



Relaxing Rational Expectations

*Lance Kent**

College of William and Mary

College of William and Mary
Department of Economics
Working Paper Number 159

Current Version: March 2, 2015
First Version: May 15, 2014

*The author is grateful for conversations with Federico Di Pace, Leonardo Melosi, Martin Eichenbaum, Christian Matthes, B.R. Gabriel, and seminar participants at the 2014 SEA meetings in Atlanta for helpful comments.

Model Misspecification and Relaxing Rational Expectations

Abstract

The assumption of rational expectations is potentially a serious source of misspecification in DSGE models. Many recent theories of expectations formation have relaxed rational expectations and improved the predictive properties of benchmark macroeconomic models. Problematically, the space of possible theoretical deviations from rational expectations is very large, especially since the aggregate consequences of deviations from rational expectations in equilibrium may not be directly measurable using existing surveys of expectations. This paper provides evidence on which small reduced-form state-contingent deviations from rational expectations yield the most improvement in replicating features of macroeconomic time series, and which aspects of model misspecification are and are not ameliorated by these small deviations from rational expectations. The findings: a) The data favor deviations from rational expectations among firms in which they over-estimate the persistence of inflation. b) Relaxing rational expectations in a New Keynesian model partially substitutes for the additional structural mechanisms in the larger Smets Wouters (2007) model. Relaxing rational expectations within the Smets Wouters (2007) model improves that model's ability to reproduce some of the spectral coherencies between output growth, investment growth, and labor supply. The mechanism is a combination of shocks to beliefs themselves and the role that deviations from rational expectations have in changing the propagation of other shocks.

Keywords: Expectations, DSGE, misspecification, Bayesian estimation.
JEL Codes: C52, D84, E17, E27, E32.

Lance Kent
College of William and Mary
Department of Economics
P.O. Box 8795
Williamsburg, VA 23185
United States

lckent@wm.edu

lkent@wm.edu

Relaxing Rational Expectations

Lance Kent*

Department of Economics, College of William and Mary

This draft: March 2, 2015

First draft: May 15, 2014

Abstract The assumption of rational expectations is potentially a serious source of misspecification in DSGE models. Many recent theories of expectations formation have relaxed rational expectations and improved the predictive properties of benchmark macroeconomic models. Problematically, the space of possible theoretical deviations from rational expectations is very large, especially since the aggregate consequences of deviations from rational expectations in equilibrium may not be directly measurable using existing surveys of expectations. This paper provides evidence on which small reduced-form state-contingent deviations from rational expectations yield the most improvement in replicating features of macroeconomic time series, and which aspects of model misspecification are and are not ameliorated by these small deviations from rational expectations. The findings: a) The data favor deviations from rational expectations among firms in which they over-estimate the persistence of inflation. b) Relaxing rational expectations in a New Keynesian model partially substitutes for the additional structural mechanisms in the larger Smets Wouters (2007) model. Relaxing rational expectations within the Smets Wouters (2007) model improves that model's ability to reproduce some of the spectral coherencies between output growth, investment growth, and labor supply. The mechanism is a combination of shocks to beliefs themselves and the role that deviations from rational expectations have in changing the propagation of other shocks.

Keywords: Expectations, DSGE, misspecification, Bayesian estimation.

JEL Classification Numbers: C52, D84, E17, E27, E32.

*Address: 115 Morton Hall, 100 Ukrop Way, Williamsburg, VA 23188, USA, telephone: 1-757-221-1389, e-mail: *lckent@wm.edu*. The author is grateful for conversations with Federico Di Pace, Leonardo Melosi, Martin Eichenbaum, Christian Matthes, B.R. Gabriel, and seminar participants at the 2014 SEA meetings in Atlanta for helpful comments.

1 Introduction

Agents' expectations about the future are central to all modern macroeconomic models. The benchmark assumption in most modern macroeconomic models is that agents' expectations are rational, that is, that agents' beliefs about the distributions of future variables coincide with the actual distributions. However, it is still an open question to what extent households and firms have rational expectations. This is an important question for understanding macroeconomic dynamics because when agents' beliefs deviate from rational expectations, the equilibrium dynamics of the model may deviate significantly as well. In fact, the literature acknowledges that there is potentially considerable explanatory power in considering situations where agents do not possess rational expectations.

The literature has proposed a wide range of alternatives to rational expectations. The crucial theoretical difficulty in considering deviations from rational expectations is that there is only one way in which expectations can be rational, and a large universe of ways in which expectations can deviate. While it is possible to assess one mechanism of expectations formation against another using traditional notions of goodness-of-fit, the fact that the space of possible mechanisms is so vast suggests that it might be useful to have additional reduced-form evidence to aid the creation and selection of mechanisms. This paper seeks to provide such evidence.

Every theory which gives rise to a deviation from rational expectations implies some departure of agents' perceived laws of motion of economic quantities from the actual laws of motion. This paper asks: For a given macroeconomic model, which perceived law of motion does the data favor the most? To answer that question, this paper proposes a new equilibrium concept, the generalized expectations equilibrium (GEE). In a GEE, the laws of motion that agents perceive to hold for various economic variables are represented by a set of functions of historical variables. The elasticities of agents' perceived laws of motion with respect to historical variables are taken to be parameters.

For a given macroeconomic model, a GEE allows agents' perceived laws of motion to differ from the actual laws of motion. For example, in a GEE, agents might believe that inflation is more persistent than it actually is. In this example, the GEE specifies a "deviation" between the actual and perceived law of motion that is larger in times when inflation is high. In equilibrium, this "deviation" forces the actual law of motion away from the law of motion under rational expectations. Consider the set of deviations between actual and perceived laws of motion. For each element in the set of deviations, there is an associated actual law of motion. If these "deviations" are zero for all aggregate states, an rational expectations equilibrium obtains as a special case. As long as the perceived law(s) of motion (allowing for heterogeneity among agents' perceptions) differ from the actual, this causes the actual law of motion in equilibrium to differ from the actual law of motion under rational expectations. This paper estimates the coefficients of the the perceived laws of motion using Bayesian methods, effectively choosing among elements in the set of equilibrium

laws of motion defined by the set of possible “deviations.”¹ By putting relatively tight priors on the elasticities, the GEE stays relatively close to a rational expectations equilibrium (REE). Estimating a GEE allows us to see which state-dependent deviations are favored by the data under a given model, once the assumption of REE is gently relaxed.

It is important to stress what the estimation identifies: it is the deviation from rational expectations favored *while maintaining the rest of a specific model as given*. In contrast to other papers, such as Milani (2007) and Milani and Rajbhandari (2012) which discuss how the estimates of the the remainder of the “deep” parameters (not governing expectations) change when expectations are no longer REE, I take the parameters not governing expectations as given. If the rest of the model is mis-specified, the data will seize upon the reduced-form specification of expectations to push against that misspecification as well.

There is a long literature testing rational expectations, both using macro and micro data. A subset of this literature exploits surveys on expectations. There are estimates of the process that drives expectations, such as Andolfatto et al. (2008) and Del Negro and Eusepi (2011), and investigations of whether survey responses conform to various theories, such as sticky information in Branch (2007) and Coibion and Gorodnichenko (2011). It’s not yet clear to what extent the expectations of various surveyed agents (including the Survey of Professional forecasters, households (such as exploited in Bryan and Palmqvist (2005)) or firms (such as in Bachmann and Elstner (2013))) are relevant for the equilibrium relationships between macroeconomic aggregates as claimed in benchmark models; if heterogeneity among households and firms is important, then there is perhaps a role for higher-order expectations, aggregation wedges, and other distortions which are not captured in the surveys themselves. While survey data on expectations can teach researchers a great deal, this paper will give evidence on how much relaxing REE can improve on REE for benchmark models using only macro data instead of working with surveys and forecasts directly.

Estimating a GEE is similar in spirit to a kind of structural VAR, in that both impose an extensive set of cross-equation restrictions. In contrast to the structural VAR, the GEE allows only agents’ beliefs to be richly parametrized functions of historical variables; the rest of the benchmark model is maintained as a set of identifying restrictions. One approach that is similar to the GEE is the DSGE-VAR of Del Negro and Schorfheide (2004) and Del Negro and Schorfheide (2009). In that approach, the tight restrictions of the DSGE model are relaxed in the direction of the richly parametrized but reduced-form specification of a VAR. The GEE also relaxes the cross-equation restrictions of a DSGE model under REE, but does so in a way that is more restrictive than a DSGE-VAR. In the DSGE-VAR, every equation in the model is relaxed towards a reduced-form lin-

¹In this way, estimating a GEE is similar in spirit to estimating the “wedges” in a “Business Cycle Accounting” exercise, as in Chari et al. (2007). While both exercises are silent on specific structural interpretations of the deviations from the benchmark model, they both seek to provide some empirical guidelines for future theories. One important difference between the two approaches is that the “deviations” in a GEE are not wedges that change the allocations at a given point in time; they are functions of lagged states that change the equilibrium saddle path for the economy.

ear relationship. In contrast, the GEE only introduces deviations where expectations enter into the DSGE model. Since the GEE does not introduce any deviations to equations which capture exclusively contemporaneous relationships, all equations without expectations are preserved. Therefore, the nature and degree of misspecification that a GEE (and hence, any theory of deviation from REE) can address is limited.

This paper works with two workhorse monetary macroeconomic models: the small New Keynesian model and the medium-scale model of Smets and Wouters (2007). First, this paper derives the GEE of the small New Keynesian model and demonstrates some equilibrium consequences of nonzero “deviations” from rational expectations. Second, this paper shows that several theories of non-rational expectations in the literature entail state-contingent deviations from REE. As illustrative examples, this paper considers sticky information of Mankiw and Reis (2002), constant-gain learning, from Evans and Honkapohja (2001), Milani (2007) and others, diverse beliefs as in Kurz (2012), and sentiments-based models of Angeletos and La’O (2009), Angeletos and La’O (2011), Angeletos and La’O (2013), and Angeletos et al. (2014). I use log-linearizations of these theories to argue that the GEE is a parsimonious representation that might capture some of the state-dependent deviations that these and other (potentially as-of-yet unexplored) theories might entail. I do not present these theories to establish a horse race, but only to stress that the reduced-form deviations in the GEE can capture a wide variety of theoretical mechanisms.

The primary contribution of the paper is to identify which misspecification can and cannot be ameliorated by entertaining deviations from rational expectations, especially in a way that can inform existing and future models. I show that the posterior distributions of some of the elasticities of the state-contingency of beliefs are comfortably away from the null of rational expectations, and that the impulse responses of output, inflation and the interest rate to shocks are different as well. Specifically, the data favor deviations from REE that are observationally equivalent to firms overestimating the persistence of inflation; once inflation is high, firms in a GEE expect it to stay higher than what will actually prevail.

In terms of reproducing spectral properties of observed time series, relaxing rational expectations in a New Keynesian model partially substitutes for the additional structural mechanisms in the larger Smets Wouters (2007) model. Additionally, relaxing rational expectations within the Smets Wouters (2007) model improves upon that model’s ability to reproduce some of the spectral coherencies between output growth, investment growth, and labor supply.

The mechanism for these improvements is a combination of shocks to beliefs themselves and the role that deviations from rational expectations have in changing the propagation of other structural shocks. A variance decomposition exercise shows that shocks to beliefs among firms who set sticky goods prices can explain between 25% and 30% of the variance of output growth, inflation, and the interest rate. Additionally, the share of total variance explained by shocks to the goods and labor markup drops dramatically under a GEE vs. REE, and the change in the propagation of technology shocks results in their explaining an larger share of output and inflation under a GEE vs. REE.

This paper demonstrates what explanatory power one might hope to gain in macroeconomic models by considering deviations from rational expectations. What is the reduced-form evidence on which deviations might achieve the greatest increase in explanatory power? What misspecification can small deviations from rational expectations address, and what misspecification can they not? This paper presents novel empirical evidence to answer these questions.

2 The Generalized Expectations Equilibrium

In this section, I define the Generalized Expectations Equilibrium (GEE) and demonstrate some of its properties within a benchmark New Keynesian model.

2.1 Notation and definitions

In what follows, I first define the GEE over a generic model, then contrast the GEE with a rational expectations equilibrium (REE) for the same generic model. The notation in this section is meant to be fairly general, encompassing a wide class of DSGE models. First, let X_t be a K -dimensional element vector, with elements indexed by $k = 1...K$, of all variables in the model, both endogenous and exogenous, and let ϵ_t by the vector of exogenous shocks. The equilibrium conditions are then a set of equations. Let $\{F^i, \}_{i=1..N}$ be a set of equations that take as inputs $(X_{t+1}, X_t, X_{t-1}, \epsilon_t)$, which includes forward-looking variables X_{t+1} . Let F^0 be a set of equations that do not include forward-looking variables. The collection $\{\{F^i, \}_{i=1..N}, F^0\}$ defines a benchmark DSGE model.

Definition 1 *A **rational expectations equilibrium** for a collection $\{\{F^i, \}_{i=1..N}, F^0\}$ is a law of motion $g^{REE}(\cdot, \cdot)$ for X and ϵ such that:*

1. $\forall t, X_t = g^{REE}(X_{t-1}, \epsilon_t)$. Let E_t denote the expectation operator with respect to the law of motion g^{REE} and the true distribution for ϵ_t .
2. $\forall i, 0 = \tilde{E}_t F^i(X_{t+1}, X_t, X_{t-1}, \epsilon_t)$
3. $0 = F^0(X_t, X_{t-1}, \epsilon_t)$

A GEE is a flexible departure from REE in the following way: for each equation i with forward-looking variables, we suppose that the actual laws of motion for these quantities are distorted by a time-varying deviation from REE Γ^i , which defines an expectation operator under the distorted law of motion \tilde{E}_t : $\tilde{E}_t f(X_{t+1}) = E_t f(\Gamma_t^i X_{t+1})$. Thus, we can write any DSGE model in the form $0 = \tilde{E}_t F^i(X_{t+1}, X_t, X_{t-1}, \epsilon_t)$ for a collection of i and $0 = F^0(X_t, X_{t-1}, \epsilon_t)$, for the set of equations F_0 that have no forward-looking variables.

Definition 2 *A **generalized expectations equilibrium** for a collection $\{\{F^i, \Gamma^i\}_{i=1..N}, F^0\}$ is a law of motion $g^{GEE}(\cdot, \cdot)$ for X and ϵ such that:*

1. $\forall t, X_t = g^{GEE}(X_{t-1}, \epsilon_t)$. Let E_t denote the expectation operator with respect to the law of motion g^{GEE} and the true distribution for ϵ_t .
2. $\forall i, 0 = \tilde{E}_t^i F^i(X_{t+1}, X_t, X_{t-1}, \epsilon_t) = E_t F^i(\Gamma_t^i X_{t+1}, X_t, X_{t-1}, \epsilon_t)$
3. $0 = F^0(X_t, X_{t-1}, \epsilon_t)$
4. $\Gamma_t^i = \Gamma^i(X_t, X_{t-1}, \Gamma_{t-1}^i, \epsilon_t^{\Gamma^i})$

The first condition of Definition 2 states that the equilibrium law of motion g^{GEE} is the “true” law of motion for the variables in the model. If an econometrician were to observe all elements of X_t simulated from the equilibrium law of motion, then with enough observations he or she could estimate and recover g^{GEE} from data. However, it is not necessarily the case that agents in the model can do this; the second condition states that for equation i , there is some either mistaken belief, or aggregation wedge, or higher-order belief such that the law of motion for X_{t+1} is distorted by factor Γ_t^i . The third condition states that the equilibrium conditions for all equations which impose contemporaneous restrictions between endogenous variables are unchanged, relative to REE. Finally, the fourth condition states that the deviation Γ_t^i follows its own law of motion.

Definition 2 achieves three goals: 1) The expectation operator E_t with respect to the equilibrium law of motion for X_t obeys the law of iterated expectations². 2) REE is a special case where $\Gamma_t^i = 1 \forall i, t$. 3) The set of possible Γ^i is observationally equivalent to a wide range of deviations from REE, encompassing both those where beliefs themselves are different in a unified way across agents, and models where heterogeneity in beliefs leads to aggregation wedges, or higher-order beliefs. Section 2.2 describes a benchmark New Keynesian model and Section 2.3 discusses several theories that are approximately observationally equivalent to Definition 2 within that New Keynesian model.

2.2 GEE within a New Keynesian model

The New Keynesian model is one of the workhorse models of modern monetary macroeconomics. Since it is so widely studied and used to understand policy, it is a natural starting point for exploring meaningful deviations from rational expectations within DSGE models.

Households supply labor L_t to firms, earn nominal wages W_t , receive nominal profits Π_t from firms, hold nominal assets B_t , and consume C_t . There is a representative household. Because there is no heterogeneity among households, all households share the same deviations in beliefs. Following Coibion et al. (2012), households are subject to a rate spread shock b_t which drives a wedge between the return households earn on their nominal assets and the central bank’s risk free rate.

²Mavroeidis et al. (2014) raise concerns that in general, non-rational expectations may not satisfy the law of iterated expectations; the assumption taken here only says that whatever process that gives rise to a deviation from REE only has to admit a set of deviations Γ^i that can be represented as function of current and lagged variables X_t .

The representative household's maximization problem is:

$$\max_{C_t, B_{t+1}, L_t} E_0 \sum_{t=0}^{\infty} \beta^t \left(\frac{C_t^{1-\sigma_c}}{1-\sigma_c} - \gamma^{(1-\sigma_c)t} \frac{L_t^{1+\sigma_l}}{1+\sigma_l} \right) \quad (1)$$

$$\text{s.t. } P_t C_t + B_{t+1} \leq B_t b_{t-1} R_{t-1} + W_t L_t + \Pi_t \quad (2)$$

A **final-good producing firm** aggregates intermediate goods into a composite final good using a CES technology. It yields the standard demand function for intermediates relating the relative price to relative output and the elasticity of substitution between intermediates. Since this problem is intratemporal, deviations in this firm's beliefs over future variables are not relevant for equilibrium.

$$\max_{P_t, P_{i,t}} P_t Y_t - \int P_{i,t} Y_{i,t} di \quad (3)$$

$$\text{s.t. } Y_t = \left(\int Y_{i,t}^{\frac{1}{1+\lambda_p}} di \right)^{1+\lambda_p} \quad (4)$$

Intermediate-good firms hire labor L_t to produce output using a linear technology, where productivity grows at a constant rate γ . They sell to the final-good producing firm taking the demand function for their good as given. They are subject to Calvo pricing frictions with persistence parameter ξ_p . Since the standard Calvo setup entails a tradeoff between profit over the entire lifetime of the stuck price, beliefs over future profits influence the price set today. The firm discounts future profits using the marginal utility of consumption of the representative household. For a firm that can set a new price $P_{i,t}$ at time t , its maximization problem is:

$$\max_{P_{i,t}} E_t \sum_{j=0}^{\infty} \xi_p^j \Xi_{t+j} (P_{it} Y_{i,t+j} - \Theta_{t+j} Y_{i,t+j}) \quad (5)$$

$$\text{s.t. } Y_{i,t+j} = Y_{t+j} \left(\frac{P_{i,t}}{P_{t+j}} \right)^{-(1+\lambda_p)/\lambda_p} \quad (6)$$

In **equilibrium**, the central bank follows a Taylor Rule, and the market-clearing conditions are standard. The equilibrium is stationary when the appropriate variables are scaled by the growing level of productivity. The full set of equilibrium conditions, including all first-order conditions from households and firms, are listed in the Appendix.

The two equations in which expectations are relevant are the household Euler equation and the intermediate firm's pricing decision. Let \tilde{p} denote the firm's desired relative price at time t , let π_t denote inflation at time t , let r_t denote the real return on bonds, and let ξ_t denote the household's marginal utility of consumption. In a log-linearization, where variables with hats denote

log-deviations from steady state, these two equations are:

$$\frac{1}{1 - \xi_p \gamma \bar{\beta}} \hat{p}_t = \hat{v}_t + \frac{\xi_p \gamma \bar{\beta}}{1 - \xi_p \gamma \bar{\beta}} E_t [\hat{p}_{t+1} + \hat{\pi}_{t+1}] \quad (7)$$

$$\hat{\xi}_t = \hat{b}_t + \hat{r}_t + \tilde{E}_t [\hat{\xi}_{t+1} - \hat{\pi}_{t+1}] \quad (8)$$

What are all the variables to which deviations in beliefs might apply? For the representative household, the relevant variables are the household's future marginal utility of consumption ξ_{t+1} and future inflation. For the representative firm, they are future inflation and its future desired price. As shown in the Appendix, it is not possible to identify deviations in beliefs over all of these objects simultaneously in a linearized model. The objects that can be identified are 1) the household's deviation from REE about its own stochastic discount factor, and 2) the firm's deviation from REE on its future nominal desired price. Denote the first by $\hat{\Gamma}_t^H$ and the second by $\hat{\Gamma}_t^F$. A log-linearization and some algebra yields the standard New Keynesian model, augmented with the terms $\hat{\Gamma}_t$, or the log-difference of the deviations from their steady states.

$$\pi_t = \frac{(1 - \xi_p \beta)(1 - \xi_p)(1 + \sigma_l)}{\xi_p} (\hat{y}_t - \hat{a}_t) + \beta E_t [\pi_{t+1}] + (1 - \xi_p) \beta \hat{\Gamma}_t^F \quad (9)$$

$$\hat{y}_t = -\hat{b}_t - \hat{r}_t + E_t [\hat{y}_{t+1} + \hat{\pi}_{t+1}] + \hat{\Gamma}_t^H \quad (10)$$

$$\hat{r}_t = \rho_r \hat{r}_{t-1} + (1 - \rho_r)(\hat{r}_t^* + \phi_\pi \hat{\pi}_t + \phi_y (\hat{y}_t - \hat{y}_t^*)) + \hat{\varepsilon}_{r,t} \quad (11)$$

$$\hat{a}_t = \rho_a \hat{a}_{t-1} + \hat{\varepsilon}_{a,t} \quad (12)$$

$$\hat{b}_t = \rho_b \hat{b}_{t-1} + \hat{\varepsilon}_{b,t} \quad (13)$$

Note that households' deviated beliefs enter in the “New Keynesian IS curve,” since that curve is derived from the household Euler equation. Firms' beliefs enter in the “Phillips curve” since it that curve derived from firms' pricing decisions. It remains then to specify the laws of motion for the deviations. Assume that in a log-linearization, these are just linear function of lagged observables and lagged realizations of the deviations themselves.

$$\hat{\Gamma}_t^H = \sum_{k=0}^K \left(\Gamma_{y,k}^H \hat{y}_{t-k} + \Gamma_{r,k}^H \hat{r}_{t-k} + \Gamma_{\pi,k}^H \hat{\pi}_{t-k} + \Gamma_{\Gamma,k}^H \hat{\Gamma}_{t-k}^H \right) \quad (14)$$

$$\hat{\Gamma}_t^F = \sum_{k=0}^K \left(\Gamma_{y,k}^F \hat{y}_{t-k} + \Gamma_{r,k}^F \hat{r}_{t-k} + \Gamma_{\pi,k}^F \hat{\pi}_{t-k} + \Gamma_{\Gamma,k}^F \hat{\Gamma}_{t-k}^F \right) \quad (15)$$

2.3 GEEs and theoretical departures from REE

When taken to the data, the state-dependency of the belief deviations Γ^i is represented parsimoniously as a linear function of the history of observed variables and an autoregressive term. In this section, I motivate such a representation by appealing to several theories of departures from

REE: near-rational equilibrium (Woodford (2010)), sticky information (Mankiw and Reis (2002)), constant-gain learning (Evans and Honkapohja (2001)), sentiments (Angeletos and La'O (2013)), and diverse beliefs (Kurz (2012)). The GEE is too simple to nest all of these theories exactly, even in a linearized approximation. But the GEE can at a minimum approximate a wide range of theories, each with different foundations and implications. As such, the theories in this section should give the reader a non-exhaustive framework in which to interpret the empirical results later in this paper.

