Structural Estimation with Unstructured Data*

Sara Casella
LUISS University, EIEF

Jesús Fernández-Villaverde
University of Pennsylvania, NBER, CEPR

Stephen Hansen
University College London, IFS, CEPR

Ryohei Oishi
University College London

Minchul Shin
Federal Reserve Bank of Philadelphia

March 8, 2026

Abstract

Standard macroeconomic data cannot separate the systematic and non-systematic components of monetary policy. We show that this distinction can be recovered by incorporating unstructured text data into the structural estimation of a DSGE model. We augment a standard state-space model with a non-core measurement block that links structural shocks to time series derived from FOMC transcripts, using a spike-and-slab prior to let the data select which series are informative. In a medium-scale New Keynesian model for the U.S., incorporating text improves predictive performance and materially alters structural inference: the estimated Taylor-rule coefficients fall and the Calvo parameter drops, implying a steeper Phillips curve.

JEL Codes: C11, C32, C55, E37, E52.

Keywords: Unstructured Data, Text as Data, DSGE Models, Spike-and-Slab Priors, Monetary Policy, Phillips Curve, FOMC Transcripts.

*The views expressed in this paper are solely those of the authors and do not necessarily reflect the views of the Federal Reserve Bank of Philadelphia or the Federal Reserve System. Emails: scasella@luiss.it, jesusfv@econ.upenn.edu, stephen.hansen@ucl.ac.uk, ryohei.oishi.22@ucl.ac.uk, minchul.shin@phil.frb.org

1 Introduction

Dynamic equilibrium models are built to answer policy questions, yet the data used to estimate them typically record outcomes that reflect both the rules agents follow and the shocks they face, without revealing which is which.[1] For instance, we observe the federal funds rate the FOMC sets, but not how much of its movement reflects the systematic component of the Taylor rule and how much reflects departures from it.

Unstructured data could help close this gap.[2] Text, images, video, audio, sensor data, and high-dimensional firm or product characteristics are now routinely used in applied economics: for forecasting (Bybee et al. 2024), for measuring latent variables such as policy uncertainty (Baker et al. 2016) and sentiment (Shapiro et al. 2022), for characterizing market structure (Hoberg and Phillips 2016), for measuring skill requirements through job-posting text (Deming and Kahn 2018), and for tracking economic activity through satellite imagery (Henderson et al. 2012). But these applications are reduced-form: text regressions, event studies, sentiment indices, and forecasting exercises that treat textual measures as exogenous regressors or as additional time series in vector autoregressions. The counterfactual experiments and welfare calculations that often only dynamic equilibrium models can deliver have remained cut off from these advances.

The problem is sharpest for DSGE models, where the stakes are highest. Central banks use these models to evaluate policy, yet the parameters that govern their policy prescriptions are among the hardest to estimate. Output, inflation, and the interest rate move the same way whether the Fed raised rates because inflation rose or because the FOMC decided conditions warranted a different course.

FOMC transcripts contain precisely the information that macroeconomic aggregates lack. The deliberations recorded in these documents reveal which rate movements reflect the systematic rule and which reflect judgment, communication, or accommodation. The question is whether this information, once brought into the estimation, changes what we learn about the structural parameters. When we do so, the answer is clear. The estimated response of interest rates to inflation weakens, and the degree of price stickiness falls, shifting the policy implications of the model.

[1] By dynamic equilibrium models, we refer in this paper to models that are explicit about preferences, technology, information sets, and the intertemporal equilibrium outcomes implied by those assumptions. By structural estimation, we mean the estimation of parameters governing preferences and technology that are invariant to the policy interventions of interest (Hurwicz 1962), as well as the structural shocks that drive the economy.

[2] By unstructured data, we refer to data that do not conform to a predefined, tabular, or relational schema and therefore cannot be readily represented as a small number of rows and columns with fixed fields in a spreadsheet or a SQL table.

To exploit this information, we need a framework that links unstructured text to the latent states of a DSGE model without a priori specifying which textual signals are informative. We write the DSGE model in state-space form, which we call the *core block*, and add a *non-core block* that links a large collection of text-derived time series to the model's structural shocks. The econometric difficulty is selection: the researcher does not know *ex ante* which of these series carry information about the shocks and which are noise. Including everything risks overfitting. Excluding useful series discards identifying variation.

To address this challenge, we place a spike-and-slab prior on the loading matrix that governs the relationship between the non-core series and structural shocks. This prior allows each loading to be exactly zero with positive probability, letting the data select which series are informative while shrinking the loadings on uninformative series toward zero. It disciplines estimation when the number of candidate series is large relative to the sample size and preserves structural interpretability by preventing the non-core block from absorbing variation that properly belongs to the core model.

We implement the framework in a medium-scale New Keynesian DSGE model for the U.S., estimated at a quarterly frequency over the period 1987Q3–2019Q4. The sample is dictated by the availability of FOMC transcripts, which extend through 2019Q4 at the time of our empirical analysis. The core block includes six standard macroeconomic observables: inflation, output, the policy interest rate, wages, hours worked, and the relative price of investment. The model, based on Fernández-Villaverde and Guerrón-Quintana (2021), is deliberately rich but conventional. The objective is not to innovate on the core block, but to show that unstructured data can improve its estimation. For the non-core block, we process FOMC transcripts using a Latent Dirichlet Allocation (LDA) topic model and extract time-varying topic shares from the Committee's discussions. As a benchmark, we also include Greenbook and Tealbook staff projections.

Our empirical analysis addresses three questions:

1. Does the inclusion of FOMC text improve the model's predictive performance for standard macroeconomic variables?
2. Which topics in FOMC policy discussions are informative about structural shocks, and which contain little or no relevant information?
3. How does the inclusion of text affect structural inference, in particular the estimated coefficients of the monetary policy rule and the degree of nominal rigidity?

Incorporating FOMC transcripts improves joint predictive performance, with reductions in root-mean-square error of roughly 5 to 15 percent for most core variables. The spike-and-slab prior selects a small number of economically interpretable topics as informative about

the monetary policy shock and shuts down the rest. The transcripts shift the posterior distributions of the Taylor-rule coefficients and the Calvo parameter in ways that connect to longstanding debates about the slope of the Phillips curve and the conduct of optimal monetary policy.

The predictive gains are concentrated in the federal funds rate and hours worked, and are not driven by a single episode: the cumulative forecast-error advantage for the interest rate widens steadily over the evaluation sample. The policy-discussion portion of the transcripts (FOMC2) is more informative than the economic-conditions portion (FOMC1), and combining both corpora reduces accuracy relative to using FOMC2 alone. Greenbook and Tealbook projections, used as an alternative non-core source, do not improve upon the no-text benchmark jointly. Among text representations, LDA topic shares outperform high-dimensional embeddings reduced via PCA.

