 1982) to adjust prices. Firms' per period profits are given by $D_t \equiv P_t(j)Y_t(j) - W_t L_t(j) - \frac{\psi_p}{2}\left(\frac{P_t(j)}{P_{t-1}(j)} - 1\right)^2 P_t Y_t$. The firm's profit maximization problem is

$$\max_{P_t(j)} \left\{ P_t(j)Y_t(j) - W_t L_t(j) - \frac{\psi_p}{2}\left(\frac{P_t(j)}{P_{t-1}(j)} - 1\right)^2 P_t Y_t + \mathbb{E}_t^{\theta}\left[\sum_{s=1}^{\infty} \beta^s Q_{t,t+s} D_{t+s}\right] \right\} \tag{30}$$

where $Q_{t,t+s}$ is the household's nominal stochastic discount factor.

4.1.3 Government

The government sets nominal interest rate with the following rule $1 + i_t = (1 + i_{ss})\Pi_t^{\phi_\pi}\left(\frac{Y_t}{Y_t^*}\right)^{\phi_x}$, where $Y_t^* = A_t$ is the natural rate allocation, $i_{ss} = \frac{1}{\beta} - 1$ is the steady state nominal interest rate, $\phi_\pi \geq 0$, $\phi_x \geq 0$, and steady state gross inflation $\Pi = 1$. Total output produced is equal to household consumption expenditure and adjustment costs spent when adjusting prices. We first consider a model where is no government spending, and nominal bonds are in zero net supply.

4.1.4 Equilibrium

Appendix D presents the equilibrium conditions.[26] In particular, it shows that the household intertemporal first order condition is equation (23). This appendix also goes over the log-linear approximation in detail. The resulting equilibrium is given by the following four equations:

$$\hat{c}_t = \mathbb{E}_t^{\theta}\left[\hat{c}_{t+1}\right] - \left(\hat{i}_t - \left(\mathbb{E}_t^{\theta}[\hat{p}_{t+1}] - \hat{p}_t\right)\right) \tag{31}$$

$$\hat{\pi}_t = \beta \mathbb{E}_t^{\theta}[\hat{\pi}_{t+1}] + \tilde{\kappa}(\hat{c}_t - \hat{a}_t) + \tilde{\kappa}\nu(\hat{y}_t - \hat{a}_t) \tag{32}$$

$$\hat{i}_t = \phi_\pi \hat{\pi}_t + \phi_x(\hat{y}_t - \hat{a}_t) \tag{33}$$

$$\hat{c}_t = \hat{y}_t \tag{34}$$

[26]The DE operator is the expectation over a continuous density, hence one gets these first-order conditions by taking derivatives, as usual.

where $\tilde{\kappa} \equiv \frac{\epsilon_p - 1}{\psi_p}$, $\hat{y}_t$, $\hat{c}_t$, $\hat{p}_t$, $\hat{i}_t$ are the log deviation of output, consumption, the price level, and the nominal interest rate respectively, and $\hat{\pi}_t$ is the log deviation of inflation from the zero-inflation steady state. The shock process is given by:

$$\hat{a}_t = \rho_a \hat{a}_{t-1} + \varepsilon_{a,t} \tag{35}$$

where $\varepsilon_{a,t} \sim i.i.d.\ N(0, \sigma_a^2)$.

As explain in the context of Example 2 in Section 3, equation (31) can be written as (25), showing that DE change the expression for the approximated Euler equation by introducing a current inflation surprise term. We obtain a similar Phillips curve (32) to the RE case using Rotemberg (1982) pricing. The key to this result is that, different than with Calvo pricing, Rotemberg pricing with DE allows one to obtain a recursion that only involves one expectation forward. This turns out to be key for tractability. The appendix presents the detailed derivation.[27]

Define $\kappa \equiv (1+\nu)\tilde{\kappa}$. We make the following assumption in order to guarantee the existence of a bounded solution (Proposition 3).[28]

Assumption 2 (Boundedness) $\theta < \phi_\pi + \kappa^{-1}(1 + \phi_x)$

We provide an explicit solution for the model in Appendix D.

4.2 Diagnostic Expectations and the Possibility of Extra Amplification

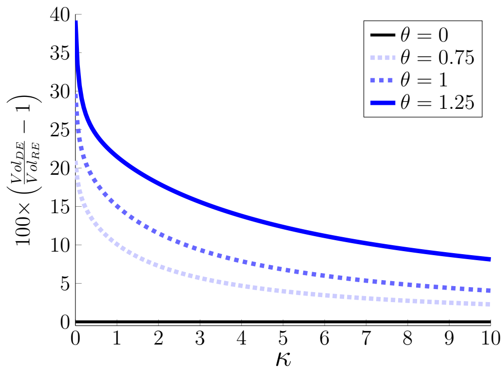
A classic challenge in macroeconomic modeling is finding ways to generate realistic business cycles with shocks of moderate size. The literature has relied on multiple types of frictions (e.g. nominal, as in Christiano, Eichenbaum, and Evans 2005, or financial, as in Bernanke and Gertler 1989; Kiyotaki and Moore 1997), interactions in the form of strong complementarities (Benhabib and Farmer 1994), or multiple shocks (Smets and Wouters 2007) to fit the data.

We demonstrate that diagnosticity provides a viable behavioral alternative to understand the large size of observed fluctuations within the NK model. Because diagnosticity leads agents to extrapolate the impact of exogenous shocks, expectations are more volatile. Intuitively, one would expect the DE-NK model to predict a higher volatility of output than under RE. Indeed, the following proposition establishes that

[27] In a loglinearized RE model with perfect inflation indexation, one can obtain identical Phillips curves using either the Calvo or the Rotemberg price setting assumption.

[28] We also assume that $\kappa(\phi_\pi - 1) + (1-\beta)\phi_x > 0$ to ensure a stable solution in the sense of Proposition 2.

Figure 2: Excess Volatility under DE, Baseline NK Model



Notes: The figure presents percentage points of volatility under DE relative to RE as a function of the slope of Phillips curve, κ, and for various values of diagnosticity parameter, θ. The model is given by equations (31)-(35). See Footnote 30 for the default calibration.

diagnosticity can generate extra endogenous volatility in the NK model. We analytically prove this result when prices are completely rigid ($\psi_p \to \infty$).[29]

Proposition 4 (Extra Volatility: NK Model) *Consider the model given by* (31)-(35)*. Assume parameters are such that Assumption 2 is satisfied. When $\psi_p \to \infty$ (rigid prices), output is more volatile under DE than under RE : $Var(\hat{y}_t^{DE}) > Var(\hat{y}_t^{RE})$. When $\psi_p \to 0$ (flexible prices), output volatility under DE is equal to that under RE.*

In the flexible price limit, we obtain the efficient benchmark where output volatility is equal to the stationary TFP process volatility. In the perfectly rigid price case, diagnosticity interacts with price rigidity to amplify fluctuations in output whenever $\theta > 0$. In the intermediate range, we numerically illustrate how excess volatility under DE varies with the degree of price rigidity, parameterized by κ. Our default calibration of the NK model is based on the textbook by Galí (2015).[30] θ is set to 1 following Bordalo, Gennaioli, Shleifer, and Terry (2021). We obtain a standard deviation of output of 2.96%, relative to 1.82% under RE. Thus, output volatility increases by 63% due to DE.

DE interact with the nominal frictions embedded in the NK model in order to generate extra output volatility. Figure 2 plots the excess volatility under DE relative to RE as a function of κ plotted on the x-axis, for different values of θ. κ is inversely related to ψ_p, the adjustment cost parameter. Given the default calibration, DE gen-

[29] Away from this limit, we can use the solution of the model presented in the appendix and obtain a condition for extra volatility, but this condition is messy and does not lend itself to any clear interpretation.

[30] We set $\beta = 0.99$, $\epsilon_p = 9$, $\phi_\pi = 1.50$, and $\phi_x = 0.5$. We set $\nu = 2$, and ψ_p such that $\kappa = 0.050$. The TFP process is calibrated with persistence 0.90 and standard deviation of 2%.

erates highest excess volatility relative to RE when prices are perfectly rigid. Excess volatility monotonically declines as prices become flexible.[31] In the flexible price limit, the excess volatility converges to zero. Also, the excess volatility is increasing in the diagnosticity parameter as long as Assumption 2 is satisfied.

In order to further demonstrate the interaction of nominal rigidities with diagnosticity, we consider the case of a frictionless real business cycle (RBC) model. The model is standard and is provided in Appendix E. There, we analytically show that output is less volatile under DE than under RE when there is full depreciation of capital ($\delta = 1$) and TFP shock process has zero persistence ($\rho_a = 0$). For a calibration away from these analytical assumptions, we find that the standard deviation of output is lower under DE than under RE.

It is useful to draw a parallel to the news shocks literature originating in the seminal work by Beaudry and Portier (2004) and Beaudry and Portier (2006) in order to understand these results. The addition of DE to the NK model can be seen as a way of generating *errors in expectations* that resemble news about the future. For instance, in the case of a positive TFP shock, agents extrapolate this shock, expecting a further positive TFP shock in the next period. Therefore, the TFP shock generates a contemporaneous raise in TFP, and an excessive increase in expectations about TFP in the next period. Shocks to expectations can be seen as shifts in aggregate demand. Whether aggregate demand can move away from aggregate supply depends on the degree of nominal rigidities. When prices are sticky, output is demand determined: The positive income effect raises consumption and in general equilibrium this effect dominates. Output ultimately increases. This explains the extra volatility afforded by the DE-NK model. Similarly, in the presence of capital, shocks to expectations also face difficulties in generating comovement in a baseline, frictionless, RBC model with flexible prices (Beaudry and Portier 2006; Jaimovich and Rebelo 2009). Indeed, in the case of a positive news shock, the implied income effect produces a fall of labor supply and hence output (Barro and King 1984). However, as shown in Blanchard, L'Huillier, and Lorenzoni (2013), nominal rigidities are also a solution to this counterfactual prediction of the RBC model. Indeed, we will return to this property of DE in the case of an estimated medium-scale DSGE models.[32]

We note that the recent important paper by Bordalo, Gennaioli, Shleifer, and Terry (2021) presents another case in which DE interact with frictions to generate extra

[31] It is possible to get lower volatility of output under DE relative to RE for different parameter configurations. For example, when persistence of the TFP process $\rho_a = 0.1$, $\kappa = 1$, and $\theta = 1$, we obtain dampening of output volatility under DE relative to under RE.

[32] To be clear, we use this parallel to news shocks only for the purposes of providing intuition. In fact, compared to news shocks, DE generate novel and different effects. Section 5 will expand more on this point.

volatility. The paper looks at an RBC model with financial frictions on the firm side. Firms are heterogeneous. The paper shows that the interaction of firms' expectations with financial frictions successfully generate amplification of investment and output dynamics, and fits a number of facts relating to credit cycles.

4.3 Keynesian Supply Shocks

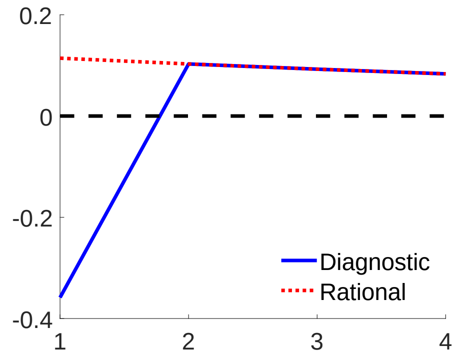
Motivated by economic crisis caused by the COVID-19 pandemic, a rapidly growing literature focuses on constructing models that have the ability to generate a demand shortfall that is fundamentally caused by a disruption on the supply side of the economy, that is, a 'Keynesian' supply shock. Thus far, some of the candidate explanations for this phenomenon include multiple consumption goods (Guerrieri, Lorenzoni, Straub, and Werning 2022), endogenous firm-entry (Bilbiie and Melitz 2020), heterogenous risk-tolerance (Caballero and Simsek 2021), and endogenous TFP growth (Fornaro and Wolf 2021). As the following proposition shows, DE present a behavioral mechanism capable of producing Keynesian supply shocks.

Proposition 5 (Keynesian Supply Shocks) *Consider the model given by* (31)-(35). *Assume that* $\psi_p \to \infty$ *and that the diagnosticity parameter is high enough, that is,* $\theta > 2(1-\rho_a)(1+\phi_x)/(\phi_x\rho_a)$. *Then, the output gap* $\hat{x}_t$ *positively co-moves with the unanticipated component of TFP:* $\frac{\partial \hat{x}_t}{\partial \varepsilon_{a,t}} > 0$.

Similar to Proposition 4, the proposition imposes completely rigid prices for tractability. The result extends to the case of moderately rigid prices, as Figure 3 shows. We use the default calibration discussed in Footnote 30. The figure plots the evolution of the output gap. Following a negative TFP shock, the economy enters a recession: the output gap falls under DE. In the RE case, the output gap moves in the opposite direction.

The key to this striking result is extrapolation: following the shock, agents extrapolate and become excessively pessimistic about future output. This leads to a large drop in consumption, which due to nominal rigidities, leads to contemporaneous fall in output. Due to diagnosticity, expectations become sufficiently pessimistic to induce a fall in output larger than the initial drop in TFP, generating a Keynesian recession. This is in contrast to the result under RE where the fall in TFP, being only transitory, does not lead to a fall in aggregate demand. Hence, there is a boom: lower TFP for the same level of aggregate demand increases the demand for labor; this generates a boom in the labor market, together with a *rise* in the output gap.

Figure 3: Output Gap Response to a Negative TFP Shock, Baseline NK Model



Notes: The figure depicts the impulse response of the output gap to a unit negative shock to TFP. The productivity shock process is given by equation (35). The blue solid line denote impulses responses with diagnostic expectations, whereas the red dotted line denote responses with rational expectations. The dynamics of employment are exactly the same as the output gap.

A noteworthy result, following BGS, is that there is a systematic *reversal* in output gap to the RE forecast when the extrapolation of current news turns out to be incorrect under DE. When at time $t-1$ the news about productivity is bad, agents become pessimistic and the output gap becomes negative. Next period, the excess pessimism subsides and the output gap is corrected upwards. The forecast errors of output gap are thus predictable: diagnostic forecasts neglect the systematic reversals in output gap.[33]

4.4 Fiscal Policy Multiplier

Here we address the implications of DE for the size of the fiscal policy multiplier. There are two reasons to do this.

First, given the recent unprecedented fiscal response to the COVID-19 crisis in the U.S. and other countries, understanding the effects of fiscal policy is central. Also, substantial empirical evidence indicates that marginal propensities to consume are large (see Fagereng, Holm, and Natvik 2021, among others), or similarly, that fiscal multipliers are large in the cross section (Nakamura and Steinsson 2014).[34] We show that DE constitute a useful addition to the NK framework, because it generates novel, rich implications for the fiscal multiplier.

Second, this exercise is a natural path for understanding the endogenous extrapolation generated by the diagnostic Fisher equation embedded in equation (23). This

[33]The solution for output gap under DE, with rigid prices, is $\hat{x}_t = \frac{\phi_x \rho_a (1+\theta) - (1+\phi_x - \rho_a)}{(1+\phi_x)(1+\phi_x - \rho_a)} \hat{a}_t - \frac{\phi_x \theta \rho_a^2}{(1+\phi_x)(1+\phi_x-\rho_a)} \hat{a}_{t-1}$. When $\theta > 0$, the second term denotes the reversal from revision in expectations.

[34]See Steinsson (2021) for a similar discussion.

endogenous extrapolation channel highlights the implication of the extra term arising due to belief path-dependence, as explained in Example 2, Section 3.

We add government spending shocks to the NK model. There is a balanced budget government spending financed by lump-sum taxes. Now, the total output in the economy is used for consumption and government expenditure. That is, we replace equation (34) with:

$$\hat{y}_t = \hat{c}_t + \hat{g}_t \tag{36}$$

where $\hat{g}_t$ is the percentage change of government spending from its steady state as fraction of steady state output. $\hat{g}_t$ follows an exogenous process:

$$\hat{g}_t = \rho_g \hat{g}_{t-1} + \varepsilon_{g,t} \tag{37}$$

where $\varepsilon_{g,t} \sim iid\ N(0, \sigma_g^2)$. The equilibrium is given by equations (31), (32), (33), and (36), for a given process (37).

For convenience, we write the diagnostic Fisher equation here:

$$\hat{r}_t = \hat{i}_t - \mathbb{E}_t[\pi_{t+1}] - \theta(\mathbb{E}_t[\pi_{t+1}] - \mathbb{E}_{t-1}[\pi_{t+1}]) - \theta(\pi_t - \mathbb{E}_{t-1}[\pi_t]) \tag{38}$$

Extrapolation implied by DE reduces the real interest rate, and hence leads to higher multipliers.

To make this point in a transparent way, we start by looking at i.i.d. government spending shocks. The reason is that with i.i.d. shocks, there is no exogenous extrapolation.[35] We obtain the following proposition.

Proposition 6 (Fiscal Policy Multiplier) *Consider the model given by equations* (31), (32), (33), (36), *and* (37). *Assume that* $\phi_x = 0$ *and that the persistence of the shock* $\rho_g = 0$. *Then:*

1. *Under rational expectations, the fiscal policy multiplier is always strictly less than 1. Under diagnostic expectations, the fiscal policy multiplier is greater than 1 if* $\theta > \phi_\pi$, *and less than 1 if* $\theta < \phi_\pi$.
2. *The fiscal policy multiplier is greater under diagnostic expectations than under rational expectations.*
3. *The fiscal policy multiplier is increasing in* θ, *and tends to infinity as* $\theta \longrightarrow \phi_\pi + \kappa^{-1}$.

[35] To see this, notice that equation (11) implies, for an AR(1) process, $\mathbb{E}_t^\theta[x_{t+1}] = \rho_x \check{x}_t + \theta \rho_x \check{\varepsilon}_t$.

Hence, when the degree of diagnosticity is above the reaction parameter of the monetary authority, the multiplier is greater than one. The intuition for this result is as follows. The diagnostic real rate moves, in response to current inflation, due to the endogenous extrapolation (governed by θ), and by the response of the central bank. In the RE benchmark, the multiplier is always smaller than 1 because the central bank moves the nominal rate to dampen the effect of fiscal policy. The condition $\theta > \phi_\pi$ ensures that endogenous extrapolation offsets this dampening.

The degree of diagnosticity parametrizes the multiplier, increasing it above the RE multiplier, and spanning the full range of values to infinity. We assume that $\phi_x = 0$ in order to get a clean and easy to interpret condition such that the multiplier is greater than 1 in the DE model.[36]

This analytical case highlights that the higher multiplier under DE is only working through the term $\theta(\pi_t - E_{t-1}[\pi_t])$ in the diagnostic Fisher equation. Extrapolation is endogenous, generating the expansionary effect discussed in Example 2 above. Given that the government spending shock is i.i.d., there is no *exogenous* extrapolation of the shock due to diagnosticity.

We move away from the default calibration to illustrate the results in the case $\phi_x = 0$. In order to illustrate a case where the multiplier is greater than 1, we consider a dovish interest rate rule ($\phi_\pi = 1.1$) and a moderately higher diagnosticity parameter of $\theta = 1.5$. Using a persistence of the government shock equal to 0.5 generates a DE multiplier of 1.04, and an RE multiplier of 0.91. Raising the diagnosticity parameter slightly generates much larger multipliers. Furthermore, using a steeper Phillips curve (say, $\kappa = 0.20$) strengthens the endogenous inflation extrapolation channel: the DE multiplier is now 1.13, for an RE multiplier of 0.73.

We conclude this section by noting that DE do not always lead to higher multipliers. When government shocks are persistent, the expectation of future spending crowds out current consumption, reducing output. With DE, expectations of future spending are exaggerated, and can considerably reduce multipliers when persistence is high. To illustrate this, we go back to our default calibration. In addition, we set the persistence of the shock to 0.9. In this case, the RE multiplier is 0.17, for a DE multiplier of -0.32. In this simulation, the exogenous extrapolation channel is so strong that it dominates the endogenous extrapolation channel, leading to a negative multiplier.

[36] The general condition is $\theta \geq \phi_\pi + \frac{\phi_x}{(1-\psi)\kappa}$.

4.5 Overreaction and Delayed Overreaction

Whether beliefs as measured by surveys feature under- or overreaction is the subject of an important debate in recent literature. Indeed, Coibion and Gorodnichenko (2012) provide evidence of underreaction of consensus forecasts, whereas Bordalo, Gennaioli, Ma, and Shleifer (2020) provide evidence of overreaction at the level of the individual forecaster. Kohlhas and Walther (2021) find that there is overreaction, in some cases, even at the aggregate level. In a complementary way, Angeletos, Huo, and Sastry (2020) stress that at the aggregate level one can observe both under- and overreation. According to them, what matters is the horizon: there is underreaction in the short run, whereas overreaction dominates in the medium run.

The importance of the horizon at which one observes the dynamics of forecasts has also been stressed in an application to stock returns by Bordalo, Gennaioli, La Porta, and Shleifer (2019). The authors stress that the key is to look at the medium-term forecast errors to find evidence of overreaction to news. The explanation is the following. A gradual arrival of news can happen some time after an anticipated event, and a buildup of the overreaction can move forecasts away from the underreaction generated by imperfect information on impact.

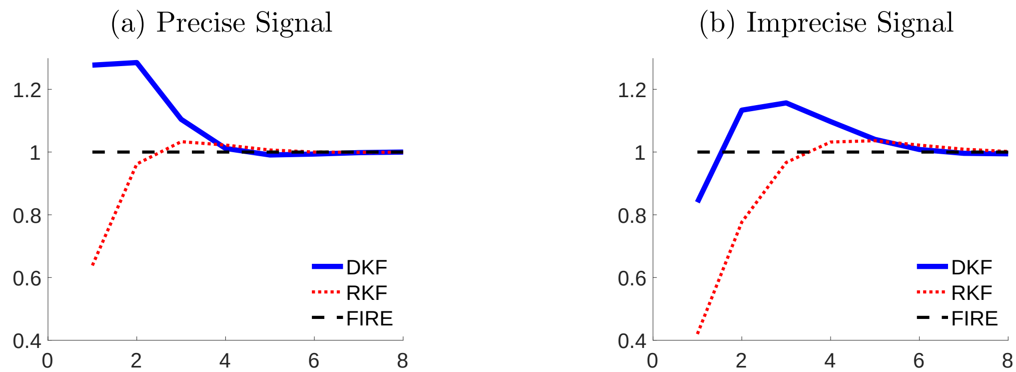
Based on the premise by Bordalo et al. (2019), our broad aim in this section is to contribute to this debate by presenting an extension of the NK model in which long-term beliefs are guided by the diagnostic Kalman filter. The key innovation of our setup compared to previous exercises in the literature is that agents form beliefs about a hidden component that features both sizeable *persistence*, and is also *permanent* (in the sense that the underlying process has a unit root.) To model the long-term nature of this hidden object, we calibrate this persistence to a high value, which conceptually connects our exercise to the long-run risks approach (Bansal and Yaron 2004). However, ours is a general equilibrium representative-agent macroeconomic model where consumers are concerned with the long run path of income.

Assume prices are completely rigid. Consumption is pinned down solely by beliefs about long-run income.[37] The information structure is as follows. TFP, in logs, now has a permanent component ζ_t and a temporary component ξ_t. Agents do not observe these components separately. Instead, they observe realized TFP and a noisy signal about the permanent component $s_t = \zeta_t + \varepsilon_{s,t}$ where $\varepsilon_{s,t} \sim i.i.d.\ N(0, \sigma_s^2)$, and form beliefs using the diagnostic Kalman filter introduced by Bordalo et al. (2020).[38]

[37] We take the limit $\phi_x \to 0$ and $\psi_p \to \infty$. For brevity we do not write down the equations more explicitly, but this conclusion can be reached by iterating forward the Euler equation.

[38] Even though the model does not explicitly have dispersed information as in Coibion and Gorodnichenko (2012), we follow Lorenzoni (2009) by using a simple representative agent model with aggregate noisy signals. The filter

Figure 4: Impulse Responses: Beliefs About the Long-Run



Notes: The panels depict the impulse responses of beliefs about long-run productivity to a one unit positive shock to the permanent component of TFP. The left-hand side panel presents the case of a precise signal ($\sigma_s = 0.01$ and $\theta = 1.0$); the right-hand side panel with the case of an imprecise signal ($\sigma_s = 0.03$ and $\theta = 1.0$).

The following analytical result offers a simple comparison of beliefs about the long-run under a) the diagnostic Kalman filter (DKF), b) the rational Kalman filter (RKF), and c) the full information RE benchmark (FIRE).

Proposition 7 (Overreaction) *Assume that $\psi_p \to \infty$, $\phi_x = 0$, and the persistence of the permanent component $\rho_\zeta = 0$. Consider a positive shock to ζ_t. Then,*

1. *Beliefs about the long-run are greater under the DKF than under the RKF.*
2. *If θ is high enough, beliefs about the long-run under the DKF are greater than under FIRE.*

When $\rho_\zeta > 0$, delayed overreaction is possible. We offer a collection of numerical results using the following calibration. In order to capture the idea that the agent is forming beliefs about a very long-run object, we calibrate the persistence of the permanent component to a high value, $\rho_\zeta = 0.98$. We normalize the standard deviation of TFP to 1. We consider two values of the standard deviation of the signal: a relatively precise signal (of standard deviation 0.01), or a relatively imprecise signal (of standard deviation 0.03). Figure 4 presents the dynamics for beliefs about long-run productivity in response to a one standard deviation permanent shock. The left-hand side (LHS) panel presents the case of a precise signal, and the right-hand side (RHS) panel presents the case of an imprecise signal.

Under FIRE, long-run beliefs jump to 1 on impact and stay there. This is because the standard deviation of TFP innovations has been normalized to 1, and beliefs im-

needs to be adapted to the particular information structure here, but the ideas are the same. For details of the model specification, see Blanchard et al. (2013), or Appendix F for a full specification in the context of the medium-scale DSGE.

mediately adjust to the long-run value of TFP after the shock. In the case of a precise signal (LHS panel), beliefs under the RKF underreact on impact, starting off at 0.70. As learning happens over time, these beliefs rise, gradually reverting back to 1 in the long run.[39] Instead, beliefs under the DKF strongly overreact on impact. This because the signal is so precise that diagnosticity overwhelms imperfect information.

Turning to the case of an imprecise signal (RHS panel), beliefs under the RKF underreact significantly, starting off at 0.41. Given that now imperfect information is more severe, DKF beliefs also slightly underreact on impact, starting off at 0.84. However, because agents receive a new signal every period, there is gradual learning. Therefore, as they gather more information, DKF implies a sizeable overreaction over periods 2 to 6, with a peak at 1.16. Notice, the RKF also slightly overreacts around period 5. This is due to a mechanical effect induced by the persistence of beliefs. However, diagnosticity induces overreaction above and beyond this mechanical effect, with a subsequent systematic reversal.

We conclude by noting that we reported results only varying the precision of the signal. By varying the degree of diagnosticity one modifies the degree of overreaction independently. For instance, increasing θ to 1.5 (which is within the range of estimates reported by Bordalo, Gennaioli, Ma, and Shleifer 2020) can generate a slight overreaction in the short run and a stronger overreaction in the medium run, leading to a hump-shaped pattern of beliefs.