2.3.1 Woodford's robust monetary policy

Several recent papers (such as Woodford (2010) and Woodford (2013)) have explored the concept of deviations from rational expectations within the setting of robust policy setting. In these papers, the private sector (households and firms) hold beliefs that are subject to “distortion factors.” These distortion factors are not the result of a completely specified theory of expectations formation, but instead represent policymakers' uncertainty regarding the private sector's beliefs. Namely, the distortion factors are random variables $m_{t+1} > 0$ such that $E_t[m_{t+1}] = 1$, and $\tilde{E}_t[X_{t+1}] = E_t[m_{t+1}X_{t+1}]$. The extent of policymakers' uncertainty is captured by the relative entropy of the distortion factors under the expectations operator E_t : $R_t = E_t[m_{t+1}\log m_{t+1}]$. The smaller R_t , the more tightly the private sector's forecasts will coincide with those under the expectations operator E_t . In this sense, for small R_t , beliefs are constrained to be “close” to rational expectations. This is done so the central bank can solve for optimal monetary policy as a robust control problem: the policymakers acknowledge that private sector beliefs may be misspecified, but are not willing to put any additional structure on why or how beliefs may be misspecified.

While it is conceivable that real world private sector's beliefs may be in fact very far from rational expectations, the exercise performed by Woodford still teaches us how robust our model-based inferences are to bounded deviations from REE. The GEE also is meant to assess the extent to which REE contributes to misspecification. For that reason, I focus on models that are close to REE. To achieve that, beliefs in a GEE are expressed in terms of deviations from REE, and the empirical estimation of those deviations entails priors over the size those deviations that can be adjusted much like the relative entropy R_t can.

2.3.2 Sticky information

The sticky information model of Mankiw and Reis (2002), explored in more detail in a wide variety of applications, including Mankiw and Reis (2007), Coibion (2010), and Coibion and Gorodnichenko (2012), has largely been proposed as a replacement for sticky prices. However, to approximately nest a sticky information model as a GEE of the New Keynesian model, it is necessary to maintain the assumed nominal rigidities and relax only the assumption of rational expectations.

The theory of sticky information assumes that agents can only update their information set

each period with a fixed exogenous probability λ . Assume that the information arrival process is independent of the Calvo process. Firms that have not updated their information set continue to make decisions under their outdated information set. The probability that a given firm will have an information set k periods old is $\lambda(1 - \lambda)^k$. Aside from the dates of the information sets, there is no other departure from rational expectations, such as learning or higher-order beliefs. In other words, a firm whose information set was formed in period $t - k$ simply evaluates its expectations according to the operator $\tilde{E}_{t-k}^F = E_{t-k}$.

The linearized belief deviation $\hat{\Gamma}_t$ is a weighted average³ of the k -period-ahead forecast errors $(E_{t-k} - E_t) \left((1 - \beta\xi_p)\hat{\vartheta}_t + \beta\xi_p(\hat{p}_{t+1} + \hat{\pi}_{t+1}) \right)$. To the extent that a finite autoregressive process for $\hat{\Gamma}_t$ is a reasonable approximation⁴ to the weighted average of forecast errors, the GEE can approximately nest the predictions of a model with sticky information.

2.3.3 Learning

The assumption that agents learn about the functioning of the economy over time as they observe more data has been explored by many authors (e.g. Evans and Honkapohja (2001), Milani (2007), and Milani and Rajbhandari (2012)). Under this assumption, agents act as econometricians: they observe only a history of observations, and form beliefs over the laws of motion for those variables which change over time as the history of observations grows.

More concretely, the log-linearized perceived laws of motion⁵ for observed variables X_t from beliefs $\gamma^i(X_t, \epsilon_{t+1})$ are:

$$\hat{X}_{t+1} = \hat{\gamma}_X^i \hat{X}_t + \hat{\epsilon}_{t+1} \quad (16)$$

Each period, the agents' beliefs are described by an estimate of the autoregressive component of the law of motion⁶, $\hat{\gamma}_{X,t}^i$, and the precision of this point estimate. Denote this precision matrix $S_{\gamma_X,t}^i$. The precision matrix can be thought of as the inverse of the covariance matrix which describes agents' Bayesian prior over the autoregressive component of the law of motion, with the estimate $\hat{\gamma}_{X,t}^i$ representing the mean of that prior.

Over time, agents observe more realizations of X_t . With each new observation, they update their prior over the autoregressive component of the law of motion. Constant-gain learning implies

³Derivations are given in the Appendix.

⁴As an example of the lagged expectations rational expectations models considered in Wang and Wen (2006), the quantity $\hat{\Gamma}_t^F$ admits an $MA(\infty)$ representation in equilibrium. In practice, Wang and Wen (2006) show that truncating this to a finite-order $MA(N)$ representation reasonably approximates the equilibrium dynamics of a sticky information model. In contrast, this paper assumes that the deviations follow an autoregressive process with a small number of lags, which is both more parsimonious and more conveniently represented in state-space form.

⁵Here, expressed as log-deviations from steady state values of X .

⁶I have suppressed the coefficient $\hat{\gamma}_\epsilon^i$, under the assumption that agents, like econometricians, may not be able to directly observe every exogenous process, and therefore may not be able to directly identify structural shocks without additional assumptions.

a law of motion for both the mean and precision of the agents' prior beliefs that is a function⁷ of both the states which are hidden to the econometrician, $(\hat{\gamma}_{X,t}^i, S_{\gamma_X,t}^i)$, and the history of observed X_t .

The implied GEE belief deviations are the differences between the time-varying coefficient on the perceived law of motion and the actual law of motion: $\hat{\Gamma}_{X,t}^i = \hat{\gamma}_{X,t}^i - g_X^{GEE}$. Since the coefficient $\hat{\gamma}_{X,t}^i$ is a function of hidden states and lagged observables, the law of motion for $\hat{\Gamma}_{X,t}^i$ can be parsimoniously approximated as a linear function of lagged state variables both observed and unobserved by the econometrician. In the GEE estimated in this paper, the belief deviations are functions of only one unobserved state each: the belief deviation itself. The inclusion of the deviation as state variable (that is, assuming that the belief deviation is a function of lagged values of itself in addition to other state variables) allows the GEE to approximate theories that include hidden states, such as models of learning.

2.3.4 Diverse beliefs

Several theories feature heterogeneity of beliefs among agents. One such example is Kurz (2012). In this approach, there is not a representational household or firm holding mis-specified beliefs over the laws of motion of economic aggregates, but instead a wide range of perceived laws of motion held by the population of households and firms.

Kurz (2012) gives a version of the New Keynesian model with diverse beliefs, and shows that belief diversity introduces two major changes relative to rational expectations to the IS curve and the Phillips Curve:

$$\hat{y}_t = \bar{E}_t(\hat{y}_{t+1} + \hat{\pi}_{t+1}) + \Phi_t(\hat{c}) - \hat{r}_t - \hat{b}_t \quad (19)$$

$$\hat{\pi}_t = \frac{(1 - \xi_p)(1 - \beta\xi_p)}{\xi_p}(1 + \sigma_l)(\hat{y}_t - \hat{a}_t) + \beta\bar{E}_t(\hat{\pi}_{t+1} + (1 - \xi_p)\Phi_t(\hat{p})) \quad (20)$$

There are two sources of deviations. First, there are wedges Φ_t . The differences between individual forecasts of aggregate conditions and individual forecasts of individual decisions gives rise to heterogeneity in individual consumption and pricing decisions. Aggregating these individual consumption and pricing decisions gives rise to the wedges Φ_t : $\Phi_t(\hat{c}) = \int_0^1 \left(\tilde{E}_t^j(\hat{c}_{t+1}^j) - \bar{E}_t^j(\hat{c}_{t+1}) \right) dj$, where $j \in [0, 1]$ is the index over the continuum of heterogeneous households, \hat{c}_t represents aggregate consumption, and \hat{c}_t^j represents individual j 's consumption⁸. Second, the average belief $\bar{E}_t = \int_0^1 \tilde{E}_t^j dj$

⁷Namely,

$$\hat{\gamma}_{X,t}^i = \hat{\gamma}_{X,t-1}^i + \bar{g}(S_{\gamma_X,t}^i)^{-1} \mathbf{X}_t(\mathbf{X}_{t-1} - (\hat{\gamma}_{X,t-1}^i)' \mathbf{X}_t)' \quad (17)$$

$$S_{\gamma_X,t}^i = S_{\gamma_X,t-1}^i + \bar{g}(\mathbf{X}_t \mathbf{X}_t' - S_{\gamma_X,t-1}^i) \quad (18)$$

where \mathbf{X}_t is the vector of observations $(X_1 \dots X_t)$, and \bar{g} is the constant-gain parameter controlling the weight agents give to additional informations in updating their beliefs.

⁸ $\Phi_t(\hat{p})$ is defined similarly, aggregating over firms and their desired prices \tilde{p} .

is not an expectation with respect to a single probability distribution, and hence does not obey the law of iterated expectations.

Theorem 2 of Kurz (2012) shows that there are finite state-space representations of both the wedges Φ_t and the difference between the mean expectation \bar{E}_t and the expectation with respect to the actual law of motion E_t . Since there is a well-defined state-space representation for these wedges, they are observationally equivalent (in a log-linear world) to the GEE's assumption of expectations by a representative household and a representative firm with respect to a mistaken law of motion, where the law of iterated expectations applies. Therefore the GEE is able to nest settings where there are a diversity of beliefs, and the deviation represents both the mis-specification of beliefs themselves and the consequences of aggregation across a heterogeneous group of agents.

2.3.5 Higher-order beliefs

As another example of heterogeneous beliefs, consider high-order beliefs, that is, settings where agents' beliefs about other agents' beliefs can drive macroeconomic dynamics (e.g. Angeletos and La'O (2009) and Angeletos et al. (2014)).

Specifically, Angeletos et al. (2014) assume imperfect knowledge about shocks to productivity. Each household observes a private signal about each innovation to productivity, but assumes that the private signals that other households get are biased. Specifically, if x_{it} is the private signal that agent i receives about TFP a_t , then that agent i believes that signal is unbiased: $x_{it} = a_t + \epsilon_{it}$, where $\epsilon_{it} \sim N(0, \sigma^2)$, but agent i also believes that the signals that other agents get are not unbiased: $\epsilon_{jt} \sim N(\psi_t, \sigma^2)$.

This particular treatment of high-order beliefs yields a tractable equilibrium, where the policy rule for the aggregate capital stock is only altered by the presence of a single additional state variable, ψ_t , which effectively summarizes the total effects of all higher-order beliefs. Furthermore, the law of motion of ψ_t is a first-order autoregressive process with a shock ζ_t that is independent to other structural shocks in the economy.

The example illustrates two points motivating the specification of the GEE. First, there may be shocks (such as ζ_t) which drive the belief deviations that are also independent of other shocks in the economy. Second, and more fundamentally, the misspecified beliefs in a GEE may not necessarily be on the laws of motion regarding observables, but also potentially on unobservable quantities such as other agents' beliefs.

2.4 GEE: altering the impacts and propagation of shocks

The belief deviations in a GEE add a new set of shocks to the benchmark model, and they change the propagation of existing shocks. This section gives a sense of the range of possible changes in the equilibrium propagation of shocks within the benchmark NK model.

Consider a simplification of the New Keynesian model introduced earlier: suppose that there

is no autocorrelation in the Taylor rule ($\rho_R = 0$), and suppose that all deviations are functions of contemporaneous inflation, contemporaneous output, and a shock alone. That is, equations (14) and (15) become:

$$\hat{\Gamma}_t^H = \Gamma_y^H \hat{y}_t + \Gamma_\pi^H \hat{\pi}_t + \epsilon_{\Gamma^H,t} \quad (21)$$

$$\hat{\Gamma}_t^F = \Gamma_y^F \hat{y}_t + \Gamma_\pi^F \hat{\pi}_t + \epsilon_{\Gamma^F,t} \quad (22)$$

Some algebraic manipulation (in the appendix) reveals a set of equilibrium conditions that are identical to that of the canonical New Keynesian model, with the exception of the deviations Γ .

$$\pi_t = \kappa(\hat{y}_t - \hat{a}_t) + \beta E_t [\pi_{t+1}] + (1 - \xi_p) \beta \hat{\Gamma}_t^F \quad (23)$$

$$\hat{y}_t = -\hat{b}_t - \hat{r}_t + E_t [\hat{y}_{t+1} + \hat{\pi}_{t+1}] + \hat{\Gamma}_t^H \quad (24)$$

$$\hat{r}_t = \rho_r \hat{r}_{t-1} + (1 - \rho_r)((\rho_a - 1)\hat{a}_t + \phi_\pi \hat{\pi}_t + \phi_y(\hat{y}_t - \hat{a}_t)) + \hat{\epsilon}_{r,t} \quad (25)$$

$$\hat{a}_t = \rho_a \hat{a}_{t-1} + \hat{\epsilon}_{a,t} \quad (26)$$

$$\hat{b}_t = \rho_b \hat{b}_{t-1} + \hat{\epsilon}_{b,t} \quad (27)$$

$$\hat{\Gamma}_t^H = \Gamma_y^H \hat{y}_t + \Gamma_\pi^H \hat{\pi}_t + \epsilon_{\Gamma^H,t} \quad (28)$$

$$\hat{\Gamma}_t^F = \Gamma_y^F \hat{y}_t + \Gamma_\pi^F \hat{\pi}_t + \epsilon_{\Gamma^F,t} \quad (29)$$

The parameter $\Gamma_\pi^H = \frac{\partial}{\partial \pi} \Gamma^H(\bar{\pi})$ is the elasticity of the belief deviation in the household Euler equation with respect to current inflation, evaluated at steady state. Holding households' beliefs on their future real consumption constant, and abstracting from aggregation wedges or higher-order beliefs, this is the deviation of households' beliefs about future inflation with respect to current inflation. Consider the case where $\Gamma_\pi^H > 0$. When households observe positive inflation, they expect more inflation tomorrow than will actually occur. Accordingly, households will believe they are facing a lower real rate of return on their saving than they actually will.

Similarly, holding firms' beliefs on their future marginal costs and households' marginal utilities of consumption constant, and likewise abstracting from aggregation wedges or higher-order beliefs, the parameter Γ_π^F is the elasticity of the deviation of firms' beliefs about future inflation with respect to current inflation. Consider the case where $\Gamma_\pi^F > 0$. When firms observe inflation today, they will expect more inflation tomorrow than what will actually happen. As a result, firms have an incentive to set a higher price today, believing that they will face higher inflation in the future than they actually will.

The solution for GEE will be a set of policy functions for $(\hat{y}_t, \hat{\pi}_t, \hat{r}_t)$ that are linear in current exogenous states. For example, $\hat{y}_t = g_a^y \hat{a}_t + g_b^y \hat{b}_t + g_r^y \hat{\epsilon}_{r,t} + g_H^y \hat{\epsilon}_{\Gamma^H,t} + g_F^y \hat{\epsilon}_{\Gamma^F,t}$. The method of undetermined coefficients yields algebraic solutions for the coefficients g^9 : The full algebraic solutions

⁹Note that there are no lagged state variables in this radically simple example. Each endogenous quantity is entirely a function of contemporaneous shocks.

for the policy functions are given in the appendix. For illustration, consider a technology shock.

$$g_a^y = 1 + (\Gamma_y^H(1 - \beta\Gamma_\pi^F(1 - \xi_p)) + \beta\Gamma_y^F(1 - \xi_p)(\Gamma_\pi^H - \phi_\pi))/c \quad (30)$$

$$g_a^\pi = (\Gamma_y^H\kappa + \beta\Gamma_y^F(1 - \xi_p)(1 + \phi_y))/c \quad (31)$$

$$g_a^r = \rho_a - 1 + \phi_y(g_a^y - 1) + \phi_\pi g_a^\pi \quad (32)$$

In the third equation above we can see the effects of an equilibrium condition that the GEE does not relax: the Taylor Rule. In a GEE, the endogenous response of the interest rate to shocks, *conditional on the responses of output and inflation*, is the same as under REE. The first and second equations, however, show how the state-dependency of the belief deviations (Γ_π^F , Γ_y^F , etc) alter the responses of output and inflation to a technology shock. Figure 1 plots the coefficients of the policy rule describing the responses of output, inflation, and the interest rate to shocks to technology, monetary policy, and beliefs as functions of the elasticities Γ_π^H , Γ_π^F , Γ_y^H , and Γ_y^F , under standard parameter values. When the parameters Γ_π^H , Γ_π^F , Γ_y^H , and Γ_y^F are all equal to zero, REE obtains.

Increasing Γ_π^H is observationally equivalent to supposing that when households observe inflation above steady state in period t , they believe that inflation in $t + 1$ will be greater than what it will actually be. Increasing Γ_π^F has the same interpretation, but for firms. Increasing Γ_y^H or Γ_y^F is observationally equivalent to a similarly mistaken view about future inflation among households or firms, but conditional on *output* being above steady state in period t .

Consider a positive technology shock. Under REE, a technology shock has no impact on inflation or the output gap, since the Taylor Rule stabilizes both. In contrast, a technology shock can induce movements in both inflation and the output gap under GEE. As long as Γ_y^H and Γ_y^F are both equal to zero, the equilibrium response of inflation will be zero, no matter the values of Γ_π^H and Γ_π^F . This result holds because the central bank never allows any amount of inflation in response to technology shocks in this case, so there is no role for the belief deviation based on observed level of inflation following a technology shock. However, once either Γ_y^H or Γ_y^F are different from zero, the increase of output in time t following a technology shock will lead either households or firms to believe there will be greater inflation in time t than what will actually prevail, regardless of how well the Taylor Rule works to stabilize inflation. If $\Gamma_y^H > 0$, then at every point in time, households consistently over-estimate inflation, given that output is high from the technology shock. Effectively, they under-estimate the real return on their savings, and increase their consumption in period t over where it would be under REE. If $\Gamma_y^F > 0$, firms over-estimate the inflation that follows an increase in output. In this case, a technology shock will lead firms to set a higher initial price than they would have under REE. This causes positive inflation contemporaneously with the technology shock. The belief deviations have introduced a time-varying wedge that the Taylor Rule cannot fully accommodate; inflation arises even in response to a technology shock.

Consider a positive monetary policy shock ϵ_r . Under REE, the higher interest rate induces

households to save more, and output falls. Inflation also falls as the interest rate increases. Since the shocks have no persistence, households would expect no inflation in the period after the shock under REE. But when $\Gamma_{\pi}^H > 0$, households observe a fall in inflation contemporaneously and believe the inflation rate will be even lower than will actually obtain. As a result, households expect further disinflation and save more than they would under REE. With greater saving, the fall in output is larger. At the same time, the contemporaneous disinflation is larger. Because the responses of both output and inflation are larger, the monetary policy shock raises the interest rate less than it would under REE.

The effects of a positive monetary policy shock also change when firms' beliefs deviate from REE. When $\Gamma_{\pi}^F > 0$, firms observe a fall in inflation contemporaneously and believe the inflation rate will be lower than will actually obtain in the next period. As a result, the firms who can update their prices immediately will set a lower price than they would under REE. Firms' belief deviations amplify the contemporaneous disinflation from a monetary policy shock. This stronger disinflation results in a weaker contemporaneous response of interest rates to the monetary policy shock through the Taylor rule, giving households less incentive to save. This causes the fall in output to be more moderate than under REE.

Finally, consider shocks to the belief deviations themselves. The intuition for these shocks is very similar to belief deviations that arise as consequences of contemporaneous changes in output or inflation: when households suddenly and exogenously believe that inflation is going to be high in the future, then they will save less and consume more for every nominal interest rate. This increases current output, driving firms to raise prices and the central bank to raise the interest rate. When firms suddenly and exogenously believe that inflation is going to be high in the future, this leads to an increase in the price set by firms who can adjust their prices; the central bank raises the interest rate in response, which leads to a decline in consumption contemporaneously.

Figure 1 illustrates the means by which the elasticities Γ are identified from the data: they both change the propagation of existing shocks, and provide a new set of shocks to the model. While the deviations do not necessarily have to represent mistaken beliefs on inflation (for example, they could be the consequence of aggregation wedges or higher-order beliefs), interpreting the deviations as mistaken beliefs directly yields changes in equilibrium dynamics that are intuitive.