The loadings that survive selection are economically interpretable: topics related to labor market conditions, unconventional monetary policy and quantitative easing, the zero lower bound and forward guidance, and inflation and cost pressures load strongly on the monetary policy shock. Many other topics, despite appearing relevant ex ante, are effectively shut down. Topics such as Trade and Risk/Uncertainty, which capture forces outside the scope of our closed-economy New Keynesian model, receive near-zero inclusion probabilities, consistent with the selection mechanism respecting the model's theoretical boundaries.

When text is excluded, variation that the standard Taylor rule cannot capture, including communication, forward guidance, and deliberation, is absorbed into the systematic feedback coefficients, inflating them. Once the transcripts are included, this variation is reallocated to the monetary policy shock, and the systematic coefficients fall. In turn, to match inflation dynamics under a less aggressive feedback rule, the estimation lowers the Calvo parameter, steepening the Phillips curve. The additional information also tightens the posterior credible intervals for the monetary policy parameters. These shifts speak to contested questions in the literature: existing DSGE estimates often imply a flat Phillips curve that is difficult to reconcile with micro-level evidence on price adjustment (Bils and Klenow 2004, Nakamura and Steinsson 2008, Mavroeidis et al. 2014), and the magnitude of the Taylor-rule response to inflation has direct implications for equilibrium determinacy (Clarida et al. 2000, Lubik and Schorfheide 2004).

Specifications relying solely on standard macroeconomic observables may suffer from an omitted-information bias. Without text, the econometrician cannot distinguish the systematic component of monetary policy from the non-systematic one. The impulse responses to the investment-specific technology shock illustrate what this means in practice: the text-

augmented model generates a markedly stronger inflation and interest-rate response, consistent with firms adjusting prices more aggressively when the Calvo parameter is lower. These differences are not marginal; the posterior mean responses of inflation and the interest rate roughly double at short horizons relative to the benchmark without text. Our framework does not resolve fundamental identification failures; the formal result is that the Fisher information with non-core data weakly dominates the information based on core data only. But the empirical shifts in posterior mass and the narrower credible intervals are consistent with text alleviating a practical identification problem.

Our paper builds on three bodies of work. The first incorporates large information sets into DSGE estimation. Boivin and Giannoni (2006) and Kryshko (2011) develop factor-augmented DSGE models that exploit many observables through latent factors, while Schorfheide et al. (2010), Gelfer (2019), and Gelfer (2021) extend these methods to forecasting. These approaches rely on structured numeric data and do not impose sparsity on the measurement loadings. We depart from them in two respects: our non-core series load directly on structural shocks rather than on reduced-form factors, and our spike-and-slab prior allows the data to shut down uninformative series. On the methodological side, our prior builds on the spike-and-slab framework of Ishwaran and Rao (2005), applied to economic forecasting by Giannone et al. (2021) and to dynamic factor models by Kaufmann and Schumacher (2017, 2019). Bryzgalova et al. (2023) use related ideas in asset pricing. We adapt these tools to a setting where the factor structure is disciplined by economic theory rather than estimated freely.

The second is the literature on text-as-data in economics, surveyed by Gentzkow et al. (2019) and Ash and Hansen (2023). In macroeconomics, text-based measures have been used to construct indices of policy uncertainty, geopolitical risk (Caldara and Iacoviello 2022), and news sentiment, and have shown incremental forecasting value beyond standard numeric controls. A related strand focuses on FOMC transcripts: Hansen et al. (2018) apply topic models to these documents, and Aruoba and Drechsel (2025) use FOMC text to help identify monetary policy shocks. A substantial literature seeks to identify monetary policy shocks using information beyond standard macroeconomic time series, including the narrative approach of Romer and Romer (2004) and high-frequency identification around FOMC announcements (Gürkaynak et al. 2005). Our approach is complementary: rather than constructing an external instrument or shock series, we embed the text directly in the model's measurement system and let the estimation determine its informational content jointly with the structural parameters.

In industrial organization, Compiani et al. (2025) use embedded representations of prod-

uct text and images in demand estimation, and Pellegrino (2025) use text-based industry classifications to parameterize substitution patterns. Relative to all of these bodies of work, our key departure is to treat text-derived series not as observed variables but as noisy proxies for latent structural shocks, embedded in a measurement equation that allows the data to determine which series are informative. Although our application focuses on FOMC transcripts and a New Keynesian DSGE model, our framework extends naturally to a large class of structural models in state-space form and to any source of unstructured data.

The remainder of the paper is structured as follows. Section 2 develops the econometric framework, including the non-core measurement equation, the sparsity structure, and the estimation algorithm. Section 3 presents the empirical application to the U.S. economy using FOMC transcripts. Section 4 concludes.

2 DSGE model with Unstructured Data

Our starting point is the state-space representation implied by a DSGE model. We augment this structure with a non-core measurement block that links unstructured data to the model's latent states and shocks. The block makes no assumptions about which series are informative about which components of the model: the unstructured data follow their own autoregressive dynamics and load on a subset of the model's states, with a spike-and-slab prior that allows the data to determine which loadings are active and which are zero.

2.1 Standard econometric framework for DSGE model

The solution of a DSGE model is summarized by the state–transition equation

$$S_t = \mathcal{T}(S_{t-1}, \varepsilon_t; \psi), \quad \varepsilon_t \sim \mathcal{N}(0, \Sigma_\varepsilon), \tag{1}$$

where S_t is an $N_s \times 1$ vector of state variables, $\mathcal{T}(\cdot)$ is the equilibrium transition mapping implied by the DSGE model, and ε_t is an $N_\varepsilon \times 1$ vector of structural shocks that follows a multivariate normal distribution with $N_\varepsilon \times N_\varepsilon$ covariance matrix Σ_ε. The transition dynamics are governed by the set of structural parameters ψ.

The state–transition equation is combined with a measurement equation that links observed variables to the underlying state vector. The core measurement equation takes the form

$$Y_t^C = H_0^C(\theta) + H_1^C(\theta) S_t + u_t, \quad u_t \sim \mathcal{N}(0, \Sigma_u^C(\theta)), \tag{2}$$

where Y_t^C denotes an $N \times 1$ vector of observed "core" macroeconomic variables, such as inflation, interest rates, or output growth, and u_t is an $N \times 1$ measurement error vector with $N \times N$ covariance matrix $\Sigma_u^C(\theta)$. The objects $H_0^C(\theta)$, $H_1^C(\theta)$, and $\Sigma_u^C(\theta)$ collect additional parameters (such as measurement error variances) that, together with the structural parameters ψ, form the broader parameter vector θ.