5 Empirical Evaluation

Given the theoretical findings of the previous sections, we undertake an empirical evaluation of diagnostic expectations using standard structural methods. The primary goal is to ask the following question. Consider a baseline, medium scale, rational expectations DSGE model. Replace rational expectations with diagnostic expectations. (The diagnostic model nests the rational expectations model via the diagnosticity parameter.) Is there evidence that diagnostic expectations improve the ability of the DSGE model to fit business cycle data?

With this formulation of the broad question that will guide our empirical investigation, four interrelated subquestions emerge: What is the estimated value of the diagnosticity parameter? Does the credible interval span the RE limit? Ultimately, is there statistical evidence that diagnosticity provides an advantage when fitting busi-

[39] There is a light overreaction in period 3 even in the case of the RKF. This is simply a mechanical implication of the persistence of beliefs inherited from the highly persistent permanent component.

ness cycle data? If so, what changes in the interpretation of the data?

Given the recent interest in the literature on survey data (see Bordalo, Gennaioli, Ma, and Shleifer 2020, Coibion and Gorodnichenko 2015b, among others), we include five survey forecast series from the Survey of Professional Forecasters (SPF) among the set of observable variables. Recently, Milani and Rajbhandari (2020) and Miyamoto and Nguyen (2020) have shown that DSGE models featuring news shocks can fit SPF data.[40] Hence, we also include news and noise shocks in the estimation, based on the specification by Blanchard, L'Huillier, and Lorenzoni (2013) (henceforth BLL). In model evaluation, Chahrour and Jurado (2018) find BLL to be the best candidate for fitting the data with shocks to rational expectations, such as news or noise.[41]

Thus, we highlight that our empirical exercise is disciplined by the addition of a host of ingredients in the baseline model: a rich set of frictions, shocks and competing channels. This includes the frictions introduced in the seminal work by Christiano, Eichenbaum, and Evans (2005). We include the exogenous driving processes introduced by Smets and Wouters (2007). We include news shocks as an alternative channel to explain expectations. We include information frictions in the form of noise shocks (included in the news and noise specification by BLL). By adding all these bells and whistles (nominal, real, and information frictions), driving processes, and the alternative expectation channel, we aim to perform a tough test of the usefulness of the behavioral friction embodied by diagnostic expectations. Indeed, we want to assess whether it provides a significant empirical advantage, *even when* all the other commonly used ingredients have been included.

We note that the inclusion of information frictions leads to a diagnostic Kalman filter, as introduced by Bordalo, Gennaioli, Ma, and Shleifer (2020). Coibion and Gorodnichenko (2015b) have also emphasized the importance of expectation underreaction in the aggregate, which our information structure is able to account for.

5.1 Medium-Scale DSGE Model

Since the model is standard (Christiano, Eichenbaum, and Evans 2005), we describe here its main ingredients and relegate the details to the appendix. The preferences of the representative household feature habit formation and differentiated labor supply.

[40]Related work by the Federal Reserve Bank of New York has included data on inflation and Federal Funds rate expectations in DSGE estimation (Del Negro et al. 2013).

[41]'News and noise' models of belief-driven fluctuations are models where rational agents receive noisy advance information about fundamental shocks hitting the economy. In a series of representation results for this type of theories of the business cycle, Chahrour and Jurado (2018) show that, generally, "news and noise" models admit both a pure news representation of beliefs and fundamentals, and a noise representation.

The capital stock is owned and rented by the representative household, and the capital accumulation features a quadratic adjustment cost in investment, as introduced by Christiano et al. (2005). The model features variable capacity utilization.

The final good is a Dixit-Stiglitz aggregate of a continuum of intermediate goods, produced by monopolistic competitive firms, with Rotemberg (1982) costs of price adjustment. Similarly, specialized labor services are supplied under monopolistic competition, with Rotemberg (1982) costs of nominal wage adjustment. The monetary authority sets the nominal interest rate following an inertial Taylor rule.

The model features eight persistent structural shocks: shocks to temporary and permanent productivity, a noise shock to the signal about permanent productivity, a shock to the marginal efficiency of investment, shocks to price and wage markups, shocks to monetary and fiscal policy. We introduce i.i.d. measurement errors for SPF forecasts.[42]

Following Smets and Wouters (2007) and Justiniano, Primiceri, and Tambalotti (2010), the model is estimated based on U.S. time series for GDP, consumption, investment, employment, the federal funds rate, inflation, and wages, for the period 1954:III-2004:IV. This sample period facilitates comparison of our results across models in the robustness section, and avoids complications arising from the zero lower bound. We also include SPF data on consumption growth, investment growth, output growth, short-term inflation and short-term interest rate forecasts. The data appendix presents more details. We set up a Kalman filter to get smoothed estimates of the permanent component of productivity and the associated agents' beliefs. Table 7 in the appendix presents the parameter prior distributions. We generate 5,000,000 draws using a Metropolis-Hastings algorithm and discard the first 20% as initial burn-in.

5.2 Results

The parameter estimates are reported in Table 1. We report mean posterior estimates, along with the 90% credible interval. We present estimates for the diagnostic model and the rational model, side-by-side. The bottom row reports the marginal likelihood for both models.

Let us first look at the estimate for the diagnosticity parameter θ. Our prior distribution is normal with mean 1 and standard deviation 0.3. The estimated posterior mean for θ is 0.7325. This estimate is close to the one obtained in the previous empirical

[42] The two productivity shocks are not separately observed by the agent. Instead, a public signal on permanent productivity is available. These three variables imply a distributed lag model for TFP and beliefs that admit a pure news representation, as shown by Chahrour and Jurado (2018).

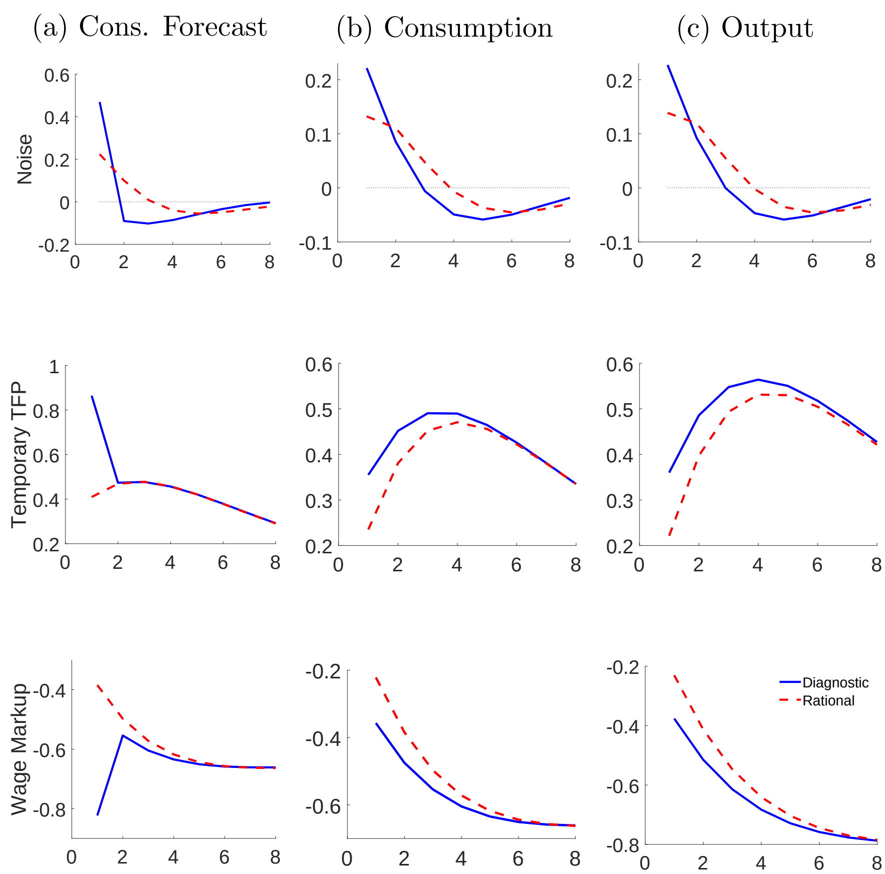
Table 1: Posterior Distribution

Parameter	Description	Diagnostic		Rational	
		Mean	[05, 95]	Mean	[05, 95]
θ	diagnosticity	0.7325	[0.5917, 0.8746]		
α	cap. share	0.1340	[0.1226, 0.1453]	0.1390	[0.1278, 0.1505]
h	habits	0.7211	[0.6922, 0.7502]	0.5803	[0.5424, 0.6178]
$\frac{\chi''(1)}{\chi'(1)}$	cap. util. costs	5.0666	[3.4432, 6.6709]	5.5929	[3.9095, 7.2242]
ψ_p	Rotemberg prices	125.58	[98.710, 152.17]	181.84	[126.66, 188.88]
ϕ_w	Rotemberg wages	582.13	[256.01, 897.76]	9710.9	[4510.5, 14712.]
ν	inv. Frisch elas.	3.8520	[2.4474, 5.2254]	1.2832	[0.5012, 1.9475]
$S''(1)$	inv. adj. costs	6.9588	[5.8400, 8.0723]	7.0701	[6.0111, 8.1332]
ρ_R	m.p. rule	0.5818	[0.5429, 0.6209]	0.5563	[0.4380, 0.6806]
ϕ_π	m.p. rule	1.5363	[1.4173, 1.6537]	1.0682	[1.0001, 1.2046]
ϕ_x	m.p. rule	0.0061	[0.0001, 0.0109]	0.0013	[0.0001, 0.0030]
Technology Shocks					
ρ	persist.	0.8573	[0.8368, 0.8780]	0.9535	[0.9352, 0.9716]
σ_a	tech. shock s.d.	1.3772	[1.2603, 1.4947]	1.5258	[1.3896, 1.6601]
σ_s	noise shock s.d.	0.5400	[0.3196, 0.7531]	1.0594	[0.3781, 1.7574]
Investment-Specific Shocks					
ρ_μ	persist.	0.3027	[0.2474, 0.3575]	0.3310	[0.2631, 0.4003]
σ_μ	s.d.	18.905	[15.017, 22.716]	20.212	[16.369, 23.989]
Markup Shocks					
ρ_p	persist.	0.8749	[0.8303, 0.9209]	0.8205	[0.7663, 0.8769]
ϕ_p	ma. comp.	0.5858	[0.4728, 0.7022]	0.5563	[0.4380, 0.6806]
σ_p	s.d.	0.1591	[0.1306, 0.1877]	0.1988	[0.1700, 0.2271]
ρ_w	persist.	0.9969	[0.9939, 0.9999]	0.6543	[0.5146, 0.7978]
ϕ_w	ma. comp.	0.5765	[0.3942, 0.7630]	0.5142	[0.2882, 0.7444]
σ_w	s.d.	0.4383	[0.3434, 0.5300]	0.4490	[0.3836, 0.5142]
Policy Shocks					
ρ_{mp}	persist.	0.0295	[0.0100, 0.0514]	0.0197	[0.0009, 0.0383]
σ_{mp}	s.d.	0.3801	[0.3440, 0.4158]	0.3283	[0.3000, 0.3556]
ρ_g	persist.	0.9341	[0.9058, 0.9626]	0.8974	[0.8682, 0.9275]
σ_g	s.d.	0.3699	[0.3384, 0.4017]	0.3706	[0.3384, 0.4022]
Measurement Errors					
σ_y^{ME}	s.d.	0.4975	[0.4467, 0.5471]	0.5034	[0.4529, 0.5533]
σ_c^{ME}	s.d.	0.4089	[0.3594, 0.4575]	0.4255	[0.3739, 0.4764]
σ_i^{ME}	s.d.	1.4320	[1.2539, 1.6039]	1.4514	[1.2692, 1.6284]
σ_r^{ME}	s.d.	0.2692	[0.2417, 0.2966]	0.2285	[0.2018, 0.2551]
σ_π^{ME}	s.d.	0.1639	[0.1432, 0.1845]	0.1482	[0.1267, 0.1693]
log Marg. Likelihood		-1812.71		-1847.38	

exercises reported by Bordalo, Gennaioli, Ma, and Shleifer (2020), and to the value used by Bordalo, Gennaioli, Shleifer, and Terry (2021). Figure 7 in the appendix shows that the posterior distribution of θ is unimodal. The 90% credible interval covers values from 0.5917 to 0.8746, away from the RE limit of zero.

In order to understand the implications of DE, we analyze the impulse response functions (IRFs) to the main driving processes.[43] Figure 5 plots IRFs for the one-step-ahead consumption forecast, and for selected quantities (consumption and output growth, specifically). Figure 6 plots these IRFs for selected prices (price inflation, nominal and real interest rates).

Figure 5: Impulse Responses: Quantities



Notes: The panels depict the impulse responses of one-step-ahead consumption forecast, consumption, and output to a one standard deviation shock to noise signal, temporary TFP, and wage markup. The blue solid lines denote impulses responses with diagnostic expectations, whereas the red dashed lines denote responses with rational expectations. See Table 1 for parameters.

Consider the IRFs to the noise shock. In this model, the noise shock raises expectations of future income. This pushes the consumer to increase demand, and firms

[43]We plot the IRFs of the shocks that explain the highest share of consumption volatility on impact.

to invest in anticipation of higher profits. Therefore, it can be intuitively thought of as an aggregate demand shock. The noise shock increases consumption and output in the DE model (full line, blue). The same happens in the rational counterfactual, obtained by shutting down diagnosticity ($\theta = 0$, dashed line, red). Focusing on the consumption forecast, we see that the behavioral consumer's beliefs overreact to the shock, exhibiting a more volatile conditional response to the noise shock, and a rapid reversal relative to RE. The combination of excess volatility and reversal results in a boom-bust in actual consumption and economic activity more broadly, above and beyond the mechanical reversal generated with the Kalman filter in the RE counterfactual. The boom-bust in consumption forecast is sharper than response of actual consumption due to the presence of habits in consumption.

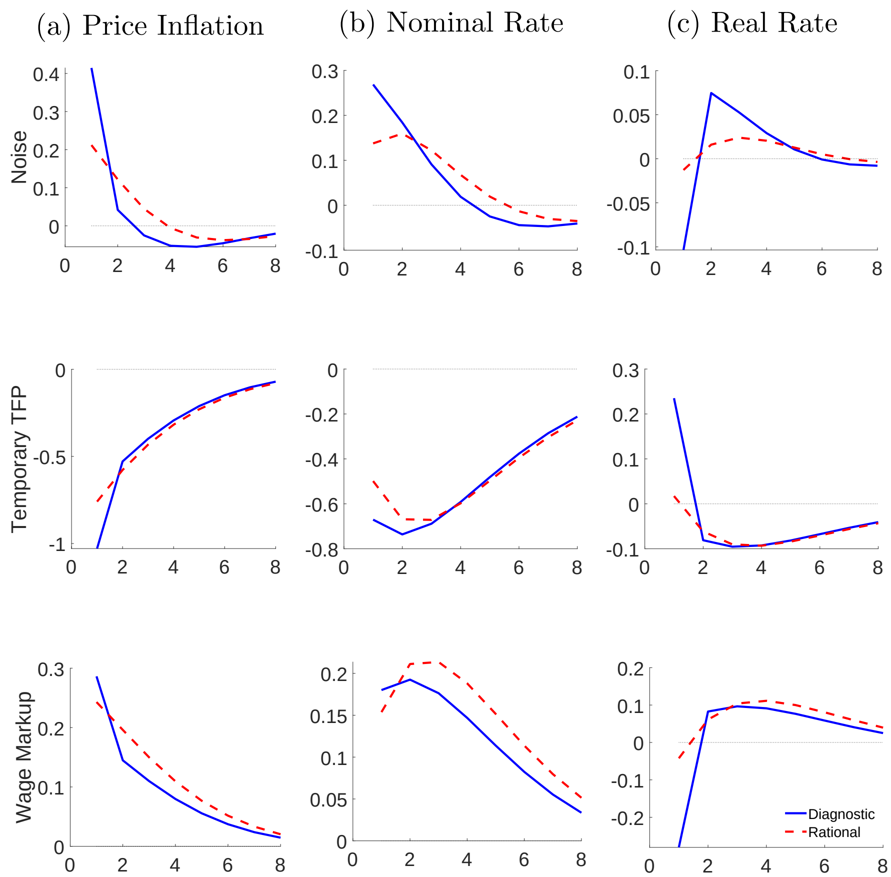
Turning to the IRFs of prices, we note that price inflation tends to feature an overreaction on impact. This is due to the forward-looking behavior of prices: By the logic of the NK Phillips curve, current inflation depends on the expectation of future inflation. For instance, following a positive TFP shock, inflation falls by a larger amount under DE than under RE. This overreaction is particularly acute in the case of the noise and the temporary TFP shock. Another noticeable difference between the DE model and the RE counterfactual comes from the responses of the real rate. As highlighted earlier in the paper, the diagnostic real rate expression features an extra term, whereby current surprise inflation is extrapolated. As in Bianchi et al. (2022), this term generates an additional propagation mechanism in the DE model.

Overall, comparing the IRFs of the DE model and the RE counterfactual suggests that the DE model affords extra volatility of endogenous variables in general equilibrium. This theme was developed analytically at the beginning of Section 4, and linked to nominal rigidities. To quantify this point we compute the excess unconditional volatility afforded by DE for consumption growth, output growth, price and wage inflation, and the real rate. Among these, the largest increase is the one of the real rate, with a 37% increase in volatility. The increase of consumption volatility is particularly strong as well, at 23%. This is consistent with the view that consumers overreact in their expectations of future consumption via the Euler equation. The volatility of other variables increase as well (see Table 8 in the appendix).

We use the Bayes factor to empirically evaluate the fit of the diagnostic model against the rational model. The log marginal likelihood of the data given the estimated diagnostic model is -1812.71. This statistic is lower, -1847.38, in the case of the rational counterpart.[44] This difference in log marginal likelihoods represents evidence in favor

[44]Following the Kass and Raftery (1995) classification, $2\log(BF) = 2 \times 34.67 = 69.34$ statistic represents "very

Figure 6: Impulse Responses: Prices



Notes: The panels depict the impulse responses of inflation, the nominal interest rate, and the real rate to a one standard deviation shock to noise signal, temporary TFP, and wage markup. The blue solid lines denote impulses responses with diagnostic expectations, whereas the red dashed lines denote responses with rational expectations. See Table 1 for parameters.

of the diagnostic model.

We look at the forecast error variance decompositions to get intuition into how the DE model fits the data and outperforms the RE model. Table 2 presents the 1Q ahead variance decomposition across all structural shocks for quantities and prices. For each, we first present the case of the DE model. Then, for comparison, we present the variance decomposition for the estimated RE model (with parameter estimates presented in Table 1). A striking finding is that the DE model relies much less on noise shocks to explain consumption fluctuations. The contribution of noise shocks to consumption volatility is only 12% in the DE model (compared to 43% in the RE model), and to output is 7% in the DE model (compared to 25% in the RE model). Other shocks explain these variables, with temporary TFP shocks explaining 30% of

strong evidence" in favor of the diagnostic model. This statistic has been used for model comparisons in the DSGE literature. See, for example, Ascari, Bonomolo, and Lopes (2019).

consumption and 18% of output volatility in the DE model (versus 15% and 7% in the RE model), and with wage markup shocks explaining 30% of consumption and 19% output volatility (versus 1% and 0% in the RE model).[45]

Table 2: Variance Decomposition: Quantities and Prices

Variable	Noise	Perm. TFP	Temp. TFP	Invest.	Price Markup	Wage Markup	Monet.	Fiscal
Consumption								
DE	0.1158	0.0432	0.2976	0.0013	0.0313	0.3010	0.1814	0.0283
RE	0.4310	0.0039	0.1509	0.0006	0.0334	0.0121	0.3680	0.0001
Investment								
DE	0.0020	0.0018	0.0279	0.9347	0.0102	0.0187	0.0035	0.0012
RE	0.0156	0.0002	0.0104	0.9585	0.0050	0.0014	0.0089	0.0001
Output								
DE	0.0707	0.0262	0.1776	0.2842	0.0373	0.1942	0.1093	0.1005
RE	0.2493	0.0021	0.0716	0.2867	0.0278	0.0059	0.2017	0.1547
Price Inflation								
DE	0.0658	0.0000	0.4055	0.0880	0.3259	0.0314	0.0656	0.0179
RE	0.0175	0.0003	0.2859	0.0025	0.5902	0.1023	0.0007	0.0006
Wage Inflation								
DE	0.1285	0.0216	0.0115	0.1120	0.4138	0.2210	0.0814	0.0101
RE	0.0046	0.0003	0.0835	0.0004	0.2449	0.6662	0.0000	0.0000
Nominal Rate								
DE	0.0279	0.0000	0.1737	0.0378	0.1350	0.0125	0.6053	0.0077
RE	0.0026	0.0000	0.0413	0.0003	0.0840	0.0146	0.8571	0.0001
Real Rate								
DE	0.0319	0.0000	0.1647	0.0431	0.0360	0.0147	0.7006	0.0090
RE	0.0077	0.0001	0.0848	0.0019	0.0006	0.0391	0.8656	0.0002

What explains this pattern? The DE model exploits the rich propagation afforded by extrapolation in the forward-looking behavior of consumers, firms, workers, and financial markets. Consumers extrapolate, generating extra volatility and reversals of consumption. Firms extrapolate, generating extra volatility and reversals of investment. Price and wage setters extrapolate, generating extra volatility and reversals

[45] Chahrour and Jurado (2018) propose a variance decomposition of news and noise in terms of fundamental shocks and noise. For ease of comparison, we retain the original BLL decomposition.

in price and wage inflation. Financial markets extrapolate, generating extrapolation of current surprise inflation when pricing nominal bonds, together with implied dynamics of the real rate. Instead, in the case of the RE model, errors in expectations arise only about future income, following the permanent income channel emphasized by BLL, who build on Lorenzoni (2009). Overall, the DE model affords a more flexible structure of errors in expectations, and is able to explain deviations from belief rationality on several dimensions. This finding can be interpreted as evidence that DE outcompete noise shocks as a preferred channel to explain fluctuations. Consistent with this view, we point to the fact that the estimated noise in the signal, σ_s, is 0.5400 in the DE model (versus 1.0594 in the RE model). The DE model fits the data with a more precise signal and therefore a lower degree of information imperfections. It explains fluctuations with the aid of other shocks, which employ diagnosticity in order to propagate internally in general equilibrium.

We also note the sharp drop in the importance of exogenous markups in explaining price and wage inflation variance. Indeed, price markup shocks explain 33% of price inflation volatility in the DE model (versus 59% in the RE model). Similarly, wage markup shocks explain 22% of wage inflation volatility in the DE model (versus 67% in the RE model). The DE model exploits the forward-looking behavior of wage setters to explain goods prices and wages, relying more on other shocks and internal propagation mechanisms. For instance, temporary TFP shocks explain 41% of price inflation volatility in the DE model (versus 29% in the RE model). Also, *price* markup shocks explain 41% of *wage* inflation volatility in the DE model, instead of 24% of wage inflation volatility in the RE model. Consistent with this finding on the variance decomposition, the wage Phillips curve is steeper in the DE model, as evidenced by a much lower Rotemberg costs parameter.

Table 3 presents variance decomposition for the forecast data, distinguishing between the contribution of structural shocks and of observation errors. We find that the structural shocks account for a higher share of the empirical volatility of forecast data in the DE model. For instance, 44% of the one-step-ahead consumption forecast is explain by structural shocks in the DE model (versus 31% in the RE model), and 56% of the one-step-ahead inflation forecast volatility is explained by structural shocks in the DE model (versus 33% in the RE model).

To sum up, this analysis points to two primary reasons that explain the empirical success of the DE model. First, diagnosticity is a powerful and flexible amplification mechanism, generating errors in expectations along several dimensions. Thus, the key role of noise shocks is diminished in the DE model. This is important, since these

Table 3: Variance Decomposition: One-Step-Ahead Forecasts

Variable	Structural Shocks	Measurement Errors
Consumption		
DE	0.44	0.56
RE	0.31	0.69
Investment		
DE	0.33	0.67
RE	0.17	0.83
Output		
DE	0.44	0.56
RE	0.30	0.70
Price Inflation		
DE	0.56	0.44
RE	0.33	0.67
Nominal Rate		
DE	0.91	0.09
RE	0.76	0.24

shocks require a fairly dovish monetary authority to propagate. Notice that, compared to the RE model, the DE model can explain fluctuations in quantities by relying less on such a dovish monetary authority. To corroborate this claim, we emphasize that the reaction parameter ϕ_π to inflation is estimated at 1.54, compared to a very low value of 1.07 in the RE model. The DE estimate is more in line with other estimates in the literature, as for example 2.04 by Smets and Wouters (2007), and 2.09 by Justiniano, Primiceri, and Tambalotti (2010).[46] Second, the model is able to explain price and wage dynamics inflation *internally*, relying less on exogenous markup drivers. Consistent with this view, we highlight that the standard deviation of price markup shock σ_μ is estimated at 0.1591 in the DE model (versus 0.1998 in the RE model).[47] We interpret the fact that price and wage fluctuations are explained internally, rather than exogenously, as an encouraging finding. This is because DSGE models could be criticized on the grounds that markup shocks constitute a rather black-boxy ingredient without a realistic counterpart (Chari, Kehoe, and McGrattan 2009).

[46] The RE estimate is actually consistent with BLL, who discuss why the news and noise model requires a fairly accommodative rule in the case of log preferences.

[47] The estimated standard deviation of wage markup shocks is also lower, 0.44 in the DE model, versus 0.45 in the RE model.

5.3 Robustness

5.3.1 Prior on the Diagnosticity Parameter θ Centered at Zero

The model can in principle fit the data with a value of θ that is either close to zero, or negative. Thus, it is important to check that imposing a prior distribution centered at a positive value (such as 1) does not importantly affect our results. Using a symmetric prior distribution around 0 ($\theta \sim N(0.0.3)$), we re-estimate the DSGE model under DE. Table 9 in the appendix presents the results. Our posterior estimate of θ is not importantly affected, with a mean posterior of 0.6537. Again, the 90% credible interval is away from zero, covering values from 0.5193 to 0.7884, away from the RE limit of zero. The log marginal likelihood is also higher for the DE model than for the RE model (-1814.82 versus -1847.38).

5.3.2 Diagnostic Expectations in Alternative Off-The-Shelf Models

Another robustness check concerns the importance of details of our implementation for the conclusion that DE provide a superior fit of business cycle data. There are two separate angles of potential concern. First, does this conclusion crucially depend on using the BLL model? Second, does this conclusion crucially depend on including SPF data among the set of observable variables?

In order to demonstrate that the answer to both these questions is negative, here, we undertake the estimation of two influential off-the-shelf DSGE models: Smets and Wouters (2007) and Justiniano, Primiceri, and Tambalotti (2010). In the estimation of each of this models, we make sure to replicate the authors' procedure as close as possible: We use the same sample 1954:III–2004:IV. We use their data set, ensuring variable construction does not cause any differences. We also set the same prior distribution. Hence, our RE results replicate their findings. The introduction of DE in this case constitutes a particularly tight test that DE are useful to explain business cycle data.