3 Estimation

3.1 Estimation strategy

The GEE for a given model is a parsimonious representation of many possible deviations from rational expectations. The parameters that vary between the various non-REE nested within the GEE are the elasticities of the belief deviations with respect to lagged observables and lagged deviations. The empirical strategy of this paper is to estimate these elasticities using Bayesian

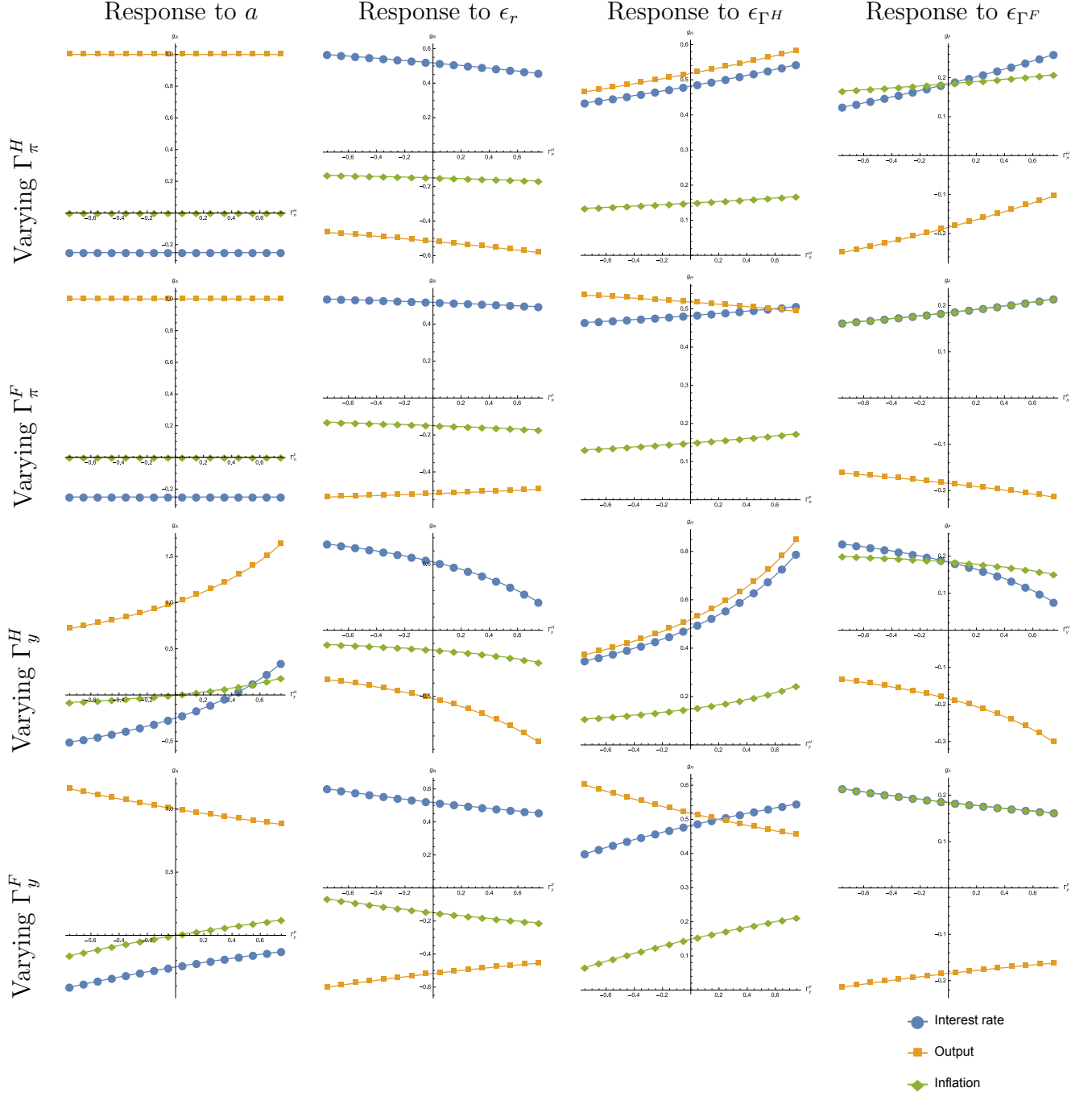


Figure 1: Coefficients on the policy functions for the interest rate \hat{r} , output \hat{y} and inflation $\hat{\pi}$ to a technology shock \hat{a} , a monetary policy shock ϵ_r , shocks to household belief deviations ϵ_{Γ^H} , and shocks to firm belief deviations ϵ_{Γ^F} . These coefficients are plotted as functions of Γ_π^H , the elasticity of the deviation of households' beliefs on future inflation with respect to current inflation, Γ_π^F , the elasticity of the deviation of firms' beliefs on future inflation with respect to current inflation, Γ_y^H , the elasticity of the deviation of households' beliefs on future inflation with respect to current output, and Γ_y^F , the elasticity of the deviation of firms' beliefs on future inflation with respect to current output, holding the other agents' belief deviations at 0. $\Gamma_\pi^H = \Gamma_\pi^F = \Gamma_y^H = \Gamma_y^F = 0$ represents rational expectations.

methods (as in Smets and Wouters (2003), Schorfheide (2011), and Guerrón-Quintana and Nason (2012)), while calibrating the rest of the model parameters to values estimated under REE.

As in any empirical exercise, identification is a central concern. Recent work (see Canova

and Sala (2009) and Schorfheide (2011)) has addressed the particular concerns that arise in the Bayesian estimation of DSGE models. I rely upon the techniques of Iskrev (2010) to numerically assess whether there is sufficient information in the density of the data under the model to identify the elasticities of the belief deviations.

3.1.1 Comparison with DSGE-VAR

The set of equations (9) through (15) defines the GEE. These seven equations combine the structural content of the New Keynesian model with the reduced form specification of the belief deviations. The reduced form specification of the belief deviations effectively relaxes the cross-equation restrictions imposed by the structural equations. Another empirical strategy, the DSGE-VAR of Del Negro and Schorfheide (2004) and Del Negro and Schorfheide (2009), also combines the tight restrictions of a structural model with a more richly parametrized specification. Since the DSGE-VAR has also been used to explore DSGE model misspecification, it is worthwhile to highlight the differences between the two approaches, both in terms of assumptions and in terms of application.

To estimate a DSGE-VAR, one estimates a VAR using Bayesian methods, while imposing a prior on these coefficients by using synthetic data simulated from a DSGE model. If the prior accords heavy weight to the data simulated from the DSGE model, then the estimated VAR coefficients will be very close to those implied by the DSGE model. In this case, the “deviations” between the DSGE-VAR and the coefficients implied by the DSGE model alone will be small. The less weight accorded by the prior to the DSGE restrictions, the larger the “deviations” will be and the more the DSGE-VAR will resemble a VAR.

The GEE also relaxes the cross-equation restrictions of a DSGE model under REE, but does so in a way that is more restrictive than a DSGE-VAR. The first difference is that the GEE fixes the parameters of the DSGE model and the prior weight on the deviations¹⁰. This is in contrast to the DSGE-VAR approach (e.g. Cole and Milani (2014)), which estimates these parameters together with the deviations. The second difference is where the deviations appear in the set of equations to be estimated. In the DSGE-VAR, cross-equation restrictions in all equations are relaxed towards a reduced-form linear relationship. The GEE approach preserves more cross-equation restrictions in two ways: a) since the GEE does not introduce any deviations to equations which capture exclusively contemporaneous relationships, all equations without expectations are preserved; and b) depending on further assumptions on how heterogeneous beliefs are allowed to be between agents, and which variables agents are allowed to have non-REE beliefs over, there could be extensive cross-equation restrictions placed on the deviations themselves. The usefulness of these additional restrictions is to give the data greater leeway only in the specific equations that would be affected by a potential departure from REE.

¹⁰I leave for future work an exploration of the data’s guidance towards the selection of the prior weight.

For example, consider the household’s Euler equation:

$$\hat{y}_t = -\hat{r}_t + E_t[\hat{y}_{t+1} + \hat{\pi}_{t+1}] + \hat{\Gamma}_t^H - \hat{\epsilon}_{b,t} \quad (33)$$

With uninformative priors over the elasticities of Γ^H with respect to observables, those elasticities are allowed to take any value, effectively overwhelming any identification from the structural parameters in the rest of the equation. As a result, under uninformative priors, every equation with an expectation in it (like the Euler equation above) would become an unrestricted linear function of lagged observables and hidden minimum state variables, with the belief wedge introduced as another hidden state. The collection of equations describing the entire model would assume the form of a kind of hybrid of the original model and a VAR with hidden states, where only the equations that lack any expectations operators would retain any structural content.

With relatively tight priors, I ensure that what I estimate is close to 1) the original model and 2) REE. Tight priors effectively put a small state-contingent noise term in every equation that has an expectation in it. The robust policymaker of Woodford (2010) puts an entropy ball around REE and performs robust control over that set; this exercise puts a prior around REE and estimates the state-contingency of the deviations.

3.2 Estimating GEE under canonical NK

The goal is to estimate the state-contingency of a wedge in beliefs that is small enough to be a “local” departure from REE. First, I estimate a set of parameters (such as the discount factor, the Calvo price survival probability etc.), under the assumption of REE. Table 3 describes this estimation. For the sake of consistency with the medium-scale DSGE model to follow, the priors for these parameters are taken from Smets and Wouters (2007). Second, I estimate the elasticities of the belief deviations in GEE, fixing the non-GEE parameters to their posterior modes in Table 3. For identification of the GEE, I calibrate the standard deviations of the shocks to the belief deviations to be the same order of magnitude as those of the other shocks.

The belief deviations could arise from a variety of sources. In the most generic interpretation, it is plausible that the deviations differ significantly between households and firms, because 1) the beliefs are over different objects (firms over their desired relative price, and households over their marginal utility of consumption) 2) the information frictions or aggregation properties between the two sets of agents might differ significantly. Recognizing the wide range of possibilities for the presence of belief deviations, I estimate the elasticities for $\hat{\Gamma}^F$ and $\hat{\Gamma}^H$ separately.

I also entertain an alternative identification of the belief deviations that supposes that the deviations are the same across the two equations. I call these deviations “unified beliefs,” and denote them with $\hat{\Gamma}^U$. One immediate interpretation of why beliefs deviations might be unified over different agents’ beliefs over different objects is that these deviations represent a shared belief over future inflation, the variable that is common to both belief deviations in the NK model.

The posteriors for the elasticities of the belief deviations are described in Table 3. The prior for each elasticity is normal, with mean 0 and standard deviation 0.20. These parameters are identified under the moment rank conditions described in Iskrev (2010) and Ratto (2008). The observation equations and data are also taken from Smets and Wouters (2007).

Among the elasticities estimated for the NK model, only a few take on values where the 90% credible interval does not include zero. Under the assumption of unified beliefs between households and firms, these elasticities are with respect to current inflation (negative), lagged inflation (positive), and twice-lagged inflation (positive). Once beliefs are allowed to differ between households and firms, the pattern changes. The elasticity of household belief deviations with respect to current inflation is significantly negative, as in the case with unified beliefs, but the same does not hold for firm belief deviations. Firm belief deviations have elasticities that are significantly different from zero for lagged inflation (positive), current output (positive), and lagged output (negative). Additionally, firm belief deviations are persistent, with an autoregressive coefficient of 0.848. Interpreted as evidence on the most promising empirical direction for a theory of departures from REE, these point estimates represent the state-contingency of the belief deviations that the data favor.

Since these deviations operate over the timespan of every shock, their net effect in equilibrium is not immediately obvious, but they do cause the equilibrium response to shocks to differ significantly from REE. Figure 10, in the Appendix, gives the impulse responses of output, the interest rate, and inflation to each of the shocks, both under the estimated GEE (under beliefs that are not unified across agents) and under the benchmark REE. The shock whose effects depart the most from REE is the technology shock, shown in Figure 2. The basic intuition for this result is that, under REE, the Taylor Rule keeps variation in the output gap, the interest rate, and inflation very small under technology shocks. However, the belief deviations introduce a time-varying wedge. As a result, technology shocks give rise to fluctuations in the interest rate and inflation that are much larger than those under REE. At the same time, the impacts of technology shocks on output are dampened.

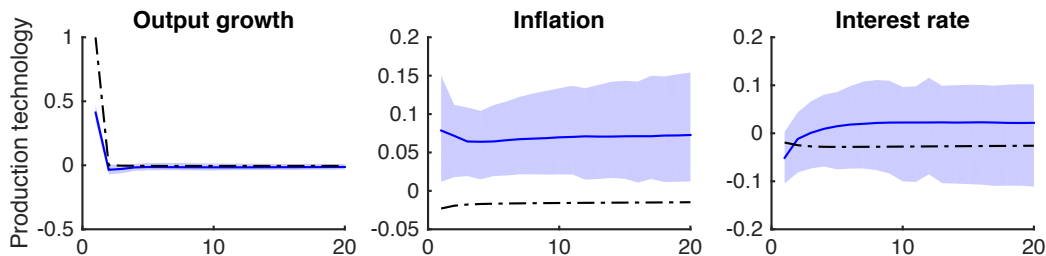


Figure 2: Impulse responses to a technology shock under GEE (with posterior credible interval shaded) and REE (dashed), New Keynesian benchmark. The estimated sensitivity of belief deviations to lagged output, combined with a very high estimated persistence to technology, yields large and persistent deviations of inflation and the interest rate from the responses under REE.

The different role of structural shocks is also evident in Table 1, which gives unconditional

variance decompositions of each variable under GEE (both with beliefs that are unified and those that are not) and REE. When beliefs are assumed to be unified across agents, the largest changes relative to REE are in the variance decompositions of inflation and the interest rate: in both cases, shocks to beliefs themselves absorb most of the variance of each variable attributed to shocks to the rate spread. The variance decomposition of output growth is effectively unchanged. In this case, the empirical significance of introduction of belief deviations has been to provide a new source of shocks.

When beliefs are allowed to differ across agents, the pattern changes. Shocks to both belief deviations explain no more than 12% of the variances of inflation and the interest rate. A much larger change comes through the belief deviations' role in the propagation of structural shocks. The change in the propagation of technology shocks causes the GEE to ascribe significantly more of the variance of both inflation and the interest rate to technology shocks (for inflation, 67.6% in GEE versus 9.5% in REE; for the interest rate, 54.4% in GEE versus 14.0% in REE). At the same time, the share of the variance in output growth attributed to technology shocks declines from 87.5% to 14.6%, with shocks to firms' beliefs taking up almost all of the difference. These variance decompositions provide evidence that the GEE approaches the data more closely than the REE *both* by allowing for additional shocks¹¹, *and* by modifying the propagation of other structural shocks.

3.3 Estimating GEE under Smets Wouters (2007)

Small New Keynesian and real business cycle models are known to lack internal propagation, and the estimated GEE seizes upon this as a source of misspecification. What role do belief deviations play when other mechanisms known to achieve internal propagation are added? To answer that question, this section estimates a GEE for a medium-scale DSGE model: Smets and Wouters (2007) (hereafter SW). This section shows the extent to which the GEE for a New Keynesian model is a substitute for the additional mechanisms in SW, and the extent to which the GEE for SW addresses some of the misspecification that persists in SW.

SW has been widely studied in the literature, and the derivation of its log-linearization under REE is widely available. However, in deriving the GEE for SW, it is necessary to pay special attention to where beliefs over future variables enter agents' problems, and how these deviations alter the widely-known log-linearized equations. For this reason, the appendix gives the exposition and linearization of the GEE of SW. The five log-linearized equations where belief deviations enter are the following:

¹¹The belief deviation shocks are identified separately from the other shocks in large part through the presence of the belief deviations as hidden states (guaranteed through their autoregressive structure).

Parameter	Mean	90% CI†	
NKGEE with unified beliefs			
$\Gamma_{\pi,0}^H$	-1.066	-1.269	-0.909
$\Gamma_{\pi,1}^H$	0.343	0.091	0.620
$\Gamma_{\pi,2}^H$	0.642	0.390	0.887
NKGEE with multiple beliefs			
Household savings beliefs			
$\Gamma_{\pi,0}^H$	-0.493	-0.811	-0.130
Good pricing beliefs			
$\Gamma_{\Gamma^F}^F$	0.848	0.556	1.144
$\Gamma_{\pi,1}^F$	0.524	0.184	0.822
$\Gamma_{y,0}^F$	0.577	0.464	0.693
$\Gamma_{y,1}^F$	-0.480	-0.696	-0.285

Parameter	Mean	90% CI†	
SWGEE with unified beliefs			
$\Gamma_{\pi,0}^H$	-1.161	-1.272	-1.043
$\Gamma_{\pi,2}^H$	0.484	0.222	0.722
$\Gamma_{c,0}^H$	-0.222	-0.415	-0.054
$\Gamma_{i,0}^H$	0.075	0.005	0.162
$\Gamma_{r,0}^H$	-0.456	-0.705	-0.211
SWGEE with multiple beliefs			
Household savings beliefs			
$\Gamma_{l,2}^H$	0.264	0.036	0.493
$\Gamma_{i,0}^H$	0.297	0.094	0.504
Good pricing beliefs			
$\Gamma_{\pi,1}^F$	0.436	0.105	0.798
$\Gamma_{\pi,2}^F$	0.424	0.195	0.677
$\Gamma_{i,0}^F$	0.297	0.058	0.458
$\Gamma_{i,1}^F$	-0.495	-0.731	-0.111
$\Gamma_{i,2}^F$	0.192	0.036	0.304
$\Gamma_{w,1}^F$	-0.248	-0.468	-0.020
Capital pricing beliefs			
Wage setting beliefs			
$\Gamma_{w,0}^W$	-0.573	-0.812	-0.277
$\Gamma_{w,2}^W$	0.275	0.074	0.476

Figure 3: Posterior distributions for belief deviations. All prior distributions are Normal, with mean 0 and standard deviation 0.2. †Credible Interval. Only shown are estimates for which 90% credible interval does not contain 0.

$$(1 + \gamma\bar{\beta}\iota_p)\hat{\pi}_t = \iota_p\hat{\pi}_{t-1} + \frac{(1 - \xi_p\gamma\bar{\beta})(1 - \xi_p)}{\xi_p} \left(\frac{\lambda_p}{1 + \lambda_p} \hat{\lambda}_{p,t} + \hat{v}_t \right) + \gamma\bar{\beta}E_t[\hat{\pi}_{t+1}] + \gamma\bar{\beta}\hat{\Gamma}_t^{F,\bar{p}+\pi} \quad (34)$$

$$-\sigma_c(1 + h/\gamma)\hat{c}_t = -\sigma_c(h/\gamma)\hat{c}_{t-1} + (1 - h/\gamma)(\hat{b}_t + \hat{r}_t - E_t[\hat{\pi}_{t+1}] - \hat{\Gamma}_t^{H,\pi-\xi}) \\ + (\sigma_c - 1)(w^h L/c)(E_t[\hat{L}_{t+1}] - \hat{L}_t) - \sigma_c E_t[\hat{c}_{t+1}] \quad (35)$$

$$\hat{i}_t = \frac{1}{S''\gamma^2}(\hat{Q}_t + \hat{\mu}) + \frac{1}{1 + \bar{\beta}\gamma}(\hat{i}_{t-1} + \bar{\beta}\gamma E_t[\hat{i}_{t+1}]) + \frac{\bar{\beta}\gamma}{1 + \bar{\beta}\gamma}\hat{\Gamma}_t^{H,i} \quad (36)$$

$$\hat{Q}_t = -\hat{b}_t - (\hat{r}_t - E_t[\hat{\pi}_{t+1}] - \hat{\Gamma}_t^{H,\pi-\xi}) + \bar{\beta}E_t[r^k(\hat{r}_{t+1}^k) + \hat{Q}_{t+1}(1 - \delta)] + \bar{\beta}\hat{\Gamma}_t^{H,Q} \quad (37)$$

$$(1 + \bar{\beta}\gamma)\hat{w}_t - \hat{w}_{t-1} = \frac{(1 - \xi_w)(1 - \xi_w\gamma\bar{\beta})}{\xi_w} \left(\frac{\lambda_w}{1 + \lambda_w} \hat{\lambda}_{w,t} + \frac{1}{1 - h/\gamma}(\hat{c}_t - (h/\gamma)\hat{c}_{t-1}) + \sigma_l \hat{L}_t - \hat{w}_t \right) \\ + \gamma\bar{\beta}E_t[\hat{w}_{t+1} + \hat{\pi}_{t+1}] + \iota_w\hat{\pi}_{t-1} - (1 + \gamma\bar{\beta}\iota_w)\hat{\pi}_t + \gamma\bar{\beta}\hat{\Gamma}_t^{L,\bar{w}+\pi} \quad (38)$$

Identification analysis by Iskrev (2010) shows that identification can be obtained by setting $\hat{\Gamma}_t^{H,i} = 0$. As in the NK model, it is generally not possible to separately identify belief deviations over individual endogenous variables. Also as with the NK model, I consider two alternatives to identifying belief deviations. The first alternative is that all belief deviations have the same state-contingency: $\hat{\Gamma}_t^U = \hat{\Gamma}_t^{F,\bar{p}+\pi} = \hat{\Gamma}_t^{H,\pi-\xi} = \hat{\Gamma}_t^{H,Q} + \hat{\Gamma}_t^{H,\pi-\xi} = \hat{\Gamma}_t^{L,\bar{w}+\pi}$. If one assumes that all agents have rational expectations on all objects besides inflation, and that there aren't any aggregation wedges or high-order beliefs, unifying beliefs implies that all agents have the same beliefs on inflation. The second alternative is that there are four different belief deviations: $\hat{\Gamma}_t^F = \hat{\Gamma}_t^{F,\bar{p}+\pi}$, $\hat{\Gamma}_t^H = \hat{\Gamma}_t^{H,\pi-\xi}$, $\hat{\Gamma}_t^Q = \hat{\Gamma}_t^{H,Q} + \hat{\Gamma}_t^{H,\pi-\xi}$, $\hat{\Gamma}_t^W = \hat{\Gamma}_t^{L,\bar{w}+\pi}$. This assumption allows agents to have non-rational expectations over more than just inflation, and allows agents to disagree.