The transition equation and the core measurement equation define a standard state–space representation. Estimation of θ proceeds either by maximizing the likelihood function implied by this system (from a frequentist perspective) or by combining the likelihood with a prior distribution for θ and sampling from the resulting posterior (from a Bayesian approach).

When the DSGE model is linearized, the likelihood can be evaluated exactly using the Kalman filter. In nonlinear settings, the likelihood function can be approximated using particle filtering or other nonlinear filtering methods (Fernández-Villaverde et al. 2016).

2.2 State-space representation with unstructured data

We now introduce unstructured data, such as large collections of text, into the standard DSGE estimation framework. In practice, an empirical researcher first transforms the raw unstructured data into a multivariate time series, denoted Y_t^{NC}. This transformation may take many forms: for example, converting text into an uncertainty or sentiment index, extracting topic shares from a corpus using a topic model, or constructing other numerical representations using natural language processing or machine learning techniques. We remain agnostic about the specific preprocessing method; the framework treats Y_t^{NC} simply as additional observable time series derived from unstructured data.

To incorporate these non-core variables into the estimation, we augment the core measurement equation with an additional "non-core" measurement equation. Importantly, we do not impose a priori which unstructured series should be informative about which elements of the DSGE model. That is, we avoid specifying a direct mapping between unstructured data and particular state variables. When such a mapping is known, the corresponding data can be included directly in the core measurement equation, but in most applications with unstructured data this link is inherently uncertain.

More formally, suppose we have M time series constructed from unstructured data or machine-learning outputs, denoted Y_t^{NC}. These series may contain information about the state of the economy, but it is unknown a priori which of them are informative. We introduce these non-core observables through the measurement equation

$$Y_t^{NC} = B_0 + B_1 Y_{t-1}^{NC} + B_2 S_t + v_t, \quad v_t \sim \mathcal{N}(0, \Sigma_v) \tag{3}$$

where B_0 is an $M \times 1$ vector of intercepts, B_1 is an $M \times M$ VAR coefficient matrix capturing the own dynamics of the non-core series, B_2 is an $M \times N_s$ loading matrix linking the non-core variables to the model's states, and Σ_v is an $M \times M$ covariance matrix of measurement errors.

This specification mirrors the structure of the core measurement equation but is intentionally more flexible. Many unstructured data series exhibit strong persistence or internal dynamics that are not fully captured by the DSGE state vector S_t. Including the lagged term Y_{t-1}^{NC} allows the model to absorb serial correlation that arises from the construction of the unstructured data themselves, rather than from economic fundamentals. Consequently, we leave B_0, B_1, and B_2 unrestricted, reflecting the fact that we do not know in advance whether any given non-core series is informative about the structural state vector.

A useful way to interpret Y_t^{NC} is as an additional proxy for the state of the economy. If the core variables do not fully measure certain economic forces (for example, if a relevant state variable does not have a direct macroeconomic counterpart), then the unstructured data series may contain information that supplements or refines what is observed through the core variables. In this sense, unstructured data provide extra measurements of the underlying economic states, functioning either as complements to the core observables or as independent signals of components that the core observables do not capture. Measurement error plays a key role here: it ensures that the non-core series can be informative without requiring them to perfectly track the structural states. It also naturally guards against misclassification and estimation errors that arise from processing large unstructured data (Battaglia et al. 2025, Ludwig et al. 2026).

Another way to view our non-core measurement equation is as a vector autoregressive representation for the non-core variables. To see this consider the linearized DSGE model in which $\mathcal{T}(S_{t-1}, \varepsilon_t; \psi) = T(\psi)S_{t-1} + R(\psi)\varepsilon_t$. Substituting this into the non-core measurement equation leads to the following measurement equation,

$$Y_t^{NC} = \Phi_0 + \Phi_1 Y_{t-1}^{NC} + \Phi_2 S_{t-1} + \Phi_3 \varepsilon_t + v_t \tag{4}$$

for suitably defined Φ_0, Φ_1, Φ_2, and Φ_3. Under this representation, the reduced-form innovation in the VAR for Y_t^{NC} is decomposed into two components. The first component, $\Phi_3\varepsilon_t$, is the structural part driven by the DSGE shocks that also govern the core macroeconomic variables. The second component, v_t, is an additional innovation that moves only the non-core variables.

Our model imposes an exclusion restriction: the non-core innovation v_t does not enter the

DSGE transition equation and does not affect the core measurement equation. As a result, the contemporaneous comovement between the core variables and the non-core variables is entirely driven by the structural shocks ε_t. From a structural-model perspective, this restriction is less stringent than it might first appear. If one believes that there exists and additional disturbance that moves both the non-core observables and the DSGE states, this simply indicates that the structural model is misspecified and that the corresponding shock should be added to the structural shock vector ε_t.

In practice, the state vector $S_{1:T}$ is inferred from the joint information in the observables. For fixed structural parameters θ and a correctly specified state-space model, adding Y_t^{NC} as extra measurements can only improve inference about the latent states: the posterior covariance of S_t conditional on θ weakly decreases when we condition on $(Y_{1:T}^C, Y_{1:T}^{NC})$ rather than on $Y_{1:T}^C$ alone. Intuitively, the non-core block provides additional noisy measurements of the same underlying state vector, so at worst these series are uninformative and leave the distribution of S_t essentially unchanged, and at best they sharpen the Kalman smoother or nonlinear filter.

From the perspective of the structural parameters, the likelihood based on the joint data (Y_t^C, Y_t^{NC}) has a score whose conditional expectation given Y_t^C equals the score based on Y_t^C alone, implying that the Fisher information matrix for θ with non-core data weakly dominates the information based only on the core variables. In that idealized setting, additional non-core measurements therefore tend to tighten inference for both the latent states and the structural parameters. In finite samples, however, these potential gains can be offset by overfitting when the non-core measurement block introduces a large number of additional parameters. We next discuss how to mitigate this risk through regularization and selection.

2.3 Incorporating a selection mechanism for unstructured data

Unlike the standard measurement equation for the DSGE model Equation (2), researchers typically do not know how unstructured data are related to the model economy. As a result, we avoid imposing strong a priori restrictions on this relationship. However, when the dimension of the non-core variables is large, leaving all loadings unrestricted gives the model too much flexibility and risks overfitting. Moreover, many variables constructed from unstructured data are unlikely to be informative about the DSGE structure. We therefore introduce a sparsity structure on the parameters that link the non-core series to the model, in particular on the loading matrix B_2. In other words, we do not assume that every unstructured series is necessarily informative for the structural states.