Table 11 in the appendix presents the results for the Smets and Wouters (2007) model. We present estimates for this model under diagnosticity, and for the baseline the rational model, side-by-side.[48] Our posterior estimate of θ is smaller than under BLL, but still positive with a mean posterior of 0.4435. The 90% credible interval is away from the RE limit of zero, covering values from 0.1822 to 0.6928. The log marginal likelihood is also higher for the DE model than for the RE model (-897.91

[48]For the RE model, we obtain a similar slope of the price and wage Phillips curves despite the use of Rotemberg adjustment costs instead of Calvo in our specification of nominal rigidities.

versus -900.69).

Table 13 in the appendix presents the results for the Justiniano, Primiceri, and Tambalotti (2010) model. We present estimates for this model under diagnosticity, and for the baseline the rational model, side-by-side. Our posterior estimate of θ is smaller than under BLL but still positive with a mean posterior of 0.4336. The 90% credible interval is away from the RE limit of zero, covering values from 0.1894 to 0.6745.[49] The log marginal likelihood is also higher for the DE model than for the RE model (-1190.86 versus -1193.78).

6 Conclusion

In this paper, we argue that diagnostic expectations constitute a behavioral mechanism that can be fruitfully incorporated into New Keynesian macroeconomics. To this end, we first considered a set of challenges encountered by researchers working with this type of models, and revisited them analytically under diagnostic expectations. We concluded that the use of diagnostic expectations opens up avenues to make significant progress in the context of these challenges. We then asked if diagnostic expectations are validated empirically. Using a rich medium-scale DSGE model with news shocks and imperfect information, we conclude that the answer to this question is yes: The diagnostic model dominates the rational counterpart in terms of fit.

Our general solution method offers opportunities to explore and revisit a number of themes in macroeconomics and international macroeconomics in the context of diagnostic expectations. For example, a challenge in open economy models has been to account for the cyclicality of the current account in emerging countries, or to improve our understanding of exchange rate predictability. We leave these explorations to future work.

References

Afrouzi, H., S. Kwon, A. Landier, Y. Ma, and D. Thesmar (2020). Overreaction and working memory. *Mimeo*.

Angeletos, G.-M., Z. Huo, and K. A. Sastry (2020). Imperfect macroeconomic expectations: Evidence and theory. In *NBER Macroeconomics Annual 2020, Vol. 35*, NBER Chapters. National Bureau of Economic Research.

[49]In unreported results, we find that diagnosticity also generates extra volatility for output, consumption, investment, price inflation, wage inflation, and the real rate, for both the Smets and Wouters (2007) and the Justiniano et al. (2010) models.

Angeletos, G.-M. and C. Lian (2018). Forward guidance without common knowledge. *American Economic Review 108*(9), 2477–2512.

Ascari, G., P. Bonomolo, and H. F. Lopes (2019). Walk on the wild side: Temporarily unstable paths and multiplicative sunspots. *American Economic Review 109*(5), 1805–42.

Bansal, R. and A. Yaron (2004). Risks for the long run: A potential resolution of asset pricing puzzles. *Journal of Finance LIX*(4), 1481–1509.

Barro, R. J. and R. G. King (1984). Time-separable preferences and intertemporal-substitution models of business cycles. *The Quarterly Journal of Economics 99*(4), 817–839.

Beaudry, P. and F. Portier (2004). An exploration into Pigou's theory of cycles. *Journal of Monetary Economics 51*(6), 1183–1216.

Beaudry, P. and F. Portier (2006). Stock prices, news and economic fluctuations. *American Economic Review 96*(4), 1293–1307.

Benhabib, J. and R. E. Farmer (1994). Indeterminacy and increasing returns. *Journal of Economic Theory 63*(1), 19–41.

Bernanke, B. and M. Gertler (1989). Agency costs, net worth, and business fluctuations. *The American Economic Review 79*(1), 14–31.

Bianchi, F., C. Ilut, and H. Saijo (2022). Diagnostic business cycles. *Mimeo*.

Bianchi-Vimercati, R., M. Eichenbaum, and J. Guerreiro (2022). Fiscal policy at the zero lower bound without rational expectations. *Mimeo*.

Bilbiie, F. O. and M. J. Melitz (2020). Aggregate-demand amplification of supply disruptions: The entry-exit multiplier. Working Paper 28258, National Bureau of Economic Research.

Blanchard, O. J., J.-P. L'Huillier, and G. Lorenzoni (2013). News, noise, and fluctuations: An empirical exploration. *American Economic Review 103*(7), 3045–3070.

Bordalo, P., N. Gennaioli, R. La Porta, and A. Shleifer (2019). Diagnostic expectations and stock returns. *Journal of Finance 74*(6), 2839–2874.

Bordalo, P., N. Gennaioli, Y. Ma, and A. Shleifer (2020). Overreaction in macroeconomic expectations. *American Economic Review 110*(9), 2748–82.

Bordalo, P., N. Gennaioli, and A. Shleifer (2018). Diagnostic expectations and credit cycles. *The Journal of Finance 73*(1), 199–227.

Bordalo, P., N. Gennaioli, A. Shleifer, and S. J. Terry (2021). Real credit cycles. *NBER Working Paper No. 28416*.

Caballero, R. J. and A. Simsek (2021). A model of endogenous risk intolerance and LSAPs: Asset prices and aggregate demand in a "COVID-19" shock. *Review of Financial Studies 34*(11), 5522–80.

Chahrour, R. and K. Jurado (2018). News or noise? The missing link. *American Economic Review 108*(7), 1702–36.

Chari, V. V., P. J. Kehoe, and E. R. McGrattan (2009). New Keynesian Models: Not yet useful for policy analysis. *American Economic Journal: Macroeconomics 1*(1), 242–66.

Christiano, L. J., M. Eichenbaum, and C. L. Evans (2005). Nominal rigidities and the dynamic effects of a shock to monetary policy. *Journal of Political Economy 113*(1), 1–45.

Coibion, O. and Y. Gorodnichenko (2012). What can survey forecasts tell us about informational rigidities? *Journal of Political Economy 120*(1), 116–159.

Coibion, O. and Y. Gorodnichenko (2015a). Information rigidity and the expectations formation process: A simple framework and new facts. *American Economic Review 105*(8).

Coibion, O. and Y. Gorodnichenko (2015b). Is the Phillips curve alive and well after all? Inflation expectations and the missing disinflation. *American Economic Journal: Macroeconomics 7*(1), 197–232.

D'Arienzo, D. (2020). Maturity increasing overreaction and bond market puzzles. Mimeo, Bocconi University.

Del Negro, M., S. Eusepi, M. P. Giannoni, A. M. Sbordone, A. Tambalotti, M. Cocci, R. Hasegawa, and M. Linder (2013). The FRBNY DSGE model. *FRB of New York Staff Report* (647).

Fagereng, A., M. B. Holm, and G. J. Natvik (2021). MPC heterogeneity and household balance sheets. *American Economic Journal: Macroeconomics 13*(4).

Farhi, E. and I. Werning (2019). Monetary policy, bounded rationality, and incomplete markets. *American Economic Review 109*(11), 3887–3928.

Farhi, E. and I. Werning (2020). Taming a Minsky cycle. Working paper, MIT.

Fornaro, L. and M. Wolf (2021). The scars of supply shocks. Universitat Pompeu Fabra Working Paper 1748.

Gabaix, X. (2020). A behavioral New Keynesian model. *American Economic Review 110*(8), 2271–2327.

Galí, J. (2015). *Monetary policy, inflation and the business cycle: An introduction to the New Keynesian framework*. Princeton University Press.

Garcia-Schmidt, M. and M. Woodford (2019). Are low interest rates deflationary? A paradox of perfect-foresight analysis. *American Economic Review 109*(1), 86–120.

Gennaioli, N. and A. Shleifer (2010). What comes to mind. *The Quarterly Journal of Economics 125*(4), 1399–1433.

Gennaioli, N. and A. Shleifer (2018). *A crisis of beliefs: Investor psychology and financial fragility*. Princeton University Press.

Guerrieri, V., G. Lorenzoni, L. Straub, and I. Werning (2022). Macroeconomic implications of COVID-19: Can negative supply shocks cause demand shortages? *American Economic Review 112*(5), 1437–74.

Iovino, L. and D. Sergeyev (2020). Central bank balance sheet policies without rational expectations. *Mimeo*.

Jaimovich, N. and S. Rebelo (2009). Can news about the future drive the business cycle? *American Economic Review 99*(4), 1097–1118.

Justiniano, A., G. Primiceri, and A. Tambalotti (2010). Investment shocks and business cycles. *Journal of Monetary Economics 57*(2), 132–145.

Kahneman, D., P. Slovic, and A. Tversky (Eds.) (1982). *Judgment under uncertainty: Heuristics and biases*. Cambridge: Cambridge University Press.

Kahneman, D. and A. Tversky (1972). Subjective probability: A judgment of representativeness. *Cognitive Psychology 3*(3), 430–454.

Kass, R. E. and A. E. Raftery (1995). Bayes factors. *Journal of the American Statistical Association 90*(430), 773–795.

Kiyotaki, N. and J. Moore (1997). Credit cycles. *Journal of Political Economy 105*(2), 211–248.

Kohlhas, A. N. and A. Walther (2021). Asymmetric attention. *American Economic Review 111*(9).

Krishnamurthy, A. and W. Li (2020). Dissecting mechanisms of financial crises: intermediation and sentiment. *Mimeo*.

Lorenzoni, G. (2009). A theory of demand shocks. *American Economic Review 99*(5), 2050–84.

Ma, Y., T. Ropele, D. Sraer, and D. Thesmar (2020). A quantitative analysis of distortions in managerial forecasts. *NBER Working Paper No. 26830*.

Mankiw, N. G. and R. Reis (2002). Sticky information versus sticky prices: A proposal to replace the New Keynesian Phillips curve. *The Quarterly Journal of Economics 117*(4), 1295–1328.

Matsuyama, K. (2007). Aggregate implications of credit market imperfections. In *NBER Macroeconomics Annual 2007, Vol. 22*, NBER Chapters. National Bureau of Economic Research.

Maxted, P. (2022). A macro-finance model with sentiment. *Mimeo*.

Milani, F. and A. Rajbhandari (2020). Observed expectations, news shocks, and the business cycle. *Research in Economics 74*(2), 95–118.

Miyamoto, W. and T. L. Nguyen (2020). The expectational effects of news in business cycles: Evidence from forecast data. *Journal of Monetary Economics 116*, 184–200.

Muth, J. F. (1961). Rational expectations and the theory of price movements. *Econometrica 29*(3), 315–335.

Nakamura, E. and J. Steinsson (2014). Fiscal stimulus in a monetary union: Evidence from US regions. *American Economic Review 104*(3), 753–92.

Rotemberg, J. J. (1982). Monopolistic price adjustment and aggregate output. *Review of Economic Studies 49*(5), 517–531.

Smets, F. and R. Wouters (2007). Shocks and frictions in US business cycles: A bayesian DSGE approach. *American Economic Review 97*(3), 586–606.

Steinsson, J. (2021). Online panel: A new macroeconomics? *EfIP Research Brief*.

Uhlig, H. (1995). A toolkit for analyzing nonlinear dynamic stochastic models easily. *Institute for Empirical Macroeconomics* (Discussion Paper 101).

Woodford, M. (2002). Imperfect common knowledge and the effects of monetary policy. P. Aghion, R. Frydman, J. Stiglitz, and M. Woodford (Eds.). Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps, Princeton Univ. Press, 2002.

Woodford, M. (2003). *Interest and prices*. Princeton University Press.

A Muth (1961) Under Diagnostic Expectations: Detailed Derivation

A.1 Muth's Over-Discounting

We show an equivalence of DE with Muth's model of deviations from rationality (Muth 1961, pp. 321-2). The shock process can be written as linear combination of current and past normally and independently distributed random variables ϵ_i with zero mean and variance σ^2:

$$u_t = \sum_{i=0}^{\infty} w_i \epsilon_{t-i} \tag{39}$$

Any desired correlogram in the u's may be obtained by an appropriate choice of the weights w_i.

As in Muth (eqn 3.7), conjecture that price is a linear function of the independent disturbances

$$P_t = \sum_{i=0}^{\infty} W_i \epsilon_{t-i} \tag{40}$$

The expected price is then

$$P_t^{\theta} = \mathbb{E}_{t-1}^{\theta}[P_t] = (1+\theta)\mathbb{E}_{t-1}[P_t] - \theta\mathbb{E}_{t-2}[P_t] = \sum_{i=1}^{\infty} W_i \epsilon_{t-i} + \theta W_1 \epsilon_{t-1} = (1+\theta) W_1 \epsilon_{t-1} + \sum_{i=2}^{\infty} W_i \epsilon_{t-i} \tag{41}$$

From the equilibrium equation (4), we obtain:

$$P_t + \frac{\gamma}{\beta}\mathbb{E}_{t-1}^{\theta}[P_t] = -\frac{1}{\beta} u_t \tag{42}$$

$$W_0 \epsilon_t + \left(1 + \frac{\gamma(1+\theta)}{\beta}\right) W_1 \epsilon_{t-1} + \left(1 + \frac{\gamma}{\beta}\right) \sum_{i=2}^{\infty} W_i \epsilon_{t-i} = -\frac{1}{\beta} \sum_{i=0}^{\infty} w_i \epsilon_{t-i} \tag{43}$$

Hence, we get

$$W_0 = -\frac{1}{\beta} w_0 \tag{44}$$

$$\left(1+\frac{\gamma(1+\theta)}{\beta}\right)W_1 = -\frac{1}{\beta}w_1 \iff W_1 = -\frac{1}{\beta+(1+\theta)\gamma}w_1 \tag{45}$$

$$\left(1+\frac{\gamma}{\beta}\right)W_i = -\frac{1}{\beta}w_i \iff W_i = -\frac{1}{\beta+\gamma}w_i \quad \forall i = 2, ... \tag{46}$$

Compare the solution to Muth's solution in equations (3.19) to see that Muth's extrapolation coefficient $f_1 = 1+\theta$.

A.2 Inventory Speculation Model (Muth 1961, Section 4, pp. 322-30)

$$Q_t^d = -\beta P_t; \beta > 0 \qquad \text{(Demand)}$$

$$Q_t^s = I_{t-1} + u_t \qquad \text{(Supply)}$$

$$I_t = \gamma \tilde{\mathbb{E}}_t P_{t+1}; \gamma > 0 \qquad \text{(Production in Advance)}$$

$$S_t = \alpha\left(\tilde{\mathbb{E}}_t[P_{t+1}] - P_t\right); \alpha > 0 \qquad \text{(Inventory Demand)}$$

$$Q_t^d + S_t = Q_t^s + S_{t-1} \qquad \text{(Market Clearing)}$$

We get the following univariate equilibrium condition:

$$-(\alpha+\beta)P_t + \alpha\mathbb{E}_t P_{t+1} = (\alpha+\gamma)\mathbb{E}_{t-1}P_t - \alpha P_{t-1} + u_t \tag{47}$$

We assume that shock process u_t has zero persistence ($\rho = 0$). That is, $u_t = \epsilon_t \sim i.i.d.N(0, \sigma_\epsilon^2)$. Guess the solution takes the following form:

$$P_t = AP_{t-1} + B\epsilon_t + C\epsilon_{t-1} \tag{48}$$

Using method of undetermined coefficients we find that the solution under RE and DE are as follows:

$$P_t^{RE} = AP_{t-1} - \frac{1}{\alpha+\beta-\alpha A}\epsilon_t \tag{49}$$

$$P_t^{DE} = AP_{t-1} + B\epsilon_t + C\epsilon_{t-1} \tag{50}$$

where A is the smaller root of the quadratic $\alpha A^2 - (2\alpha+\beta+\gamma)A + \alpha = 0$,

$$B = -\frac{(\alpha+\gamma)(1+\theta)+\alpha+\beta-\alpha A}{[(\alpha+\gamma)(1+\theta)+\alpha+\beta]\left(\alpha+\beta-\alpha(1+\theta)A\right) - \alpha A\left(\alpha+\beta-\alpha(1+\theta)A\right) + \alpha(1+\theta)(\alpha+\gamma)\theta A} \tag{51}$$

$$C = -\frac{(\alpha+\gamma)\theta AB}{(\alpha+\gamma)(1+\theta)+\alpha+\beta-\alpha A} \tag{52}$$

We can see that, for $\theta > 0$, B is a function of θ (whereas the coefficient $-1/(\alpha+\beta-\alpha A)$ is not), and that $C \neq 0$. This shows that the DE and RE solutions are different, even for i.i.d. shocks.[50]

B Diagnosticy, RE Representation, and More on Predetermined Variables

This appendix collects all proofs for the results stated in Section 3. Standard matrix operations to obtain the solution, and associated proofs (needed once the RE representation has been obtained), are discussed in Appendix C.

B.1 Diagnostic Expectation of Future Variables

Suppose that x_t follows an univariate AR(1) process, $x_t = \rho_x x_{t-1} + \varepsilon_t$, with $\varepsilon_t \sim$ *i.i.d.* $N(0, \sigma_\varepsilon^2)$. Given (realized) states $\check{x}_t$ and $\check{x}_{t-1}$, the diagnostic probability distribution function of x_{t+1} is

$$f_t^\theta(x_{t+1}) = f(x_{t+1}|x_t = \check{x}_t) \cdot \left[\frac{f(x_{t+1}|x_t = \check{x}_t)}{f(x_{t+1}|x_t = \rho_x \check{x}_{t-1})}\right]^\theta \cdot C \tag{53}$$

When looking at equation (53), it is important to notice that, generically, $\check{x}_t \neq \rho_x \check{x}_{t-1}$ (due to the realization of the shock ε_t.) However, since ε_t is fixed at 0 by the NNA, then

$$f(x_{t+1}|x_t = \rho_x \check{x}_{t-1}) \propto \varphi\left(\frac{x_{t+1} - \rho_x^2 \check{x}_{t-1}}{\sigma_\varepsilon}\right) \tag{54}$$

Thanks to the NNA, the variance of this pdf is σ_ε^2, which is the same as the variance of the true pdf of x_{t+1}. Thus, the true and the reference distributions have the same variance. This ensures tractability.

We now prove that the diagnostic expectation of a univariate variable can be expressed in terms of rational expectations.

[50] If one were to drop the P_{t-1} term in the equilibrium equation, it is straightforward to verify that the DE and RE solutions would coincide when $\rho = 0$.

Lemma 1 (Univariate RE Representation) *Suppose that x_t follows an AR(1) process and that the NNA holds. Then,*

$$\mathbb{E}_t^\theta[x_{t+1}] = \mathbb{E}_t[x_{t+1}] + \theta(\mathbb{E}_t[x_{t+1}] - \mathbb{E}_{t-1}[x_{t+1}]) \tag{55}$$

Proof (Lemma 1.) The diagnostic expectation of x_{t+1} is given by

$$\mathbb{E}_t^\theta[x_{t+1}] = \int_{-\infty}^{\infty} x f_t^\theta(x) dx \tag{56}$$

The diagnostic pdf is given by

$$f_t^\theta(x) = \frac{\left[\frac{1}{\sigma_\varepsilon}\varphi\left(\frac{x-\rho_x\check{x}_t}{\sigma_\varepsilon}\right)\right]^{1+\theta}}{\left[\frac{1}{\sigma_\varepsilon}\varphi\left(\frac{x-\rho_x^2\check{x}_{t-1}}{\sigma_\varepsilon}\right)\right]^{\theta}} C \tag{57}$$

where C is a normalizing constant given by

$$C = \exp\left\{ -\frac{1}{2}\left(\frac{\theta(1+\theta)\rho_x^2\check{x}_t^2 + \theta(\theta+1)\rho_x^4\check{x}_{t-1}^2 - 2(1+\theta)\theta\rho_x^3\check{x}_t\check{x}_{t-1}}{\sigma_\varepsilon^2}\right)\right\} \tag{58}$$

in which case

$$\begin{aligned}
\mathbb{E}_t^\theta[x_{t+1}] &= \int_{-\infty}^{\infty} x f_t^\theta(x) dx \\
&= \int_{-\infty}^{\infty} x \frac{1}{\sigma_\varepsilon}\varphi\left(\frac{x - (\rho_x\check{x}_t + \theta(\rho_x\check{x}_t - \rho_x^2\check{x}_{t-1}))}{\sigma_\varepsilon}\right) dx
\end{aligned} \tag{59}$$

Thus, the diagnostic distribution $f_t^\theta(x_{t+1})$ is normal with variance σ_ε^2 and mean

$$\mathbb{E}_t^\theta[x_{t+1}] = \mathbb{E}_t[x_{t+1}] + \theta(\mathbb{E}_t[x_{t+1}] - \mathbb{E}_{t-1}[x_{t+1}]) \tag{60}$$

■

In formula (55), the lagged expectation $\mathbb{E}_{t-1}[x_{t+1}]$ is the expectation conditional on information available at $t-1$, that is, conditional on $\check{x}_{t-1}$. Thus, $\mathbb{E}_t[x_{t+1}] = \rho_x\check{x}_t$ and $\mathbb{E}_{t-1}[x_{t+1}] = \rho_x^2\check{x}_{t-1}$. For a given realized $\check{\varepsilon}_t$, this proof implies that:

$$\mathbb{E}_t^\theta[x_{t+1}] = E_t[x_{t+1}] + \theta\rho_x\check{\varepsilon}_t > \mathbb{E}_t[x_{t+1}] \tag{61}$$

if and only if $\check{\varepsilon}_t > 0$, that is diagnostic expectations indeed extrapolate the past shock into future beliefs.

Extension to Multivariate Case. Assume that the vector $\mathbf{z}_t$ follows a multivariate AR(1) process, $\mathbf{z}_t = \mathbf{A_z}\mathbf{z}_{t-1} + \mathbf{w}_t$, where $\mathbf{w}_t \sim N(0, \Sigma_\mathbf{w})$, and $\mathbf{A_z}$ is a persistence matrix. (Notice that we do not require orthogonality.)

We make the following multivariate no-news assumption (henceforth NNA) for any Gaussian AR(1) vector $\mathbf{z}_{t+1}$.

Assumption 3 (Multivariate No-News Assumption)

$$f(\mathbf{z}_{t+1}| - G_t) = f(\mathbf{z}_{t+1}|\mathbf{z}_t = \mathbf{A_z}\check{\mathbf{z}}_{t-1}) \tag{62}$$

The extension to the multivariate case is based on the fact that each element of the vector is univariate normal.

Lemma 2 (Multivariate DE Formula) *Assume that the vector $\mathbf{z}_t$ follows a multivariate AR(1) process. Then,*

$$\mathbb{E}^\theta_t[\mathbf{z}_{t+1}] = \mathbb{E}_t[\mathbf{z}_{t+1}] + \theta(\mathbb{E}_t[\mathbf{z}_{t+1}] - \mathbb{E}_{t-1}[\mathbf{z}_{t+1}]) \tag{63}$$

Proof (Lemma 2). The proof proceeds element-by-element of the vector $\mathbf{z}_{t+1}$. Without loss of generality, consider the first element $\mathbf{z}_{1,t+1}$. The marginal distribution of $\mathbf{z}_{1,t+1}$ is also normal, and thus, under the NNA,

$$\mathbb{E}^\theta_t[\mathbf{z}_{i,t+1}] = \mathbb{E}_t[\mathbf{z}_{i,t+1}] + \theta(\mathbb{E}_t[\mathbf{z}_{i,t+1}] - \mathbb{E}_{t-1}[\mathbf{z}_{i,t+1}]) \tag{64}$$

The proof for the other elements of $\mathbf{z}_t$ is identical. ■

Equation (11) follows from this lemma.

B.2 Existence and Uniqueness of the Rational Expectations Representation for the General Linear Model

The proof will use the fact that endogenous variables of the DSGE model are normally distributed, allowing to use the multivariate BGS formula (11), together with the linearity of the RE operator. First, we need to prove the following lemma.

Lemma 3 (Distribution of the Linear Combination $\mathbf{F}\mathbf{y}_{t+1} + \mathbf{G_1}\mathbf{y}_t + \mathbf{M}\mathbf{x}_{t+1} + \mathbf{N_1}\mathbf{x}_t$) *Consider the multivariate process (7) and model (12). The vector $\mathbf{F}\mathbf{y}_{t+1} + \mathbf{G_1}\mathbf{y}_t + \mathbf{M}\mathbf{x}_{t+1} + \mathbf{N_1}\mathbf{x}_t$ follows a multivariate normal distribution.*

Proof (Lemma 3). First,

$$\mathbf{M}\mathbf{x}_{t+1} \sim N(\mathbf{M}\mathbf{A}\check{\mathbf{x}}_t, \mathbf{M}\Sigma_{\mathbf{v}}\mathbf{M}') \tag{65}$$

Also,

$$\mathbf{F}\mathbf{y}_{t+1} \sim N(\mathbf{F}\mathbf{P}\check{\mathbf{y}}_t + \mathbf{F}\mathbf{Q}\mathbf{A}\check{\mathbf{x}}_t, \mathbf{F}(\mathbf{Q}+\mathbf{R})\Sigma_{\mathbf{v}}(\mathbf{Q}+\mathbf{R})'\mathbf{F}') \tag{66}$$

Finally,

$$\begin{aligned}\mathbf{F}\mathbf{y}_{t+1} + \mathbf{G_1}\mathbf{y}_t + \mathbf{M}\mathbf{x}_{t+1} + \mathbf{N_1}\mathbf{x}_t \sim N(&(\mathbf{F}\mathbf{P}+\mathbf{G_1})\check{\mathbf{y}}_{\mathbf{t}} + ((\mathbf{F}\mathbf{Q}+\mathbf{M})\mathbf{A}+\mathbf{N_1})\check{\mathbf{x}}_{\mathbf{t}},\\ &(\mathbf{F}(\mathbf{Q}+\mathbf{R})+\mathbf{M})\Sigma_{\mathbf{v}}(\mathbf{F}(\mathbf{Q}+\mathbf{R})+\mathbf{M})')\end{aligned} \tag{67}$$

■

Proof (Proposition 1). Lemma 3 shows $\mathbf{F}\mathbf{y}_{t+1} + \mathbf{G_1}\mathbf{y}_t + \mathbf{M}\mathbf{x}_{t+1} + \mathbf{N_1}\mathbf{x}_t$ is multivariate Gaussian. As a consequence of this fact, we can can evaluate the DE on the multivariate model using Lemma 2. Re-writing equation (14):

$$\begin{aligned}\mathbb{E}_t^\theta[\mathbf{F}\mathbf{y}_{t+1} + \mathbf{G_1}\mathbf{y}_t + \mathbf{M}\mathbf{x}_{t+1} + \mathbf{N_1}\mathbf{x}_t] = {}& \mathbb{E}_t[\mathbf{F}\mathbf{y}_{t+1} + \mathbf{G_1}\mathbf{y}_t + \mathbf{M}\mathbf{x}_{t+1} + \mathbf{N_1}\mathbf{x}_t]\\ &+\theta\big(\mathbb{E}_t[\mathbf{F}\mathbf{y}_{t+1} + \mathbf{G_1}\mathbf{y}_t + \mathbf{M}\mathbf{x}_{t+1} + \mathbf{N_1}\mathbf{x}_t]\\ &-\mathbb{E}_{t-1}[\mathbf{F}\mathbf{y}_{t+1} + \mathbf{G_1}\mathbf{y}_t + \mathbf{M}\mathbf{x}_{t+1} + \mathbf{N_1}\mathbf{x}_t]\big)\end{aligned} \tag{68}$$

Using the linearity of the RE operator:

$$\begin{aligned}\mathbb{E}_t^\theta[\mathbf{F}\mathbf{y}_{t+1} + \mathbf{G_1}\mathbf{y}_t + \mathbf{M}\mathbf{x}_{t+1} + \mathbf{N_1}\mathbf{x}_t] = {}& \mathbf{F}\mathbb{E}_t[\mathbf{y}_{t+1}] + \mathbf{G_1}\mathbb{E}_t[\mathbf{y}_t] + \mathbf{M}\mathbb{E}_t[\mathbf{x}_{t+1}] + \mathbf{N_1}\mathbb{E}_t[\mathbf{x}_t]\\ &+\theta\big(\mathbf{F}\mathbb{E}_t[\mathbf{y}_{t+1}] + \mathbf{G_1}\mathbb{E}_t[\mathbf{y}_t] + \mathbf{M}\mathbb{E}_t[\mathbf{x}_{t+1}] + \mathbf{N_1}\mathbb{E}_t[\mathbf{x}_t]\\ &-\mathbf{F}\mathbb{E}_{t-1}[\mathbf{y}_{t+1}] + \mathbf{G_1}\mathbb{E}_{t-1}[\mathbf{y}_t] + \mathbf{M}\mathbb{E}_{t-1}[\mathbf{x}_{t+1}] + \mathbf{N_1}\mathbb{E}_{t-1}[\mathbf{x}_t]\big)\end{aligned} \tag{69}$$

Since $\mathbb{E}_t[\mathbf{y}_t] = \mathbf{y_t}$ and $\mathbb{E}_t[\mathbf{x}_t] = \mathbf{x}_t$, and using the definitions of $\mathbf{G}$ and $\mathbf{N}$ in the statement of the proposition, we find that equation (12) implies equation (15).