3.3.1 Coefficient estimates

One of the goals of this paper is to see to what extent NK GEE substitutes for the mechanisms of SW REE. To ensure the two models are as comparable as possible, I use a version of the NK model that is just a special case of the SW model, and I estimate the non-GEE structural parameters of the two models under the same prior distributions. Some structural parameters appear in SW but not in NK, such as the elasticity of the adjustment cost on investment. I do not estimate these parameters for use in the NK model. As a result, the values I estimate for the parameters that appear in both models differ between models. Table 3 gives the priors and posteriors for the structural parameters estimated under REE.

Table 3 gives the posteriors for the GEE elasticities in the SW model. One might expect many of these posteriors to differ from those estimated in the NK GEE, since the NK GEE is partially taking up some of the misspecification that SW model was engineered to address. But some coefficients are similar. Under the assumption of unified beliefs, the elasticity of the deviation with respect to

current inflation is negative, and with respect to twice-lagged inflation is positive. This pattern is also observed in the NK GEE under unified beliefs. This suggests that whatever misspecification those coefficients were ameliorating in the NK GEE under unified beliefs, that misspecification was not entirely ameliorated with the introduction of the additional mechanisms in SW that give rise to internal persistence.

One major difference between the NK GEE and the SW GEE is the greater range of elasticities that can be identified. For example, investment does not appear in the NK model, but in the SW model, the estimated elasticity of unified beliefs with respect to current investment is significantly greater than zero (though small). As an additional example, the estimated elasticity of unified beliefs with respect to current consumption is significantly less than zero. In the NK model, consumption and output are identical, so the elasticities with respect to output and consumption cannot be separately identified.

In the SW GEE with multiple beliefs, the elasticities that are estimated to be significantly different from zero are generally different from those in the SW GEE with unified beliefs, and both NK GEE models. This suggests that the introduction of the SW mechanisms, combined with the ability to estimate each belief deviation independently of the rest, allows the GEE to address a more specific source of misspecification than in the other specifications. Namely, in the NK model, the elasticity with respect to current output is positive and lagged output is negative; in the SW model, the corresponding elasticities for investment, and not output, exhibit the same signs. This is likely a consequence of the fact that the NK GEE does not exploit variation in observed investment, and instead attributes this variation in the belief deviation to variation in output instead.

However, the SW GEE with multiple beliefs *does* resemble the NK model with multiple beliefs for the elasticities estimated for firms' beliefs. Specifically, the coefficients estimated for lagged inflation are significantly positive in both specifications. Abstracting from aggregation wedges and heterogeneous beliefs, this could amount to evidence that the data favor deviations from REE where agents think inflation will be more persistent than it actually is.

3.3.2 Spectral Analysis

There are many ways to assess how well a model approaches the data. In this section, I follow a number of authors and compare the implications of the model with the data in the frequency domain. Watson (1993), King and Watson (1996) show that a baseline RBC model's lack of endogenous propagation mechanisms implies that it requires a volatile and persistent exogenous process for technology to replicate the spectral properties of the data. Wen (1998) shows that adding habit to labor supply and an externality to labor in the production function can provide an endogenous mechanism to help ameliorate this shortcoming. SW contains many additional mechanisms in this vein. Tkachenko and Qu (2012) assess the SW model specifically and find that while the SW model performs well in the spectral densities of individual series, the pairwise spectral coherences are not always reproduced well by the model. I turn to the spectral methodologies used by these

authors to answer two questions: 1) To what extent does relaxing REE in a NK model substitute for the structural mechanisms which distinguish SW from NK? 2) To what extent does relaxing REE improve upon the SW model's ability to reproduce features of the data?

The statistical objects of interest, both in the data and in the model, are the spectral density and spectral coherence. For a general introduction to spectral approaches to time series, see Hamilton (1994). For greatest comparability with similar studies, I follow Tkachenko and Qu (2012) both in notation and in the code to calculate the objects of interest. The Appendix gives expanded notes on the derivation of these objects.

The **spectral density** of a time series is an estimate of how the total variance of a series can be decomposed over frequencies. Figure 6 in the appendix indicates the nonparametric (smoothed) spectral density estimates for each series in the data and its 95% confidence band. Superimposed are the theoretical spectral densities for the NK REE, NK GEE, SW REE, and SW GEE models at the posterior modes of the estimated parameters. Since the NK models do not give a prediction for investment growth, this series is omitted for these models. Business cycle frequencies are indicated with vertical dashed lines. The NK models (both REE and GEE) have severe problems replicating the observed spectral densities of all series (in particular, consumption growth, output growth, and wage growth, reproduced in excerpt in Figure 4), and the SW models (including REE) largely ameliorate them. The performance of SW REE was discussed by Tkachenko and Qu (2012); the GEE improves neither NK nor SW here. One possible reason for this lack of improvement for NK is the tight implications of that model which force consumption growth and output growth to be collinear. This is a structural consequence of NK that the GEE cannot relax.

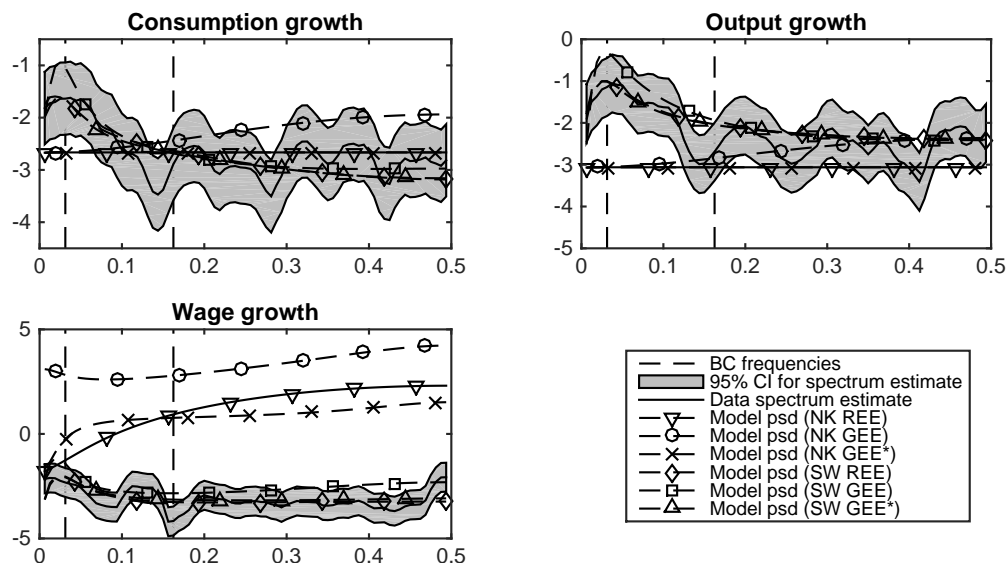


Figure 4: Spectral densities (log scale), data vs. models. *Unified beliefs.

The **spectral coherence** of two time series is an estimate of how the correlation between two time series can be decomposed over frequencies. Figures 7, 8, and 9 in the appendix display the

spectral coherences for each pair of observed series. Five examples are given below. The predicted coherence from each model is superimposed onto the nonparametric estimate with its confidence interval. The NK model performs very poorly here, not just because it has no prediction for investment, but also because it predicts that consumption growth, output growth, and labor hours are all perfectly correlated at all frequencies. One consistent feature of NK REE is its failure to predict observed coherences at business-cycle frequencies, for example between output growth and the interest rate, and between consumption growth and inflation. The version of NK GEE where beliefs are unified across agent types (households and firms) does not substantially improve the performance of the model on this dimension. However, once beliefs are allowed to differ between households and firms, the model-implied coherences improve substantially and in the direction of the data. The SW REE model entails predictions for these coherencies that are much closer to the data as well. On this dimension, the GEE relaxes NK into the direction of SW. This implies that relaxing REE can substitute for the additional structural mechanisms of SW in terms of replicating the observed spectral coherencies between output growth, inflation, and the interest rate.

As Tkachenko and Qu (2012) document, SW REE falls short of the data in terms of several of its implied spectral coherencies. For example, SW REE predicts coherencies between consumption growth and investment growth, between investment growth and labor hours, and between investment growth and output growth that are all lower than the estimates from the data at business cycle frequencies. The version of SW GEE with beliefs that differ across agents improves upon SW REE for those coherencies, placing the model's implications inside the confidence band of the estimates from the data in every case. Interestingly, the SW GEE with unified beliefs resembles SW REE more than the SW GEE with multiple beliefs; the data exploit the ability to distinguish between mis-specification in different equations. However, this is not enough to overcome all of the shortcomings of SW REE; in some cases, such as the coherence between consumption growth and labor hours, SW GEE is closer to the data than SW REE but does not attain the confidence bands of the data. In other cases, such as between output growth and the interest rate, the SW GEE hardly improves upon SW REE at all. This implies that SW GEE is not simply relaxing the model all the way to a VAR; there are sources of misspecification that relaxing REE cannot ameliorate.

3.4 Impulse responses and variance decompositions

Figures 11, 12, and 13 show impulse responses to structural and belief shocks. Tables 1 and 2 show unconditional variance decompositions with respect to the same. In both the impulse responses and the variance decompositions, there is evidence for a different role for a production technology shock in GEE as opposed to REE: it gives rise to more inflation, less consumption growth, and a more severe labor response. The firm belief shock also absorbs a substantial fraction (between a quarter to a third on average) of the variance of all endogenous variables. This shows the data favor a role for belief deviations that both propagate structural shocks and are a source of shocks

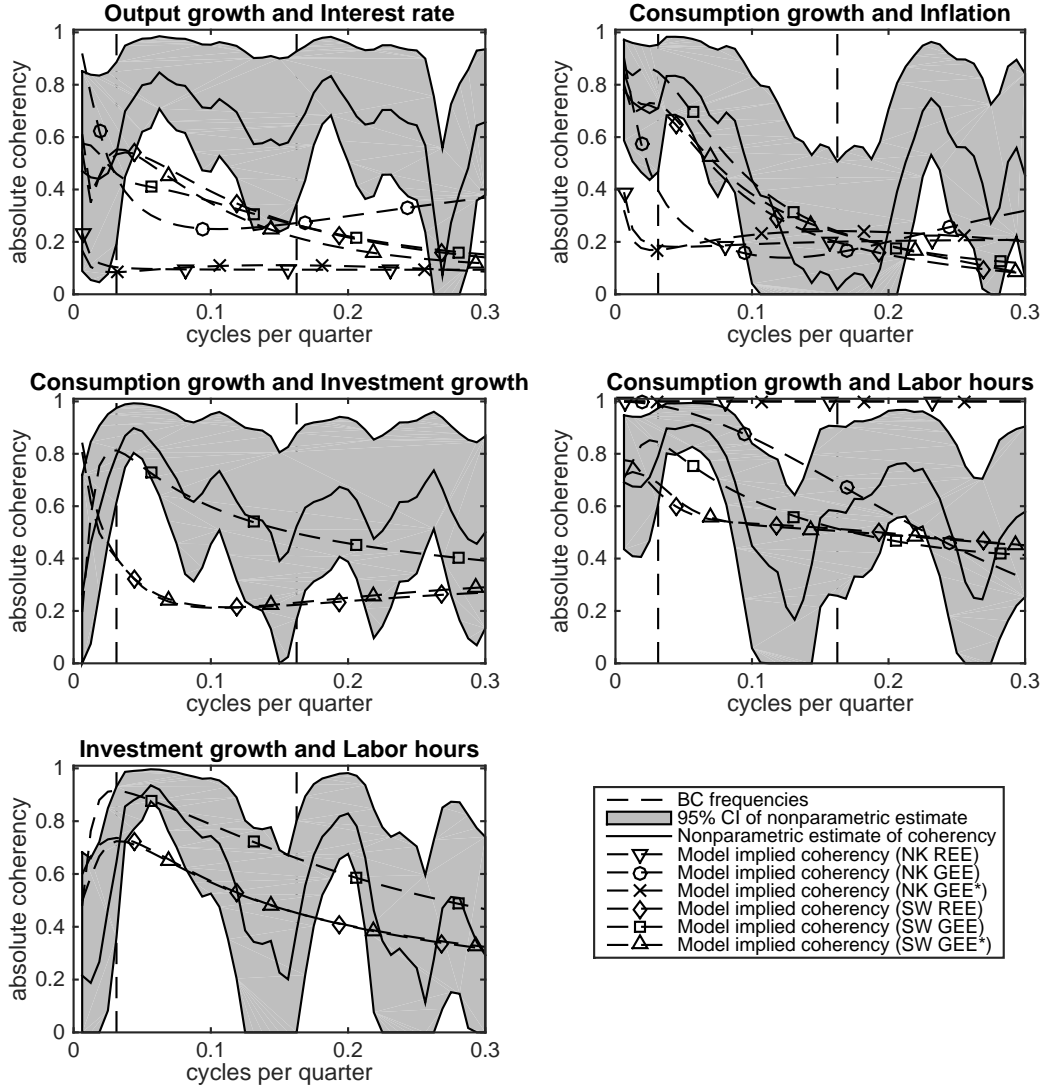


Figure 5: Spectral coherence, data vs. models. *Unified beliefs.

themselves.

A few of the impulse responses to structural shocks differ significantly between GEE and REE. For example, the response of investment to a positive TFP shock under GEE is severely dampened relative to REE, and the response of inflation under GEE to a TFP shock becomes significantly positive after 15 quarters. These results highlight the fact that relaxing REE may fundamentally change the predicted responses to structural shocks. Future work on estimating GEE may use information on impulse responses as an input to the formation of the priors over belief deviation elasticities, to the extent that the responses to certain structural shocks can be identified in other ways.

Shock	NK REE	NK GEE*	NK GEE	SW REE	SW GEE*	SW GEE
Output growth						
Production technology	87.49	84.34	14.55	17.40	15.25	9.08
Rate spread	3.68	3.38	2.58	17.18	23.17	14.00
Monetary policy	8.83	10.71	3.39	5.97	7.95	4.43
Government spending				28.23	29.93	21.14
Investment technology				15.48	14.52	12.48
Goods markup				3.30	1.26	1.04
Labor markup				12.45	3.69	0.97
Unified beliefs		1.56			4.23	
Household savings beliefs			13.50			1.87
Good pricing beliefs			65.97			33.36
Capital pricing beliefs						0.15
Wage setting beliefs						1.49
Inflation						
Production technology	9.48	15.92	67.62	3.10	6.27	17.49
Rate spread	66.97	19.78	11.50	0.56	2.43	2.12
Monetary policy	23.55	13.06	3.63	3.43	1.91	1.62
Government spending				0.70	1.34	8.05
Investment technology				3.27	9.59	23.31
Goods markup				15.09	3.24	1.99
Labor markup				73.84	24.73	10.97
Unified beliefs		51.24			50.49	
Household savings beliefs			10.50			0.28
Good pricing beliefs			6.75			28.36
Capital pricing beliefs						0.44
Wage setting beliefs						5.37
Interest rate						
Production technology	13.95	19.24	54.41	8.39	12.71	22.34
Rate spread	80.26	30.29	14.80	5.73	3.58	3.60
Monetary policy	5.79	7.94	14.25	11.36	19.05	10.25
Government spending				2.80	3.29	10.81
Investment technology				16.24	26.86	9.01
Goods markup				4.55	1.27	2.67
Labor markup				50.94	17.98	10.75
Unified beliefs		42.53			15.27	
Household savings beliefs			11.79			0.33
Good pricing beliefs			4.75			27.78
Capital pricing beliefs						0.12
Wage setting beliefs						2.34

Table 1: Unconditional variance decompositions, holding shock processes and structural parameters constant. *Beliefs are unified over all agent types.

Shock	NK REE	NK GEE*	NK GEE	SW REE	SW GEE*	SW GEE
Consumption growth						
Production technology				6.12	4.72	5.86
Rate spread				47.39	62.83	42.28
Monetary policy				11.57	13.28	8.55
Government spending				2.13	1.58	5.25
Investment technology				1.46	1.30	7.31
Goods markup				3.37	1.04	0.61
Labor markup				27.96	8.78	2.51
Unified beliefs					6.47	
Household savings beliefs						8.00
Good pricing beliefs						16.93
Capital pricing beliefs						0.08
Wage setting beliefs						2.64
Investment growth						
Production technology				6.13	6.24	1.09
Rate spread				2.27	3.18	3.70
Monetary policy				2.37	4.17	2.05
Government spending				0.82	0.87	2.36
Investment technology				80.26	79.78	29.52
Goods markup				2.70	0.89	1.85
Labor markup				5.46	1.62	0.76
Unified beliefs					3.26	
Household savings beliefs						0.33
Good pricing beliefs						56.44
Capital pricing beliefs						0.54
Wage setting beliefs						1.36
Wage growth						
Production technology				3.56	4.51	10.71
Rate spread				0.91	1.31	1.35
Monetary policy				1.37	2.67	0.75
Government spending				0.09	0.26	2.20
Investment technology				1.60	1.60	8.20
Goods markup				14.12	18.02	0.61
Labor markup				78.36	66.60	11.30
Unified beliefs					5.02	
Household savings beliefs						0.31
Good pricing beliefs						19.05
Capital pricing beliefs						0.33
Wage setting beliefs						45.20
Labor						
Production technology				1.69	5.09	28.56
Rate spread				1.61	4.51	2.59
Monetary policy				2.12	7.51	3.12
Government spending				8.32	13.06	9.72
Investment technology				6.64	8.13	12.03
Goods markup				3.35	1.07	1.68
Labor markup				76.27	53.85	10.72
Unified beliefs					6.79	
Household savings beliefs						0.29
Good pricing beliefs						28.67
Capital pricing beliefs						0.05
Wage setting beliefs						2.57

Table 2: Unconditional variance decompositions, holding shock processes and structural parameters constant. *Beliefs are unified over all agent types.

4 Conclusions

As Woodford (2013) argues, there is a large number of plausible departures from rational expectations equilibrium (REE), but as of yet little empirical guidance between them. This paper proposes an equilibrium concept, the generalized expectations equilibrium (GEE), in which deviations from REE are generic functions of current and lagged observables. Starting from a benchmark structural model, these deviations change the propagation of shocks in equilibrium in a way that can be identified from the data. While similar to a DSGE-VAR, the GEE is more restrictive because it relaxes only the equations where expectations enter. In posing relatively tight priors on the elasticities of the belief deviations, this paper has shown the direction in which the aggregate data push when the assumption of REE is gently relaxed within both a canonical three-equation New Keynesian model and the medium-scale Smets Wouters (2007) model: namely, the data favor deviations observationally equivalent to those where firms believe lagged inflation and current investment are more strongly associated with future inflation than they really are in equilibrium. Not surprisingly, with greater flexibility, the model approaches the data more closely, particularly in terms of the spectral coherences between output growth, investment growth, and the interest rate. The GEE within the simpler NK model substitutes for some of the structural mechanisms in the SW model. The GEE also improves upon SW, through a combination of shocks to beliefs themselves and the role that deviations from REE change the propagation of other structural shocks. But the GEE does not ameliorate all misspecification, even within the more sophisticated SW model. Adding more observations on survey data (particularly of firms, as in Bachmann and Elstner (2013)) may help refine the estimation, at the expense of limiting the role for higher-order beliefs or aggregation wedges. Adding more prior information on the propagation of structural shocks identified in other ways may also help refine the estimates. Finally, determining the “optimal” prior over the size of the deviations from REE remains for future work.

References

- Andolfatto, D., Hendry, S., and Moran, K. (2008). Are inflation expectations rational? *Journal of Monetary Economics*, 55(2):406–422.
- Angeletos, G.-M., Collard, F., and Dellas, H. (2014). Quantifying confidence. Technical report, MIT working paper.
- Angeletos, G.-M. and La’O, J. (2009). Incomplete information, higher-order beliefs and price inertia. *Journal of Monetary Economics*, 56:S19–S37.
- Angeletos, G.-M. and La’O, J. (2011). Optimal monetary policy with informational frictions. Technical report, National Bureau of Economic Research.
- Angeletos, G.-M. and La’O, J. (2013). Sentiments. *Econometrica*, 81(2):739–779.
- Bachmann, R. and Elstner, S. (2013). Firms optimism and pessimism. Technical report, National Bureau of Economic Research.

- Branch, W. A. (2007). Sticky information and model uncertainty in survey data on inflation expectations. *Journal of Economic Dynamics and Control*, 31(1):245–276.
- Bryan, M. F. and Palmqvist, S. (2005). Testing near-rationality using detailed survey data. Technical report, Sveriges Riksbank Working Paper Series.
- Canova, F. and Sala, L. (2009). Back to square one: identification issues in dsge models. *Journal of Monetary Economics*, 56(4):431–449.
- Chari, V. V., Kehoe, P. J., and McGrattan, E. R. (2007). Business cycle accounting. *Econometrica*, 75(3):781–836.
- Coibion, O. (2010). Testing the sticky information phillips curve. *The Review of Economics and Statistics*, 92(1):87–101.
- Coibion, O. and Gorodnichenko, Y. (2011). What can survey forecasts tell us about informational rigidities? Technical report, National Bureau of Economic Research.
- Coibion, O. and Gorodnichenko, Y. (2012). Information Rigidity and the Expectations Formation Process; A Simple Framework and New Facts. Imf working papers, International Monetary Fund.
- Coibion, O., Gorodnichenko, Y., and Wieland, J. (2012). The optimal inflation rate in new keynesian models: should central banks raise their inflation targets in light of the zero lower bound? *The Review of Economic Studies*, page rds013.
- Cole, S. and Milani, F. (2014). The misspecification of expectations in new keynesian models: A dsge-var approach.
- Del Negro, M. and Eusepi, S. (2011). Fitting observed inflation expectations. *Journal of Economic Dynamics and Control*, 35(12):2105–2131.
- Del Negro, M. and Schorfheide, F. (2004). Priors from general equilibrium models for vars*. *International Economic Review*, 45(2):643–673.
- Del Negro, M. and Schorfheide, F. (2009). Monetary policy analysis with potentially misspecified models. *American Economic Review*, 99(4):1415–50.
- Evans, G. W. and Honkapohja, S. (2001). *Learning and expectations in macroeconomics*. Princeton University Press.
- Guerrón-Quintana, P. A. and Nason, J. M. (2012). *Bayesian estimation of DSGE models*. Federal Reserve Bank of Philadelphia.
- Hamilton, J. D. (1994). *Time series analysis*, volume 2. Princeton university press Princeton.
- Iskrev, N. (2010). Local identification in dsge models. *Journal of Monetary Economics*, 57(2):189–202.
- King, R. G. and Watson, M. W. (1996). Money, prices, interest rates and the business cycle. *The Review of Economics and statistics*, pages 35–53.
- Kurz, M. (2012). A new keynesian model with diverse beliefs.
- Mankiw, N. G. and Reis, R. (2002). Sticky information versus sticky prices: a proposal to replace the new keynesian phillips curve. *Quarterly Journal of Economics*, pages 1295–1328.