In our empirical analysis, we implement this idea in two steps. First, we assume that only

a subset of the elements of S_t enters the non-core measurement equation. For example, some state variables may be highly correlated (or, even redundant), in which case it is preferable to select only one representative component a priori. To formalize this, we parameterize the loading matrix as $B_2 = \Lambda R$, where R is a deterministic $N_r \times N_s$ selection matrix that picks N_r components of the N_s-dimensional state vector, and Λ is an $M \times N_r$ loading matrix that links the selected state variables to the M non-core series. Under this notation, the non-core measurement equation can be written as

$$Y_t^{NC} = B_0 + B_1 Y_{t-1}^{NC} + \Lambda(RS_t) + v_t, \quad v_t \sim \mathcal{N}(0, \Sigma_v). \tag{5}$$

Second, for the remaining loading parameters in Λ, we allow many elements to be exactly zero. Unlike in the first step, where the selection matrix R is fixed ex ante, we do not specify in advance which entries of Λ should be zero. Instead, we place a spike-and-slab prior on each element $\lambda_{m,i}$ of Λ: conditional on a variance parameter V_m and an inclusion probability q_m,

$$\lambda_{m,i}|V_m, q_m \sim \begin{cases} \mathcal{N}(0, V_m) & \text{with probability } q_m \\ 0 & \text{with probability } 1 - q_m \end{cases} \tag{6}$$

where $\lambda_{m,i}$ denotes the (m,i)-th element of Λ and V_m is a series-specific prior variance. This prior assigns positive probability to $\lambda_{m,i}$ being exactly zero, allowing the model to mute the contribution of the m-th non-core series if the data suggest that it is uninformative about the selected states. As a result, the non-core block behaves as a sparse factor model: some series derived from unstructured data are actively linked to the model's states, while others are effectively set to be irrelevant, which helps prevent contamination from noisy or irrelevant measurements. Our approach differs from data-rich DSGE estimations such as Boivin and Giannoni (2006) and Kryshko (2011), which do not allow for such sparsity and therefore implicitly treat all included series as informative to some extent. The hyperparameter q_m governs the degree of sparsity for the m-th observable: smaller values of q_m place more prior mass on $\lambda_{m,i} = 0$. In our empirical analysis, we follow the parameterization proposed by Giannone et al. (2021) and also estimate the hyperparameters, including q_m; implementation details are provided in Appendix B.

In summary, our framework balances the benefits of incorporating unstructured data into DSGE estimation against the risk of model misspecification. When the mapping between structural model variables and observables is well understood (as for standard macroeconomic variables such as inflation or output) it is efficient to encode that structure directly in the core measurement equation, even at the cost of some vulnerability to misspecifica-

tion. By contrast, for unstructured data the relevant mapping is typically unclear a priori, and many unstructured series are only weakly or not at all informative about the structural states.

Our approach is designed for this latter case. By treating unstructured data as non-core measurements with flexible dynamics, and by using a sparse loading structure that lets the model and the data jointly determine which series are informative, we avoid committing to a potentially incorrect structural mapping while still exploiting useful signals. This yields a practical and robust way to incorporate unstructured data into the structural estimation of dynamic macroeconomic models. Although our empirical application focuses on text data in our application, the same framework can be applied to other types of unstructured information such as satellite imagery, audio data, or high-dimensional firm- or product-level signals and to other structural models used in different fields of economics, including IO (e.g., Compiani et al. 2025).

2.4 Estimation and Inference

In our empirical application, we consider the joint state-space model implied by the linearized DSGE system and the non-core measurement equation:

$$\begin{aligned} Y_t^C &= H_0^C(\theta) + H_1^C(\theta) S_t + u_t, \quad u_t \sim \mathcal{N}(0, \Sigma_u^C(\theta)) \\ Y_t^{NC} &= B_0 + B_1 Y_{t-1}^{NC} + \Lambda(R S_t) + v_t, \quad v_t \sim \mathcal{N}(0, \Sigma_v) \\ S_t &= T(\psi) S_{t-1} + R(\psi)\varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \Sigma_\varepsilon) \end{aligned} \tag{7}$$

This representation involves a large number of unknown parameters. In our empirical application, there are 22 structural parameters in the DSGE block, $M \times 2$ parameters for the intercept and lag coefficients in the non-core measurement equation, $M \times 6$ factor loadings since the non-core observation loads on six structural shocks, M non-core measurement-error variances, and $M \times 2$ hyperparameters for the spike-and-slab prior. For the baseline specification with $M = 30$ topic-share series from the FOMC transcripts, this yields 352 parameters in total.

To address the challenges of posterior sampling in this high-dimensional setting, we use a Metropolis-Hastings-within-Gibbs algorithm. This approach allows us to exploit conditional structure in the model. In particular, we sample the structural parameters of the DSGE model using a random-block random-walk Metropolis-Hasting algorithm (Chib et al. 2023), draw the latent states via a simulation smoother, and update the parameters governing the non-core measurement equation using Gibbs steps combined with discrete updates for the

inclusion indicators.

More concretely, we partition the unknowns into three blocks: θ, $S_{1:T}$, and φ, where θ collects the DSGE structural parameters, $S_{1:T}$ denotes the full history of latent states, and φ contains the parameters appearing in the non-core measurement equation, $(B_0, B_1, \Lambda, \Sigma_v)$. Our target is the joint posterior

$$p(\theta, S_{1:T}, \varphi | Y^C_{1:T}, Y^{NC}_{1:T}). \tag{8}$$

Each MCMC iteration consists of the following steps:

1. Update $(\theta, S_{1:T})$ conditional on φ: We draw from

$$p(S_{1:T}|\theta, \varphi, Y^C_{1:T}, Y^{NC}_{1:T}) p(\theta|\varphi, Y^C_{1:T}, Y^{NC}_{1:T})$$

   by iterating:

   1-1 Sample θ from $p(\theta|\varphi, Y^C_{1:T}, Y^{NC}_{1:T})$ using a random-block random-walk Metropolis-Hasting step. This step is analogous to standard MCMC estimation of a DSGE model based on core observables, but with the likelihood augmented by the non-core measurement equation for Y^{NC}_t.

   1-2 Sample $S_{1:T}$ from $p(S_{1:T}|\theta, \varphi, Y^C_{1:T}, Y^{NC}_{1:T})$ using a simulation smoother (e.g., Durbin and Koopman 2002). This step corresponds to drawing from the smoothing distribution in the joint state-space model with both core and non-core measurements.

2. Update φ conditional on $(\theta, S_{1:T})$: Conditional on $(\theta, S_{1:T})$, this block is independent of the DSGE dynamics. Therefore, the non-core measurement equation becomes a separate regression system with spike-and-slab priors, in which Y^{NC}_t is regressed on Y^{NC}_{t-1} and S_t. The parameters φ are then updated from $p(\varphi|\theta, S_{1:T}, Y^C_{1:T}, Y^{NC}_{1:T})$ using Gibbs sampling and discrete updates for the inclusion indicators.