Uniqueness follows from the fact that the DE can only be evaluated in a unique way once NNA on the multivariate model (Assumption 3) has been assumed.

Hence, model (12) is equivalent to model (15).

■

B.3 Formalization of the Limit Argument for Predetermined Variables Presented in Section 3.3

Following the notation of Section 3.3, the predetermined variable is x_t. At time-t, x_t has been realized and therefore it is degenerate. Based on the fact that, in the DSGE model, predetermined variables are in combination of future variables, the limit argument in Section 3.3 proceeded by introducing an arbitrarily amount of uncertainty around x_t. We considered the linear combination $x_t + \delta_{t+1}$, where δ_{t+1} is white noise. We are interested in the DE of $x_t + \delta_{t+1}$:

$$\mathbb{E}_t^\theta[x_t] \equiv \lim_{\sigma_\delta \to 0^+} \mathbb{E}_t^\theta[x_t + \delta_{t+1}] \tag{70}$$

We formalize this idea using the Dirac delta function, defined as follows. Suppose that $\check{x}_t$ is the realization of x_t. Since x_t is degenerate, it can be represented by a cumulative distribution function (cdf) with vanishing uncertainty:

$$Pr(x_t \leq \check{x} | x_t = \check{x}_t) = \lim_{\sigma_\varepsilon \to 0^+} \frac{1}{\sigma_\varepsilon} \Phi\left(\frac{\check{x} - \check{x}_t}{\sigma_\varepsilon}\right) \tag{71}$$

This is the probability that x_t is below any given value $\check{x}$, where $\Phi(x)$ is the cumulative distribution function (cdf) of a standard normal random variable:

$$\Phi\left(\check{x}\right) = \int_{-\infty}^{\check{x}} \varphi(x) dx \tag{72}$$

This implies that $Pr(x_t = \check{x}_t) = 1$ and $Pr(x_t \neq \check{x}_t) = 0$, also denoted using the Dirac delta function $\delta(x)$:

$$\delta(x) = \lim_{a \to 0^+} \frac{1}{a} \varphi\left(\frac{x}{a}\right) \tag{73}$$

with the requirement that $\delta(x)$ is a pdf. Using this notation, $\delta(x_t - \check{x}_t)$ is the pdf of x_t, and thus

$$Pr(x_t \leq \check{x} | x_t = \check{x}_t) = \int_{-\infty}^{\check{x}} \delta(x - \check{x}_t) dx \tag{74}$$

is equal to 1 for $\check{x} \geq \check{x}_t$ and equal to 0 otherwise.

Lemma 4 will now formalize the limit argument presented in the body. It computes the time-t diagnostic expectation of x_t under the NNA. In this case, the reference distribution of x_t is degenerate, with expectation $\rho_x \check{x}_{t-1}$, where $\check{x}_{t-1}$ is the past realization. We represent this reference distribution by a cdf with vanishing uncertainty, as follows

$$Pr(x_t \leq \check{x} | x_t = \rho_x \check{x}_{t-1}) = \lim_{\sigma_\varepsilon \to 0^+} \frac{1}{\sigma_\varepsilon} \Phi\left(\frac{\check{x} - \rho_x \check{x}_{t-1}}{\sigma_\varepsilon}\right) \tag{75}$$

Lemma 4 (DE of a Degenerate Random Variable) *Suppose that x_t is an AR(1) process. Assume that*

$$Pr(x_t \leq \check{x} | x_t = \check{x}_t) = \lim_{\sigma_\varepsilon \to 0^+} \frac{1}{\sigma_\varepsilon} \Phi\left(\frac{\check{x} - \check{x}_t}{\sigma_\varepsilon}\right) \tag{76}$$

and that the NNA holds. Then,

$$\mathbb{E}_t^\theta[x_t] = \check{x}_t + \theta(\check{x}_t - \rho_x \check{x}_{t-1}) \tag{77}$$

Proof (Lemma 4). The diagnostic expectation of x_t is given by

$$\mathbb{E}_t^\theta[x_t] = \int_{-\infty}^{\infty} x f_t^\theta(x) dx \tag{78}$$

In order to get the diagnostic pdf of x_t, we start by looking at the diagnostic cdf, which by virtue of the NNA is

$$Pr_t^\theta(x_t \leq \check{x}) = \lim_{\sigma_\varepsilon \to 0^+} \int_{-\infty}^{\check{x}} \frac{\left[\frac{1}{\sigma_\varepsilon}\varphi\left(\frac{x - \check{x}_t}{\sigma_\varepsilon}\right)\right]^{1+\theta}}{\left[\frac{1}{\sigma_\varepsilon}\varphi\left(\frac{y - \rho_x \check{x}_{t-1}}{\sigma_\varepsilon}\right)\right]^{\theta}} C \; dx \tag{79}$$

First, note that

$$\begin{aligned}
\frac{\left[\frac{1}{\sigma_\varepsilon}\varphi\left(\frac{x - \check{x}_t}{\sigma_\varepsilon}\right)\right]^{1+\theta}}{\left[\frac{1}{\sigma_\eta}\varphi\left(\frac{y - \rho_y \check{x}_{t-1}}{\sigma_\varepsilon}\right)\right]^{\theta}} &= \frac{1}{\sqrt{2\pi}\sigma_\varepsilon} \exp\left\{-\frac{1}{2}\left[(1+\theta)\left(\frac{x - \check{x}_t}{\sigma_\varepsilon}\right)^2 - \theta\left(\frac{x - \rho_x \check{x}_{t-1}}{\sigma_\varepsilon}\right)^2\right]\right\} \\
&= \frac{1}{\sqrt{2\pi}\sigma_\varepsilon} \exp\left\{-\frac{1}{2} \frac{\left[y - \left((1+\theta)\check{x}_t - \theta\rho_t \check{x}_{t-1}\right)\right]^2}{\sigma_\varepsilon^2}\right\} \times \frac{1}{C}
\end{aligned} \tag{80}$$

where the value of C must be

$$C = \exp\left\{-\frac{1}{2}\left[\frac{\theta(1+\theta)\check{x}_t^2 + \theta(1+\theta)\rho_x^2 \check{x}_{t-1}^2 - 2\theta(1+\theta)\rho_x \check{x}_t \check{x}_{t-1}}{\sigma_\varepsilon^2}\right]\right\} \tag{81}$$

so that

$$Pr_t^\theta(x_t \leq \check{x}) = \lim_{\sigma_\varepsilon \to 0^+} \frac{1}{\sigma_\varepsilon} \Phi\left(\frac{\check{x} - (\check{x}_t + \theta(\check{x}_t - \rho_x \check{x}_{t-1}))}{\sigma_\varepsilon}\right) \tag{82}$$

and

$$f_t^\theta(x_t) = \delta(x_t - (\check{x}_t + \theta(\check{x}_t - \rho_x \check{x}_{t-1}))) \tag{83}$$

Hence, in order to look for the DE of x_t, we write

$$\begin{aligned}
\mathbb{E}_t^\theta[x_t] &= \lim_{\sigma_\varepsilon \to 0^+} \lim_{u\to\infty} \int_{-\infty}^{u} x \, \frac{\left[\frac{1}{\sigma_\varepsilon}\varphi\left(\frac{x-\check{x}_t}{\sigma_\varepsilon}\right)\right]^{1+\theta}}{\left[\frac{1}{\sigma_\varepsilon}\varphi\left(\frac{x-\rho_x\check{x}_{t-1}}{\sigma_\varepsilon}\right)\right]^{\theta}} \, C dx \\
&= \lim_{\sigma_\varepsilon \to 0^+} \lim_{u\to\infty} \int_{-\infty}^{u} x \, \frac{1}{\sigma_\varepsilon}\varphi\Big(\frac{x - \big((1+\theta)\check{x}_t - \theta\rho_x\check{x}_{t-1}\big)}{\sigma_\varepsilon}\Big) dx \\
&= \lim_{\sigma_\varepsilon \to 0^+} \lim_{u\to\infty} \Bigg\{ \int_{-\infty}^{u} \frac{x - \big((1+\theta)\check{x}_t - \theta\rho_x\check{x}_{t-1}\big)}{\sigma_\varepsilon} \varphi\Big(\frac{x - \big((1+\theta)\check{x}_t - \theta\rho_x\check{x}_{t-1}\big)}{\sigma_\varepsilon}\Big) dx \\
&+ \big((1+\theta)\check{x}_t - \theta\rho_x\check{x}_{t-1}\big) \int_{-\infty}^{u} \frac{1}{\sigma_\varepsilon}\varphi\Big(\frac{x - \big((1+\theta)\check{x}_t - \theta\rho_x\check{x}_{t-1}\big)}{\sigma_\varepsilon}\Big) dx \Bigg\}
\end{aligned} \tag{84}$$

We will evaluate the integral by change of variables. To this end, define $z \equiv \frac{x-((1+\theta)\check{x}_t-\theta\rho_x\check{x}_{t-1})}{\sigma_\varepsilon}$ such that

$$\mathbb{E}_t^\theta[x_t] = \lim_{\sigma_\varepsilon \to 0^+} \lim_{u\to\infty} \left\{ \sigma_\varepsilon \int_{-\infty}^{\frac{u-((1+\theta)\check{x}_t-\theta\rho_x\check{x}_{t-1})}{\sigma_\varepsilon}} z\varphi(z)dz + ((1+\theta)\check{x}_t - \theta\rho_x\check{x}_{t-1}) \int_{-\infty}^{\frac{u-((1+\theta)\check{x}_t-\theta\rho_x\check{x}_{t-1})}{\sigma_\varepsilon}} \varphi(z)dz \right\} \tag{85}$$

Since $\lim_{\sigma_\varepsilon \to 0^+} \frac{u-((1+\theta)\check{x}_t-\theta\rho_x\check{x}_{t-1})}{\sigma_\varepsilon} = +\infty$ when $u > (1+\theta)\check{x}_t - \theta\rho_x\check{x}_{t-1}$, we have

$$\lim_{\sigma_\varepsilon \to 0^+} \int_{-\infty}^{\frac{u-((1+\theta)\check{x}_t-\theta\rho_x\check{x}_{t-1})}{\sigma_\varepsilon}} z\varphi(z)dz = 0 \quad \text{and} \quad \lim_{\sigma_\varepsilon \to 0^+} \int_{-\infty}^{\frac{u-((1+\theta)\check{x}_t-\theta\rho_x\check{x}_{t-1})}{\sigma_\varepsilon}} \varphi(z)dz = 1 \tag{86}$$

Therefore,

$$\mathbb{E}_t^\theta[x_t] = \check{x}_t + \theta(\check{x}_t - \rho_x\check{x}_{t-1}) \tag{87}$$

■

We already mentioned above that agents in the model do not compute the DE of predetermined variables in isolation. This is important to keep in mind when interpreting expression (77). In the model, predetermined variables are exclusively in linear combination of future variables, which is why diagnosticity is activated for these variables. This is the correct way of thinking why there is a transformation of the expectation of the predetermined variables as well.

The expression (77) allows to obtain additivity in the following sense, extending

BGS's result to settings with predetermined variables.[51] (The case of backward looking expectations is trivial by the absence of diagnosticity.)

Proposition 8 (Strong Additivity of the DE) *Suppose that x_t and y_t are orthogonal AR(1) processes and that the NNA holds. Then, for any integers r, s:*

$$\mathbb{E}_t^\theta[x_{t+r} + y_{t+s}] = \mathbb{E}_t^\theta[x_{t+r}] + \mathbb{E}_t^\theta[y_{t+s}] \tag{88}$$

Proof (Proposition 8). Suppose first that both $r, s \geq 0$. The case $s, r \geq 1$ follows from the fact that both x_{t+r} and y_{t+s} are normal. The case of $s = 0$ or $r = 0$ follows from Lemma 4.

Suppose now that one of $r, s < 0$, and assume, without loss of generality, that this is r. Then, x_{t+r} is a constant, and there is no diagnosticity associated with this variable (i.e., by the NNA, the value under the reference event is the same as the realized value, $G_t = -G_t$). Then, clearly, $\mathbb{E}_t^\theta[x_{t+r}] = x_{t+r}$ and

$$\begin{aligned}\mathbb{E}_t^\theta[x_{t+r} + y_{t+s}] &= (1+\theta)(x_{t+r} + \mathbb{E}_t[y_{t+s}]) - \theta(x_{t+r} + \mathbb{E}_{t-1}[y_{t+s}]) \\ &= x_{t+r} + (1+\theta)\mathbb{E}_t[y_{t+s}] - \theta\mathbb{E}_{t-1}[y_{t+s}] \end{aligned} \tag{89}$$

since x_{t+r} is known at $t-1$. Hence,

$$\mathbb{E}_t^\theta[x_{t+r} + y_{t+s}] = \mathbb{E}_t^\theta[x_{t+r}] + \mathbb{E}_t^\theta[y_{t+s}] \tag{90}$$

The proof when both $r, s < 0$ is similar. ■

This additivity result clarifies that we now have a consistent definition of the DE at all leads and lags.

Case of No Residual Uncertainty, $\sigma_\varepsilon = 0$. For completeness of this discussion, let us now suppose that there is no residual uncertainty in the observation of x_t. In other words, there is no diagnosticity associated with computing DE for x_t at time t (the NNA is dropped). Since x_t is degenerate, its true distribution is $Pr(x_t = \check{x} | x_t = \check{x}_t) = 1$ if $\check{x} = \check{x}_t$ and 0 otherwise. The absence of diagnosticity implies $G_t = -G_t$, and so the diagnostic distribution is equal to the true distribution. As a consequence,

$$\mathbb{E}_t^\theta\left[x_t | \sigma_\varepsilon = 0\right] = \check{x}_t \tag{91}$$

[51] See BGS, proof of Corollary 1, Online Appendix, for the case $r, s \geq 1$.

[FOR ONLINE PUBLICATION]

Online Appendix For: "Incorporating Diagnostic Expectations into the New Keynesian Framework"

Jean-Paul L'Huillier, Sanjay R. Singh, and Donghoon Yoo

August 2022

C Detailed Solution Procedure, Stability, and Boundedness of the Solution: Supplementary Materials and Proofs

Detailed Solution Procedure. We solve for the recursive equilibrium law of motion of a linear diagnostic-expectations DSGE model using the method of undetermined coefficients.

Suppose that there is a unique stable solution of the model:

$$\mathbf{y}_t = \mathbf{P}\mathbf{y}_{t-1} + \mathbf{Q}\mathbf{x}_t + \mathbf{R}\mathbf{v}_t \tag{92}$$

We write the model in the rational expectations representation as

$$\begin{aligned}0 = &\mathbf{F}\mathbf{P}^2\mathbf{y}_{t-1} + \mathbf{F}\mathbf{P}\mathbf{Q}\mathbf{x}_t + \theta\mathbf{F}\mathbf{P}\mathbf{Q}\mathbf{v}_t + (1+\theta)\mathbf{F}\mathbf{P}\mathbf{R}\mathbf{v}_t + \mathbf{F}\mathbf{Q}\mathbf{A}\mathbf{x}_t + \theta\mathbf{F}\mathbf{Q}\mathbf{A}\mathbf{v}_t + \mathbf{G_1}\mathbf{P}\mathbf{y}_{t-1} + ...\\ &+ \mathbf{G_1}\mathbf{Q}\mathbf{x}_t + \theta\mathbf{G_1}\mathbf{Q}\mathbf{v}_t + (1+\theta)\mathbf{G_1}\mathbf{R}\mathbf{v}_t + \mathbf{G_2}\mathbf{P}\mathbf{y}_{t-1} + \mathbf{G_2}\mathbf{Q}\mathbf{x}_t + \mathbf{G_2}\mathbf{R}\mathbf{v}_t + \mathbf{M}\mathbf{A}\mathbf{x}_t + ...\\ &+ \theta\mathbf{M}\mathbf{A}\mathbf{v}_t + \mathbf{N_1}\mathbf{x}_t + \theta\mathbf{N_1}\mathbf{v}_t + \mathbf{H}\mathbf{y}_{t-1} + \mathbf{N}_2\mathbf{x}_t\end{aligned} \tag{93}$$

It is now straightforward to proceed by the method of undetermined coefficients to find a solution of the form (92), and the matrices $\mathbf{P}, \mathbf{Q}, \mathbf{R}$ can be found solving the following matrix equations.

$$\mathbf{F}\mathbf{P}^2 + \mathbf{G}\mathbf{P} + \mathbf{H} = 0 \tag{94}$$

$$\mathbf{F}\mathbf{P}\mathbf{Q} + \mathbf{F}\mathbf{Q}\mathbf{A} + \mathbf{G}\mathbf{Q} + \mathbf{M}\mathbf{A} + \mathbf{N} = 0 \tag{95}$$

$$\theta\mathbf{F}\mathbf{P}\mathbf{Q} + (1+\theta)\mathbf{F}\mathbf{P}\mathbf{R} + \theta\mathbf{F}\mathbf{Q}\mathbf{A} + \theta\mathbf{G_1}\mathbf{Q} + \mathbf{G}\mathbf{R} + \theta\mathbf{G_1}\mathbf{R} + \theta\mathbf{M}\mathbf{A} + \theta\mathbf{N_1} = 0 \tag{96}$$

where $\mathbf{G} = \mathbf{G_1} + \mathbf{G_2}$ and $\mathbf{N} = \mathbf{N_1} + \mathbf{N_2}$.

We can use the techniques discussed in Uhlig (1995) to solve the quadratic matrix equation (94) in $\mathbf{P}$. The solution of the other two equations is straightforward as they are linear in $\mathbf{Q}$ and $\mathbf{R}$: After vectorization, equation (95) becomes

$$(\mathbf{I}_m \otimes \mathbf{FP})vec(\mathbf{Q}) + (\mathbf{A}^T \otimes \mathbf{F})vec(\mathbf{Q}) + (\mathbf{I}_m \otimes \mathbf{G})vec(\mathbf{Q}) + vec(\mathbf{MA}) + vec(\mathbf{N}) = 0 \tag{97}$$

such that

$$vec(\mathbf{Q}) = -\Big((\mathbf{I}_m \otimes \mathbf{FP}) + (\mathbf{A}^T \otimes \mathbf{F}) + (\mathbf{I}_m \otimes \mathbf{G})\Big)^{-1} \times (vec(\mathbf{MA}) + vec(\mathbf{N})) \tag{98}$$

$\mathbf{R}$ can be found from (96):

$$\mathbf{R} = -\big((1+\theta)\mathbf{FP} + \mathbf{G} + \theta\mathbf{G_1}\big)^{-1}\big(\theta\mathbf{FPQ} + \theta\mathbf{FQA} + \theta\mathbf{G_1Q} + \theta\mathbf{MA} + \theta\mathbf{N_1}\big) \tag{99}$$

Observe that solution for matrices $\mathbf{P}$ and $\mathbf{Q}$ does not depend on diagnosticity parameter.

The Solution under Rational Expectations. Consider the model under rational expectations:

$$\mathbf{F}\mathbb{E}_t[\mathbf{y}_{t+1}] + \mathbf{G}\mathbf{y}_t + \mathbf{H}\mathbf{y}_{t-1} + \mathbf{M}\mathbb{E}_t[\mathbf{x}_{t+1}] + \mathbf{N}\mathbf{x}_t = 0 \tag{100}$$

where $\mathbf{G} = \mathbf{G_1} + \mathbf{G_2}$ and $\mathbf{N} = \mathbf{N_1} + \mathbf{N_2}$ and, as above, $\mathbf{y}_t$ and $\mathbf{x}_t$ denote vectors of endogenous variables (including controls and states) $(m \times 1)$ and of exogenous states $(n \times 1)$. $\mathbb{E}_t$ denotes the rational expectation operator, and the exogenous process is given by (7).

Suppose that there is a unique stable solution of the model:

$$\mathbf{y}_t = \widetilde{\mathbf{P}}\mathbf{y}_{t-1} + \widetilde{\mathbf{Q}}\mathbf{x}_t \tag{101}$$

then, we can rewrite the stochastic difference equation (100) as follows:

$$\mathbf{F}\mathbb{E}_t\big[\widetilde{\mathbf{P}}\mathbf{y}_t + \widetilde{\mathbf{Q}}\mathbf{x}_{t+1}\big] + \mathbf{G}\widetilde{\mathbf{P}}\mathbf{y}_{t-1} + \mathbf{G}\widetilde{\mathbf{Q}}\mathbf{x}_t + \mathbf{H}\mathbf{y}_{t-1} + \mathbf{MA}\mathbf{x}_t + \mathbf{N}\mathbf{x}_t = 0 \tag{102}$$

We can simplify the above equation to

$$\mathbf{F}\widetilde{\mathbf{P}}^2\mathbf{y}_{t-1} + \mathbf{F}\widetilde{\mathbf{P}}\widetilde{\mathbf{Q}}\mathbf{x}_t + \mathbf{F}\widetilde{\mathbf{Q}}\mathbf{A}\mathbf{x}_t + \mathbf{G}\widetilde{\mathbf{P}}\mathbf{y}_{t-1} + \mathbf{G}\widetilde{\mathbf{Q}}\mathbf{x}_t + \mathbf{H}\mathbf{y}_{t-1} + \mathbf{M}\mathbf{A}\mathbf{x}_t + \mathbf{N}\mathbf{x}_t = 0 \tag{103}$$

and can solve similarly for the recursive equilibrium law of motion via the method of undetermined coefficients. Specifically, the matrices $\widetilde{\mathbf{P}}$ and $\widetilde{\mathbf{Q}}$ can be found solving the following matrix equations.

$$\mathbf{F}\widetilde{\mathbf{P}}^2 + \mathbf{G}\widetilde{\mathbf{P}} + \mathbf{H} = 0 \tag{104}$$

$$\mathbf{F}\widetilde{\mathbf{P}}\widetilde{\mathbf{Q}} + \mathbf{F}\widetilde{\mathbf{Q}}\mathbf{A} + \mathbf{G}\widetilde{\mathbf{Q}} + \mathbf{M}\mathbf{A} + \mathbf{N} = 0 \tag{105}$$

Comparison of these equations with their counterpart under DE immediately shows that $\mathbf{P} = \widetilde{\mathbf{P}}$ and $\mathbf{Q} = \widetilde{\mathbf{Q}}$.

Stability Conditions. Given the quadratic matrix equation (94)

$$\mathbf{F}\mathbf{P}^2 + \mathbf{G}\mathbf{P} + \mathbf{H} = \mathbf{0} \tag{106}$$

for the $m \times m$ matrix $\mathbf{P}$ and $m \times m$ matrices $\mathbf{G}$ and $\mathbf{H}$, define the $2m \times 2m$ matrices $\boldsymbol{\Xi}$ and $\boldsymbol{\Delta}$:

$$\boldsymbol{\Xi} = \begin{bmatrix} -\mathbf{G} & -\mathbf{H} \\ \mathbf{I}_m & \mathbf{0}_m \end{bmatrix} \tag{107}$$

and

$$\boldsymbol{\Delta} = \begin{bmatrix} -\mathbf{F} & \mathbf{0}_m \\ \mathbf{0}_m & \mathbf{I}_m \end{bmatrix} \tag{108}$$

where $\mathbf{I}_m$ is the identity matrix of size m and $\mathbf{0}_m$ is the $m \times m$ matrix with only zero entries.

Uhlig (1995) shows that if (a) s is a generalized eigenvector and λ is the corresponding generalized eigenvalue of $\boldsymbol{\Xi}$ with respect to $\boldsymbol{\Delta}$, then s can be written as $s' = \left[\lambda x', x'\right]$ for some $x \in \mathsf{R}^m$, and (b) there are m generalized eigenvalues $\lambda_1, ..., \lambda_m$ together with generalized eigenvectors $s_1, ..., s_m$ of $\boldsymbol{\Xi}$ with respect to $\boldsymbol{\Delta}$, written as $s_i' = [\lambda_i x_i', x_i']$ for some $x_i \in \mathsf{R}^m$, and if $(x_1, ..., x_m)$ is linearly dependent, then

$$\mathbf{P} = \boldsymbol{\Omega}\boldsymbol{\Lambda}\boldsymbol{\Omega}' \tag{109}$$

is a solution to the matrix quadratic equation, where $\boldsymbol{\Omega} = [x_1, ..., x_m]$ and $\boldsymbol{\Lambda} = diag(\lambda_1, ..., \lambda_m)$.

The stability conditions are given as follows.[52]

Theorem 1 *The solution* $\mathbf{P}$ *is stable if* $|\lambda_i| < 1$ *for all* $i = 1, ..., m$.

Thus, we can easily show that the stability conditions for both models are the same.