- Mankiw, N. G. and Reis, R. (2007). Sticky information in general equilibrium. *Journal of the European Economic Association*, 5(2-3):603–613.
- Mavroeidis, S., Plagborg-Møller, M., and Stock, J. H. (2014). Empirical evidence on inflation expectations in the new keynesian phillips curve. *Journal of Economic Literature*, 52(1):124–188.
- Milani, F. (2007). Expectations, learning and macroeconomic persistence. *Journal of monetary Economics*, 54(7):2065–2082.
- Milani, F. and Rajbhandari, A. (2012). Expectation formation and monetary dsge models: Beyond the rational expectations paradigm. *Advances in Econometrics*, 28:253–288.
- Ratto, M. (2008). Analysing dsge models with global sensitivity analysis. *Computational Economics*, 31(2):115–139.
- Schorfheide, F. (2011). Estimation and evaluation of dsge models: progress and challenges. Technical report, National Bureau of Economic Research.
- Smets, F. and Wouters, R. (2003). An estimated dynamic stochastic general equilibrium model of the euro area. *Journal of the European economic association*, 1(5):1123–1175.
- Smets, F. and Wouters, R. (2007). Shocks and frictions in us business cycles: A bayesian dsge approach. *American Economic Review*, 97(3):586–606.
- Tkachenko, D. and Qu, Z. (2012). Frequency domain analysis of medium scale dsge models with application to smets and wouters (2007). *Advances in Econometrics*, 28:319–385.
- Wang, P. and Wen, Y. (2006). Solving linear difference systems with lagged expectations by a method of undetermined coefficients. *Federal Reserve Bank of St. Louis Working Paper Series*, (2006-003).
- Watson, M. W. (1993). Measures of fit for calibrated models. *Journal of Political Economy*, 101(6):pp. 1011–1041.
- Wen, Y. (1998). Can a real business cycle model pass the watson test? *Journal of Monetary Economics*, 42(1):185–203.
- Woodford, M. (2010). Robustly optimal monetary policy with near-rational expectations. *American Economic Review*, 100(1):274–303.
- Woodford, M. (2013). Macroeconomic analysis without the rational expectations hypothesis. Technical report, National Bureau of Economic Research.

A Appendix for Online Publication

A.1 Derivation of belief deviation for sticky information model

First, consider a firm allowed to choose a price in period t , whose information was acquired in period $t - k$. This firm faces the same problem as a firm in the standard New Keynesian setup; it knows that its future prices and quantities are subject to Calvo nominal rigidities. Once this firm's price is stuck, the quantity it produces is determined by the demand curve for its good. The only difference between this firm's problem and the standard problem is the date of the information set at the time of its pricing decision.

$$\max_{P_{i,k,t}} E_{t-k} \sum_{j=0}^{\infty} \xi_p^j \Xi_{t+j} (P_{it} Y_{i,t+j} - \Theta_{t+j} Y_{i,t+j}) \quad (39)$$

$$\text{s.t. } Y_{i,t+j} = Y_{t+j} \left(\frac{P_{i,t}}{P_{t+j}} \right)^{-(1+\lambda_p)/\lambda_p} \quad (40)$$

The first-order condition of this problem is the same as that of the standard problem, with the only difference being the date of the information set. The firm's optimal price $\tilde{p}_{k,t}$ will be the ratio of $K_{k,t}$ and $F_{k,t}$, where

$$K_{k,t} = E_{t-k} (1 + \lambda_p) \Xi_t Y_t \Theta_t + \beta \xi_p E_{t-k} K_{k,t+1} \pi_{t+1}^{(1+\lambda_p)/\lambda_p} \quad (41)$$

$$F_{k,t} = E_{t-k} \Xi_t Y_t + \beta \xi_p E_{t-k} F_{k,t+1} \pi_{t+1}^{1/\lambda_p} \quad (42)$$

In log-deviations,

$$\hat{p}_{k,t} = E_{t-k} \left((1 - \beta \xi_p) \hat{\vartheta}_t + \beta \xi_p (\hat{p}_{t+1} + \hat{\pi}_{t+1}) \right) \quad (43)$$

In log-deviations, the average desired price among all firms who can update their price at time t is a weighted average of the desired prices of firms whose information sets were updated $k = 0, \dots, \infty$ periods ago.

$$\hat{p}_t = \lambda \sum_{k=0}^{\infty} (1 - \lambda)^k \hat{p}_{k,t} \quad (44)$$

$$= \lambda \sum_{k=0}^{\infty} (1 - \lambda)^k E_{t-k} \left((1 - \beta \xi_p) \hat{\vartheta}_t + \beta \xi_p (\hat{p}_{t+1} + \hat{\pi}_{t+1}) \right) \quad (45)$$

In a GEE, the following relationship holds:

$$\hat{p}_t = (1 - \beta \xi_p) \hat{\vartheta}_t + \beta \xi_p E_t (\hat{p}_{t+1} + \hat{\pi}_{t+1}) + \hat{\Gamma}_t^F \quad (46)$$

$$(47)$$

Therefore, solving for the deviations,

$$\hat{\Gamma}_t^F = \lambda \sum_{k=0}^{\infty} (1 - \lambda)^k (E_{t-k} - E_t) \left((1 - \beta \xi_p) \hat{\vartheta}_t + \beta \xi_p (\hat{p}_{t+1} + \hat{\pi}_{t+1}) \right) \quad (48)$$

A.2 Estimation of spectrum and spectral coherence

The exposition here is largely taken from Hamilton (1994) and Tkachenko and Qu (2012).

Let the sample time series X_t and Y_t run from $t = 1$ to $t = T$. Let $M = \frac{T-1}{2}$, and the series $\omega_j = \frac{2j\pi}{T}$ for $j = 1 \dots M$. Let $\mu_Y = \frac{1}{T} \sum_{t=1}^T Y_t$. Let

$$\hat{\gamma}_{X,Y}(j) = \frac{1}{T} \sum_{t=j+1}^T [(X_t - \mu_X)(Y_{t-j} - \mu_Y)]$$

be an estimate of the cross-covariance of X_t and Y_t at lag j . Calculate the sample cospectrum as

$$\hat{c}_{X,Y}(\omega) = \frac{1}{2\pi} \left[\hat{\gamma}_{X,Y}(0) + \sum_{j=1}^{T-1} (\hat{\gamma}_{X,Y}(j) + \hat{\gamma}_{Y,X}(j)) \cos(\omega_j) \right]$$

and the quadrature sample spectrum as

$$\hat{q}_{X,Y}(\omega) = -\frac{1}{2\pi} \left[\sum_{j=1}^{T-1} (\hat{\gamma}_{X,Y}(j) - \hat{\gamma}_{Y,X}(j)) \sin(\omega_j) \right]$$

Let the smoothed analogues of $\hat{c}_{X,Y}(\omega)$ and $\hat{q}_{X,Y}(\omega)$, respectively, be

$$\hat{c}_{X,Y}^s(\omega) = \sum_{k=-\kappa}^{\kappa} W(k) \hat{c}_{X,Y}(\omega_{j+k})$$

and

$$\hat{q}_{X,Y}^s(\omega) = \sum_{k=-\kappa}^{\kappa} W(k) \hat{q}_{X,Y}(\omega_{j+k})$$

where the weight function $W(k)$ is a Daniell kernel.¹²

The quantity $\hat{c}_{X,X}^s(\omega)$ is the estimate of the population spectrum of X . The 95% confidence intervals for the log of $\hat{c}_{X,X}^s(\omega)$ are

$$\log(\hat{c}_{X,X}^s(\omega)) \pm 1.96 \left(\sum_{k=-\kappa}^{\kappa} W(k)^2 \right)^{1/2}$$

Calculate the spectral coherence between X and Y as

$$|\hat{r}_{X,Y}(\omega)| = \frac{(\hat{c}_{X,Y}^s(\omega)^2 + \hat{q}_{X,Y}^s(\omega)^2)^{1/2}}{(\hat{c}_{X,X}^s(\omega) \hat{c}_{Y,Y}^s(\omega))^{1/2}}$$

The 95% confidence intervals for $|\hat{r}_{X,Y}(\omega)|$ are

$$|\hat{r}_{X,Y}(\omega)| \pm \frac{1.96}{\sqrt{2}} (1 - |\hat{r}_{X,Y}(\omega)|^2) \left(\sum_{k=-\kappa}^{\kappa} W(k)^2 \right)^{1/2}$$

¹²Tkachenko and Qu (2012) also use a this kernel.

A.3 Full statement of GEE for a New Keynesian Model

A.3.1 Full equilibrium conditions

$$P_t^{-1/\lambda_p} = (1 - \xi_p)\tilde{P}_t^{-1/\lambda_p} + \xi_p(\pi P_{t-1})^{-1/\lambda_p} \quad (49)$$

$$Y_t = a_t \gamma^t L_t \quad (50)$$

$$\Theta_t = W_t \gamma^{-(1-\alpha)t} (a_t)^{-1} \quad (51)$$

$$\pi_t = P_t/P_{t-1} \quad (52)$$

$$X_{t,t+s} = 1 \text{ if } s = 0, \Pi_{l=1}^s \pi_* \text{ if } s > 0 \quad (53)$$

$$\tilde{Y}_{t+s} = Y_{t+s} \left(\frac{\tilde{P}_{it} X_{t,t+s}}{P_{t+s}} \right)^{-(1+\lambda_p, t+s)/\lambda_p, t+s} \quad (54)$$

$$0 = \tilde{E}_t^F \sum_{s=0}^{\infty} (\xi_p \beta)^s \frac{\Xi_{t+s} P_t}{\Xi_t P_{t+s}} \frac{\tilde{Y}_{t+s}}{\lambda_p} \left[X_{t,t+s} \tilde{P}_t - (1 + \lambda_p) \Theta_{t+s} \right] \quad (55)$$

$$\Xi_t = C_t^{-\sigma_c} \quad (56)$$

$$\Xi_t \frac{W_t}{P_t} = \gamma^{(1-\sigma_c)t} L_t^{\sigma_l} \quad (57)$$

$$\Xi_t = \beta b_t R_t \tilde{E}_t^H \left[\frac{\Xi_{t+1}}{\pi_{t+1}} \right] \quad (58)$$

$$R_t = (R_{t-1})^{\rho_r} \left[R_t^* \left(\frac{\pi_t}{\pi_*} \right)^{\phi_\pi} \left(\frac{Y_t}{Y_t^*} \right)^{\phi_y} \right]^{1-\rho_r} \left(\frac{Y_t/Y_{t-1}}{Y_t^*/Y_{t-1}^*} \right)^{\phi_{\Delta y}} \varepsilon_{r,t} \quad (59)$$

$$Y_t = C_t \quad (60)$$

$$\text{exg processes } (a, b) \quad (61)$$

A.3.2 Scaling for stationarity

$$1 = (1 - \xi_p)\tilde{p}_t^{-1/\lambda_p, t} + \xi_p(\pi_* \pi_t^{-1})^{-1/\lambda_p, t} \quad (62)$$

$$y_t = a_t L_t \quad (63)$$

$$\vartheta_t = w_t (a_t)^{-1} \quad (64)$$

$$\tilde{y}_{t+s} = y_{t+s} \left(\frac{\tilde{p}_{it} \Pi_{l=1}^s \pi_*}{\Pi_{l=1}^s \pi_{t+l}} \right)^{-(1+\lambda_p, t+s)/\lambda_p} \quad (65)$$

$$0 = \tilde{E}_t^F \sum_{s=0}^{\infty} (\xi_p \beta \gamma^{1-\sigma_c})^s \frac{\xi_{t+s}}{\xi_t} \frac{\tilde{y}_{t+s}}{\lambda_p} \left[\frac{\Pi_{l=1}^s \pi_*}{\Pi_{l=1}^s \pi_{t+l}} \tilde{p}_t - (1 + \lambda_p) \vartheta_{t+s} \right] \quad (66)$$

$$\xi_t = c_t^{-\sigma_c} \quad (67)$$

$$\xi_t w_t = L_t^{\sigma_l} \quad (68)$$

$$\xi_t = \beta \gamma^{-\sigma_c} b_t R_t \tilde{E}_t^H \left[\frac{\xi_{t+1}}{\pi_{t+1}} \right] \quad (69)$$

$$r_t = (r_{t-1})^{\rho_r} \left[r_t^* \left(\frac{\pi_t}{\pi_*} \right)^{\phi_\pi} \left(\frac{y_t}{y_t^*} \right)^{\phi_y} \right]^{1-\rho_r} \left(\frac{y_t/y_{t-1}}{y_t^*/y_{t-1}^*} \right)^{\phi_{\Delta y}} \varepsilon_{r,t} \quad (70)$$

$$y_t = c_t \quad (71)$$

$$\text{exg processes } (a, b) \quad (72)$$

A.3.3 Steady-state

$$\bar{\beta} = \beta\gamma^{-\sigma_c} \quad (73)$$

$$a = 1 \quad (74)$$

$$\pi = \pi_* \quad (75)$$

$$\tilde{p} = 1 \quad (76)$$

$$y = L \quad (77)$$

$$\vartheta = w \quad (78)$$

$$\tilde{y} = y \quad (79)$$

$$\tilde{p} = (1 + \lambda_p)\vartheta \quad (80)$$

$$\xi = ((1 - h/\gamma)c)^{-\sigma_c} \quad (81)$$

$$\xi w = L^{\sigma_l} \quad (82)$$

$$1 = \bar{\beta}R \left[\frac{1}{\pi_*} \right] \quad (83)$$

$$y = c \quad (84)$$

A.3.4 Log-linearization

$$0 = (1 - \xi_p)\hat{p}_t - \xi_p\pi_t \quad (85)$$

$$\hat{y}_t = \hat{a}_t + \hat{L}_t \quad (86)$$

$$\hat{\vartheta}_t = \hat{w}_t - \hat{a}_t \quad (87)$$

$$\frac{1}{1 - \xi_p\gamma\bar{\beta}}\hat{p}_t = \hat{\vartheta}_t + \frac{\xi_p\gamma\bar{\beta}}{1 - \xi_p\gamma\bar{\beta}}E_t \left[\hat{p}_{t+1} + \hat{\pi}_{t+1} \right] + \frac{\xi_p\gamma\bar{\beta}}{1 - \xi_p\gamma\bar{\beta}}\hat{\Gamma}_t^{F, \bar{p} + \pi} \quad (88)$$

$$\hat{\xi}_t = -\sigma_c\hat{c}_t \quad (89)$$

$$\hat{\xi}_t + \hat{w}_t = \sigma_l\hat{L}_t \quad (90)$$

$$\hat{\xi}_t = \hat{b}_t + \hat{r}_t + E_t \left[\hat{\xi}_{t+1} - \hat{\pi}_{t+1} \right] - \hat{\Gamma}_t^{H, \pi - \xi} \quad (91)$$

$$\hat{r}_t = \rho_r(\hat{r}_{t-1}) + (1 - \rho_r)(\hat{r}_t^* + \phi_\pi\hat{\pi}_t + \phi_y(\hat{y}_t - \hat{y}_t^*)) + \phi_{\Delta y}(\hat{y}_t - \hat{y}_t^* - \hat{y}_{t-1} + \hat{y}_{t-1}^*) + \varepsilon_{r,t} \quad (92)$$

$$\hat{y}_t = \hat{c}_t \quad (93)$$

$$\text{exg processes } (a, b) \quad (94)$$

A.3.5 Log-linearization as in code (a modification of SW 2007)

$$\hat{y}_t = \hat{a}_t + \hat{L}_t \quad (95)$$

$$\hat{r}_t^k = 0 \quad (96)$$

$$\hat{v}_t = \hat{w}_t - \hat{a}_t \quad (97)$$

$$\hat{\pi}_t = \frac{(1 - \xi_p \gamma \bar{\beta})(1 - \xi_p)}{\xi_p} (\hat{v}_t) + \gamma \bar{\beta} E_t [\hat{\pi}_{t+1}] + \gamma \bar{\beta} \hat{\Gamma}_t^{F, \bar{p} + \pi} \quad (98)$$

$$\hat{k}_t = 0 \quad (99)$$

$$\hat{k}_t^s = 0 \quad (100)$$

$$-\sigma_c \hat{c}_t = (\hat{b}_t + \hat{r}_t - E_t[\hat{\pi}_{t+1}] - \hat{\Gamma}_t^{H, \pi - \xi}) - \sigma_c E_t[\hat{c}_{t+1}] \quad (101)$$

$$\hat{l}_t = 0 \quad (102)$$

$$\hat{Q}_t = 0 \quad (103)$$

$$\hat{U}_t = 0 \quad (104)$$

$$0 = \sigma_c \hat{c}_t + \sigma_l \hat{L}_t - \hat{w}_t \quad (105)$$

$$\hat{r}_t = \rho_r \hat{r}_{t-1} + (1 - \rho_r)(\phi_\pi \hat{\pi}_t + \phi_y(\hat{y}_t - \hat{y}_t^*)) + \phi_{\Delta y}(\hat{y}_t - \hat{y}_t^* - \hat{y}_{t-1} + \hat{y}_{t-1}^*) + \varepsilon_{r,t} \quad (106)$$

$$\hat{y}_t = \hat{c}_t \quad (107)$$

$$\text{exg processes } (a, b) \quad (108)$$

A.3.6 Log-linearization in canonical form

$$\pi_t = \frac{(1 - \xi_p \gamma \bar{\beta})(1 - \xi_p)}{\xi_p} ((\sigma_c + \sigma_l) \hat{y}_t - (1 + \sigma_l) \hat{a}_t) + \gamma \bar{\beta} E_t [\pi_{t+1}] + (1 - \xi_p) \gamma \bar{\beta} \hat{\Gamma}_t^{F, \bar{p} + \pi} \quad (109)$$

$$-\sigma_c \hat{y}_t = \hat{b}_t + \hat{r}_t + E_t [-\sigma_c \hat{y}_{t+1} - \hat{\pi}_{t+1}] - \hat{\Gamma}_t^{H, \pi - \xi} \quad (110)$$

$$\hat{r}_t = \rho_r (\hat{r}_{t-1}) + (1 - \rho_r)(\hat{r}_t^* + \phi_\pi \hat{\pi}_t + \phi_y(\hat{y}_t - \hat{y}_t^*)) + \phi_{\Delta y}(\hat{y}_t - \hat{y}_t^* - \hat{y}_{t-1} + \hat{y}_{t-1}^*) + \varepsilon_{r,t} \quad (111)$$

$$\text{exg processes } (a, b) \quad (112)$$

Identification alternatives:

1. $\hat{\Gamma}_t^U = \hat{\Gamma}_t^{F, \bar{p} + \pi} = \hat{\Gamma}_t^{H, \pi - \xi}$: All agents share the same belief deviation, and the deviation is identified to be on inflation.
2. $\hat{\Gamma}_t^F = \hat{\Gamma}_t^{F, \bar{p} + \pi}$, $\hat{\Gamma}_t^H = \hat{\Gamma}_t^{H, \pi - \xi}$: Two different belief deviations, potentially differing across agents, but where the targets of the mistaken beliefs are not identified among (\bar{p} , π , and ξ).