Full details on the conditional posterior, proposal distributions, and prior specification for the non-core measurement equation are provided in Appendix B.

In our baseline estimation, we run four parallel MCMC chains of length 500,000 each, for a total of 2 million draws. To reduce autocorrelation and conserve memory, we retain every 500th draw, and discard the first 10 percent of retained draws as burn-in. All reported point estimates and credible intervals are computed from the resulting posterior sample.

3 Empirical Application

In this section, we implement the framework in a medium-scale New Keynesian DSGE model using U.S. data. We combine standard U.S. macroeconomic time series with two sources of high-dimensional information: topic-share series extracted from FOMC transcripts and forecast information from the Greenbook/Tealbook. Our empirical analysis is organized around three questions:

1. **Does text improve forecasting performance?** We compare a benchmark DSGE model estimated only with core macro variables to specifications augmented with text-based non-core measurements and with Greenbook/Tealbook information, using a pseudo-out-of-sample forecasting exercise.
2. **Which topics from the FOMC transcripts are informative for the DSGE model?** Conditional on the preferred text-augmented specification, we use the sparse loading structure to identify which FOMC topics are selected as informative for the structural shocks and which are effectively discarded as noise.
3. **How does text affect structural inference?** Taking the best-performing specification as our baseline, we study how the inclusion of non-core unstructured data alters the posterior distribution of structural parameters and the implied impulse responses of key macroeconomic variables.

3.1 New Keynesian DSGE Model

We estimate a medium-scale New Keynesian DSGE model closely following Fernández-Villaverde and Guerrón-Quintana (2021). The economy features a representative household that chooses consumption, saving, real money balances, and labor supply. The final output is manufactured by a final good producer, which uses as inputs a continuum of intermediate goods manufactured by monopolistic competitors. The intermediate good producers rent capital and labor to manufacture their good. These intermediate good producers face the constraint that they can only change prices following a Calvo's rule. The representative household owns firms and receives their profits. Monetary policy sets the one-period nominal interest rate according to an inertial Taylor rule, and government spending follows an exogenous process financed by lump-sum transfers/taxes. Long-run growth is induced by the presence of two unit roots, one in the level of neutral technology and one in the investment-specific technology.

The representative household

The representative household maximizes the following lifetime utility function, which is separable in consumption, c_t, real money balances, m_t/p_t (where p_t is the price level), and hours worked, l_t:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t d_t \left\{ \log\left(c_t - h c_{t-1}\right) + \upsilon \log\left(\frac{m_t}{p_t}\right) - \phi_t \psi \frac{l_t^{1+\vartheta}}{1+\vartheta} \right\}$$

where β is the discount factor, h controls habit persistence, ϑ is the inverse of Frisch labor supply elasticity, d_t is an intertemporal preference shock with law of motion:$\log d_t = \rho_d \log d_{t-1} + \sigma_d \varepsilon_{d,t}$ where $\varepsilon_{d,t} \sim \mathcal{N}(0,1)$, and ϕ_t is a labor supply shock with law of motion:$\log \phi_t = \rho_\phi \log \phi_{t-1} + \sigma_\phi \varepsilon_{\phi,t}$ where $\varepsilon_{\phi,t} \sim \mathcal{N}(0,1)$. The preference shock d_t will change the intertemporal first order conditions, while the preference shock ϕ_t will change the static first order conditions.

The household's budget constraint is given by:

$$\begin{aligned} & c_t + x_t + \frac{m_t}{p_t} + \frac{b_{t+1}}{p_t} \\ &= w_t l_t + \left(r_t u_t - \mu_t^{-1} \Phi\left[u_t\right]\right) k_{t-1} + \frac{m_{t-1}}{p_t} + R_{t-1} \frac{b_t}{p_t} + T_t + \digamma_t \end{aligned}$$

where w_t is the real wage, r_t the real rental price of capital, $u_t > 0$ the intensity of use of capital, $\mu_t^{-1}\Phi\left[u_t\right]$ is the physical cost of use of capital in resource terms, μ_t is an investment-specific technological shock, T_t is a lump-sum transfer, and $\digamma_t$ are the profits of the firms in the economy. We assume that $\Phi\left[1\right] = 0$, Φ' and $\Phi'' > 0$.

Investment x_t induces a law of motion for capital: $k_t = (1-\delta) k_{t-1} + \mu_t \left(1 - S\left[\frac{x_t}{x_{t-1}}\right]\right) x_t$ where δ is the depreciation rate and $S\left[\cdot\right]$ is an adjustment cost function such that $S\left[\Lambda_x\right] = 0$, $S'\left[\Lambda_x\right] = 0$, and $S''\left[\cdot\right] > 0$ where Λ_x is the growth rate of investment along the balance growth path. The investment-specific technological shock follows an autoregressive process: $\mu_t = \mu_{t-1} \exp\left(\Lambda_\mu + z_{\mu,t}\right)$ where $z_{\mu,t} = \sigma_\mu \varepsilon_{\mu,t}$ and $\varepsilon_{\mu,t} \sim \mathcal{N}(0,1)$. The value of μ_t is in equilibrium the inverse of the relative price of new capital in consumption terms.

For future reference, we define λ_t as the Lagrangian multiplier associated with the budget constraint; Q_t as the Lagrangian multiplier associated with installed capital, and the marginal Tobin's Q as $q_t = \frac{Q_t}{\lambda_t}$

The Final Good Producer

There is one final good produced using intermediate goods with the technology:$y_t^d = \left(\int_0^1 y_{it}^{\frac{\varepsilon-1}{\varepsilon}} di\right)^{\frac{\varepsilon}{\varepsilon-1}}$. where ε is the elasticity of substitution.

Final good producers are perfectly competitive and maximize profits subject to the production function above, taking as given all intermediate goods prices p_{ti} and the final good price p_t. As a consequence, their maximization problem is:

$$\max_{y_{it}} p_t y_t^d - \int_0^1 p_{it} y_{it} di$$

The input demand functions associated with this problem are:

$$y_{it} = \left(\frac{p_{it}}{p_t}\right)^{-\varepsilon} y_t^d \qquad \forall i,$$

where y_t^d is the aggregate demand and the zero profit condition $p_t y_t^d = \int_0^1 p_{it} y_{it} di$ to deliver:

$$p_t = \left(\int_0^1 p_{it}^{1-\varepsilon} di\right)^{\frac{1}{1-\varepsilon}}.$$

Intermediate Good Producers

Each intermediate good producer i has access to technology:

$$y_{it} = A_t k_{it-1}^{\alpha} \left(l_{it}^d\right)^{1-\alpha} - f z_t$$

where k_{it-1} is the capital rented by the firm, l_{it}^d is the amount of the "packed" labor input rented by the firm, and where A_t follows the following process:

$$A_t = A_{t-1} \exp\left(\Lambda_A + z_{A,t}\right) \text{ where } z_{A,t} = \sigma_A \varepsilon_{A,t} \text{ and } \varepsilon_{A,t} \sim \mathcal{N}(0,1)$$

The parameter f, which corresponds to the fixed cost of production, and $z_t = A_t^{\frac{1}{1-\alpha}} \mu_t^{\frac{\alpha}{1-\alpha}}$ guarantee that economic profits are roughly equal to zero in the steady state. We rule out the entry and exit of intermediate good producers.