Proof (Proposition 2). The solutions $\mathbf{P}$ and $\widetilde{\mathbf{P}}$ are the same since they involve identical matrices $\mathbf{F}$, $\mathbf{G}$, and $\mathbf{H}$. Thus, the stability conditions stated in Theorem 1 are the same for both solutions. ■

Proof (Proposition 3). Let's consider the RE model presented in equation (100) where the exogenous variables are stacked in a $(n \times 1)$ vector $\mathbf{x}_t$ that is assumed to follow the AR(1) stochastic process

$$\mathbf{x}_t = \mathbf{A}\mathbf{x}_{t-1} + \mathbf{v}_t \tag{110}$$

where $\mathbf{v}_t$ is a $(k \times 1)$ vector of Gaussian and orthogonal exogenous shocks:

$$\mathbf{v}_t \sim N(0, \Sigma_{\mathbf{v}}) \tag{111}$$

and A is a diagonal matrix of persistence parameters.

Suppose that there is a unique stable solution of the model:

$$\mathbf{y}_t = \mathbf{P}\mathbf{y}_{t-1} + \mathbf{Q}\mathbf{x}_t \tag{112}$$

Assume, without loss of generality, that any unanticipated shocks or news only hit the economy at date 1. The economy is in steady state at date 0 or before. Then, the solution of the DE model from date 2 onwards coincides with the RE model solution. We prove this statement by considering the RE representation of the DE model derived in equation (15), reproduced here:

$$\begin{aligned}\mathbf{F}\mathbb{E}_t[\mathbf{y}_{t+1}] + \mathbf{G}\mathbf{y}_t + \mathbf{H}\mathbf{y}_{t-1} + \mathbf{M}\mathbb{E}_t[\mathbf{x}_{t+1}] + \mathbf{N}\mathbf{x}_t&\\ +\mathbf{F}\theta\big(\mathbb{E}_t[\mathbf{y}_{t+1}] - \mathbb{E}_{t-1}[\mathbf{y}_{t+1}]\big)&\\ +\mathbf{M}\theta\big(\mathbb{E}_t[\mathbf{x}_{t+1}] - \mathbb{E}_{t-1}[\mathbf{x}_{t+1}]\big)&\\ +\mathbf{G}_1\theta\big(\mathbf{y}_t - \mathbb{E}_{t-1}[\mathbf{y}_t]\big)&\\ +\mathbf{N}_1\theta\big(\mathbf{x}_t - \mathbb{E}_{t-1}[\mathbf{x}_t]\big) = 0&\end{aligned} \tag{113}$$

[52] See Section 6.3 of Uhlig (1995) for a detailed discussion.

Since no news or shocks are assumed to happen for $t \geq 2$, we get that

$$\mathbb{E}_t[\mathbf{y}_{t+1}] - \mathbb{E}_{t-1}[\mathbf{y}_{t+1}] = \mathbb{E}_t[\mathbf{x}_{t+1}] - \mathbb{E}_{t-1}[\mathbf{x}_{t+1}] = \mathbf{y}_t - \mathbb{E}_{t-1}[\mathbf{y}_t] = \mathbf{x}_t - \mathbb{E}_{t-1}[\mathbf{x}_t] = 0; \quad \forall t \geq 2 \tag{114}$$

The system from date $t \geq 2$ then simplifies to the RE model, the solution of which is given by equation (112) for $t \geq 2$. Date 1 solution for the DE model can then be found from (note the assumption that the economy is in steady state before date 1):

$$\begin{aligned} &\mathbf{F}\mathbb{E}_1[\mathbf{y}_2] + \mathbf{G}\mathbf{y}_1 + \mathbf{M}\mathbb{E}_1[\mathbf{x}_2] + \mathbf{N}\mathbf{x}_1 \\ &+\theta\big(\mathbf{F}\mathbb{E}_1[\mathbf{y}_2] + \mathbf{M}\mathbb{E}_1[\mathbf{x}_2] + \mathbf{G}_1\mathbf{y}_1 + \mathbf{N}_1\mathbf{x}_1\big) = 0 \end{aligned} \tag{115}$$

Notice that $\mathbb{E}_1[\mathbf{y}_2]$ and $\mathbb{E}_1[\mathbf{x}_2]$ are known at date 1 from the RE solution.

$$\mathbb{E}_1[\mathbf{y}_2] = \mathbf{P}\mathbf{y}_1 + \mathbf{Q}\mathbf{A}\mathbf{x}_1; \quad \mathbb{E}_1[\mathbf{x}_2] = \mathbf{A}\mathbf{x}_1 \tag{116}$$

After substituting these values and rearranging, we get:

$$\big((1+\theta)\mathbf{F}\mathbf{P} + \mathbf{G} + \theta\mathbf{G}_1\big)\mathbf{y}_1 + \big((1+\theta)(\mathbf{F}\mathbf{Q} + \mathbf{M})\mathbf{A} + \mathbf{N} + \theta\mathbf{N}_1\big)\mathbf{x}_1 = 0 \tag{117}$$

Then, it follows that a bounded solution for the DE model exists if $(1+\theta)\mathbf{F}\mathbf{P}+\mathbf{G}+\theta\mathbf{G}_1$ is full-rank. ■

General Condition for Extra Volatility We establish a general result about when DE generate extra volatility over RE. Specific examples are provided in Section 4. As a reminder, in the case of DE, the solution of a general linear model takes the form:

$$\mathbf{y}_t = \mathbf{P}\mathbf{y}_{t-1} + \mathbf{Q}\mathbf{x}_t + \mathbf{R}\mathbf{v}_t \tag{118}$$

Instead, in the case of RE, the solution of model takes the form:

$$\mathbf{y}_t = \widetilde{\mathbf{P}}\mathbf{y}_{t-1} + \widetilde{\mathbf{Q}}\mathbf{x}_t \tag{119}$$

Comparing these two immediately leads to conjecture that, under DE, there should be extra volatility due to the presence of the extra term $\mathbf{R}\mathbf{v}_t$. However, whether this conjecture is true for a given set of parameters will depend on the covariance of the matrix Q with the other matrices of parameters in the solution. This is what the following proposition makes precise.

Proposition 9 (Extra Volatility) *Let $\mathbf{y}_t^{DE}$ and $\mathbf{y}_t^{RE}$ denote the vectors of endogenous variables under DE and RE, respectively. Let y_{it}^{DE} and y_{it}^{RE} respectively denote the i-th component of the vector of endogenous variables $\mathbf{y}_t^{DE}$ and $\mathbf{y}_t^{RE}$ and $Var(y_{it}^{DE})$ and $Var(y_{it}^{RE})$ denote the variance of the variable y_{it}^{DE} and of the variable y_{it}^{RE}. Then, $Var(y_{it}^{DE})$ is larger than $Var(y_{it}^{RE})$ if and only if:*

$$diag(\mathbf{R}\mathbf{\Sigma}_{\mathbf{v}}\mathbf{R}' + 2\mathbf{Q}\mathbf{\Sigma}_{\mathbf{v}}\mathbf{R}')_i > 0 \tag{120}$$

where $\mathbf{\Sigma}_{\mathbf{v}}$ is the variance-covariance matrix of $\mathbf{v}_t$.

Proof. We have already shown that $\mathbf{P}$ and $\widetilde{\mathbf{P}}$ are the same and that $\mathbf{Q}$ and $\widetilde{\mathbf{Q}}$ are the same. Thus, given the exogenous process $\mathbf{x}_t$, the solution for the model with diagnostic expectations and for the model with rational expectations can be formulated as

$$\mathbf{y}_t^{DE} = \mathbf{P}\mathbf{y}_{t-1} + \mathbf{Q}\mathbf{x}_t + \mathbf{R}\mathbf{v}_t \tag{121}$$

$$\mathbf{y}_t^{RE} = \mathbf{P}\mathbf{y}_{t-1} + \mathbf{Q}\mathbf{x}_t \tag{122}$$

such that the variance of the vector of endogenous variables under diagnostic expectations, $\mathbf{y}_t^{DE}$, is given by

$$\begin{aligned} Var(\mathbf{y}_t^{DE}) &= Var(\mathbf{P}\mathbf{y}_{t-1}) + Var(\mathbf{Q}\mathbf{x}_t) + Var(\mathbf{R}\mathbf{v}_t) \\ &\quad + 2\ Cov(\mathbf{P}\mathbf{y}_{t-1}, \mathbf{Q}\mathbf{x}_t) + 2\ Cov(\mathbf{P}\mathbf{y}_{t-1}, \mathbf{R}\mathbf{v}_t) + 2\ Cov(\mathbf{Q}\mathbf{x}_t, \mathbf{R}\mathbf{v}_t) \end{aligned} \tag{123}$$

Similarly, the variance of the vector of endogenous variables under rational expectations, $\mathbf{y}_t^{RE}$ is given by

$$Var(\mathbf{y}_t^{RE}) = Var(\mathbf{P}\mathbf{y}_{t-1}) + Var(\mathbf{Q}\mathbf{x}_t) + 2\ Cov(\mathbf{P}\mathbf{y}_{t-1}, \mathbf{Q}\mathbf{x}_t) \tag{124}$$

Since $cov(\mathbf{P}\mathbf{y}_{t-1}, \mathbf{R}\mathbf{v}_t) = 0$, (123) is simplified to

$$Var(\mathbf{y}_t^{DE}) = Var(\mathbf{P}\mathbf{y}_{t-1}) + Var(\mathbf{Q}\mathbf{x}_t) + Var(\mathbf{R}\mathbf{v}_t) + 2\ Cov(\mathbf{P}\mathbf{y}_{t-1}, \mathbf{Q}\mathbf{x}_t) + 2\ Cov(\mathbf{Q}\mathbf{x}_t, \mathbf{R}\mathbf{v}_t) \tag{125}$$

such that by taking the difference of the two variances, we have

$$\begin{aligned} Var(\mathbf{y}_t^{DE}) - Var(\mathbf{y}_t^{RE}) &= Var(\mathbf{R}\mathbf{v}_t) + 2\ Cov(\mathbf{Q}\mathbf{x}_t, \mathbf{R}\mathbf{v}_t) \\ &= Var(\mathbf{R}\mathbf{v}_t) + 2\ Cov(\mathbf{Q}\mathbf{A}\mathbf{x}_{t-1} + \mathbf{Q}\mathbf{v}_t, \mathbf{R}\mathbf{v}_t) \\ &= \mathbf{R}\mathbf{\Sigma}_{\mathbf{v}}\mathbf{R}' + 2\mathbf{Q}\mathbf{\Sigma}_{\mathbf{v}}\mathbf{R}' \end{aligned} \tag{126}$$

Thus, for an endogenous variable y_{it} to have extra volatility with diagnostic expectations, the i-th diagonal component of the matrix $\mathbf{R}\mathbf{\Sigma}_{\mathbf{v}}\mathbf{R}' + 2\mathbf{Q}\mathbf{\Sigma}_{\mathbf{v}}\mathbf{R}'$ must be greater than zero. ■

We conclude by making a parallel to the work by Matsuyama (2007), who highlights, in the context of financial frictions, that equilibrium properties change non-monotonically with parameter values in such models. Looking at the expression for the matrix $\mathbf{R}$ reveals that it is a non-linear function of θ. Hence, even values of θ close to zero have the potential to (discontinuously) induce large volatility in linear models.

D Diagnostic New Keynesian Model: Detailed Derivation and Proofs

There are three sets of agents in the economy: households, firms and government. Total output produced is equal to consumption expenditure made by the households and adjustment costs spent in adjusting prices.

D.1 Diagnostic Distribution

We first generalize the concept of diagnostic distribution to non-linear models with exogenous shocks.

Let $\mathcal{D}_t$ be a vector of variables, endogenous and exogenous. Assume there is a non-linear transformation $\mathcal{D}_t \equiv \mathbf{T}\left(\mathcal{D}_{t-1}, \mathbf{v}_t\right)$, that maps time-$t-1$ variables, $\mathcal{D}_{t-1}$, and a given realization of the exogenous shock process $\check{\mathbf{v}}_t$, where $\mathbf{v}_t$ is a vector of i.i.d. exogenous Gaussian shocks $N(0, \Sigma_{\mathbf{v}})$. Notice that this can accommodate the AR(1) of exogenous variables as in Section 3. Let $f\left(\mathcal{D}_{t+1}|\mathcal{D}_t = \mathbf{T}\left(\mathcal{D}_{t-1}, \check{\mathbf{v}}_t\right)\right)$ denote the true distribution of $\mathcal{D}_{t+1}$ at time $t+1$, conditional on current state variables. Let $f\left(\mathcal{D}_{t+1}|\mathcal{D}_t = \mathbf{T}\left(\mathcal{D}_{t-1}, \mathbb{E}_{t-1}[\mathbf{v}_t]\right)\right)$ denote the true distribution of $\mathcal{D}_{t+1}$ at time $t+1$ conditional on a reference set of the state vector $\mathbf{T}\left(\mathcal{D}_{t-1}, \mathbb{E}_{t-1}[\mathbf{v}_t]\right)$. As in the no-news assumption, the agent has not observed the current realization of the shocks $\mathbf{v}_t$ in the reference set, and hence forms forecast of $\mathcal{D}_{t+1}$ assuming a counterfactual path for state vector given by the expectation of the shocks. Following Maxted (2022), BGS, Gennaioli and Shleifer (2010), the "representativeness" of future states $\mathcal{D}_{t+1}$ is given by the likelihood ratio:

$$\frac{f\left(\mathcal{D}_{t+1}|\mathcal{D}_t = \mathbf{T}\left(\mathcal{D}_{t-1}, \check{\mathbf{v}}_t\right)\right)}{f\left(\mathcal{D}_{t+1}|\mathcal{D}_t = \mathbf{T}\left(\mathcal{D}_{t-1}, \mathbb{E}_{t-1}[\mathbf{v}_t]\right)\right)} \tag{127}$$

Diagnostic expectations overweight states that are representative of recent news, the ones exhibiting the largest increase in likelihood based on recent information. This diagnostic distribution is formalized by assuming that agents evaluate future levels of state vector evolve as follows:

$$f^{\theta}\left(\mathcal{D}_{t+1}|\mathcal{D}_t = \mathbf{T}\left(\mathcal{D}_{t-1}, \check{\mathbf{v}}_t\right)\right) = f\left(\mathcal{D}_{t+1}|\mathcal{D}_t = \mathbf{T}\left(\mathcal{D}_{t-1}, \check{\mathbf{v}}_t\right)\right) \cdot \left[\frac{f\left(\mathcal{D}_{t+1}|\mathcal{D}_t = \mathbf{T}\left(\mathcal{D}_{t-1}, \check{\mathbf{v}}_t\right)\right)}{f\left(\mathcal{D}_{t+1}|\mathcal{D}_t = \mathbf{T}\left(\mathcal{D}_{t-1}, \mathbb{E}_{t-1}[\mathbf{v}_t]\right)\right)}\right]^{\theta} \frac{1}{Z} \tag{128}$$

The extent to which representativeness distorts expectations is governed by the parameter θ.

D.2 Households

Households have the following lifetime utility

$$\log C_t - \omega \frac{L_t^{1+\nu}}{1+\nu} + \mathbb{E}_t^{\theta}\left[\Sigma_{s=t+1}^{\infty} \beta^{s-t}\left[\log(C_s) - \frac{\omega}{1+\nu} L_s^{1+\nu}\right]\right] \tag{129}$$

subject to budget constraint:

$$P_t C_t + \frac{B_{t+1}}{(1+i_t)} = B_t + W_t L_t + D_t + T_t\,, \tag{130}$$

$P_t C_t$ is nominal expenditure on final consumption good, B_{t+1} denotes purchase of nominal bonds that pay off $1 + i_t$ interest rate in the following period, $W_t L_t$ denotes labor income, D_t and T_t denote dividends from firm-ownership and lump-sum government transfers respectively. $\mathbb{E}_t^{\theta}$ is the diagnostic expectations operator with diagnosticity parameter θ.

Let $\log C_t \equiv u(C_t)$. The consumption Euler equation is given by:

$$\frac{u'(C_t)}{P_t} = \beta(1+i_t)\mathbb{E}_t^{\theta}\left[\frac{u'(C_{t+1})}{P_{t+1}}\right] \tag{131}$$

Substitute the resource constraint $Y_t = C_t$, and loglinearizing:

$$\hat{y}_t = \mathbb{E}_t^{\theta}\left[\hat{y}_{t+1}\right] - (\hat{i}_t - (\mathbb{E}_t^{\theta}[\hat{p}_{t+1}] - \hat{p}_t)) \tag{132}$$

where $\{\hat{y}_t, \hat{i}_t, \hat{p}_t\}$ denote loglinear deviations of output, the nominal interest rate from their respective steady states, and of the price level from an initial price level, respectively. We can show that $(\mathbb{E}_t^{\theta}[\hat{p}_{t+1}] - \hat{p}_t)$ can be rewritten as $\mathbb{E}_t^{\theta}[\hat{\pi}_{t+1}] + \theta(\hat{\pi}_t - \mathbb{E}_{t-1}[\hat{\pi}_t])$.

Using the BGS formula (11) presented in the main text, we can get:

$$\mathbb{E}_t^{\theta}[\hat{p}_{t+1}] - \hat{p}_t = (1+\theta)\mathbb{E}_t[\hat{p}_{t+1}] - \theta\mathbb{E}_{t-1}[\hat{p}_{t+1}] - \hat{p}_t \tag{133}$$

Adding and subtracting $(1+\theta)\hat{p}_t$, we get:

$$\mathbb{E}_t^{\theta}[\hat{p}_{t+1}] - \hat{p}_t = (1+\theta)\mathbb{E}_t[\hat{\pi}_{t+1}] - \theta\mathbb{E}_{t-1}[\hat{p}_{t+1}] + \theta\hat{p}_t \tag{134}$$

Adding and subtracting $\theta\mathbb{E}_{t-1}[\hat{p}_t]$, we get

$$\mathbb{E}_t^{\theta}[\hat{p}_{t+1}] - \hat{p}_t = (1+\theta)\mathbb{E}_t[\hat{\pi}_{t+1}] - \theta\mathbb{E}_{t-1}[\hat{\pi}_{t+1}] - \theta\mathbb{E}_{t-1}[\hat{p}_t] + \theta\hat{p}_t \tag{135}$$

Adding and subtracting $\theta\hat{p}_{t-1}$, we get

$$\mathbb{E}_t^{\theta}[\hat{p}_{t+1}] - \hat{p}_t = (1+\theta)\mathbb{E}_t[\hat{\pi}_{t+1}] - \theta\mathbb{E}_{t-1}[\hat{\pi}_{t+1}] + \theta(\hat{\pi}_t - \mathbb{E}_{t-1}[\hat{\pi}_t]) \tag{136}$$

Recognize that $(1+\theta)\mathbb{E}_t[\hat{\pi}_{t+1}] - \theta\mathbb{E}_{t-1}[\hat{\pi}_{t+1}] \equiv \mathbb{E}_t^{\theta}[\hat{\pi}_{t+1}]$, we get that

$$\mathbb{E}_t^{\theta}[\hat{p}_{t+1}] - \hat{p}_t = \mathbb{E}_t^{\theta}[\hat{\pi}_{t+1}] + \theta(\hat{\pi}_t - \mathbb{E}_{t-1}[\hat{\pi}_t]) \tag{137}$$

D.3 Firms

Monopolistically competitive firms, indexed by $j \in [0,1]$, produce a differentiated good, $Y_t(j)$. We assume a Dixit-Stiglitz aggregator that aggregates intermediate goods into a final good, Y_t. Intermediate goods demand given by:

$$Y_t(j) = \left(\frac{P_t(j)}{P_t}\right)^{-\epsilon_p} Y_t \tag{138}$$

where $\epsilon_p > 1$ is the elasticity of substitution across intermediate goods' varieties, $P_t(j)$ is price of intermediate good j, and P_t is the price of final good Y_t. Each intermediate good is produced using the technology:

$$Y_t(j) = A_t L_t(j) \tag{139}$$

where $\log(A_t)$ is an aggregate TFP process that follows an AR(1) process with persistence coefficient ρ_a:

$$\log A_t = \rho_a \log A_{t-1} + \varepsilon_{a,t} \tag{140}$$

where $\epsilon_{a,t} \sim iid\ N(0, \sigma_a^2)$. Firm pays a quadratic adjustment cost in units of final good (Rotemberg 1982) to adjust prices:

$$\frac{\psi_p}{2}\left(\frac{P_t(j)}{P_{t-1}(j)} - 1\right)^2 P_t Y_t \tag{141}$$

Firm's per period profits are given by:

$$D_t \equiv P_t(j)Y_t(j) - W_t L_t(j) - \frac{\psi_p}{2}\left(\frac{P_t(j)}{P_{t-1}(j)} - 1\right)^2 P_t Y_t \tag{142}$$

Firm's profit maximization problem

$$\max_{P_t(j)} \left\{ P_t(j)Y_t(j) - W_t L_t(j) - \frac{\psi_p}{2}\left(\frac{P_t(j)}{P_{t-1}(j)} - 1\right)^2 P_t Y_t + \mathbb{E}_t^\theta \left[\sum_{s=1}^{\infty} \beta^s Q_{t,t+s} D_{t+s}\right]\right\} \tag{143}$$

where $Q_{t,t+s}$ is the nominal stochastic discount factor of the household. Substitute in the demand for intermediate goods to get:

$$\max_{P_t(j)} \left\{ P_t(j)\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon_p} Y_t - \frac{W_t}{A_t}\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon_p} Y_t - \frac{\psi_p}{2}\left(\frac{P_t(j)}{P_{t-1}(j)} - 1\right)^2 P_t Y_t + \mathbb{E}_t^\theta \left[\sum_{s=1}^{\infty} \beta^s Q_{t,t+s} D_{t+s}\right]\right\} \tag{144}$$

Notice that $P_t(j)$ appears in period t profits and period $t+1$ adjustment costs. It doesn't appear anywhere else in the problem. So we can "ignore" the remaining terms as we take the first-order condition. The monopolistically competitive firm solves the following problem:

$$\max_{P_t(j)} \left\{ P_t(j)\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon_p} Y_t - \frac{W_t}{A_t}\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon_p} Y_t - \frac{\psi_p}{2}\left(\frac{P_t(j)}{P_{t-1}(j)} - 1\right)^2 P_t Y_t - \mathbb{E}_t^\theta \left[\beta Q_{t,t+1} \frac{\psi_p}{2}\left(\frac{P_{t+1}(j)}{P_t(j)} - 1\right)^2 P_{t+1} Y_{t+1}\right]\right\}$$
$$+ \text{ other terms} \tag{145}$$

First order condition:

$$\begin{aligned}(1-\epsilon_p)\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon_p} Y_t + \epsilon_p \frac{W_t}{A_t P_t}\left(\frac{P_t(j)}{P_t}\right)^{-\epsilon_p - 1} Y_t - \psi_p \left(\frac{P_t(j)}{P_{t-1}(j)} - 1\right)\frac{P_t}{P_{t-1}(j)} Y_t \\ -\psi_p \beta \mathbb{E}_t^\theta \left[\frac{u'(C_{t+1}}{u'(C_t)}\left(\frac{P_{t+1}(j)}{P_t(j)} - 1\right)\frac{P_{t+1}(j)}{P_t(j)}\frac{P_t}{P_t(j)} Y_{t+1}\right] = 0\end{aligned} \tag{146}$$

Symmetry across all firms implies that reset price equals the aggregate price level.

Define $\Pi_t = \frac{P_t}{P_{t-1}}$:

$$(1-\epsilon_p)Y_t + \epsilon_p \frac{W_t}{A_t P_t} Y_t - \psi_p(\Pi_t - 1)\Pi_t Y_t + \psi_p \beta \mathbb{E}_t^\theta \left[\frac{u'(C_{t+1}}{u'(C_t)} (\Pi_{t+1} - 1)\Pi_{t+1} Y_{t+1} \right] = 0 \tag{147}$$

Divide by Y_t:

$$(1-\epsilon_p) + \epsilon_p \frac{W_t}{A_t P_t} - \psi_p(\Pi_t - 1)\Pi_t + \frac{\psi_p}{Y_t} \beta \mathbb{E}_t^\theta \left[\frac{u'(C_{t+1}}{u'(C_t)} (\Pi_{t+1} - 1)\Pi_{t+1} Y_{t+1} \right] = 0 \tag{148}$$

Log-linearize around the deterministic steady state such that $A = 1$, $w = \frac{W}{P} = \omega C Y^\nu = \frac{\epsilon_p - 1}{\epsilon_p}$, $\Pi = 1$, and $Y_t = Y$. Let $w_t = \frac{W_t}{P_t}$

$$\epsilon_p w(\hat{w}_t - \hat{a}_t) - \psi_p \hat{\pi}_t + \psi_p \beta \mathbb{E}_t^\theta \hat{\pi}_{t+1} = 0 \tag{149}$$

Rearrange to get

$$\hat{\pi}_t = \beta \mathbb{E}_t^\theta [\hat{\pi}_{t+1}] + \frac{\epsilon_p w}{\psi_p} (\hat{w}_t - \hat{a}_t) \tag{150}$$

From the intra-temporal labor supply first order condition, we have:

$$\hat{w}_t = \hat{c}_t + \nu(\hat{y}_t - \hat{a}_t) \tag{151}$$

Use the resource constraint $\hat{c}_t = \hat{y}_t$, to rewrite the new Keynesian Phillips Curve (NKPC):

$$\hat{\pi}_t = \beta \mathbb{E}_t^\theta [\hat{\pi}_{t+1}] + \frac{\epsilon_p w}{\psi_p} (1 + \nu)\hat{y}_t \tag{152}$$

Note that $\frac{\epsilon_p \ w}{\psi_p} = \frac{\epsilon_p - 1}{\psi_p}$. Then, the NKPC is given by

$$\hat{\pi}_t = \beta \mathbb{E}_t^\theta [\hat{\pi}_{t+1}] + \kappa(\hat{y}_t - \hat{a}_t) \tag{153}$$

where $\kappa \equiv \frac{\epsilon_p - 1}{\psi_p}(1 + \nu)$.

D.4 Policy Rule

The government sets nominal interest rate with the following rule:

$$\frac{1+i_t}{1+i_{ss}} = \Pi_t^{\phi_\pi} \left(\frac{Y_t}{Y_t^*}\right)^{\phi_x} \tag{154}$$

where $Y_t^* = A_t$ is the natural rate allocation, $i_{ss} = \frac{1}{\beta} - 1$ is the steady state nominal interest rate, $\phi_\pi \geq 0$, $\phi_x \geq 0$, and steady state inflation $\Pi = 1$. Log-linearized policy rule is given by:

$$\hat{i}_t = \phi_\pi \hat{\pi}_t + \phi_x(\hat{y}_t - \hat{a}_t) \tag{155}$$

We assume that nominal bonds are in net zero supply. There is no government spending.