A.3.7 Simplification and closed form solutions

Simplify as in the text: set $\gamma = 1$, $\sigma_c = 1$, $\phi_{\Delta y} = 0$, and assume that deviations are functions of current output, current inflation, and a shock:

$$\pi_t = \frac{(1 - \xi_p \beta)(1 - \xi_p)(1 + \sigma_l)}{\xi_p} (\hat{y}_t - \hat{a}_t) + \beta E_t [\pi_{t+1}] + (1 - \xi_p) \beta \hat{\Gamma}_t^F \quad (113)$$

$$\hat{y}_t = -\hat{b}_t - \hat{r}_t + E_t [\hat{y}_{t+1} + \hat{\pi}_{t+1}] + \hat{\Gamma}_t^H \quad (114)$$

$$\hat{r}_t = \rho_r \hat{r}_{t-1} + (1 - \rho_r) (\hat{r}_t^* + \phi_\pi \hat{\pi}_t + \phi_y (\hat{y}_t - \hat{y}_t^*)) + \hat{\varepsilon}_{r,t} \quad (115)$$

$$\hat{a}_t = \rho_a \hat{a}_{t-1} + \hat{\varepsilon}_{a,t} \quad (116)$$

$$\hat{b}_t = \rho_b \hat{b}_{t-1} + \hat{\varepsilon}_{b,t} \quad (117)$$

$$\hat{\Gamma}_t^H = \Gamma_y^H \hat{y}_t + \Gamma_\pi^H \hat{\pi}_t + \epsilon_{\Gamma^H,t} \quad (118)$$

$$\hat{\Gamma}_t^F = \Gamma_y^F \hat{y}_t + \Gamma_\pi^F \hat{\pi}_t + \epsilon_{\Gamma^F,t} \quad (119)$$

Suppose the natural rates are determined as the counterfactual where $\xi_p = 0$, there are no rate spread shocks, and there are no belief deviations. Then $\hat{y}_t^* = \hat{a}_t$ and $\hat{r}_t^* = E_t(\hat{y}_{t+1}^*) - \hat{y}_t^* = (\rho_a - 1)\hat{a}_t$. Let $\kappa = \frac{(1 - \xi_p)(1 - \beta \xi_p)(1 + \sigma_l)}{\xi_p}$. This implies:

$$\pi_t = \kappa (\hat{y}_t - \hat{a}_t) + \beta E_t [\pi_{t+1}] + (1 - \xi_p) \beta \hat{\Gamma}_t^F \quad (120)$$

$$\hat{y}_t = -\hat{b}_t - \hat{r}_t + E_t [\hat{y}_{t+1} + \hat{\pi}_{t+1}] + \hat{\Gamma}_t^H \quad (121)$$

$$\hat{r}_t = \rho_r \hat{r}_{t-1} + (1 - \rho_r) ((\rho_a - 1)\hat{a}_t + \phi_\pi \hat{\pi}_t + \phi_y (\hat{y}_t - \hat{a}_t)) + \hat{\varepsilon}_{r,t} \quad (122)$$

$$\hat{a}_t = \rho_a \hat{a}_{t-1} + \hat{\varepsilon}_{a,t} \quad (123)$$

$$\hat{b}_t = \rho_b \hat{b}_{t-1} + \hat{\varepsilon}_{b,t} \quad (124)$$

$$\hat{\Gamma}_t^H = \Gamma_y^H \hat{y}_t + \Gamma_\pi^H \hat{\pi}_t + \epsilon_{\Gamma^H,t} \quad (125)$$

$$\hat{\Gamma}_t^F = \Gamma_y^F \hat{y}_t + \Gamma_\pi^F \hat{\pi}_t + \epsilon_{\Gamma^F,t} \quad (126)$$

To solve for equilibrium, postulate the following policy functions and use the method of undetermined coefficients:

$$\hat{y}_t = g_a^y \hat{a}_t + g_b^y \hat{b}_t + g_r^y \hat{\varepsilon}_{r,t} + g_H^y \hat{\varepsilon}_{\Gamma^H,t} + g_F^y \hat{\varepsilon}_{\Gamma^F,t} \quad (127)$$

$$\hat{\pi}_t = g_a^\pi \hat{a}_t + g_b^\pi \hat{b}_t + g_r^\pi \hat{\varepsilon}_{r,t} + g_H^\pi \hat{\varepsilon}_{\Gamma^H,t} + g_F^\pi \hat{\varepsilon}_{\Gamma^F,t} \quad (128)$$

$$\hat{r}_t = g_a^r \hat{a}_t + g_b^r \hat{b}_t + g_r^r \hat{\varepsilon}_{r,t} + g_H^r \hat{\varepsilon}_{\Gamma^H,t} + g_F^r \hat{\varepsilon}_{\Gamma^F,t} \quad (129)$$

The solution:

$$c = 1 - \Gamma_y^H - \Gamma_\pi^H \kappa + \phi_y + \kappa \phi_\pi - \beta(1 - \xi_p)(\Gamma_\pi^F(1 - \Gamma_y^H + \phi_y) + \Gamma_y^F(\Gamma_\pi^H - \phi_\pi)) \quad (130)$$

$$g_a^y = 1 + (\Gamma_y^H(1 - \beta\Gamma_\pi^F(1 - \xi_p)) + \beta\Gamma_y^F(1 - \xi_p)(\Gamma_\pi^H - \phi_\pi))/c \quad (131)$$

$$g_a^\pi = (\Gamma_y^H \kappa + \beta\Gamma_y^F(1 - \xi_p)(1 + \phi_y))/c \quad (132)$$

$$g_a^r = \rho_a - 1 + \phi_y(g_a^y - 1) + \phi_\pi g_a^\pi \quad (133)$$

$$g_b^y = (\beta\Gamma_\pi^F(1 - \xi_p) - 1)/c \quad (134)$$

$$g_b^\pi = -(\kappa + \beta\Gamma_y^F(1 - \xi_p))/c \quad (135)$$

$$g_b^r = \phi_y g_b^y + \phi_\pi g_b^\pi \quad (136)$$

$$g_r^y = (\beta\Gamma_\pi^F(1 - \xi_p) - 1)/c \quad (137)$$

$$g_r^\pi = -(\kappa + \beta\Gamma_y^F(1 - \xi_p))/c \quad (138)$$

$$g_r^r = 1 + \phi_y g_r^y + \phi_\pi g_r^\pi \quad (139)$$

$$g_H^y = (1 - \beta\Gamma_\pi^F(1 - \xi_p))/c \quad (140)$$

$$g_H^\pi = -(\kappa + \beta\Gamma_y^F(1 - \xi_p))/c \quad (141)$$

$$g_H^r = \phi_y g_H^y + \phi_\pi g_H^\pi \quad (142)$$

$$g_F^y = \beta(1 - \xi_p)(\Gamma_\pi^H - \phi_\pi)/c \quad (143)$$

$$g_F^\pi = \beta(1 - \xi_p)(1 - \Gamma_y^H + \phi_y)/c \quad (144)$$

$$g_F^r = \phi_y g_F^y + \phi_\pi g_F^\pi \quad (145)$$

When the belief distortions are functions of shocks alone, and have no state-contingency ($\Gamma_y^H = \Gamma_\pi^H = \Gamma_y^F = \Gamma_\pi^F = 0$), the policy functions reduce to more familiar forms:

$$c = 1 + \phi_y + \kappa \phi_\pi \quad (146)$$

$$g_a^y = 1 \quad (147)$$

$$g_a^\pi = 0 \quad (148)$$

$$g_a^r = \rho_a - 1 \quad (149)$$

$$g_b^y = -1/c \quad (150)$$

$$g_b^\pi = -\kappa/c \quad (151)$$

$$g_b^r = \phi_y g_b^y + \phi_\pi g_b^\pi \quad (152)$$

$$g_r^y = -1/c \quad (153)$$

$$g_r^\pi = -\kappa/c \quad (154)$$

$$g_r^r = 1 + \phi_y g_r^y + \phi_\pi g_r^\pi \quad (155)$$

$$g_H^y = 1/c \quad (156)$$

$$g_H^\pi = -\kappa/c \quad (157)$$

$$g_H^r = \phi_y g_H^y + \phi_\pi g_H^\pi \quad (158)$$

$$g_F^y = -\beta(1 - \xi_p)\phi_\pi/c \quad (159)$$

$$g_F^\pi = \beta(1 - \xi_p)(1 + \phi_y)/c \quad (160)$$

$$g_F^r = \phi_y g_F^y + \phi_\pi g_F^\pi \quad (161)$$

A.4 Full statement of GEE for Smets Wouters (2007)

A.4.1 Full equilibrium conditions

$$P_t^{-1/\lambda_{p,t}} = (1 - \xi_p) \tilde{P}_t^{-1/\lambda_{p,t}} + \xi_p (\pi_{t-1}^{\iota_p} \pi^{1-\iota_p} P_{t-1})^{-1/\lambda_{p,t}} \quad (162)$$

$$Y_t(i) = a_t K_t^s(i) \alpha (\gamma^t L_t(i))^{1-\alpha} - \gamma^t \Phi \quad (163)$$

$$(1 - \alpha) R_t^k K_t^s = \alpha W_t L_t \quad (164)$$

$$\Theta_t = \alpha^{-\alpha} (1 - \alpha)^{-(1-\alpha)} W_t^{1-\alpha} R_t^{\alpha} \gamma^{-(1-\alpha)t} (a_t)^{-1} \quad (165)$$

$$\pi_t = P_t / P_{t-1} \quad (166)$$

$$X_{t,t+s} = 1 \text{ if } s = 0, \Pi_{l=1}^s \pi_{t+l-1}^{\iota_p} \pi_*^{1-\iota_p} \text{ if } s > 0 \quad (167)$$

$$\tilde{Y}_{t+s} = Y_{t+s} \left(\frac{\tilde{P}_{it} X_{t,t+s}}{P_{t+s}} \right)^{-(1+\lambda_{p,t+s})/\lambda_{p,t+s}} \quad (168)$$

$$0 = \tilde{E}_t^F \sum_{s=0}^{\infty} (\xi_p \beta)^s \frac{\Xi_{t+s} P_t}{\Xi_t P_{t+s}} \frac{\tilde{Y}_{t+s}}{\lambda_{p,t+s}} \left[X_{t,t+s} \tilde{P}_t - (1 + \lambda_{p,t+s}) \Theta_{t+s} \right] \quad (169)$$

$$K_t = (1 - \delta) K_{t-1} + \epsilon_t^i \left[1 - S \left(\frac{I_t}{I_{t-1}} \right) \right] I_t \quad (170)$$

$$K_t^s = U_t K_{t-1} \quad (171)$$

$$\Xi_t = \exp \left(\frac{\sigma_c - 1}{1 + \sigma_l} L_t^{1+\sigma_l} \right) (C_t - \lambda C_{t-1})^{-\sigma_c} \quad (172)$$

$$\frac{W_t^h}{P_t} = (C_t - h C_{t-1}) L_t^{\sigma_l} \quad (173)$$

$$\Xi_t = \beta b_t R_t \tilde{E}_t^H \left[\frac{\Xi_{t+1}}{\pi_{t+1}} \right] \quad (174)$$

$$\Xi_t = \Xi_t^k \mu_t \left(1 - S \left(\frac{I_t}{I_{t-1}} \right) - S' \left(\frac{I_t}{I_{t-1}} \right) \left(\frac{I_t}{I_{t-1}} \right) \right) + \beta \tilde{E}_t^H \left[\Xi_{t+1}^k \mu_{t+1} S' \left(\frac{I_{t+1}}{I_t} \right) \left(\frac{I_{t+1}}{I_t} \right)^2 \right] \quad (175)$$

$$\Xi_t^k = \beta \tilde{E}_t^H \left[\Xi_{t+1} \left(\frac{R_{t+1}^k}{P_{t+1}} U_{t+1} - a(U_{t+1}) \right) + \Xi_{t+1}^k (1 - \delta) \right] \quad (176)$$

$$\frac{R_t^k}{P_t} = a'(U_t) \quad (177)$$

$$X_{t,s}^w = 1 \text{ if } s = 0, \Pi_{l=1}^s \gamma \pi_{t+l-1}^{\iota_w} \pi_*^{1-\iota_w} \text{ if } s > 0 \quad (178)$$

$$0 = \tilde{E}_t^W \sum_{s=0}^{\infty} \xi_w^s \frac{\beta^s \Xi_{t+s} P_t}{\Xi_t P_{t+s}} L_{t+s}(i) \frac{1}{\lambda_{w,t+s}} \left[X_{t,s} \tilde{W}_t - (1 + \lambda_{w,t+s}) W_{t+s}^h \right] \quad (179)$$

$$W_t^{1/\lambda_{w,t}} = (1 - \xi_w) \tilde{W}_t^{1/\lambda_{w,t}} + \xi_w (\gamma \pi_{t-1}^{\iota_w} \pi^{1-\iota_w} W_{t-1})^{1/\lambda_{w,t}} \quad (180)$$

$$\frac{R_t}{R^*} = \left(\frac{R_{t-1}}{R^*} \right)^{\rho_r} \left[\left(\frac{\pi_t}{\pi^*} \right)^{\phi_\pi} \left(\frac{Y_t}{Y_t^*} \right)^{\phi_y} \right]^{1-\rho_r} \left(\frac{Y_t/Y_{t-1}}{Y_t^*/Y_{t-1}^*} \right)^{\phi_{\Delta y}} r_t \quad (181)$$

$$Y_t = C_t + I_t + G_t + a(U_t) \bar{K}_{t-1} \quad (182)$$

$$\text{exg processes } (\lambda^p, \lambda^w, a, b, \mu, g) \quad (183)$$

$$(184)$$

A.4.2 Scaling for stationarity

$$k_t = K_t / \gamma^t, w_t = W_t / (P_t \gamma^t), r_t^k = R_t^k / P_t, \xi_t = \Xi_t \gamma^{\sigma_c t}$$

$$\pi, y, w, r^k, \vartheta, \tilde{p}, k, k^s, \xi, L, r, i, \xi^k, U, \tilde{w}, w^h, c$$

$$1 = (1 - \xi_p) \tilde{p}_t^{-1/\lambda_{p,t}} + \xi_p (\pi_{t-1}^{\iota_p} \pi_t^{1-\iota_p} \pi_t^{-1})^{-1/\lambda_{p,t}} \quad (185)$$

$$y_t = a_t k_t^{s\alpha} L_t^{1-\alpha} - \Phi \quad (186)$$

$$(1 - \alpha) r_t^k k_t^s = \alpha w_t L_t \quad (187)$$

$$\vartheta_t = \alpha^{-\alpha} (1 - \alpha)^{-(1-\alpha)} w_t^{1-\alpha} r_t^{k\alpha} (a_t)^{-1} \quad (188)$$

$$\tilde{y}_{t+s} = y_{t+s} \left(\frac{\tilde{p}_{it} \Pi_{l=1}^s \pi_{t+l-1}^{\iota_p} \pi_*^{1-\iota_p}}{\Pi_{l=1}^s \pi_{t+l}} \right)^{-(1+\lambda_{p,t+s})/\lambda_{p,t+s}} \quad (189)$$

$$0 = \tilde{E}_t^F \sum_{s=0}^{\infty} (\xi_p \beta \gamma^{1-\sigma_c})^s \frac{\xi_{t+s}}{\xi_t} \frac{\tilde{y}_{t+s}}{\lambda_{p,t+s}} \left[\frac{\Pi_{l=1}^s \pi_{t+l-1}^{\iota_p} \pi_*^{1-\iota_p}}{\Pi_{l=1}^s \pi_{t+l}} \tilde{p}_t - (1 + \lambda_{p,t+s}) \vartheta_{t+s} \right] \quad (190)$$

$$k_t = (1 - \delta) k_{t-1} / \gamma + \mu_t \left[1 - S \left(\frac{i_t \gamma}{i_{t-1}} \right) \right] i_t \quad (191)$$

$$k_t^s = U_t k_{t-1} / \gamma \quad (192)$$

$$\xi_t = \exp \left(\frac{\sigma_c - 1}{1 + \sigma_l} L_t^{1+\sigma_l} \right) (c_t - h c_{t-1} / \gamma)^{-\sigma_c} \quad (193)$$

$$w_t^h = (c_t - h c_{t-1} / \gamma) L_t^{\sigma_l} \quad (194)$$

$$\xi_t = \beta \gamma^{-\sigma_c} b_t R_t \tilde{E}_t^H \left[\frac{\xi_{t+1}}{\pi_{t+1}} \right] \quad (195)$$

$$\xi_t = \xi_t^k \mu_t \left(1 - S \left(\frac{i_t \gamma}{i_{t-1}} \right) - S' \left(\frac{i_t \gamma}{i_{t-1}} \right) \left(\frac{i_t \gamma}{i_{t-1}} \right) \right) + \beta \gamma^{-\sigma_c} \tilde{E}_t^H \left[\xi_{t+1}^k \mu_{t+1} S' \left(\frac{i_{t+1} \gamma}{i_t} \right) \left(\frac{i_{t+1} \gamma}{i_t} \right)^2 \right] \quad (196)$$

$$\xi_t^k = \beta \gamma^{-\sigma_c} \tilde{E}_t^H \left[\xi_{t+1}^k (r_{t+1}^k U_{t+1} - a(U_{t+1})) + \xi_{t+1}^k (1 - \delta) \right] \quad (197)$$

$$r_t^k = a'(U_t) \quad (198)$$

$$0 = \tilde{E}_t^W \sum_{s=0}^{\infty} (\xi_w \beta \gamma^{1-\sigma_c})^s \frac{\xi_{t+s}}{\xi_t} \frac{L_{t+s}}{\lambda_{w,t+s}} \left[\frac{\Pi_{l=1}^s \pi_{t+l-1}^{\iota_w} \pi_*^{1-\iota_w}}{\Pi_{l=1}^s \pi_{t+l}} \tilde{w}_t - (1 + \lambda_{w,t+s}) w_{t+s}^h \right] \quad (199)$$

$$(w_t)^{1/\lambda_{w,t}} = (1 - \xi_w) (\tilde{w}_t)^{1/\lambda_{w,t}} + \xi_w (\pi_{t-1}^{\iota_p} \pi_t^{1-\iota_p} w_{t-1} \pi_t^{-1})^{1/\lambda_{w,t}} \quad (200)$$

$$\frac{r_t}{r^*} = \left(\frac{r_{t-1}}{r^*} \right)^{\rho_r} \left[\left(\frac{\pi_t}{\pi^*} \right)^{\phi_\pi} \left(\frac{y_t}{y_t^*} \right)^{\phi_y} \right]^{1-\rho_r} \left(\frac{y_t/y_{t-1}}{y_t^*/y_{t-1}^*} \right)^{\phi_{\Delta y}} \varepsilon_{r,t} \quad (201)$$

$$y_t = c_t + i_t + g_t + a(U_t) k_{t-1} / \gamma \quad (202)$$

$$\text{exg processes } (\lambda^p, \lambda^w, a, b, \mu, g) \quad (203)$$

A.4.3 Steady-state

$$\begin{aligned}
1 &= \tilde{p} \\
\bar{\beta} &= \beta \gamma^{-\sigma_c} \\
y &= a(k^s)^\alpha L^{1-\alpha} - \Phi \\
w &= (1-\alpha)\vartheta a k^{s\alpha} L^{-\alpha} \\
(1-\alpha)r^k k^s &= \alpha w L \\
\vartheta &= \alpha^{-\alpha} (1-\alpha)^{-(1-\alpha)} w^{1-\alpha} r^{k\alpha} (a)^{-1} \\
\tilde{y} &= y \\
\tilde{p} &= (1+\lambda_p)\vartheta \\
k &= (1-\delta)k/\gamma + i \\
k^s &= U k / \gamma \\
\xi &= \exp\left(\frac{\sigma_c - 1}{1 + \sigma_l} L^{1+\sigma_l}\right) (c - hc/\gamma)^{-\sigma_c} \\
w^h &= (1 - h/\gamma) c L^{\sigma_l} \\
1 &= \beta \gamma^{-\sigma_c} r \left[\frac{1}{\pi} \right] \\
\xi &= \xi^k \\
1 &= \beta \gamma^{-\sigma_c} [r^k + (1-\delta)] \\
r^k &= a'(U) \\
a(U) &= 0 \\
a &= 1 \\
b &= 1 \\
S'(\gamma) &= 0 \\
\mu &= 1 \\
\tilde{w} &= (1 + \lambda_w) w^h \\
w &= \tilde{w} \\
y &= c + i + g \\
&\text{exg processes } (\lambda^p, \lambda^w, a, b, \mu, g)
\end{aligned}$$

A.4.4 Log-linearization

$$0 = (1 - \xi_p)\hat{p}_t + \xi_p(\iota_p\hat{\pi}_{t-1} - \pi_t) \quad (204)$$

$$\hat{y}_t = (\hat{a}_t + \alpha\hat{k}_t^s + (1 - \alpha)\hat{L}_t) \left(\frac{y + \Phi}{y} \right) \quad (205)$$

$$r^{\hat{k}}_t + \hat{k}_t^s = \hat{w}_t + \hat{L}_t \quad (206)$$

$$\hat{\vartheta}_t = (1 - \alpha)\hat{w}_t + \alpha r^{\hat{k}}_t - \hat{a}_t \quad (207)$$

$$\frac{1}{1 - \xi_p\gamma\bar{\beta}}\hat{p}_t = \frac{\lambda_p}{1 + \lambda_p}\hat{\lambda}_{p,t} + \hat{\vartheta}_t + \frac{\xi_p\gamma\bar{\beta}}{1 - \xi_p\gamma\bar{\beta}}E_t \left[\hat{p}_{t+1} - (\iota_p\hat{\pi}_t - \hat{\pi}_{t+1}) \right] + \frac{\xi_p\gamma\bar{\beta}}{1 - \xi_p\gamma\bar{\beta}}\hat{\Gamma}_t^{F,\bar{p}+\pi} \quad (208)$$

$$\hat{k}_t = \frac{1 - \delta}{\gamma}\hat{k}_{t-1} + \frac{\gamma - 1 + \delta}{\gamma}(\mu_t + i_t) \quad (209)$$

$$\hat{k}_t^s = \hat{U}_t + \hat{k}_{t-1} \quad (210)$$

$$(1 - h/\gamma)\hat{\xi}_t = (\sigma_c - 1)(w^h L/c)\hat{L}_t - \sigma_c\hat{c}_t + \sigma_c(h/\gamma)\hat{c}_{t-1} \quad (211)$$

$$(1 - h/\gamma)\hat{w}^h_t = \hat{c}_t - (h/\gamma)\hat{c}_{t-1} + \sigma_l(1 - h/\gamma)\hat{L}_t \quad (212)$$

$$\hat{\xi}_t = \hat{b}_t + \hat{r}_t + E_t \left[\hat{\xi}_{t+1} - \hat{\pi}_{t+1} \right] - \hat{\Gamma}_t^{H,\pi-\xi} \quad (213)$$

$$\hat{i}_t = \frac{1}{S''\gamma^2}(\hat{\xi}_t^k - \hat{\xi}_t + \hat{\mu}) + \frac{1}{1 + \bar{\beta}\gamma}(\hat{i}_{t-1} + \bar{\beta}\gamma E_t[\hat{i}_{t+1}]) + \frac{\bar{\beta}\gamma}{1 + \bar{\beta}\gamma}\hat{\Gamma}_t^{H,i} \quad (214)$$

$$\hat{\xi}_t^k = \bar{\beta}E_t \left[r^k(\hat{\xi}_{t+1} + \hat{r}_{t+1}^k) + \hat{\xi}_{t+1}^k(1 - \delta) \right] + \bar{\beta}\hat{\Gamma}_t^{H,Q} \quad (215)$$

$$r^k r^{\hat{k}}_t = a''\hat{U}_t \quad (216)$$

$$\frac{\tilde{w}}{1 - \xi_w\gamma\bar{\beta}}\hat{w}_t = \lambda_w w^h \hat{\lambda}_{w,t} + (1 + \lambda_w)w^h \hat{w}_t^h + \frac{\tilde{w}\xi_w\gamma\bar{\beta}}{1 - \xi_w\gamma\bar{\beta}}E_t \left[\hat{w}_{t+1} - (\iota_w\hat{\pi}_t - \hat{\pi}_{t+1}) \right] + \frac{\tilde{w}\xi_w\gamma\bar{\beta}}{1 - \xi_w\gamma\bar{\beta}}\hat{\Gamma}_t^{L,\tilde{w}+\pi} \quad (217)$$