Since $z_t = A_t^{\frac{1}{1-\alpha}} \mu_t^{\frac{\alpha}{1-\alpha}}$, we have that

$$z_t = z_{t-1} \exp\left(\Lambda_z + z_{z,t}\right) \text{ where } z_{z,t} = \frac{z_{A,t} + \alpha z_{\mu,t}}{1-\alpha} \text{ and } \Lambda_z = \frac{\Lambda_A + \alpha \Lambda_\mu}{1-\alpha}.$$

Intermediate goods producers solve a two-stage problem. In the first stage, taken the input

prices w_t and r_t as given, firms rent l_{it}^d and k_{it-1} in perfectly competitive factor markets in order to minimize real cost. that simplifies to:

$$mc_t = \left(\frac{1}{1-\alpha}\right)^{1-\alpha}\left(\frac{1}{\alpha}\right)^{\alpha}\frac{w_t^{1-\alpha}r_t^{\alpha}}{A_t}$$

Note that the marginal cost does not depend on i: all firms receive the same technology shocks and all firms rent inputs at the same price.

In the second stage, intermediate good producers choose the price that maximizes discounted real profits subject to Calvo pricing: in each period, a fraction $1-\theta_p$ of firms can change their prices. All other firms can only index their prices by past inflation. Indexation is controlled by the parameter $\chi \in [0,1]$, where $\chi = 0$ is no indexation and $\chi = 1$ is total indexation.

Then, we can write the g's recursively:

$$g_t^1 = \lambda_t mc_t y_t^d + \beta\theta_p \mathbb{E}_t \left(\frac{\Pi_t^{\chi}}{\Pi_{t+1}}\right)^{-\varepsilon} g_{t+1}^1$$

and

$$g_t^2 = \lambda_t \Pi_t^* y_t^d + \beta\theta_p \mathbb{E}_t \left(\frac{\Pi_t^{\chi}}{\Pi_{t+1}}\right)^{1-\varepsilon}\left(\frac{\Pi_t^*}{\Pi_{t+1}^*}\right) g_{t+1}^2$$

where:

$$\Pi_t^* = \frac{p_t^*}{p_t}.$$

Given Calvo's pricing, the price index evolves:

$$1 = \theta_p\left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{1-\varepsilon} + (1-\theta_p)\,\Pi_t^{*1-\varepsilon}$$

Public sector

We model a consolidated public sector that conducts monetary policy and fiscal policy. We assume there is no interaction between monetary and fiscal policy, that is, the results of open market operations are distributed in a lump-sum fashion to households and not transferred to the general revenue of the government.

Monetary Policy: The government sets the nominal interest rates according to the Taylor rule:

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\gamma_R}\left(\left(\frac{\Pi_t}{\Pi}\right)^{\gamma_\Pi}\left(\frac{\frac{y_t^d}{y_{t-1}^d}}{\Lambda_{y^d}}\right)^{\gamma_y}\right)^{1-\gamma_R}\exp\left(m_t\right)$$

through open market operations that are financed through lump-sum transfers T_t. Those transfers ensure that the deficit from monetary policy is equal to zero:

$$T_t = \frac{m_t}{p_t} - \frac{m_{t-1}}{p_t} + \frac{b_{t+1}}{p_t} - R_{t-1}\frac{b_t}{p_t}$$

The variables Π represents the target level of inflation (equal to inflation in the steady state), R steady state nominal gross return of capital, and Λ_{y^d} the steady state gross growth rate of y_t^d. The term m_t is a random shock to monetary policy that follows $m_t = \sigma_m \varepsilon_{mt}$ where ε_{mt} is distributed according to $\mathcal{N}(0,1)$. The presence of the previous period interest rate, R_{t-1}, is justified because we want to match the smooth profile of the interest rate over time observed in U.S. data. Note that R is beyond the control of the monetary authority, since it is equal to the steady state real gross returns of capital plus the target level of inflation.

Fiscal Policy: Government consumption is given by $g_t = \widetilde{g}_t z_t$, where $\widetilde{g}_t$ follows an autoregressive process in logs:

$$\log \widetilde{g}_t = (1-\rho_g)\log \widetilde{g} + \rho_g \log \widetilde{g}_{t-1} + \sigma_g \varepsilon_{g,t} \quad \text{where } \varepsilon_{g,t} \sim \mathcal{N}(0,1)$$

and

$$z_t = A_t^{\frac{1}{1-\alpha}} \mu_t^{\frac{\alpha}{1-\alpha}}$$

Since the economy is growing, the level of real government consumption is multiplied by the term z_t defined above to keep it a stationary share of output. Government consumption is also financed by lump-sum taxes that ensure that the deficit is zero at all points in time.

Aggregation

Applying the definition of transfers above, the aggregated budget constraint of households is equal to: $c_t + x_t = w_t l_t^d + \left(r_t u_t - \mu_t^{-1}\Phi\left[u_t\right]\right) k_{t-1} + \digamma_t$.. Aggregate demand is given by: $y_t^d = c_t + g_t + x_t + \mu_t^{-1}\Phi\left[u_t\right] k_{t-1}$. Aggregate supply is: $y_t^s = \frac{A_t (u_t k_{t-1})^{\alpha} \left(l_t^d\right)^{1-\alpha} - f z_t}{v_t^p}$, where $v_t^p = \int_0^1 \left(\frac{p_{it}}{p_t}\right)^{-\varepsilon} di$ is the price dispersion generated by the price rigidities. By the properties of Calvo's pricing, v_t^p is given by: $v_t^p = \theta_p \left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{-\varepsilon} v_{t-1}^p + (1-\theta_p)\,\Pi_t^{*-\varepsilon}$.

Equilibrium and Solution

A definition of equilibrium in this economy is standard and can be found in the Supplementary Appendix A. In the appendix, we also describe how we make the model stationary by rescaling the equilibrium conditions, the steady state of the rescaled equilibrium condi-

tions, and the log-linearization of the equilibrium conditions around the steady state. After log-linearization, we finally cast the model in a state-space representation summarized in Appendix A.