D.5 Market Clearing

Total output produced is used for consumption expenditure.

$$Y_t = C_t \tag{156}$$

D.6 Equilibrium

The log-linearized equilibrium in the New Keynesian model with diagnostic expectations is given by following three equations in three unknowns $\{\hat{y}_t, \hat{\pi}_t, \hat{i}_t\}$ for a given shock process $\{\hat{a}_t\}$.

$$\hat{y}_t = \mathbb{E}_t^\theta\left[\hat{y}_{t+1}\right] - (\hat{i}_t - \mathbb{E}_t^\theta[\hat{\pi}_{t+1}]) + \theta(\hat{\pi}_t - E_{t-1}[\hat{\pi}_t]) \tag{157}$$

$$\hat{\pi}_t = \beta \mathbb{E}_t^\theta[\hat{\pi}_{t+1}] + \kappa(\hat{y}_t - \hat{a}_t) \tag{158}$$

$$\hat{i}_t = \phi_\pi \hat{\pi}_t + \phi_x(\hat{y}_t - \hat{a}_t) \tag{159}$$

where $\kappa \equiv \frac{\epsilon_p - 1}{\psi_p}(1+\nu)$, and the shock processes are given by:

$$\hat{a}_t = \rho_a \hat{a}_{t-1} + \varepsilon_{a,t} \tag{160}$$

where $\varepsilon_{a,t} \sim iid\ N(0, \sigma_a^2)$.

D.7 Solution

D.7.1 Rational Expectations

Under RE, the solution of the model with TFP shocks is given by:

$$\hat{y}_t = \rho_a \frac{\phi_x(1-\beta\rho_a)+\kappa(\phi_\pi-\rho_a)}{(1+\phi_x-\rho_a)(1-\beta\rho_a)+\kappa(\phi_\pi-\rho_a)}\hat{a}_{t-1} + \frac{\phi_x(1-\beta\rho_a)+\kappa(\phi_\pi-\rho_a)}{(1+\phi_x-\rho_a)(1-\beta\rho_a)+\kappa(\phi_\pi-\rho_a)}\varepsilon_{a,t} \tag{161}$$

$$\hat{\pi}_t = \frac{-\rho_a\kappa(1-\rho_a)}{(1+\phi_x-\rho_a)(1-\beta\rho_a)+\kappa(\phi_\pi-\rho_a)}\hat{a}_{t-1} - \frac{\kappa(1-\rho_a)}{(1+\phi_x-\rho_a)(1-\beta\rho_a)+\kappa(\phi_\pi-\rho_a)}\varepsilon_{a,t} \tag{162}$$

D.7.2 Diagnostic Expectations

Guess the solution takes the following form:

$$\hat{y}_t = \alpha_1\hat{a}_{t-1} + \gamma_1\varepsilon_{a,t}; \quad \hat{\pi}_t = \alpha_2\hat{a}_{t-1} + \gamma_2\varepsilon_{a,t} \tag{163}$$

Using method of undetermined coefficients, we can solve for the coefficients. The coefficients α_1 and α_2 are identical under DE and RE.

$$\alpha_1 = \rho_a \frac{\phi_x(1-\beta\rho_a)+\kappa(\phi_\pi-\rho_a)}{(1+\phi_x-\rho_a)(1-\beta\rho_a)+\kappa(\phi_\pi-\rho_a)} \tag{164}$$

$$\alpha_2 = \frac{-\rho_a\kappa(1-\rho_a)}{(1+\phi_x-\rho_a)(1-\beta\rho_a)+\kappa(\phi_\pi-\rho_a)} \tag{165}$$

Coefficients γ_1 and γ_2 depend on the DE parameter.

$$\gamma_1 = \frac{(1+\theta)\alpha_1 + (1+\theta)\alpha_2\left[1-\beta(\phi_\pi-\theta)\right]+\kappa(\phi_\pi-\theta)}{1+\phi_x+\kappa(\phi_\pi-\theta)}; \quad \gamma_2 = \beta(1+\theta)\alpha_2 + \kappa(\gamma_1-1) \tag{166}$$

D.8 Proof of Propositions 4 and 5

Because there are no government shocks, $\hat{c}_t = \hat{y}_t$. The equilibrium with completely rigid prices, i.e. $\psi_p \to \infty$, given by:

$$\hat{y}_t = \mathbb{E}_t^\theta\left[\hat{y}_{t+1}\right] - \hat{i}_t \tag{167}$$

$$\hat{i}_t = \phi_x(\hat{y}_t - \hat{a}_t) \tag{168}$$

where $\hat{a}_t = \rho_a \hat{a}_{t-1} + \varepsilon_{a,t}$, $\rho_a \in [0,1)$, and $\epsilon_{a,t} \sim iid\ N(0, \sigma_a^2)$. Substituting the policy rule into the Euler equation, we get:

$$\hat{y}_t = \frac{1}{1+\phi_x}\mathbb{E}_t^\theta\left[\hat{y}_{t+1}\right] + \frac{\phi_x}{1+\phi_x}\hat{a}_t \tag{169}$$

By forward iteration, and using the law of iterated expectations under the no-news assumption,

$$\hat{y}_t = \lim_{T\to\infty}\frac{\mathbb{E}_t^\theta\left[\hat{y}_{T+1}\right]}{(1+\phi_x)^{T+1}} + \sum_{i=1}^{\infty}\frac{\phi_x \mathbb{E}_t^\theta\left[\hat{a}_{t+i}\right]}{(1+\phi_x)^{i+1}} + \frac{\phi_x}{1+\phi_x}\hat{a}_t \tag{170}$$

The system is locally determinate if and only if $\phi_x > 0$. Let $\phi_x > 0$. Then,

$$\hat{y}_t = \sum_{i=1}^{\infty}\frac{\phi_x \mathbb{E}_t^\theta\left[\hat{a}_{t+i}\right]}{(1+\phi_x)^{i+1}} + \frac{\phi_x}{1+\phi_x}\hat{a}_t \tag{171}$$

From the definition of the shock process, we know that, $\forall\ i > 0$

$$\mathbb{E}_t^\theta\left[\hat{a}_{t+i}\right] = \rho_a^i(1+\theta)\hat{a}_t - \theta\rho_a^{i+1}\hat{a}_{t-1} = \rho_a^i\left((1+\theta)\hat{a}_t - \theta\rho_a\hat{a}_{t-1}\right) \tag{172}$$

We can then derive the solution for output:

$$\hat{y}_t = \frac{\phi_x\rho_a(1+\theta) + \phi_x(1+\phi_x-\rho_a)}{(1+\phi_x)(1+\phi_x-\rho_a)}\hat{a}_t - \frac{\phi_x\theta\rho_a^2}{(1+\phi_x)(1+\phi_x-\rho_a)}\hat{a}_{t-1} \tag{173}$$

This solution can be rewritten as:

$$\hat{y}_t = \rho\frac{\phi_x(1+\phi_x)}{(1+\phi_x)(1+\phi_x-\rho_A)}\hat{a}_{t-1} + \frac{\phi_x\rho_A\theta + \phi_x(1+\phi_x)}{(1+\phi_x)(1+\phi_x-\rho_A)}\varepsilon_{a,t} \tag{174}$$

Volatility of output is then

$$Var(\hat{y}_t) = \left(\rho\frac{\phi_x(1+\phi_x)}{(1+\phi_x)(1+\phi_x-\rho_A)}\right)^2 Var(\hat{a}_{t-1}) + \left(\frac{\phi_x\rho_A\theta + \phi_x(1+\phi_x)}{(1+\phi_x)(1+\phi_x-\rho_A)}\right)^2\sigma^2 \tag{175}$$

The first coefficient is same under rational and diagnostic expectations. Volatility is higher under diagnostic expectations relative to rational expectations if and only if

$$(\phi_x\rho_A\theta + \phi_x(1+\phi_x))^2 > \phi_x^2(1+\phi_x)^2 \tag{176}$$

$$\iff \theta > 0 \tag{177}$$

In the flexible price limit, $\kappa \to \infty$, output under DE and RE follows (from D.7):

$$\hat{y}_t = \rho_a \hat{a}_{t-1} + \varepsilon_{a,t} \equiv \hat{a}_t \tag{178}$$

Hence, DE and RE have the same output variability when $\kappa \to \infty$ (or $\psi_p \to 0$).

This completes the proof for Proposition 4.

The solution for output gap $\hat{x}_t \equiv \hat{y}_t - \hat{a}_t$ is given by:

$$\hat{x}_t = \frac{-\rho_a(1-\rho_a)(1+\phi_x)}{(1+\phi_x)(1+\phi_x-\rho_a)}\hat{a}_{t-1} + \frac{\theta\phi_x\rho_a - (1-\rho_a)(1+\phi_x)}{(1+\phi_x)(1+\phi_x-\rho_a)}\varepsilon_{a,t} \tag{179}$$

In response to an unanticipated improvement in productivity, output gap can be positive on impact if and only

$$\theta\phi_x\rho_a - (1-\rho_a)(1+\phi_x) > 0 \tag{180}$$

When $\theta = 0$, that is rational expectations, output gap negatively co-moves with productivity shock. Under diagnostic expectations, productivity improvements can be expansionary on impact. This completes the proof for Proposition 5.

D.9 Proof of Proposition 6

Note that after a one-time unanticipated shock, the solution under DE and RE coincide at subsequent dates since there is no news. (This was shown formally in the context of the general linear model in the previous appendix, proof of Proposition 3.) At date $t = 1$, we can derive the solution under DE as follows. From the RE solution, we know the expectations of forward looking variables :

$$\mathbb{E}_1\hat{y}_2 = \rho_g \frac{(1-\beta\rho_g)(1-\rho_g) + \kappa\psi(\phi_\pi - \rho_g)}{(1-\beta\rho_g)(1-\rho_g+\phi_x) + \kappa(\phi_\pi - \rho)}\varepsilon_{g,t}; \tag{181}$$

$$\mathbb{E}_1\hat{\pi}_2 = \rho_g \frac{\kappa(1-\psi)(1-\rho_g) - \kappa\psi\phi_x}{(1-\beta\rho_g)(1-\rho_g+\phi_x) + \kappa(\phi_\pi - \rho_g)}\varepsilon_{g,t} \tag{182}$$

$$\mathbb{E}_0\hat{y}_2 = \mathbb{E}_0\hat{\pi}_2 = \mathbb{E}_0\hat{\pi}_1 = 0 \tag{183}$$

We can thus construct the diagnostic expectation terms that enter the DE model, and simplify the model to

$$\hat{y}_1 = (1+\theta)\mathbb{E}_1\left[\hat{y}_2 + \hat{\pi}_2 - \hat{g}_2\right] - \hat{i}_1 + \theta\hat{\pi}_1 + \hat{g}_1 \tag{184}$$

$$\hat{\pi}_1 = \beta(1+\theta)\mathbb{E}_1\left[\hat{\pi}_2\right] + \kappa\hat{y}_1 - \kappa\psi\hat{g}_1 \tag{185}$$

$$\hat{i}_1 = \phi_\pi\hat{\pi}_1 + \phi_x\hat{y}_1 \tag{186}$$

Substituting the latter two equations into the Euler equation, and rearranging we get

$$\hat{y}_1 = \frac{(1+\theta)\mathbb{E}_1\left[\hat{y}_2 + (1+\beta\theta-\beta\phi_\pi)\pi_2\right] + \left[1 + (\phi_\pi-\theta)\kappa\psi - (1+\theta)\rho_g\right]\varepsilon_{g,1}}{1+\phi_x+(\phi_\pi-\theta)\kappa} \tag{187}$$

The corresponding RE solution can be seen with $\theta = 0$.

We study three scenarios with analytical results:

1. When the shocks are iid ($\rho_g = 0$), the solution is:

$$\hat{y}_1 = \frac{1+(\phi_\pi-\theta)\kappa\psi}{1+\phi_x+(\phi_\pi-\theta)\kappa}\varepsilon_{g,1} \tag{188}$$

   For a bounded solution (and continuity with RE solution), we assume that $\theta < \phi_\pi + \kappa^{-1}(1+\phi_x)$. There are two cases for the fiscal multiplier:

   - $\underline{\phi_x < \nu}$: The fiscal multiplier under DE is larger than under RE. The multiplier is increasing in θ, exceeds one for values of $\theta > \phi_\pi + \frac{\phi_x}{(1-\psi)\kappa}$. As $\theta \to \phi_\pi + \kappa^{-1}(1+\phi_y)$, the fiscal multiplier $\to \infty$.
   - $\underline{\phi_x > \nu}$: The fiscal multiplier under DE is smaller than under RE.

2. When the government spending shocks are iid <u>and</u> $\phi_x = 0$, the solution for output under DE is :

$$\hat{y}_1 = \frac{1+(\phi_\pi-\theta)\kappa\psi}{1+(\phi_\pi-\theta)\kappa}\varepsilon_{g,1} \tag{189}$$

   For the solution to be continuous in the RE limit and bounded, we assume that $\theta < \phi_\pi + \kappa^{-1}$. Since $\psi = \frac{1}{1+\nu} < 1$, the fiscal multiplier is increasing in θ. Under the RE limit, $\theta = 0$, the fiscal multiplier is strictly less than one. For $\theta > \phi_\pi$, the multiplier is larger than one. Finally, the multiplier explodes to infinity as $\theta \to \phi_\pi + \kappa^{-1}$.

3. When prices are perfectly rigid, that is $\kappa \to 0$, the solution for output is given

by:

$$\hat{y}_1 = \frac{(1-\rho_g)(1+\phi_x) - \theta\rho_g\phi_x}{(1+\phi_x)(1-\rho_g+\phi_x)}\varepsilon_{g,1} \tag{190}$$

Fiscal multiplier under DE is smaller than under RE. Fiscal multiplier is decreasing in θ. For $\theta > \frac{(1-\rho_g)(1+\phi_x)}{\rho_g\phi_x}$, assuming $\rho_g > 0$, output falls under DE with increase in government spending.

D.10 Proof of Proposition 7

1. When $\psi_p \to \infty$, $\phi_x = 0$, and $\rho_\zeta = 0$, beliefs about the long-run (BLR) under the DKF are given by

$$\zeta^\theta_{t|t} \equiv \zeta_{t|t} + \theta(\zeta_{t|t} - \zeta_{t|t-1}) \tag{191}$$

where $\zeta_{t|t} \equiv \mathbb{E}_t[\zeta_t]$ and $\zeta_{t|t-1} \equiv \mathbb{E}_{t-1}[\zeta_t]$. Also, BLR under the RKF are given by $\zeta_{t|t}$. From the rational Kalman filter, we have

$$\zeta_{t|t} = \zeta_{t|t-1} + Gain_t(s_t - s_{t|t-1}) \tag{192}$$

where $s_{t|t-1} \equiv \mathbb{E}_{t-1}[s_t]$ and $Gain_t$ is the Kalman gain. Thus, BLR under the DKF is simplified to

$$\zeta^\theta_{t|t} \equiv \zeta_{t|t} + \theta \times Gain_t(s_t - s_{t|t-1}) \tag{193}$$

and as $s_t - s_{t|t-1} > 0$ with a positive shock to ζ_t, BLR are greater under the DKF than under the RKF.

2. We can also rewrite (193) as

$$\zeta^\theta_{t|t} = \zeta_{t-1} + (1+\theta) \times Gain_t(s_t - s_{t|t-1}) \tag{194}$$

given that $\zeta_{t|t-1} = \zeta_{t-1}$. As BLR under FIRE are simply $\zeta_t = \zeta_{t-1} + \epsilon_{\zeta,t}$, BLR under the DKF are greater than under FIRE if $(1+\theta) \times Gain_t(s_t - s_{t|t-1}) > \epsilon_{\zeta,t}$ where $\epsilon_{\zeta,t}$ is a shock to ζ_t. Thus, if

$$\theta \geq \frac{\epsilon_{\zeta,t}}{Gain_t(s_t - s_{t|t-1})} - 1 \tag{195}$$

beliefs about the long-run under the DKF are greater than under FIRE.

E Real Business Cycle Model

We list the equilibrium conditions for a standard RBC model. Equilibrium is given by a sequence of seven unknowns $\{C_t, K_{t+1}, Y_t, I_t, N_t, R_t^k, \tilde{W}_t.\}$ that satisfy the following seven equations for a given exogenous process A_t and an initial value of capital stock K_0.

$$\frac{1}{C_t} = \beta \mathbb{E}_t^\theta \left[\frac{R_{t+1}^k + 1 - \delta}{C_{t+1}} \right] \tag{196}$$

$$\tilde{W}_t = \omega C_t N_t^\nu \tag{197}$$

$$K_{t+1} = (1-\delta)K_t + I_t \tag{198}$$

$$Y_t = K_t^\alpha \left(A_t N_t\right)^{1-\alpha} \tag{199}$$

$$Y_t = C_t + I_t \tag{200}$$

$$R_t^k = \alpha \frac{Y_t}{K_t} \tag{201}$$

$$\tilde{W}_t = (1-\alpha)\frac{Y_t}{N_t} \tag{202}$$

β is the discount rate, δ is depreciation rate, ν is inverse of the Frisch elasticity of labor supply, α is the capital share, and ω is a normalizing constant in the steady state. $\theta > 0$ is the diagnosticity parameter. The system of log-linearized equations is as follows (where the lower case letters denote the log-deviations form the respective steady state values)[53]:

$$\tilde{w}_t = c_t + \nu n_t \tag{203}$$

$$c_t = \mathbb{E}_t^\theta \left[c_{t+1} - \frac{R^k}{R^k + 1 - \delta} r_{t+1}^k \right] \tag{204}$$

$$k_{t+1} = \delta \hat{I}_t + (1-\delta)k_t \tag{205}$$

$$y_t = (1-\alpha)a_t + \alpha k_t + (1-\alpha)n_t \tag{206}$$

$$y_t = s_c c_t + (1-s_c)\hat{I}_t \tag{207}$$

$$r_t^k = y_t - k_t \tag{208}$$

$$\tilde{w}_t = y_t - n_t \tag{209}$$

where R^k is the steady state rental rate, and s_c is the steady state share of consumption in output. The economy starts in the steady state. There is a one-time unanticipated iid shock a_1 at time 1.

[53] $\hat{I}_t$ is also log-deviations of investment I_t from its steady state value.

E.1 Rational Expectations and Full Depreciation, $\delta = 1$

We derive analytical result assuming full depreciation, that is $\delta = 1$. The Euler equation under rational expectations and full depreciation is given by:

$$c_t - k_{t+1} = \mathbb{E}_t \left[c_{t+1} - y_{t+1} \right] \tag{210}$$

From the labor supply and labor demand conditions, we obtain

$$(1 + \nu) n_t = y_t - c_t \tag{211}$$

When $\delta = 1$, $\hat{I}_t = k_{t+1}$. Use the above equation into the Euler equation, along with investment equation to get

$$\hat{I}_t - y_t + (1 + \nu) n_t = (1 + \nu) \mathbb{E}_t [n_{t+1}] \tag{212}$$

Substitute in the resource constraint,

$$\frac{1}{1 - s_c} \left[y_t - s_c c_t \right] - y_t + (1 + \nu) n_t = (1 + \nu) \mathbb{E}_t [n_{t+1}] \tag{213}$$

$$\iff \frac{s_c}{1 - s_c} \left[y_t - c_t \right] + (1 + \nu) n_t = (1 + \nu) \mathbb{E}_t [n_{t+1}] \tag{214}$$

$$\iff \left(1 + \frac{s_c}{1 - s_c} \right) n_t = \mathbb{E}_t [n_{t+1}] \tag{215}$$

Solution for employment is

$$n_t = 0, \quad \forall t \geq 0 \tag{216}$$

We can solve for the solution for other variables at dates 1 and 2:

$$c_1 = y_1 = \hat{I}_1 = k_2 = (1 - \alpha) a_1; \tag{217}$$

$$c_2 = y_2 = \hat{I}_2 = k_3 = \alpha (1 - \alpha) a_1 \tag{218}$$

and so on.

E.2 Diagnostic Expectations and full depreciation, $\delta = 1$

The Euler equation is

$$c_t = \mathbb{E}_t^{\theta} \left[c_{t+1} - y_{t+1} + k_{t+1} \right] \tag{219}$$

As before, the economy starts in the steady state. There is a one-time unanticipated iid shock a_1 at time 1. From Date 2, the solution is same as rational expectations model. Since, we have iid shocks, the solution at date 1 is:

$$c_1 = (1+\theta)k_2 \tag{220}$$

Substitute into the resource constraint to get

$$y_1 = (1+\theta s_c)k_2 \tag{221}$$

From labor supply and labor demand,

$$(1+\nu)n_1 = y_1 - c_1 = -\theta(1-s_c)k_2 \tag{222}$$

Finally, from the production function

$$y_1 = (1-\alpha)a_1 + (1-\alpha)n_1 \tag{223}$$

$$\iff (1+\theta s_c)k_2 = (1-\alpha)a_1 + (1-\alpha)n_1 \tag{224}$$

$$\iff -\frac{(1+\theta s_c)}{\theta(1-s_c)}(1+\nu)n_1 = (1-\alpha)a_1 + (1-\alpha)n_1 \tag{225}$$

$$n_1 = -\frac{\theta(1-s_c)(1-\alpha)a_1}{(1+\theta s_c)(1+\nu)+(1-\alpha)\theta(1-s_c)} \tag{226}$$

Solution is

$$n_1 = -\frac{\theta(1-s_c)(1-\alpha)a_1}{(1+\theta s_c)(1+\nu)+(1-\alpha)\theta(1-s_c)} \tag{227}$$

$$k_2 = \frac{(1-\alpha)(1+\nu)a_1}{(1+\theta s_c)(1+\nu)+(1-\alpha)\theta(1-s_c)} \tag{228}$$

$$c_1 = \frac{(1+\theta)(1-\alpha)(1+\nu)a_1}{(1+\theta s_c)(1+\nu)+(1-\alpha)\theta(1-s_c)} \tag{229}$$

$$y_1 = \frac{(1+\theta s_c)(1-\alpha)(1+\nu)a_1}{(1+\theta s_c)(1+\nu)+(1-\alpha)\theta(1-s_c)} \tag{230}$$

Date 2 solution is :

$$n_2 = 0 \tag{231}$$

$$y_2 = \alpha k_2 = \frac{\alpha(1-\alpha)(1+\nu)a_1}{(1+\theta s_c)(1+\nu)+(1-\alpha)\theta(1-s_c)} \tag{232}$$

E.3 Analytical Proposition for RBC model

Proposition 10 (Extra Volatility: RBC Model) *Consider the model given by* (35), (203)-(209). *Assume that the depreciation rate* $\delta = 1$ *and that* $\rho_a = 0$. *Output is less volatile under DE than under RE:* $Var(\hat{y}_t)_{DE} < Var(\hat{y}_t)_{RE}$.

Volatility of output at date 1 is lower under DE compared to RE if and only if

$$\frac{(1+\theta s_c)(1+\nu)}{(1+\theta s_c)(1+\nu)+(1-\alpha)\theta(1-s_c)} < 1 \tag{233}$$

which is true. Further, note that volatility of output at date 1 under DE is decreasing in ν. Similarly, we can show that volatility of output under diagnostic expectations is lower at all future horizons as well. For example, Volatility of output at date 2 is lower under DE compared to RE if and only if

$$\frac{(1+\nu)}{(1+\theta s_c)(1+\nu)+(1-\alpha)\theta(1-s_c)} < 1 \tag{234}$$

which is true since $1+\theta s_c > 1$ and $(1-\alpha)\theta(1-s_c) > 0$.

E.4 Numerical Results on Extra Volatility: NK and RBC Models

To numerically demonstrate the excess volatility in the NK model, we use the calibration discussed in Table 4, which is our standard calibration. Stationary TFP follows an AR(1) process with persistence 0.9 and standard deviation 0.02. We set the discount factor β to 0.99. For the RBC model, we set the capital share α to 0.2 and the capital depreciation rate δ to 0.025. For the NK model, we set $\phi_\pi = 1.5$, $\phi_x = 0.5$, and $\kappa = 0.05$. We also set the diagnosticity parameter θ to one.

Panel a) in Table 5 shows unconditional volatilities of output growth, and consumption growth under diagnostic and rational expectations. Since there is no government spending or investment, output growth and consumption growth are equivalent in the NK model. We find that the output gap under diagnostic expectations exhibits 63 percent higher standard deviation relative to the output gap under rational expectations.

Panel b) in Table 5 shows unconditional volatilities of output growth, consumption growth, and investment growth, both under diagnostic and rational expectations in the baseline RBC model. Consumption growth is twice as volatile under diagnostic expectations than under rational expectations. On the other hand, investment growth and output growth are dampened under diagnostic expectations due to the general

Table 4: Calibration: The NK and RBC models

	Parameter	Value
Common to Both Models		
θ	Diagnosticity	1
β	Discount factor	0.99
NK model		
ν	Inv. Frisch elasticity	2
ϕ_π	Taylor rule inflation	1.5
ϕ_x	Taylor rule output gap	0.5
κ	Slope of the Phillips curve	0.05
RBC model		
α	Capital share	0.2
δ	Capital depreciation rate	0.025
Shock Process		
ρ_a	Shock persistence (stationary TFP)	0.9
σ_a	Standard dev. (stationary TFP)	0.02

Table 5: Model-Implied Volatilities with Stationary TFP Shocks

(a) New Keynesian Model

Variable	Rational Expectations	Diagnostic Expectations	Percentage Increase
Output	0.0182	0.0296	63%
Consumption	0.0182	0.0296	63%
Investment	–	–	–

(b) Real Business Cycle Model

Variable	Rational Expectations	Diagnostic Expectations	Percentage Increase
Output	0.0204	0.0188	-8%
Consumption	0.0052	0.0103	98%
Investment	0.1147	0.0816	-29%

Notes: The table reports the standard deviations of output growth, consumption growth and investment growth in the New Keynesian (NK) model and the RBC model in Panels (a) and (b) respectively. Final column titled "Percentage Increase" shows the percentage increase in standard deviation under the diagnostic expectations model relative to the rational expectations benchmark. There is one shock process in the two models. See Table 4 for the parameters.

equilibrium adjustment of the interest rate. Diagnosticity, therefore, does not always generate extra amplification.

F A Medium-Scale DSGE model

F.1 Model Ingredients

The model follows the exposition in BLL. The economy comprises of following agents: a continuum of households supplying differentiated labor, a continuum of firms producing differentiated goods, a perfectly competitive final goods firm, a perfectly competitive labor agency that provides the composite labor input demanded by firms, and a government in charge of fiscal and monetary policy.