$$\hat{w}_t = (1 - \xi_w)\hat{w}_t + \xi_w(\iota_w\hat{\pi}_{t-1} - \hat{\pi}_t + \hat{w}_{t-1}) \quad (218)$$

$$\hat{r}_t = \rho_r\hat{r}_{t-1} + (1 - \rho_r)(\phi_\pi\hat{\pi}_t + \phi_y(\hat{y}_t - \hat{y}_t^*)) + \phi_{\Delta y}(\hat{y}_t - \hat{y}_t^* - \hat{y}_{t-1} + \hat{y}_{t-1}^*) + \varepsilon_{r,t} \quad (219)$$

$$y\hat{y}_t = c\hat{c}_t + i\hat{i}_t + g\hat{g}_t + r^k k^s \hat{U}_t \quad (220)$$

$$\text{exg processes } (\lambda^p, \lambda^w, a, b, \mu, g) \quad (221)$$

A.4.5 Log-linearization, solved form given in SW 2007

$$\hat{y}_t = (\hat{a}_t + \alpha \hat{k}_t^s + (1 - \alpha) \hat{L}_t) \left(\frac{y + \Phi}{y} \right) \quad (222)$$

$$r^{\hat{k}}_t + \hat{k}_t^s = \hat{w}_t + \hat{L}_t \quad (223)$$

$$\hat{\vartheta}_t = (1 - \alpha) \hat{w}_t + \alpha r^{\hat{k}}_t - \hat{a}_t \quad (224)$$

$$(1 + \gamma \bar{\beta} \iota_p) \hat{\pi}_t = \iota_p \hat{\pi}_{t-1} + \frac{(1 - \xi_p \gamma \bar{\beta})(1 - \xi_p)}{\xi_p} \left(\frac{\lambda_p}{1 + \lambda_p} \hat{\lambda}_{p,t} + \hat{\vartheta}_t \right) + \gamma \bar{\beta} E_t [\hat{\pi}_{t+1}] + \gamma \bar{\beta} \hat{\Gamma}_t^{F, \bar{p} + \pi} \quad (225)$$

$$\hat{k}_t = \frac{1 - \delta}{\gamma} \hat{k}_{t-1} + \frac{\gamma - 1 + \delta}{\gamma} (\mu_t + i_t) \quad (226)$$

$$\hat{k}_t^s = \hat{U}_t + \hat{k}_{t-1} \quad (227)$$

$$-\sigma_c (1 + h/\gamma) \hat{c}_t = -\sigma_c (h/\gamma) \hat{c}_{t-1} + (1 - h/\gamma) (\hat{b}_t + \hat{r}_t - E_t [\hat{\pi}_{t+1}] - \hat{\Gamma}_t^{H, \pi - \xi}) \quad (228)$$

$$+ (\sigma_c - 1) (w^h L/c) (E_t [\hat{L}_{t+1}] - \hat{L}_t) - \sigma_c E_t [\hat{c}_{t+1}] \quad (229)$$

$$\hat{i}_t = \frac{1}{S'' \gamma^2} (\hat{Q}_t + \hat{\mu}) + \frac{1}{1 + \bar{\beta} \gamma} (\hat{i}_{t-1} + \bar{\beta} \gamma E_t [\hat{i}_{t+1}]) + \frac{\bar{\beta} \gamma}{1 + \bar{\beta} \gamma} \hat{\Gamma}_t^{H, i} \quad (230)$$

$$\hat{Q}_t = -\hat{b}_t - (\hat{r}_t - E_t [\hat{\pi}_{t+1}] - \hat{\Gamma}_t^{H, \pi - \xi}) + \bar{\beta} E_t [r^k (\hat{r}_{t+1}^k) + \hat{Q}_{t+1} (1 - \delta)] + \bar{\beta} \hat{\Gamma}_t^{H, Q} \quad (231)$$

$$r^k \hat{r}_t^k = a'' \hat{U}_t \quad (232)$$

$$(1 + \bar{\beta} \gamma) \hat{w}_t - \hat{w}_{t-1} = \frac{(1 - \xi_w)(1 - \xi_w \gamma \bar{\beta})}{\xi_w} \left(\frac{\lambda_w}{1 + \lambda_w} \hat{\lambda}_{w,t} + \frac{1}{1 - h/\gamma} (\hat{c}_t - (h/\gamma) \hat{c}_{t-1}) + \sigma_l \hat{L}_t - \hat{w}_t \right) \quad (233)$$

$$+ \gamma \bar{\beta} E_t [\hat{w}_{t+1} + \hat{\pi}_{t+1}] + \iota_w \hat{\pi}_{t-1} - (1 + \gamma \bar{\beta} \iota_w) \hat{\pi}_t + \gamma \bar{\beta} \hat{\Gamma}_t^{L, \bar{w} + \pi} \quad (234)$$

$$\hat{r}_t = \rho_r \hat{r}_{t-1} + (1 - \rho_r) (\phi_\pi \hat{\pi}_t + \phi_y (\hat{y}_t - \hat{y}_t^*)) + \phi_{\Delta y} (\hat{y}_t - \hat{y}_t^* - \hat{y}_{t-1} + \hat{y}_{t-1}^*) + \varepsilon_{r,t} \quad (235)$$

$$y \hat{y}_t = c \hat{c}_t + i \hat{i}_t + g \hat{g}_t + r^k \hat{k}_t^s \hat{U}_t \quad (236)$$

$$\text{exg processes } (\lambda^p, \lambda^w, a, b, \mu, g) \quad (237)$$

Identification alternatives:

1. $\hat{\Gamma}_t^U = \hat{\Gamma}_t^{F, \bar{p} + \pi} = \hat{\Gamma}_t^{H, \pi - \xi} = \hat{\Gamma}_t^{H, Q} + \hat{\Gamma}_t^{H, \pi - \xi} = \hat{\Gamma}_t^{L, \bar{w} + \pi}$: All agents share the same belief deviation, and the deviation is identified to be on inflation.
2. $\hat{\Gamma}_t^F = \hat{\Gamma}_t^{F, \bar{p} + \pi}$, $\hat{\Gamma}_t^H = \hat{\Gamma}_t^{H, \pi - \xi}$, $\hat{\Gamma}_t^Q = \hat{\Gamma}_t^{H, Q} + \hat{\Gamma}_t^{H, \pi - \xi}$, $\hat{\Gamma}_t^W = \hat{\Gamma}_t^{L, \bar{w} + \pi}$: Four different belief deviations, potentially differing across agents, but where the targets of the mistaken beliefs are not identified among $(\bar{p}, \pi, \xi, r^k, Q, \text{ and } \bar{w})$.

A.5 Full estimation results

	Distr	Prior		Posterior (NK)		Posterior (SW)	
		Mean	St. Dev.	Mean	St. Dev.	Mean	St. Dev.
φ	Normal	4.00	1.50			5.80	1.03
σ_c	Normal	1.50	0.37	2.29	0.27	1.39	0.14
h	Beta	0.70	0.10			0.71	0.04
ξ_w	Beta	0.50	0.10			0.70	0.07
σ_l	Normal	2.00	0.75	1.67	0.78	1.82	0.61
ξ_p	Beta	0.50	0.10	0.68	0.05	0.65	0.06
ι_w	Beta	0.50	0.15			0.58	0.13
ι_p	Beta	0.50	0.15			0.23	0.09
psi	Beta	0.50	0.15			0.56	0.12
Phi	Normal	1.25	0.12			1.62	0.08
r_π	Normal	1.50	0.25	1.75	0.18	2.05	0.18
ρ	Beta	0.75	0.10	0.58	0.08	0.81	0.02
r_y	Normal	0.12	0.05	0.14	0.05	0.09	0.02
$r_{\Delta y}$	Normal	0.12	0.05	0.14	0.05	0.22	0.03
π_*	Gamma	0.62	0.10	0.92	0.12	0.77	0.11
$100(\beta^{-1} - 1)$	Gamma	0.25	0.10	0.14	0.05	0.17	0.06
\bar{l}	Normal	0.00	2.00			0.49	1.07
$\bar{\gamma}$	Normal	0.40	0.10	0.31	0.05	0.43	0.01
α	Normal	0.30	0.05			0.20	0.02
ρ_a	Beta	0.50	0.20	0.99	0.00	0.96	0.01
ρ_b	Beta	0.50	0.20	0.90	0.03	0.24	0.08
ρ_g	Beta	0.50	0.20			0.97	0.01
ρ_i	Beta	0.50	0.20			0.72	0.06
ρ_r	Beta	0.50	0.20	0.35	0.08	0.15	0.06
ρ_p	Beta	0.50	0.20			0.89	0.05
ρ_w	Beta	0.50	0.20			0.97	0.01
μ_p	Beta	0.50	0.20			0.71	0.09
μ_w	Beta	0.50	0.20			0.85	0.05
ρ_{ga}	Beta	0.50	0.20			0.53	0.09
σ_a	InvGamma	0.10	2.00	0.97	0.06	0.45	0.03
σ_b	InvGamma	0.10	2.00	0.08	0.02	0.23	0.02
σ_g	InvGamma	0.10	2.00			0.53	0.03
σ_i	InvGamma	0.10	2.00			0.45	0.05
σ_r	InvGamma	0.10	2.00	0.39	0.04	0.24	0.01
σ_p	InvGamma	0.10	2.00			0.14	0.02
σ_w	InvGamma	0.10	2.00			0.25	0.02

Table 3: Non-GEE parameter estimation.

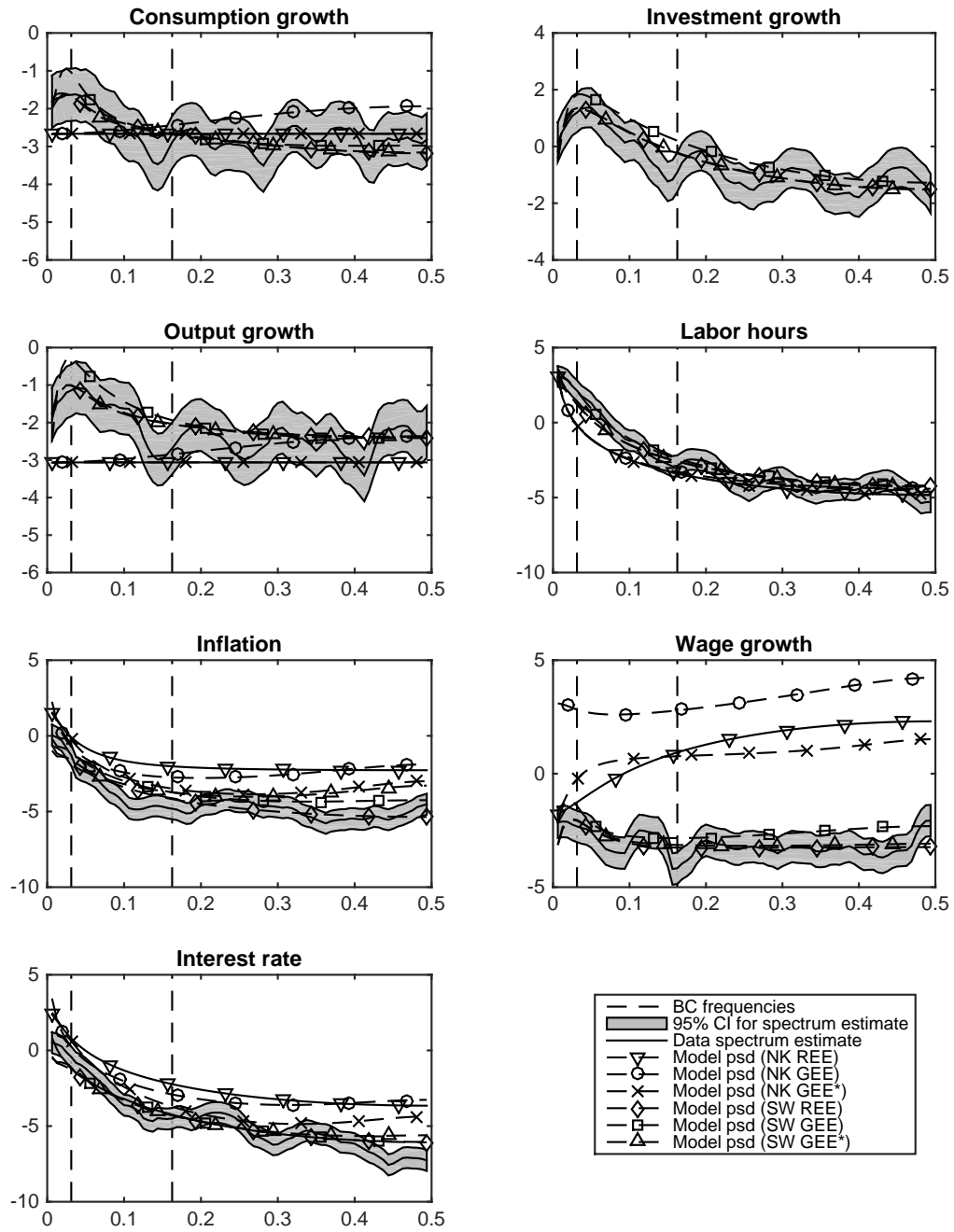


Figure 6: Spectral densities (log scale), data vs. models. *Unified beliefs.

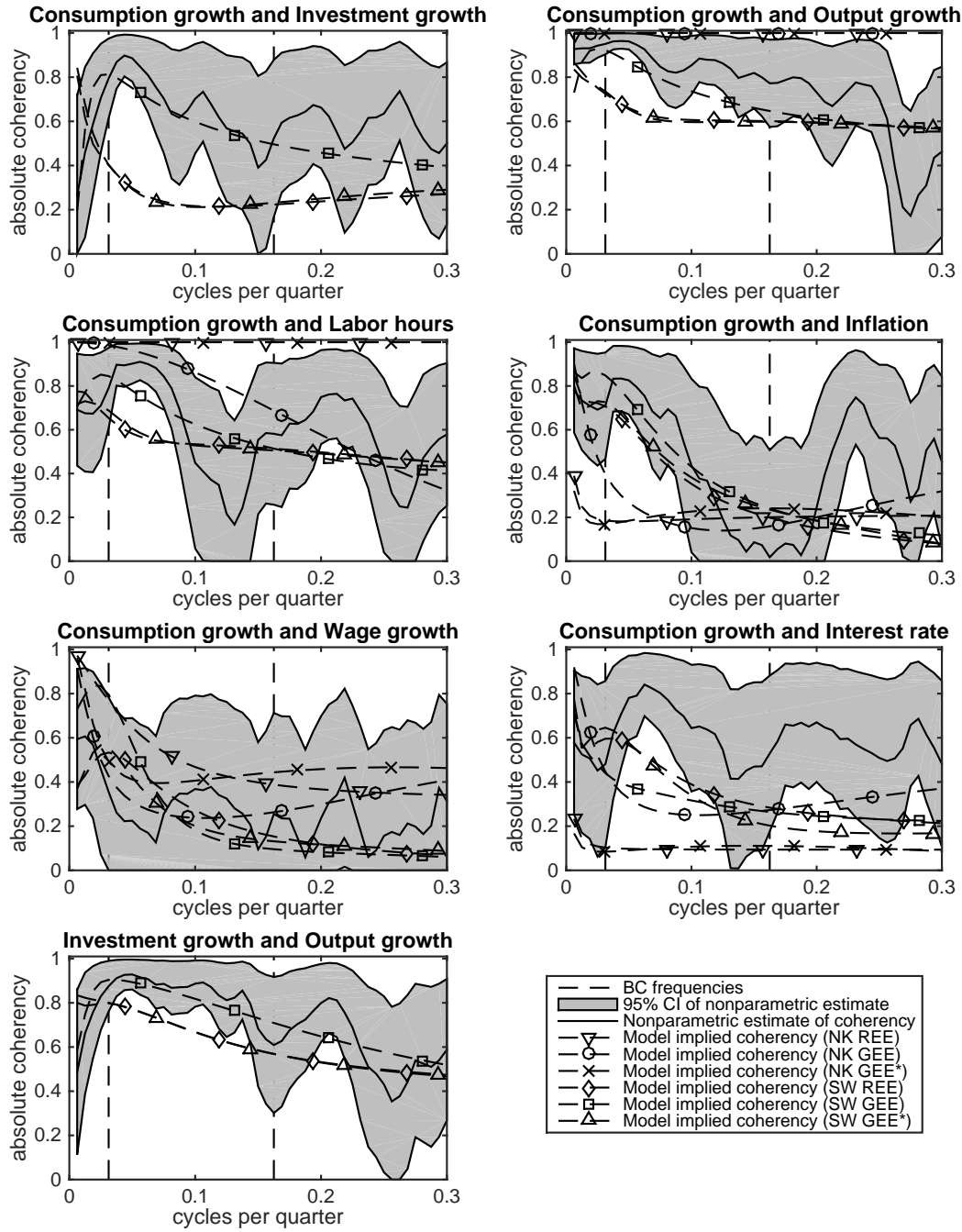


Figure 7: Spectral coherence, data vs. models. *Unified beliefs.

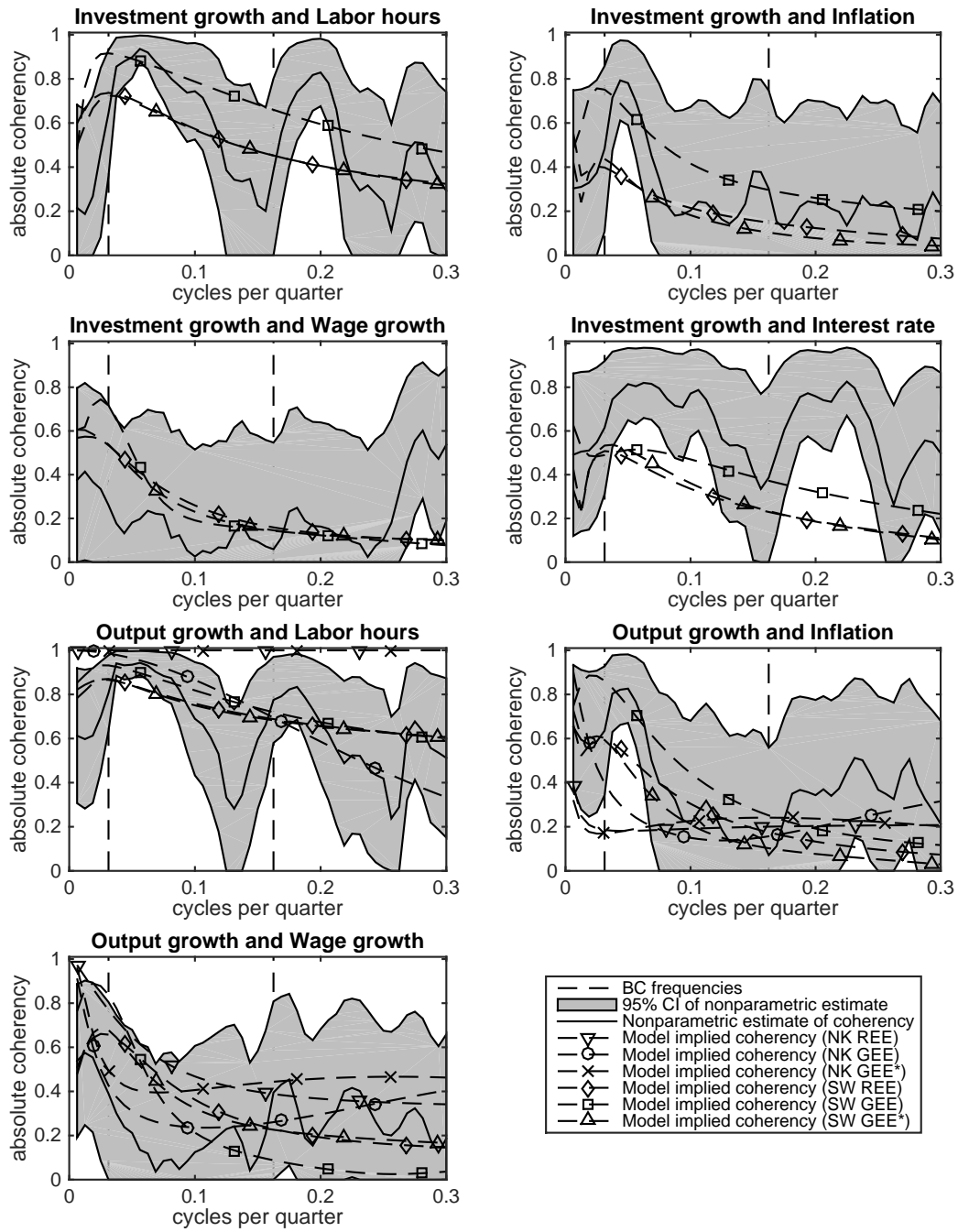


Figure 8: Spectral coherence, data vs. models. *Unified beliefs.