3.2 Data

We use six U.S. macroeconomic variables in the core measurement equation. The core vector Y_t^C includes inflation ($\log \Pi_t$), the short-term nominal interest rate ($\log R_t$), real wage growth ($\triangle \log w_t$), real output growth ($\triangle \log y_t$), hours worked ($\log l_t$), and the growth rate of the relative price of investment goods ($\triangle \log \mu_t^{-1}$),

$$Y_t^C = \left(\log \Pi_t, \log R_t, \triangle \log w_t, \triangle \log y_t, \log l_t, \triangle \log \mu_t^{-1}\right)'.$$

In this application, we consider two types of non-core variables. The first, and conceptually central to our paper, is based on unstructured text data. Specifically, we use the Federal Open Market Committee (FOMC) meeting transcripts. We apply a latent Dirichlet allocation (LDA) topic model to the transcripts and obtain a time series of topic shares. These topic shares provide a numerical representation of the high-dimensional text, measuring the fraction of each (latent) topic discussed in a given period.

Second, to assess whether the transcripts contain information beyond other numerical summaries of the FOMC's information set, we construct an alternative non-core block from the Greenbook (now Tealbook) nowcasts and forecasts. These data provide a conventional numerical representation of the FOMC staff's assessment of current and future economic conditions, which we use as a benchmark against which to compare the incremental value of the text-based measures.

All series are sampled at a quarterly frequency. The estimation sample runs from 1987Q3 to 2019Q4, reflecting the availability of FOMC transcripts up to 2019Q4 at the time we conducted the empirical analysis. Further details on data construction and transformations are provided in Supplementary Appendix B.

3.2.1 FOMC transcripts

We use the official FOMC meeting transcripts made available by the Board of Governors of the Federal Reserve System. For each FOMC meeting, staff transcribe the recorded proceedings, producing a near-verbatim written record that provides a nearly complete account of discussion. The transcripts are released with a substantial lag, typically about five years after the meeting took place. In our empirical work, we use the set of transcripts covering

FOMC meetings from 1987 through 2019, which were publicly available at the time of our analysis.

In our analysis, we consider three different transcript datasets. In the first, which we label FOMC1, we use only the "economic situation" part of each meeting, where staff present the outlook and participants discuss current and projected economic conditions. In the second, FOMC2, we use only the "policy discussion" part of each meeting, where policy alternatives are presented and members debate the appropriate monetary policy stance. In the third, FOMC-both, we combine FOMC1 and FOMC2. We use these three alternatives because it is unclear *ex ante* which part of the FOMC meetings contains the most information lies for model estimation. Below we propose an approach for selecting the most informative part(s).

From each dataset, our aim is to obtain a structured, quantitative set of quarterly time series. As is typically the case with unstructured datasets, some form of dimensionality reduction is necessary for generating these. We use two alternative algorithms to this end. The first—and the main focus of our analysis—is the Latent Dirichlet Allocation (LDA) topic model. The LDA model assumes that each document is a mixture of a fixed number of topics, and that each word in the document is generated by first drawing a topic from the document's topic distribution and then drawing a word from the topic-specific word distribution. Different words therefore have different probabilities of occurrence across topics, and each transcript is summarized by a vector of topic shares. This structure is particularly useful in our context because it both uncovers the main themes discussed in the FOMC transcripts and quantifies how prominently each topic is discussed in each document.

The second algorithm is the `text-embedding-3-large` model provided by OpenAI. This model takes a sequence of input text and converts it to a numeric vector (of dimensionality 3,072) by passing it through a Transformer model. Unlike LDA, which operates on individual term counts, Transformers are context-aware and produce different embeddings for different sequences even if term counts remain unchanged.[3] Loosely speaking, the model is trained so that documents with similar meaning have similar embeddings. While Transformer models are more powerful for capturing the semantic structure of documents, they are less interpretable than LDA. As we show below, at least without further fine-tuning on our dataset, they also appear to contain less information for model estimation. For that reason, we focus on LDA in the main analysis.

[3]For example, LDA treats the documents 'apples are better than oranges' and 'oranges are better than apples' as equivalent because they have the same term counts. But they would produce different vectors in the embedding model.

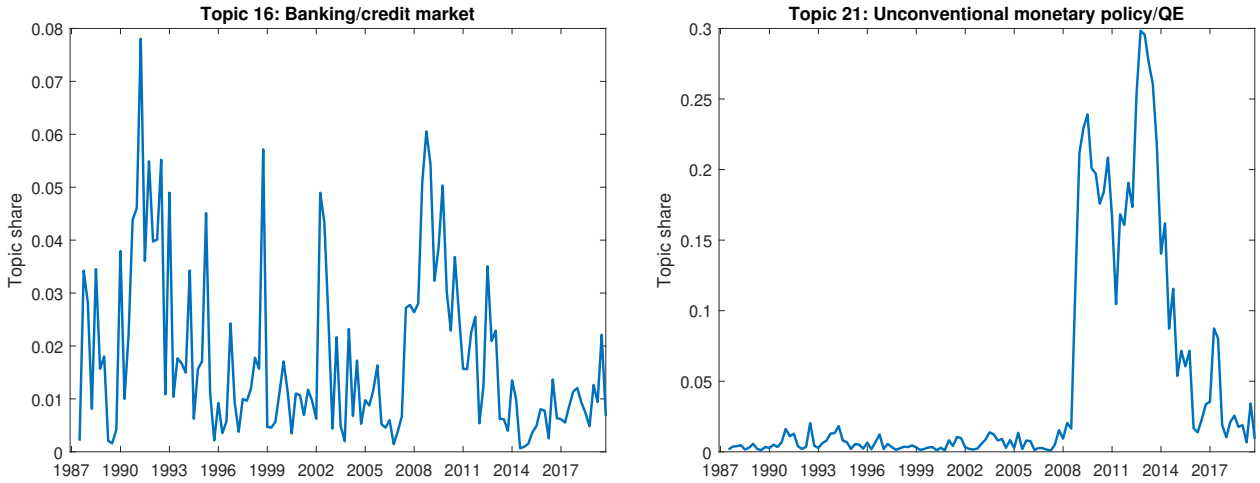
LDA estimation After pre-processing the FOMC transcripts, we proceed in two steps to obtain quarterly time series for each transcript dataset. First, we estimate LDA at the level of individual speaker utterances for different numbers of topics ($K = 20, 30, 40$) using the MCMC estimation strategy of Griffiths and Steyvers (2004) implemented in the R `topicmodels` package.[4] For each dataset and each number of topics, we initialize three different Markov chains from randomly drawn seeds, and retain the chain that achieves highest average log-likelihood across draws following a burn-in period.