F.1.1 Monopolistically Competitive Producers

Assume there is a continuum of differentiated intermediate good producers that sell the intermediate good Y_{jt}. A perfectly competitive firm aggregates intermediate goods into a final composite good $Y_t = \left[\int_0^1 Y_{jt}^{\frac{\epsilon_{p,t}-1}{\epsilon_{p,t}}} dj\right]^{\frac{\epsilon_{p,t}}{\epsilon_{p,t}-1}}$, where $\epsilon_p > 1$ is time-varying elasticity of demand. The iso-elastic demand for intermediate good j is given by: $Y_{jt} = \left(\frac{P_{jt}}{P_t}\right)^{-\epsilon_{p,t}} Y_t$, where P_t is the aggregate price index and P_{jt} is the price of intermediate goods j. Each intermediate good j is produced by a price-setting monopolistically competitive firm using labor L_{jt} and physical capital K_{jt}:

$$Y_{jt} = (A_t L_{jt})^{1-\alpha} \;\; K_{jt}^{\alpha} \tag{235}$$

where the TFP process A_t is the sum of two components (in logs):

$$\log A_t = \log Z_t + \log \Xi_t \tag{236}$$

The variable Z_t denotes a non-stationary TFP series that evolves according to:

$$\frac{Z_t}{Z_{t-1}} = \left(\frac{Z_{t-1}}{Z_{t-2}}\right)^{\rho_\zeta} G_\zeta^{1-\rho_\zeta} \exp(\varepsilon_{\zeta,t}); \quad \varepsilon_{\zeta,t} \sim iid\ N(0, \sigma_\zeta^2) \tag{237}$$

where ρ_ζ is the persistence of the shock process, and $\varepsilon_{\zeta,t}$ is a random disturbance that causes deviations of the TFP growth from its balanced growth rate G_ζ. The stationary TFP evolves as follows:

$$\log \Xi_t = \rho_\xi \log \Xi_{t-1} + \varepsilon_{\xi,t}; \quad \varepsilon_{\xi,t} \sim iid\ N(0, \sigma_\xi^2) \tag{238}$$

where ρ_ξ is the persistence of the shock process, and $\varepsilon_{\xi,t}$ is an i.i.d shock with variance σ_ξ^2. (We define $a_t \equiv \log A_t$, $\zeta_t \equiv \log Z_t$, $\xi_t \equiv \log \Xi_t$, $G_{a,t} \equiv A_t/A_{t-1}$, and $G_{\zeta,t} \equiv$

Z_t/Z_{t-1}.)

Following BLL, we assume that

$$\rho_\zeta = \rho_\xi \equiv \rho \tag{239}$$

and that the variances satisfy the following restriction[54]

$$\rho\sigma_\zeta^2 = (1-\rho)^2\sigma_\xi^2 \tag{240}$$

While agents observe the TFP process as a whole, they do not observe two components ζ_t and ξ_t separately. Considering the idea that agents have more information than merely about productivity, agents observe a noisy signal s_t about the permanent component of TFP:

$$s_t = \zeta_t + \varepsilon_{s,t}; \quad \varepsilon_{s,t} \sim iid\ N(0, \sigma_s^2) \tag{241}$$

where $\varepsilon_{s,t}$ is an i.i.d. normal shock, which affects agents' beliefs but is independent of fundamentals. This noisy signal relates to the additional informative signal that agents receive which is a straightforward interpretation of Equation (241). Ultimately, the presence of this noisy information helps the econometrician make inferences about the (unobserved) long-term productivity trend by looking at the behavior of consumption.

Firms choose inputs to minimize total cost each period. Marginal cost, independent of firm-specific variables, is given by $mc_t = \frac{1}{A_t^{1-\alpha}}\left(\frac{R_t^k/P_t}{\alpha}\right)^\alpha\left(\frac{W_t/P_t}{1-\alpha}\right)^{1-\alpha}$, where $\frac{R_t^k}{P_t}$ and $\frac{W_t}{P_t}$ denote aggregate rental rate of capital and real wage. A firm j pays a quadratic adjustment cost in units of final good (Rotemberg 1982) to adjust its price P_{jt}. The cost is given by $\frac{\psi_p}{2}\left(\frac{P_{jt}}{\tilde{\Pi}_{t-1}P_{jt-1}} - 1\right)^2 P_tY_t$, where $\psi_p \geq 0$ regulates the adjustment costs. Price change is indexed to $\tilde{\Pi}_{t-1} = \bar{\Pi}^{1-\iota_p}\Pi_{t-1}^{\iota_p}$, where ι_p governs indexation between previous period inflation rate Π_{t-1} and steady state inflation rate $\bar{\Pi}$. Firm's per period profits are given by: $D_{jt} \equiv P_{jt}Y_{jt} - P_tmc_tY_{jt} - \frac{\psi_p}{2}\left(\frac{P_{jt}}{\tilde{\Pi}_{t-1}P_{jt-1}} - 1\right)^2 P_tY_t$. Each period, the firm chooses P_{jt} to maximize present discounted value of real profits:

$$\max_{P_{jt}}\left\{\frac{\Lambda_tD_{jt}}{P_t} + \mathbb{E}_t^\theta\left[\sum_{s=1}^{\infty}\frac{\Lambda_{t+s}D_{jt+s}}{P_{t+s}}\right]\right\} \tag{242}$$

where Λ_t is the marginal utility of consumption in period t, and $\mathbb{E}_t^\theta[\ \cdot\ ]$ is the diagnostic expectation operator regulated by parameter θ. Notice that we write dynamic

[54] As shown in BLL, these restrictions imply that the univariate process for a_t is a random walk with variance σ_a^2.

maximization problems by explicitly separating time t choice variables from the expectation of future choice variables. This separation is crucial for solving the model with diagnostic expectations, and is a consequence of the technical issues discussed in Section 3.

F.1.2 Households

There is a continuum of monopolistically competitive households, indexed by $i \in [0,1]$, supplying a differentiated labor input $L_{i,t}$. A perfectly competitive employment agency aggregates various labor types into a composite labor input L_t supplied to firms, in a Dixit-Stiglitz aggregator: $L_t = \left[\int_0^1 L_{i,t}^{\frac{\epsilon_{w,t}-1}{\epsilon_{w,t}}} di\right]^{\frac{\epsilon_{w,t}}{\epsilon_{w,t}-1}}$, where $\epsilon_{w,t} > 1$ is time-varying elasticity of demand. The iso-elastic demand for labor input i is given by: $L_{i,t} = \left(\frac{W_{i,t}}{W_t}\right)^{-\epsilon_{w,t}} L_t$, where $W_{i,t}$ is household i's wage rate, and W_t is the aggregate wage rate that the household takes as given.

The household i has following lifetime-utility at time t:

$$\left(\log(C_{i,t} - h\tilde{C}_{t-1}) - \frac{\omega}{1+\nu}L_{i,t}^{1+\nu} - \psi_{i,t}^{w}\right) + \mathbb{E}_t^{\theta}\left[\Sigma_{s=t+1}^{\infty}\beta^{s-t}\left(\log(C_{i,s} - h\tilde{C}_{s-1}) - \frac{\omega}{1+\nu}L_{i,s}^{1+\nu} - \psi_{i,s}^{w}\right)\right] \tag{243}$$

where h is the degree of habit formation on external habits over aggregate consumption $\tilde{C}_{t-1}$, which the household takes as given, $\nu > 0$ is inverse of the Frisch elasticity of labor supply, $\omega > 0$ is a parameter that pins down the steady-state level of hours, and the discount factor β satisfies $0 < \beta < 1$. $\psi_{i,t}^{w}$ is the loss in utility in adjusting wages. We assume a quadratic adjustment cost given by $\psi_{i,t}^{w} = \frac{\psi_w}{2}\left[\frac{W_{it}}{\tilde{\Pi}_{t-1}^{w} W_{it-1}} - 1\right]^2$, where $\psi_w \geq 0$ is a parameter, and wage contracts are indexed to productivity and price inflation. We assume $\tilde{\Pi}_{t-1}^{w} = G_a \bar{\Pi}^{1-\iota_w}\left(\exp(\varepsilon_{\zeta,t}\varepsilon_{\xi,t})\Pi_{t-1}\right)^{\iota_w}$ with $0 \leq \iota_w < 1$.

The household's budget constraint in period t is given by

$$P_t C_{i,t} + P_t I_{i,t} + \frac{B_{i,t+1}}{1+i_t} = B_{i,t} + W_{i,t}L_{i,t} + D_t + T_t + R_t^K u_{i,t}K_{i,t}^u - P_t a(u_{i,t})K_{i,t}^u \tag{244}$$

where $I_{i,t}$ is investment, $W_{i,t}L_{i,t}$ is labor income, and $B_{i,t}$ is income from nominal bonds paying nominal interest rate i_t. Households own an equal share of all firms, and thus receive D_t dividends from profits. Finally, each household receives a lump-sum government transfer T_t.

The households own capital, $K_{i,t}^u$, and choose the utilization rate, $u_{i,t}$. The amount of effective capital, $K_{i,t}$, that the households rent to the firms at nominal rate R_t^K is given by $K_{i,t} = u_{i,t} K_{i,t}^u$. The (nominal) cost of capital utilization is $P_t \chi(u_{i,t})$ per unit of physical capital. As in the literature, we assume $\chi(1) = 0$ in the steady state and $\chi'' > 0$. Following GHLS, we assume investment adjustment costs, $S\left(\frac{I_{i,t}}{G_a I_{i,t-1}}\right)$, in the production of capital, where G_a is the steady state growth rate of A_t. Law of motion for capital is as follows:

$$K_{i,t+1}^u = \mu_t \left[1 - S\left(\frac{I_{i,t}}{G_a I_{i,t-1}}\right)\right] I_{i,t} + (1 - \delta_k) K_{i,t}^u \tag{245}$$

where δ_k denotes depreciation rate, and μ_t is an exogenous disturbance to the marginal efficiency of investment that follows:

$$\log \mu_t = \rho_\mu \log \mu_{t-1} + \varepsilon_{\mu,t}; \quad \varepsilon_{\mu,t} \sim iid\ N(0, \sigma_\mu^2) \tag{246}$$

As in the literature, we assume that $S(1) = S'(1) = 0$, and calibrate $S''(1) > 0$.

F.1.3 Government

The central bank follows a Taylor rule in setting the nominal interest rate i_t. It responds to deviations in (gross) inflation rate Π_t from its target rate $\bar{\Pi}$ and output.

$$\frac{1 + i_t}{1 + i_{ss}} = \left(\frac{1 + i_{t-1}}{1 + i_{ss}}\right)^{\rho_R} \left[\left(\frac{\Pi_t}{\bar{\Pi}}\right)^{\phi_\pi} Y_t^{\phi_y}\right]^{1-\rho_R} \exp(\lambda_t^{mp}) \tag{247}$$

with $0 < \rho_R < 1$, $\phi_\pi \geq 0$, and $\phi_y \geq 0$. i_{ss} is the steady state nominal interest rate, and λ_t^{mp} follows the process

$$\log \lambda_t^{mp} = \rho_{mp} \log \lambda_{t-1}^{mp} + \varepsilon_{mp,t}; \quad \varepsilon_{mp,t} \sim N(0, \sigma_{mp}^2) \tag{248}$$

We assume government balances budget every period $P_t T_t = P_t G_t$, where G_t is the government spending. G_t is determined exogenously as as a fraction of GDP: $G_t = \left(1 - \frac{1}{\lambda_t^g}\right) Y_t$ where the government spending shock follows the process:

$$\log \lambda_t^g = (1 - \rho_g) \log \lambda^g + \rho_g \log \lambda_{t-1}^g + \varepsilon_{g,t}; \quad \varepsilon_{g,t} \sim N(0, \sigma_g^2) \tag{249}$$

λ^g is the steady state share of government spending in final output.

F.1.4 Market Clearing

We focus on a symmetric equilibrium where all intermediate goods producing firms and households make the same decisions. Therefore, we can drop subscripts i and j. The aggregate production function, in the symmetric equilibrium, is then given by: $Y_t = (A_t L_t)^{1-\alpha} K_t^{\alpha}$, since $K_t = K_{i,t} = K_{jt}$ and $N_t = N_{i,t} = N_{jt}$. The market clearing for the final good, in the symmetric equilibrium, requires that

$$Y_t = C_t + I_t + \chi(u_t) K_t^u + G_t + \frac{\psi_p}{2}\left[\frac{\Pi_t}{\tilde{\Pi}_{t-1}} - 1\right]^2 Y_t \tag{250}$$

This completes the presentation of the DSGE model.

F.2 Stationary Allocation

We normalize the following variables :

$$y_t = Y_t/A_t \tag{251}$$
$$c_t = C_t/A_t \tag{252}$$
$$k_t = K_t/A_t \tag{253}$$
$$k_t^u = K_t^u/A_{t-1} \tag{254}$$
$$\mathbb{I}_t = I_t/A_t \tag{255}$$
$$w_t = W_t/(A_t P_t) \tag{256}$$
$$r_t^k = R_t^k/P_t \tag{257}$$
$$\lambda_t = \Lambda_t A_t \tag{258}$$

Definition 1 (Normalized Equilibrium) *18 endogenous variables* $\{\lambda_t, i_t, c_t, y_t, \Pi_t, mc_t, \tilde{\Pi}_{t-1}, \Pi_t^w, \tilde{\Pi}_{t-1}^w, w_t, L_t, k_{t+1}^u, r_t^K, \mathbb{I}_t, q_t, u_t, k_t, G_{a,t}\}$*, 8 endogenous shock processes* $\{G_{\zeta,t}, \Xi_t, s_t, \mu_t, \lambda_t^p, \lambda_t^w, \lambda_t^{mp}, \lambda_t^g\}$*, 8 exogenous shocks* $\{\varepsilon_{\zeta,t}, \varepsilon_{\xi,t}, \varepsilon_{s,t}, \varepsilon_{\mu,t}, \varepsilon_{p,t}, \varepsilon_{w,t}, \varepsilon_{mp,t}, \varepsilon_{g,t}\}$ *given initial values of* k_{t-1}^u.

Consumption Euler Equation

$$\frac{\lambda_t}{G_{a,t}\Pi_t} = \beta(1+i_t)\mathbb{E}_t^{\theta}\left[\frac{\lambda_{t+1}}{G_{a,t}G_{a,t+1}}\frac{1}{\Pi_t \Pi_{t+1}}\right] \tag{259}$$

$$\lambda_t = \frac{1}{c_t - \frac{hc_{t-1}}{G_{a,t}}} \tag{260}$$

Price-setting

$$(1-\epsilon_{p,t})+\epsilon_{p,t}\ mc_t-\psi_p\left(\frac{\Pi_t}{\tilde{\Pi}_{t-1}}-1\right)\frac{\Pi_t}{\tilde{\Pi}_{t-1}}+\psi_p\frac{\beta\Pi_t}{\lambda_t y_t}\mathbb{E}_t^\theta\left[\lambda_{t+1}\left(\frac{\Pi_{t+1}}{\tilde{\Pi}_t}-1\right)\frac{\Pi_{t+1}}{\tilde{\Pi}_t}\frac{y_{t+1}}{\Pi_t}\right]=0 \tag{261}$$

$$\tilde{\Pi}_{t-1}=\bar{\Pi}^{1-\iota_p}\Pi_{t-1}^{\iota_p} \tag{262}$$

Wage-setting

$$\psi_w\left[\frac{\Pi_t^w}{\tilde{\Pi}_{t-1}^w}-1\right]\frac{\Pi_t^w}{\tilde{\Pi}_{t-1}^w}=\psi_w\beta\mathbb{E}_t^\theta\left[\frac{\Pi_{t+1}^w}{\tilde{\Pi}_t^w}-1\right]\frac{\Pi_{t+1}^w}{\tilde{\Pi}_t^w}+L_t\lambda_t\epsilon_{w,t}\left[\omega\frac{L_t^\nu}{\lambda_t}-\frac{\epsilon_{w,t}-1}{\epsilon_{w,t}}w_t\right] \tag{263}$$

$$\tilde{\Pi}_{t-1}^w=G_a\bar{\Pi}^{1-\iota_w}(\exp(\varepsilon_{\zeta,t})\exp(\varepsilon_{\xi,t})\Pi_{t-1})^{\iota_w} \tag{264}$$

$$\Pi_t^w=\frac{w_t}{w_{t-1}}\Pi_t G_{a,t} \tag{265}$$

Capital Investment

$$k_{t+1}^u=\mu_t\left[1-S\left(\frac{\mathbb{I}_t}{\mathbb{I}_{t-1}}\frac{G_{a,t}}{G_a}\right)\right]\mathbb{I}_t+(1-\delta_k)\frac{k_t^u}{G_{a,t}} \tag{266}$$

$$q_t=\frac{\beta G_{a,t}}{\lambda_t}\mathbb{E}_t^\theta\left[\frac{\lambda_{t+1}}{G_{a,t}G_{a,t+1}}\left(r_{t+1}^K u_{t+1}-\chi(u_{t+1})+q_{t+1}(1-\delta_k)\right)\right] \tag{267}$$

$$\begin{aligned}&q_t\mu_t\left[1-S\left(\frac{\mathbb{I}_t}{\mathbb{I}_{t-1}}\frac{G_{a,t}}{G_a}\right)-S'\left(\frac{\mathbb{I}_t}{\mathbb{I}_{t-1}}\frac{G_{a,t}}{G_a}\right)\frac{\mathbb{I}_t}{\mathbb{I}_{t-1}}\frac{G_{a,t}}{G_a}\right]\\ &+\frac{\beta G_{a,t}}{\lambda_t}\mathbb{E}_t^\theta\left[\mu_{t+1}\frac{\lambda_{t+1}}{G_{a,t}}q_{t+1}\frac{G_{a,t+1}}{G_a}\left(\frac{\mathbb{I}_{t+1}}{\mathbb{I}_t}\right)^2S'\left(\frac{\mathbb{I}_{t+1}}{\mathbb{I}_t}\frac{G_{a,t+1}}{G_a}\right)\right]=1\end{aligned} \tag{268}$$

Capital Utilization Rate

$$k_t=u_t\frac{k_t^u}{G_{a,t}} \tag{269}$$

$$r_t^K=\chi'(u_t) \tag{270}$$

Production Technologies

$$y_t=k_t^\alpha L_t^{1-\alpha} \tag{271}$$

$$\frac{k_t}{L_t}=\frac{w_t}{r_t^k}\frac{\alpha}{1-\alpha} \tag{272}$$

$$mc_t=\frac{(r_t^k)^\alpha w_t^{1-\alpha}}{\alpha^\alpha(1-\alpha)^{1-\alpha}} \tag{273}$$

Government

$$\frac{1+i_t}{1+i_{ss}} = \left(\frac{1+i_{t-1}}{1+i_{ss}}\right)^{\rho_R} \left[\left(\frac{\Pi_t}{\bar{\Pi}}\right)^{\phi_\pi} Y_t^{\phi_y}\right]^{1-\rho_R} \exp(\lambda_t^{mp}) \tag{274}$$

Market Clearing

$$y_t = c_t + \mathbb{I}_t + \chi(u_t)\frac{k_t^u}{G_{a,t}} + \left(1 - \frac{1}{\lambda_t^g}\right) y_t \tag{275}$$

TFP Growth Rate

$$\log G_{a,t} = \log G_{\zeta,t} + (\log \Xi_t - \log \Xi_{t-1}) \tag{276}$$

Law of Motion of Shocks

$$\log G_{\zeta,t} = (1-\rho)\log G_\zeta + \rho_\zeta \log G_{\zeta,t-1} + \varepsilon_{\zeta,t} \tag{277}$$

$$\log \Xi_t = \rho_\xi \log \Xi_{t-1} + \varepsilon_{\xi,t} \tag{278}$$

$$s_t = \log Z_t + \varepsilon_{s,t} \tag{279}$$

$$\log \mu_t = \rho_\mu \log(\mu_{t-1}) + \varepsilon_{\mu,t} \tag{280}$$

$$\log \lambda_t^{mp} = \rho_{mp} \log \lambda_{t-1}^{mp} + \varepsilon_{mp,t} \tag{281}$$

$$\log \lambda_t^g = \rho_g \log \lambda_{t-1}^g + \varepsilon_{g,t} \tag{282}$$

Disturbances

$$\text{TFP growth shock} \quad \varepsilon_{\zeta,t} \sim N(0, \sigma_\zeta^2) \tag{283}$$

$$\text{Stationary TFP shock} \quad \varepsilon_{\xi,t} \sim N(0, \sigma_\xi^2) \tag{284}$$

$$\text{Noise shock} \quad \varepsilon_{s,t} \sim N(0, \sigma_s^2) \tag{285}$$

$$\text{MEI shock} \quad \varepsilon_{\mu,t} \sim N(0, \sigma_\mu^2) \tag{286}$$

$$\text{Monetary policy shock} \quad \varepsilon_{mp,t} \sim N(0, \sigma_{mp}^2) \tag{287}$$

$$\text{Government spending shock} \quad \varepsilon_{g,t} \sim N(0, \sigma_g^2) \tag{288}$$

F.3 Steady State

$$1 = \beta \frac{1}{G_a} \frac{1+i}{\Pi} \tag{289}$$

$$\lambda = \frac{G_a}{c(G_a - h)} \tag{290}$$

$$mc = \frac{\epsilon_p}{\epsilon_p - 1} \tag{291}$$

$$\frac{\omega L^\nu}{\lambda} = \frac{\epsilon_w - 1}{\epsilon_w} w \tag{292}$$

$$\Pi^w = \Pi G_a \tag{293}$$

$$\Pi = \bar{\Pi} \tag{294}$$

$$q = 1 \tag{295}$$

$$u = 1 \tag{296}$$

$$(1 - \frac{1 - \delta_k}{G_a}) k^u = \mathbb{I} \tag{297}$$

$$1 = \beta \left[\frac{1}{G_a} \left(r^K + (1 - \delta_k) \right) \right] \tag{298}$$

$$k = \frac{k^u}{G_a} \tag{299}$$

$$r^K = \chi'(1) \tag{300}$$

$$y = k^\alpha L^{1-\alpha} \tag{301}$$

$$r^k = \frac{\epsilon_p}{\epsilon_p - 1} \alpha \frac{y}{k} \tag{302}$$

$$w = \frac{\epsilon_p}{\epsilon_p - 1} (1 - \alpha) \frac{y}{L} \tag{303}$$

$$y = c + \mathbb{I} + \left(1 - \frac{1}{\lambda^g} \right) y \tag{304}$$

$$S(1) = S'(1) = 0; S" > 0 \tag{305}$$

$$G_a = G_\zeta \tag{306}$$

F.4 Log-linearized Model

Consumption Euler Equation

$$\hat{\lambda}_t - \hat{G}_{a,t} - \pi_t = \hat{i}_t + \mathbb{E}_t^\theta \left[\hat{\lambda}_{t+1} - \hat{G}_{a,t} - \hat{G}_{a,t+1} - \pi_t - \pi_{t+1} \right] \tag{307}$$

$$\hat{\lambda}_t + \frac{G_a}{G_a - h} \hat{c}_t - \frac{h}{G_a - h} \left(\hat{c}_{t-1} - \hat{G}_{a,t} \right) = 0 \tag{308}$$

Price-setting

$$\pi_t = \beta \mathbb{E}_t^\theta \left[\pi_{t+1} - \iota_p \pi_t\right] + \iota_p \pi_{t-1} + \frac{\epsilon_p - 1}{\psi_p} \hat{mc}_t + \hat{\lambda}_t^{p,*} \tag{309}$$

where $\hat{\lambda}_t^{p,*}$ is the normalized price-markup shock process. Let the un-normalized process be denoted with $\hat{\lambda}_t^p$. Then $\hat{\lambda}_t^{p,*} = \frac{\epsilon_p - 1}{\psi_p} \hat{\lambda}_t^p$. In steady state $\lambda^p = \frac{\epsilon_p}{\epsilon_p - 1}$

Wage-setting

$$\pi_t^w = \beta \mathbb{E}_t^\theta \left[\pi_{t+1}^w - \iota_w \pi_t - \iota_w \hat{G}_{a,t+1}\right] + \iota_w \pi_{t-1} + \iota_w \hat{G}_{a,t} + \frac{\epsilon_w \omega L^{1+\nu}}{\psi_w} \left[\nu \hat{L}_t - \hat{w}_t - \hat{\lambda}_t\right] + \hat{\lambda}_t^{w,*} \tag{310}$$

where $\hat{\lambda}_t^{w,*}$ is the normalized wage-markup shock process. Let the un-normalized wage markup process be denoted with $\hat{\lambda}_t^w$. Then $\hat{\lambda}_t^{w,*} = \frac{\epsilon_w \omega L^{1+\nu}}{\psi_w} \hat{\lambda}_t^w$. In steady state $\lambda^w = \frac{\epsilon_w}{\epsilon_w - 1}$

$$\pi_t^w = \hat{w}_t - \hat{w}_{t-1} + \pi_t + \hat{G}_{a,t} \tag{311}$$

Capital Investment

$$\hat{k}_{t+1}^u = \frac{\mathbb{I}}{k^u} \left(\hat{I}_t + \hat{\mu}_t\right) + \frac{1 - \delta_k}{G_a} \left(\hat{k}_t^u - \hat{G}_{a,t}\right) \tag{312}$$

$$\hat{q}_t - \hat{G}_{a,t} + \hat{\lambda}_t = \mathbb{E}_t^\theta \left[\hat{\lambda}_{t+1} - \hat{G}_{a,t} - \hat{G}_{a,t+1} + \frac{r^K}{r^K + 1 - \delta_k} \hat{r}_{t+1}^K + \frac{1 - \delta_k}{r^K + 1 - \delta_k} \hat{q}_{t+1}\right] \tag{313}$$

$$\hat{q}_t + \hat{\mu}_t - S"(1) \left(\hat{I}_t - \hat{I}_{t-1} + \hat{G}_{a,t}\right) + \beta S"(1) \mathbb{E}_t^\theta \left(\hat{I}_{t+1} - \hat{I}_t + \hat{G}_{a,t+1}\right) = 0 \tag{314}$$

Capital Utilization Rate

$$\hat{k}_t = \hat{u}_t + \hat{k}_t^u - \hat{G}_{a,t} \tag{315}$$

$$\hat{r}_t^K = \frac{\chi"(1)}{\chi'(1)} \hat{u}_t \tag{316}$$

Production Technologies

$$\hat{y}_t = \alpha \hat{k}_t + (1 - \alpha) \hat{L}_t \tag{317}$$

$$\hat{r}_t^K = \hat{w}_t + \hat{L}_t - \hat{k}_t \tag{318}$$

$$\hat{mc}_t = \alpha \hat{r}_t^K + (1 - \alpha) \hat{w}_t \tag{319}$$

Government

$$\hat{i}_t = \rho_R \hat{i}_{t-1} + (1-\rho_R)\left(\phi_\pi \pi_t + \phi_y \hat{y}_t\right) + \varepsilon_{mp,t} \tag{320}$$

Market Clearing

$$\frac{1}{\lambda^g}\hat{y}_t = \frac{c}{y}\hat{c}_t + \frac{\mathbb{I}}{y}\hat{I}_t + \frac{\chi'(1)k}{y}\hat{u}_t + \frac{1}{\lambda^g}\hat{\lambda}^g_t \tag{321}$$

TFP Growth Rate

$$\hat{G}_{a,t} = \hat{G}_{\zeta,t} + \hat{\xi}_t - \hat{\xi}_{t-1} \tag{322}$$

$$\hat{a}_t = \hat{\zeta}_t + \hat{\xi}_t \tag{323}$$

where $\hat{a}_t$ and $\hat{\zeta}_t$ are defined as log deviations of A_t and Z_t from their initial values.