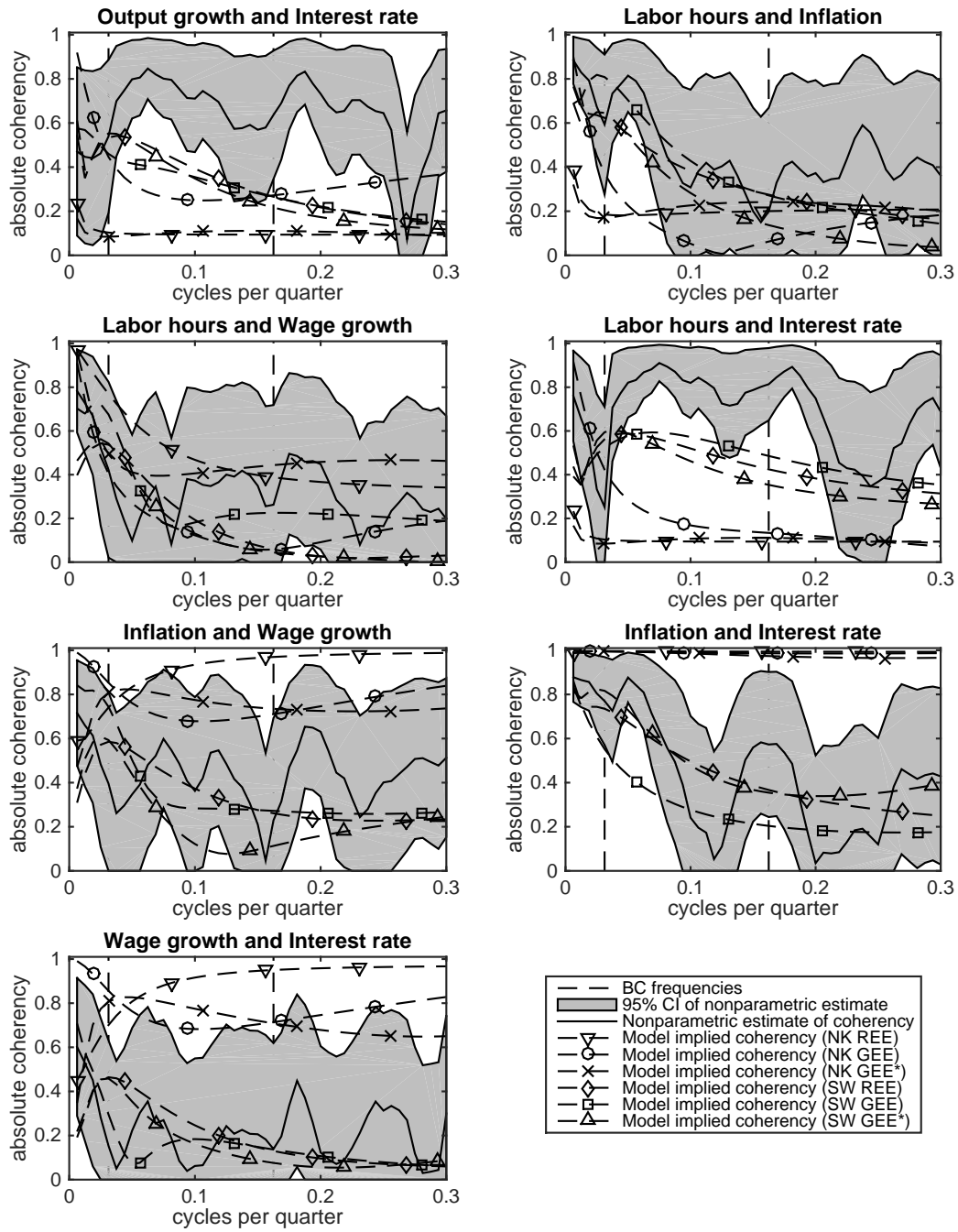


Figure 9: Spectral coherence, data vs. models. *Unified beliefs.

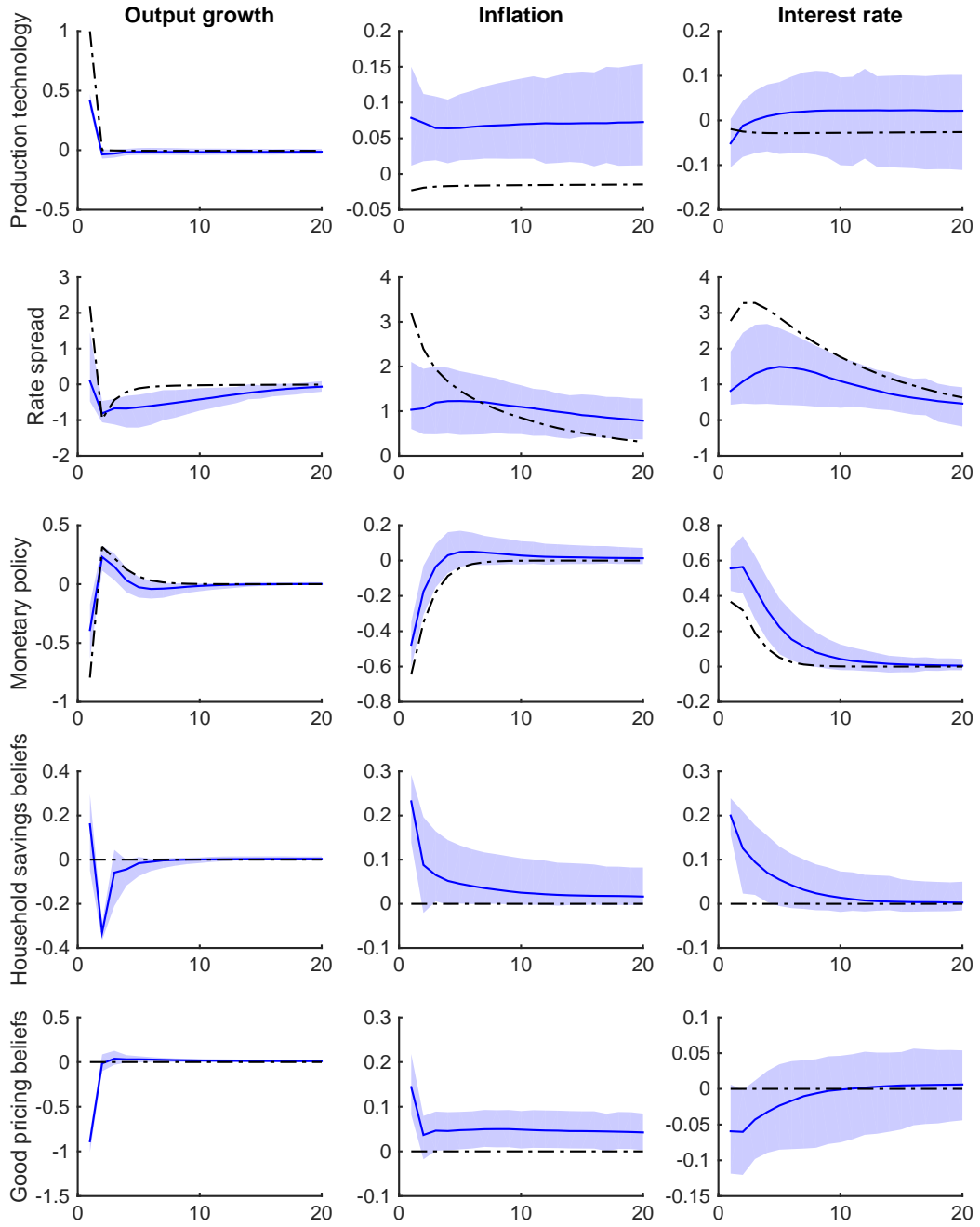


Figure 10: Impulse responses to shocks under GEE (with posterior credible interval shaded) and REE (dashed), New Keynesian benchmark.

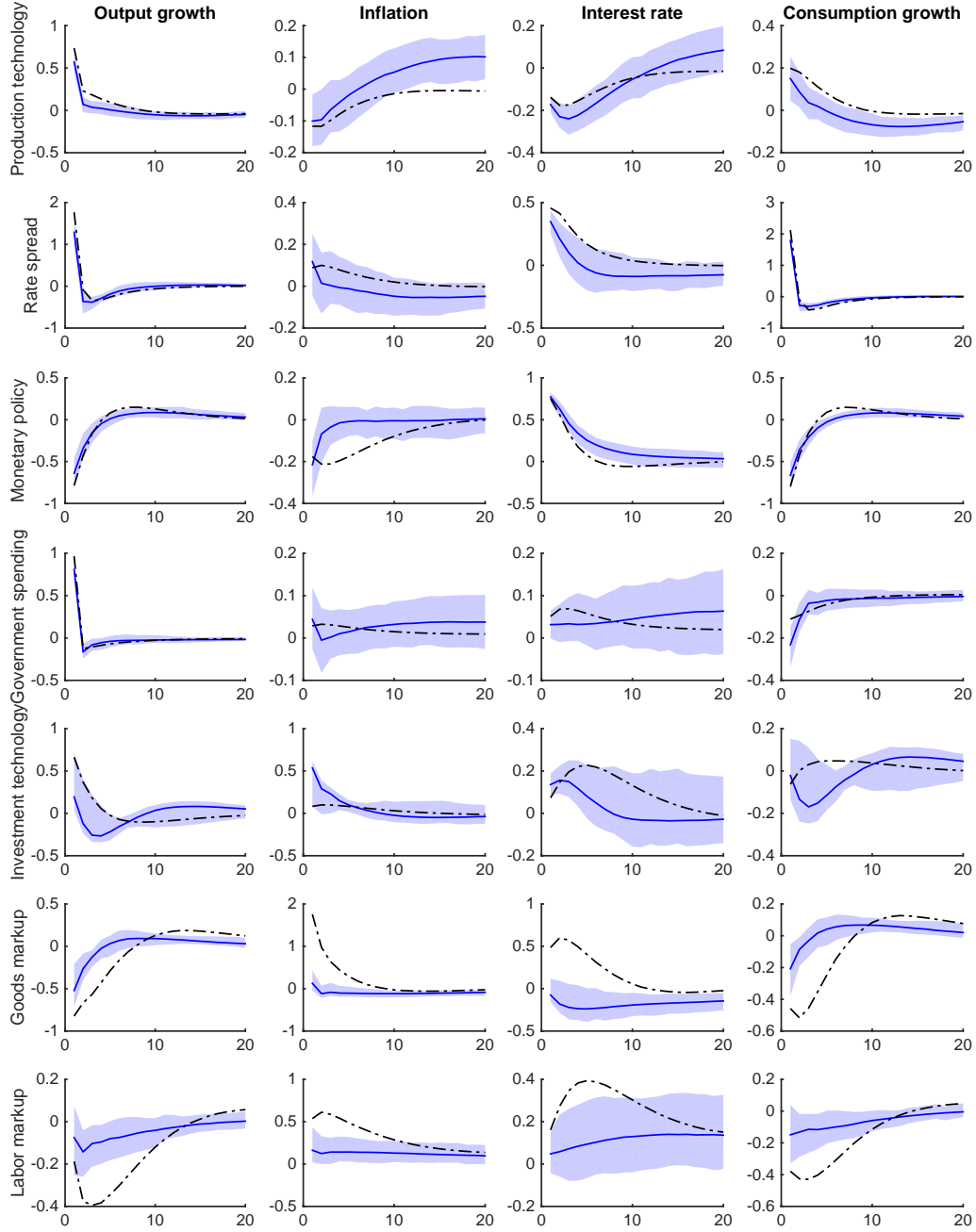


Figure 11: Impulse responses to shocks under GEE (with posterior credible interval shaded) and REE (dashed), Smets-Wouters benchmark.

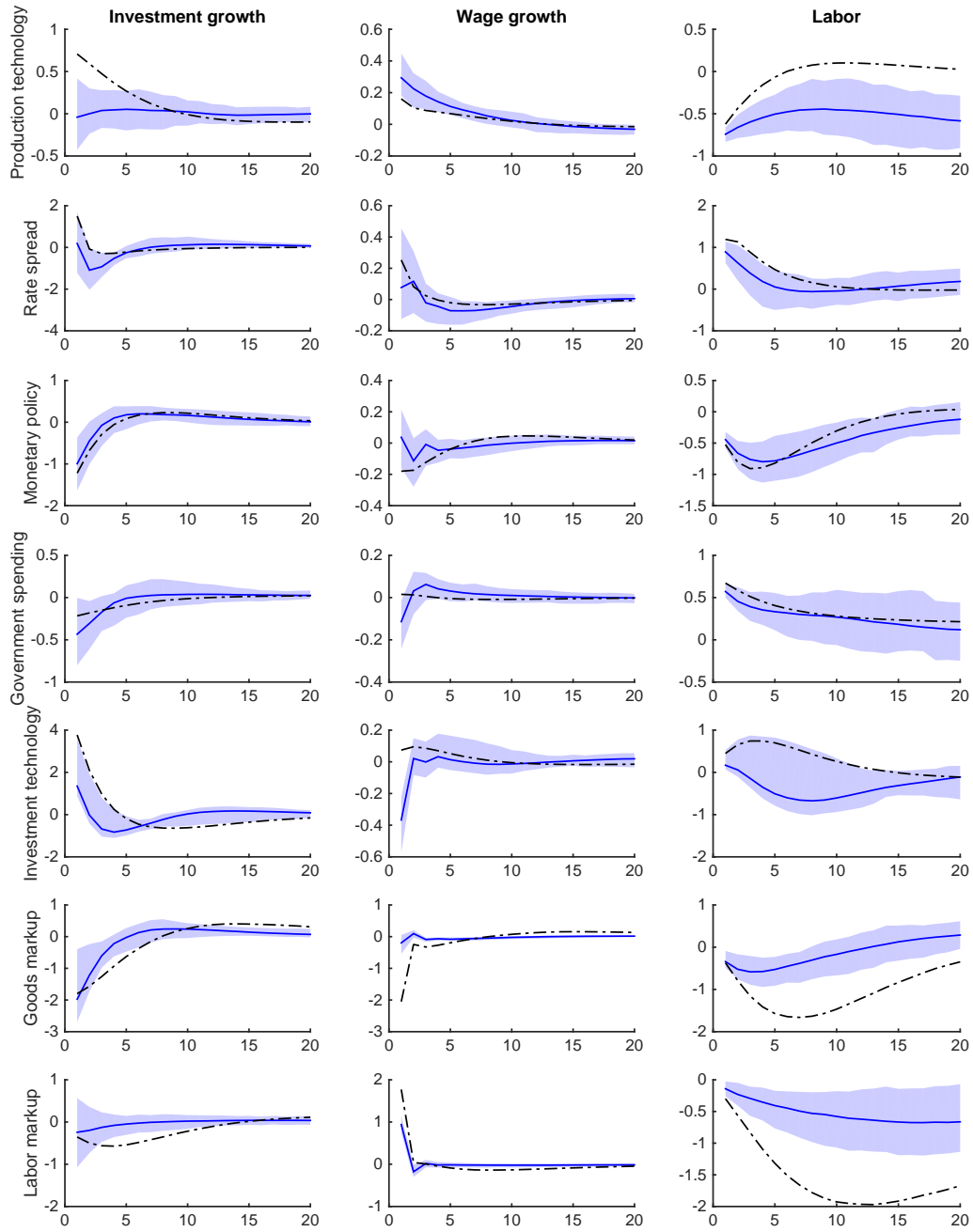


Figure 12: Impulse responses to shocks under GEE (with posterior credible interval shaded) and REE (dashed), Smets-Wouters benchmark.

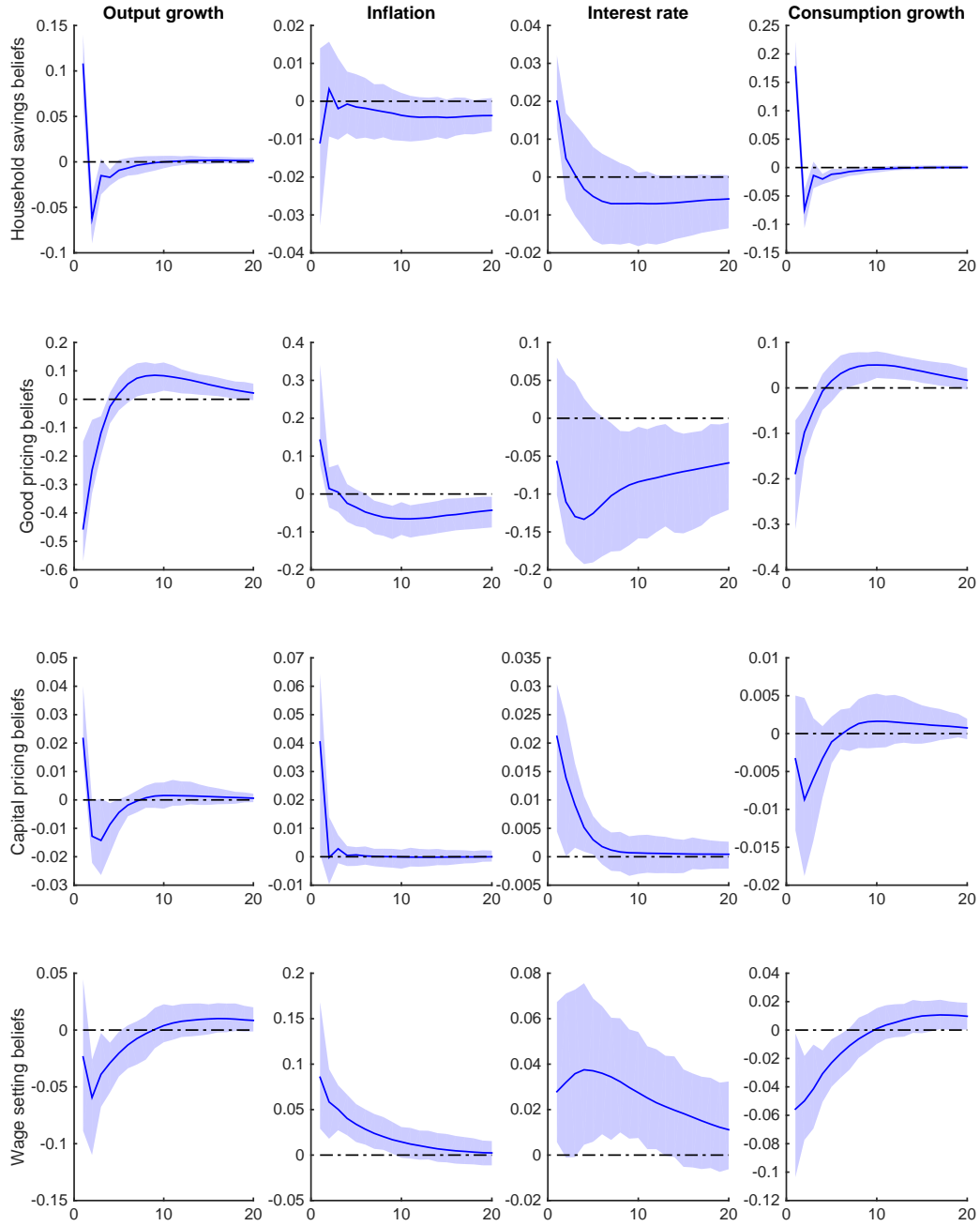


Figure 13: Impulse responses to shocks under GEE (with posterior credible interval shaded) and REE (dashed), Smets-Wouters benchmark.

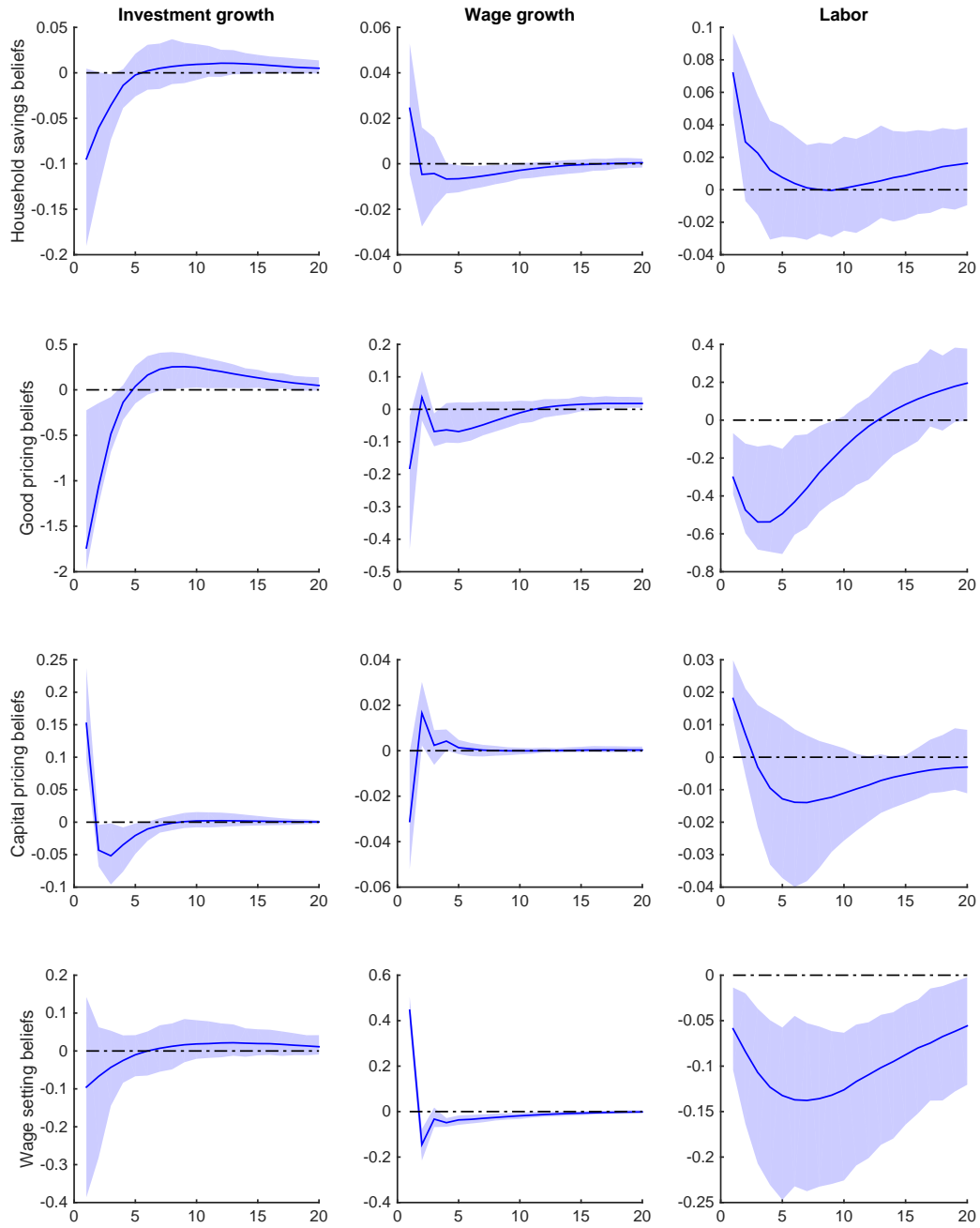


Figure 14: Impulse responses to shocks under GEE (with posterior credible interval shaded) and REE (dashed), Smets-Wouters benchmark.