Law of Motion of Shocks

$$\hat{G}_{\zeta,t} = \rho_\zeta \hat{G}_{\zeta,t-1} + \varepsilon_{\zeta,t} \tag{324}$$

$$\hat{\xi}_t = \rho_\xi \hat{\xi}_{t-1} + \varepsilon_{\xi,t} \tag{325}$$

$$\hat{\mu}_t = \rho_\mu \hat{\mu}_{t-1} + \varepsilon_{\mu,t} \tag{326}$$

$$\hat{\lambda}^{mp}_t = \rho_{mp}\hat{\lambda}^{mp}_{t-1} + \varepsilon_{mp,t} \tag{327}$$

$$\hat{\lambda}^g_t = \rho_g \hat{\lambda}^g_{t-1} + \varepsilon_{g,t} \tag{328}$$

$$\hat{\lambda}^{p,*}_t = \rho_p \hat{\lambda}^{p,*}_{t-1} + \varepsilon_{p,t} - \phi_p \varepsilon_{p,t-1} \tag{329}$$

$$\hat{\lambda}^{w,*}_t = \rho_w \hat{\lambda}^{w,*}_{t-1} + \varepsilon_{w,t} - \phi_w \varepsilon_{w,t-1} \tag{330}$$

Disturbances

$$\text{TFP growth shock} \quad \varepsilon_{\zeta,t} \sim N(0,\sigma^2_\zeta) \tag{331}$$

$$\text{Stationary TFP shock} \quad \varepsilon_{\xi,t} \sim N(0,\sigma^2_\xi) \tag{332}$$

$$\text{Noise shock} \quad \varepsilon_{s,t} \sim N(0,\sigma^2_s) \tag{333}$$

$$\text{MEI shock} \quad \varepsilon_{\mu,t} \sim N(0,\sigma^2_\mu) \tag{334}$$

$$\text{Monetary policy shock} \quad \varepsilon_{mp,t} \sim N(0,\sigma^2_{mp}) \tag{335}$$

$$\text{Government spending shock} \quad \varepsilon_{g,t} \sim N(0,\sigma^2_g) \tag{336}$$

$$\text{Price markup shock} \quad \varepsilon_{p,t} \sim N(0,\sigma^2_p) \tag{337}$$

$$\text{Wage markup shock} \quad \varepsilon_{w,t} \sim N(0,\sigma^2_w) \tag{338}$$

F.5 Prior Distribution of the Parameters

The following parameters are fixed in the estimation procedure as shown in Table 6. The depreciation rate δ_k is fixed at 0.025, and the discount factor β is set to 0.99. The Dixit-Stiglitz aggregator for the goods (ϵ_p) and for labor services (ϵ_w) are fixed at 6. The parameter affecting the level of disutility from working (ω) is set to 1, and the steady-state share of government spending to final output is fixed at 1.2.

Table 6: Fixed Parameters

	Parameter	Value
β	Discount factor	0.99
δ_k	Capital depreciation rate	0.025
$1-\frac{1}{\lambda_g}$	Government spending share	0.20
ω	Labor preference	1
ϵ_p	Elasticity of goods demand	6
ϵ_w	Elasticity of labor demand	6

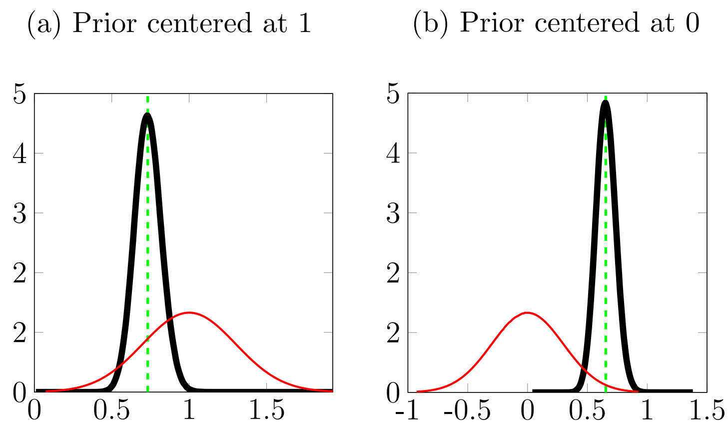
Notes: The table reports parameters fixed in the estimation procedure for both DE and RE.

Table 7: Prior Distribution

Parameter	Description	Distribution	Mean	Std. dev
θ	diagnosticity	Normal	1	0.3
α	cap. share	Normal	0.3	0.05
h	habits	Beta	0.5	0.1
$\frac{\chi''(1)}{\chi'(1)}$	cap. util. costs	Gamma	5	1
ψ_p	Rotemberg prices	Normal	100	25
ψ_w	Rotemberg wages	Normal	3000	5000
ν	inv. Frisch elas.	Gamma	2	0.75
$S''(1)$	inv. adj. costs	Normal	4	1
ρ_R	m.p. rule	Beta	0.5	0.2
ϕ_π	m.p. rule	Normal	1.5	0.3
ϕ_x	m.p. rule	Normal	0.005	0.005
Technology Shocks				
ρ	persist.	Beta	0.6	0.2
σ_a	tech. shock s.d.	Inv. Gamma	0.5	1
σ_s	noise shock s.d.	Inv. Gamma	1	1
Investment-Specific Shocks				
ρ_μ	persist.	Beta	0.6	0.2
σ_μ	s.d.	Inv. Gamma	5	1.5
Markup Shocks				
ρ_p	persist.	Beta	0.6	0.2
ϕ_p	ma. comp.	Beta	0.5	0.2
σ_p	s.d.	Inv. Gamma	0.15	1
ρ_w	persist.	Beta	0.6	0.2
ϕ_w	ma. comp.	Beta	0.5	0.2
σ_w	s.d.	Inv. Gamma	0.15	1
Policy Shocks				
ρ_{mp}	persist.	Beta	0.4	0.2
σ_{mp}	s.d.	Inv. Gamma	0.15	1
ρ_g	persist.	Beta	0.6	0.2
σ_g	s.d.	Inv. Gamma	0.5	1
Measurement Errors				
σ_y^{me}	s.d.	Inv. Gamma	0.5	1
σ_c^{me}	s.d.	Inv. Gamma	0.5	1
σ_i^{me}	s.d.	Inv. Gamma	0.5	1
σ_π^{me}	s.d.	Inv. Gamma	0.5	1
σ_r^{me}	s.d.	Inv. Gamma	0.5	1

Notes: The table reports the prior distribution of structural parameters in the estimation procedure. The diagnosticity parameter θ is fixed at 0 under RE.

Figure 7: Posterior Distribution of Diagnosticity Parameter



Notes: Panels a) and b) depict the prior and posterior density of θ when the prior is centered at 1 and 0, respectively. The red, solid lines denote the prior distribution of θ, which follows a Normal distribution with standard deviation 0.3. The black, solid lines (the green, dashed vertical line) denote the posterior distribution (the posterior mean) of θ.

Table 8: Model-Implied Volatilities in the Medium-Scale DSGE Model

Variable	Diagnostic Expectations	Rational Expectations	Percentage Increase
Consumption	0.7939	0.6445	23%
Output	1.0055	0.8928	13%
Price Inflation	0.5308	0.4682	13%
Wage Inflation	0.9411	0.8498	11%
Real Rate	0.7532	0.5516	37%

Notes: The table reports the standard deviations of consumption growth, output growth, price inflation, wage inflation, and the real rate in the medium-scale DSGE model. The final column entitled "Percentage Increase" shows the percentage increase in standard deviation under the DE model relative to the RE benchmark (setting $\theta = 0$ along with parameter estimates in Table 1). There are eight structural shocks in the model, as in Blanchard et al. (2013): the TFP growth shock, TFP level shock, noise shock, marginal efficiency of investment (MEI) shock, price markup shock, wage markup shock, monetary policy shock, and government spending shock.

G Robustness

G.1 Prior on the Diagnosticity Parameter θ Centered at Zero

Table 9: Posterior Distribution: Prior Centered at Zero

		Diagnostic		Rational	
Parameter	Description	Mean	[05, 95]	Mean	[05, 95]
θ	diagnosticity	0.6537	[0.5193, 0.7884]		
α	cap. share	0.1340	[0.1227, 0.1455]	0.1390	[0.1278, 0.1505]
h	habits	0.7110	[0.6810, 0.7415]	0.5803	[0.5424, 0.6178]
$\frac{\chi''(1)}{\chi'(1)}$	cap. util. costs	5.0273	[3.4169, 6.6350]	5.5929	[3.9095, 7.2242]
ψ_p	Rotemberg prices	124.51	[97.470, 151.19]	181.84	[126.66, 188.88]
ψ_w	Rotemberg wages	538.73	[231.71, 833.33]	9710.9	[4510.5, 14712.]
ν	inv. Frisch elas.	3.6778	[2.4841, 5.0289]	1.2832	[0.5012, 1.9475]
$S''(1)$	inv. adj. costs	6.9600	[5.8331, 8.0849]	7.0701	[6.0111, 8.1332]
ρ_R	m.p. rule	0.5920	[0.5541, 0.6304]	0.5563	[0.4380, 0.6806]
ϕ_π	m.p. rule	1.5297	[1.4093, 1.6481]	1.0682	[1.0001, 1.2046]
ϕ_x	m.p. rule	0.0062	[0.0001, 0.0111]	0.0013	[0.0001, 0.0030]
Technology Shocks					
ρ	persist.	0.8584	[0.8381, 0.8786]	0.9535	[0.9352, 0.9716]
σ_a	tech. shock s.d.	1.4050	[1.2824, 1.5249]	1.5258	[1.3896, 1.6601]
σ_s	noise shock s.d.	0.5375	[0.3182, 0.7481]	1.0594	[0.3781, 1.7574]
Investment-Specific Shocks					
ρ_μ	persist.	0.3066	[0.2493, 0.3630]	0.3310	[0.2631, 0.4003]
σ_μ	s.d.	18.947	[15.038, 22.845]	20.2121	[16.369, 23.989]
Markup Shocks					
ρ_p	persist.	0.8748	[0.8303, 0.9205]	0.8205	[0.7663, 0.8769]
ϕ_p	ma. comp.	0.5874	[0.4748, 0.7023]	0.5563	[0.4380, 0.6806]
σ_p	s.d.	0.1623	[0.1337, 0.1905]	0.1988	[0.1700, 0.2271]
ρ_w	persist.	0.9969	[0.9940, 0.9999]	0.6543	[0.5146, 0.7978]
ϕ_w	ma. comp.	0.5708	[0.3867, 0.7587]	0.5142	[0.2882, 0.7444]
σ_w	s.d.	0.4449	[0.3514,0.5354]	0.4490	[0.3836, 0.5142]
Policy Shocks					
ρ_{mp}	persist.	0.0296	[0.0100, 0.0516]	0.0197	[0.0009, 0.0383]
σ_{mp}	s.d.	0.3751	[0.3394, 0.4099]	0.3283	[0.3000, 0.3556]
ρ_g	persist.	0.9332	[0.9051, 0.9619]	0.8974	[0.8682, 0.9275]
σ_g	s.d.	0.3699	[0.3376, 0.4011]	0.3706	[0.3384, 0.4022]
Measurement Errors					
σ_y^{ME}	s.d.	0.4968	[0.4464, 0.5467]	0.5034	[0.4529, 0.5533]
σ_c^{ME}	s.d.	0.4107	[0.3607, 0.4594]	0.4255	[0.3739, 0.4764]
σ_i^{ME}	s.d.	1.4291	[1.2541, 1.6033]	1.4514	[1.2692, 1.6284]
σ_r^{ME}	s.d.	0.2681	[0.2406, 0.2949]	0.2285	[0.2018, 0.2551]
σ_π^{ME}	s.d.	0.1614	[0.1409, 0.1817]	0.1482	[0.1267, 0.1693]
log Marg. Likelihood		-1814.82		-1847.38	

G.2 Smets and Wouters (2007)

Table 10: Prior Distribution: Smets and Wouters (2007)

Parameter	Description	Distribution	Mean	Std. dev
θ	diagnosticity	Normal	0.5	0.3
α	cap. share	Normal	0.3	0.05
h	habits	Beta	0.7	0.1
$\frac{\chi''(1)}{\chi'(1)}$	cap. util. costs	Gamma	0.5	0.15
ψ_p	Rotemberg prices	Normal	350	75
ψ_w	Rotemberg wages	Normal	1000	1000
ν	inv. Frisch elas.	Normal	2	0.75
$S''(1)$	inv. adj. costs	Normal	4	1.5
ρ_R	m.p. rule	Beta	0.75	0.1
ϕ_π	m.p. rule	Normal	1.5	0.25
ϕ_x	m.p. rule	Normal	0.125	0.05
ϕ_{dx}	m.p. rule	Normal	0.125	0.05
ι_p	index. prices	Beta	0.5	0.15
ι_w	index. wages	Beta	0.5	0.15
$100G_a$	s.s. growth rate	Normal	0.4	0.1
$\log L$	s.s. hours	Normal	0	2
$100(\pi - 1)$	s.s. infl.	Gamma	0.625	0.1
$100(\beta^{-1} - 1)$	disc. factor	Gamma	0.25	0.1
F	share of fixed costs	Normal	1.25	0.125
σ_c	cons. curvature	Normal	1.5	0.375
Shocks				
ρ_a	persist. tech.	Beta	0.5	0.2
σ_a	s.d. tech.	Inv. Gamma	0.1	2
ρ_μ	persist. inv.	Beta	0.5	0.2
σ_μ	s.d. inv.	Inv. Gamma	0.1	2
ρ_b	persist. pref.	Beta	0.5	0.2
σ_b	s.d. pref.	Inv. Gamma	0.1	2
ρ_p	persist. prices	Beta	0.5	0.2
ϕ_p	ma. comp. prices	Beta	0.5	0.2
σ_p	s.d. prices	Inv. Gamma	0.1	2
ρ_w	persist. wages	Beta	0.5	0.2
ϕ_w	ma. comp. wages	Beta	0.5	0.2
σ_w	s.d. wages	Inv. Gamma	0.1	2
ρ_{mp}	persist. mon.	Beta	0.5	0.2
σ_{mp}	s.d. mon.	Inv. Gamma	0.1	2
ρ_g	persist. fisc.	Beta	0.5	0.2
σ_g	s.d. fisc	Inv. Gamma	0.1	2

Table 11: Posterior Distribution: Smets and Wouters (2007)

		Diagnostic		Rational	
Parameter	Description	Mean	[05, 95]	Mean	[05, 95]
θ	diagnosticity	0.4435	[0.1822, 0.6928]		
α	cap. share	0.1874	[0.1575, 0.2169]	0.1884	[0.1588, 0.2178]
h	habits	0.7100	[0.6385, 0.7839]	0.7027	[0.6334, 0.7725]
$\frac{\chi''(1)}{\chi'(1)}$	cap. util costs	0.6241	[0.4539, 0.8013]	0.5785	[0.4016, 0.7549]
ψ_p	Rotemberg prices	399.36	[292.11, 506.07]	383.13	[272.30, 490.97]
ψ_w	Rotemberg wages	2266.5	[1083.3, 3407.9]	2265.1	[1092.8, 3375.0]
ν	inv. Frisch elas.	1.9577	[1.0626, 2.7971]	2.0293	[1.1717, 2.8701]
$S''(1)$	inv. adj. costs	5.6924	[3.9384, 7.3949]	5.7666	[4.0637, 7.4253]
ρ_R	m.p. rule	0.7962	[0.7560, 0.8381]	0.8132	[0.7754, 0.8515]
ϕ_π	m.p. rule	2.0801	[1.7974, 2.3631]	2.0199	[1.7277, 2.3092]
ϕ_x	m.p. rule	0.0836	[0.0450, 0.1220]	0.0839	[0.0478, 0.1199]
ϕ_{dx}	m.p. rule	0.2412	[0.1943, 0.2886]	0.2327	[0.1862, 0.2790]
ι_p	index. prices	0.3075	[0.1491, 0.4647]	0.2268	[0.0905, 0.3584]
ι_w	index. wages	0.6287	[0.4343, 0.8238]	0.5712	[0.3695, 0.7756]
$100G_a$	s.s. growth rate	0.4206	[0.3950, 0.4467]	0.4226	[0.3982, 0.4465]
$\log L$	s.s. hours	0.6699	[-1.169, 2.5050]	0.6560	[-1.147, 2.4377]
$100(\pi-1)$	s.s. infl.	0.7775	[0.6156, 0.9427]	0.7543	[0.5932, 0.9219]
$100(\beta^{-1}-1)$	disc. factor	0.1640	[0.0708, 0.2523]	0.1671	[0.0731, 0.2576]
F	share of fixed costs	1.5447	[1.4160, 1.6777]	1.5845	[1.4549, 1.7142]
σ_c	cons. curvature	1.3804	[1.1204, 1.6347]	1.3740	[1.1540, 1.5844]
Shocks					
ρ_a	persist. tech.	0.9592	[0.9384, 0.9806]	0.9528	[0.9331, 0.9731]
σ_a	s.d. tech.	0.4658	[0.4174, 0.5145]	0.4623	[0.4152, 0.5101]
ρ_μ	persist. inv.	0.7815	[0.6861, 0.8806]	0.7129	[0.6197, 0.8095]
σ_μ	s.d. inv.	0.3405	[0.2507, 0.4282]	0.4528	[0.3726, 0.5315]
ρ_η	persist. pref.	0.3829	[0.1746, 0.6025]	0.2429	[0.0848, 0.3919]
σ_η	s.d. pref.	0.1747	[0.1104, 0.2358]	0.2359	[0.1944, 0.2778]
ρ_p	persist. prices	0.8709	[0.7877, 0.9529]	0.8706	[0.7914, 0.9506]
ϕ_p	ma. comp. prices	0.6564	[0.4567, 0.8613]	0.6710	[0.4991, 0.8443]
σ_p	s.d. prices	0.1044	[0.0671, 0.1409]	0.1407	[0.1116, 0.1695]
ρ_w	persist. wages	0.9600	[0.9327, 0.9882]	0.9672	[0.9455, 0.9900]
ϕ_w	ma. comp. wages	0.8620	[0.7775, 0.9500]	0.8817	[0.8188, 0.9482]
σ_w	s.d. wages	0.1899	[0.1430,0.2370]	0.2432	[0.2070, 0.2793]
ρ_{mp}	persist. mon.	0.1216	[0.0287, 0.2064]	0.1389	[0.0383, 0.2316]
σ_{mp}	s.d. mon.	0.2520	[0.2252, 0.2780]	0.2467	[0.2217, 0.2713]
ρ_g	persist. fisc.	0.9806	[0.9681, 0.9933]	0.9802	[0.9673, 0.9935]
σ_g	s.d. fisc.	0.5224	[0.4720, 0.5731]	0.5260	[0.4746, 0.5755]
ρ_{ga}	corr.	0.5255	[0.3800, 0.6717]	0.5202	[0.3745, 0.6670]
log Marg. Likelihood		-897.91		-900.69	

G.3 Justiniano, Primiceri, and Tambalotti (2010)

Table 12: Prior Distribution: Justiniano, Primiceri, and Tambalotti (2010)

Parameter	Description	Distribution	Mean	Std. dev
θ	diagnosticity	Normal	0.5	0.3
α	cap. share	Normal	0.3	0.05
h	habits	Beta	0.5	0.1
$\frac{\chi''(1)}{\chi'(1)}$	cap. util. costs	Gamma	5	1
ψ_p	Rotemberg prices	Normal	100	25
ψ_w	Rotemberg wages	Normal	3000	5000
ν	inv. Frisch elas.	Gamma	2	0.75
$S''(1)$	inv. adj. costs	Normal	4	1
ρ_R	m.p. rule	Beta	0.6	0.2
ϕ_π	m.p. rule	Normal	1.7	0.3
ϕ_x	m.p. rule	Normal	0.13	0.05
ϕ_{dx}	m.p. rule	Normal	0.13	0.05
ι_p	index. prices	Beta	0.5	0.15
ι_w	index. wages	Beta	0.5	0.15
$100G_a$	s.s. growth rate	Normal	0.5	0.03
λ_p	s.s. markup prices	Normal	0.15	0.05
λ_w	s.s. markup wages	Normal	0.15	0.05
$\log L$	s.s. log hours	Normal	0	0.5
$100(\pi - 1)$	s.s. infl.	Normal	0.5	0.1
$100(\beta^{-1} - 1)$	disc. factor	Gamma	0.25	0.1
Shocks				
ρ_a	persist. tech.	Beta	0.6	0.2
σ_a	s.d. tech.	Inv. Gamma	0.5	1
ρ_μ	persist. inv.	Beta	0.6	0.2
σ_μ	s.d. inv.	Inv. Gamma	0.5	1
ρ_b	persist. pref.	Beta	0.6	0.2
σ_b	s.d. pref.	Inv. Gamma	0.1	1
ρ_p	persist. prices	Beta	0.6	0.2
ϕ_p	ma. comp. prices	Beta	0.5	0.2
σ_p	s.d. prices	Inv. Gamma	0.1	1
ρ_w	persist. wages	Beta	0.6	0.2
ϕ_w	ma. comp. wages	Beta	0.5	0.2
σ_w	s.d. wages	Inv. Gamma	0.1	1
ρ_{mp}	persist. mon.	Beta	0.4	0.2
σ_{mp}	s.d. mon.	Inv. Gamma	0.1	1
ρ_g	persist. fisc.	Beta	0.6	0.2
σ_g	s.d. fisc.	Inv. Gamma	0.5	1

Table 13: Posterior Distribution: Justiniano, Primiceri, and Tambalotti (2010)

		Diagnostic		Rational	
Parameter	Description	Mean	[05, 95]	Mean	[05, 95]
θ	diagnosticity	0.4336	[0.1894, 0.6745]		
α	cap. share	0.1702	[0.1603, 0.1800]	0.1700	[0.1602, 0.1800]
h	habits	0.8788	[0.8443, 0.9142]	0.8270	[0.7655, 0.8902]
$\frac{\chi''(1)}{\chi'(1)}$	cap. util. costs	5.3160	[3.6696, 6.9322]	5.2978	[3.6521, 6.9145]
ψ_p	Rotemberg prices	123.01	[91.513, 154.15]	116.43	[84.65, 147.570]
ψ_w	Rotemberg wages	2863.31	[594.68, 5275.6]	3204.29	[720.56, 5835.5]
ν	inv. Frisch elas.	4.3961	[2.9554, 5.7777]	4.2917	[2.8854, 5.6762]
$S''(1)$	inv. adj. costs	2.9689	[2.0722, 3.8461]	2.7528	[1.8821, 3.6124]
ρ_R	m.p. rule	0.8064	[0.7681, 0.8445]	0.8193	[0.7822, 0.8567]
ϕ_π	m.p. rule	2.1751	[1.8764, 2.4631]	2.0782	[1.7792, 2.3655]
ϕ_x	m.p. rule	0.0559	[0.0269, 0.0847]	0.0600	[0.0306, 0.0887]
ϕ_{dx}	m.p. rule	0.2425	[0.1983, 0.2860]	0.2389	[0.1974, 0.2801]
ι_p	index. prices	0.2589	[0.1266, 0.3888]	0.1964	[0.0821, 0.3062]
ι_w	index. wages	0.1477	[0.0862, 0.2085]	0.1127	[0.0595, 0.1655]
$100G_a$	s.s. growth rate	0.4675	[0.4237, 0.5108]	0.4695	[0.4256, 0.5139]
λ_p	s.s. markup prices	0.2340	[0.1791, 0.2890]	0.2419	[0.1847, 0.2982]
λ_w	s.s. markup wages	0.1347	[0.0525, 0.2127]	0.1360	[0.0543, 0.2130]
$\log L$	s.s. log hours	0.1827	[-0.600, 0.9579]	0.2032	[-0.571, 0.9877]
$100(\pi-1)$	s.s. infl.	0.7877	[0.6831, 0.8934]	0.7677	[0.6557, 0.8782]
$100(\beta^{-1}-1)$	disc. factor	0.1379	[0.0604, 0.2119]	0.1404	[0.0611, 0.2154]
Shocks					
ρ_a	persist. tech.	0.2145	[0.1240, 0.3047]	0.2518	[0.1522, 0.3508]
σ_a	s.d. tech.	0.8828	[0.8032, 0.9591]	0.8908	[0.8121, 0.9695]
ρ_μ	persist. inv.	0.7650	[0.6965, 0.8352]	0.7352	[0.6598, 0.8125]
σ_μ	s.d. inv.	5.1618	[3.9758, 6.3096]	5.8481	[4.3020, 7.3632]
ρ_b	persist. pref.	0.3595	[0.2202, 0.4971]	0.5161	[0.3394, 0.6957]
σ_b	s.d. pref.	0.0650	[0.0379, 0.0911]	0.0566	[0.0317, 0.0818]
ρ_p	persist. prices	0.9363	[0.9003, 0.9739]	0.9276	[0.8882, 0.9680]
ϕ_p	ma. comp. prices	0.6515	[0.4918, 0.8136]	0.6790	[0.5503, 0.8127]
σ_p	s.d. prices	0.0989	[0.0692, 0.1276]	0.1344	[0.1126, 0.1559]
ρ_w	persist. wages	0.9808	[0.9652, 0.9975]	0.9773	[0.9586, 0.9972]
ϕ_w	ma. comp. wages	0.9164	[0.8734, 0.9603]	0.9202	[0.8809, 0.9606]
σ_w	s.d. wages	0.1679	[0.1355, 0.2006]	0.2115	[0.1847, 0.2387]
ρ_{mp}	persist. mon.	0.1630	[0.0623, 0.2610]	0.1627	[0.0576, 0.2636]
σ_{mp}	s.d. mon.	0.2336	[0.2109, 0.2560]	0.2262	[0.2050, 0.2472]
ρ_g	persist. fisc.	0.9970	[0.9947, 0.9990]	0.9967	[0.9941, 0.9990]
σ_g	s.d. fisc.	0.3477	[0.3187, 0.3756]	0.3476	[0.3191, 0.3764]
log Marg. Likelihood		-1190.86		-1193.78	

H Data Appendix

Our data set spans the period 1954:III to 2004:IV. The series for Real GDP, Real Personal Consumption Expenditures, Real Personal Durable Consumption Expenditures, Real Gross Private Domestic Investment, Wages and the GDP Implicit Price Deflator are from the Bureau of Economic Analysis. Population and employment series are from the Bureau of Labor Statistics online database (series IDs LNS10000000Q and LNS12000000Q respectively). The Federal Funds Rate series is from the Federal Reserve Board online database (series ID H15/H15/RIFSPFF N.M).

The GDP series is constructed by dividing Real GDP by population. The consumption series is constructed by subtracting Real Personal Durable Consumption from Real Personal Consumption and dividing by population. The investment series is constructed by dividing the sum of Real Gross Investment and Real Personal Durable Consumption by population. The labor input series is constructed by dividing Employment by Population. Inflation is constructed by computing the quarterly log difference of the Price Deflator. The real wage is constructed by dividing Real Wages by the Price Deflator. The nominal interest rate is the effective Federal Funds Rate.

For SPF forecast data, we use the median forecast (across individual forecasters) as the consensus forecast. All forecasts we use are one quarter ahead forecasts. For the output (series ID RGDP), consumption (series ID RCONSUM), and investment (series ID RNRESIN) growth rate, we subtract these growth rate forecasts by actual population growth rate to obtain per capita forecasts. Inflation and the nominal interest rate are obtained from the GDP Price Deflator and the Treasury bill rate (series IDs PGDP and TBILL).