Second, to obtain topic share distributions at the quarterly frequency, we re-estimate LDA by aggregating individual speaker utterances into a single document per quarter. During MCMC estimation, we hold fixed the topic-term distributions at their posterior mean values from the best-fit chains in step one. The resulting quarterly topic shares $\{\varphi_{1,t}, ..., \varphi_{K,t}\}$ are the posterior means computed from the MCMC draws. For example, $\varphi_{k,t}$ measures the popularity of topic k relative to the other topics in quarter t.

Estimated topic share and top terms. To illustrate the topic-term probabilities, Table 1 shows the fifteen terms with the highest probability for each topic for the FOMC2 dataset and $K = 30$. In some cases, the top-probability words in a topic convey a coherent theme which we use to manually assign a label to the topic. For instance, Topic 17 contains "economi", "risk", and "uncertainti", so we interpret the topic as representing economic risk and uncertainty. Topic 21 contains words related to unconventional monetary policy such as "purchas", "program", and "balance sheet". At the same time, some topics are difficult to interpret or contain mainly conversational words. For instance, Topic 15 contains "yes", "governor", and "laughter". We expect that our econometric framework disregards conversational topics, while capturing information from economically relevant topics, such as Topic 21.

The time-series dynamics of these topics support the validity of our labels. Figure 1 presents the time-series behavior of Topics 16 and 21, estimated using the topic model with $K = 30$ and the FOMC2 dataset. Topic 16 (Banking/Credit Market) exhibits a sharp increase following the Global Financial Crisis. Topic 21 (Unconventional Monetary Policy/QE) remained near zero prior to the crisis but surged immediately afterward, coinciding with the introduction of various unconventional policy tools. Its share subsequently declines after the rate hikes began in 2016. These movements line up closely with standard narratives of U.S. economic and policy conditions over the sample.

[4] We use a uniform prior over the document-topic share distribution and a symmetric Dirichlet prior with hyper-parameter 0.5 over the topic-term distribution.



Note: This figure shows the topic share series for Topics 16 and 21, estimated from the topic model with $K = 30$ using FOMC2 transcript data.

Figure 1: Time series dynamics of Topics 16 and 21

Bringing topic shares into non-core measurement. At each time t, the topic-share vector $[\varphi_{1,t}, ..., \varphi_{K,t}]$ lies on the simplex: the elements are non-negative and sum to one. To use these series in our linear Gaussian non-core measurement equation, we first map them from the simplex into $\mathbb{R}^K$. Specifically, we apply the centered log transformation

$$a_{k,t} = \log(\varphi_{k,t}) - \frac{1}{K}\sum_{k=1}^{K}\log(\varphi_{k,t}),$$

and include the transformed series $a_{k,t}$ as non-core observables. In total, we have nine separate sets of quarterly time series: $K = 20, 30, 40$ models for each of FOMC1, FOMC2, and FOMC-both. We then use the pseudo out-of-sample forecasting exercise to select the specification (that is, the combination of corpus and K) that delivers the best predictive performance.

3.2.2 Greenbook/Tealbook forecasts

Because our text data come from FOMC meeting transcripts, it is natural to ask whether they contain information beyond other numerical summaries of the FOMC's information set, in particular the nowcasts and forecasts in the Greenbook (now Tealbook). We therefore construct an alternative non-core block based on the Greenbook/Tealbook projection data.

We obtain the Greenbook numerical data from the Federal Reserve Bank of Philadelphia's

Table 1: Top-Probability Words by Topic (30-Topic Model for Policy Discussion FOMC2)

Label	Topic	Top1	Top2	Top3	Top4	Top5	Top6	Top7	Top8	Top9	Top10	Top11	Top12	Top13	Top14	Top15
Labor	Topic1	unemploy	labor	employ	unemploymentratexx	job	labormarketxx	trend	wage	worker	forc	declin	past	slack	level	particip
Model analysis	Topic2	model	use	estim	gap	differ	rule	shock	effect	staff	measur	output	analysi	simul	can	actual
?	Topic3	seem	may	economi	market	bit	quit	much	strong	reason	certain	data	point	suggest	perhap	come
Policy stance - tighting	Topic4	might	economi	polici	tighten	may	eas	inflat	seem	expect	effect	market	demand	level	suggest	need
Business conditions / Beige Book	Topic5	report	busi	contact	compani	sale	percent	increas	product	firm	price	industri	hear	plan	retail	cost
Conversational / transition language	Topic6	realli	thing	actual	want	talk	lot	idontxx	much	someth	look	term	peopl	right	tri	know
Inflation (expectations)	Topic7	inflat	percent	expect	inflationexpectationsxx	low	measur	target	run	around	lower	core	coreinflationxx	level	unemploymentratexx	remain
Policy communication / statement drafting	Topic8	committe	altern	might	polici	may	federalfundsratexx	econom	indic	consist	statement	note	meet	object	period	expect
Conversation	Topic9	question	ask	want	let	comment	thought	issu	discuss	start	know	answer	interest	point	rais	talk
Conversation	Topic10	know	laughter	can	peopl	world	believ	happen	work	come	cannot	great	worri	around	hope	mean
Policy communication/statement	Topic11	statement	languag	chang	support	word	altern	meet	suggest	paragraph	communic	today	use	want	alternativebxx	sentenc
Fiscal policy	Topic12	fiscal	economi	recoveri	recess	effect	state	tax	europ	govern	uncertainti	confid	econom	much	may	budget
Financial market expectation	Topic13	market	expect	-	point	period	yield	low	recess	relat	probabl	term	chang	curv	move	basi
?	Topic14	move	point	market	direct	meet	tighten	support	eas	need	today	toward	believ	agre	recommend	asymmetr
Conversation	Topic15	yes	governor	okay	laughter	right	liftoff	vice	lacker	fisher	bullard	evan	hoenig	stern	break	parri
Banking/Credit Market	Topic16	bank	credit	market	loan	fund	financi	mortgag	debt	capit	lend	asset	borrow	liquid	price	increas
Risk and Uncertainty	Topic17	risk	concern	market	uncertainti	economi	financi	downsid	balanc	outlook	potenti	econom	possibl	signific	weak	develop
Household consumption/investment	Topic18	spend	invest	household	consum	busi	incom	inventori	hous	sector	consumpt	product	growth	declin	expect	capit
Data / statistics	Topic19	percent	quarter	data	month	revis	estim	last	project	gdp	expect	averag	june	septemb	surpris	march
Regional economy	Topic20	district	nation	continu	area	region	economi	manufactur	sector	construct	employ	report	strong	industri	sale	activ
Unconventional policy / QE	Topic21	purchas	program	market	end	condit	expect	billion	provid	effect	septemb	improv	cost	balancesheetxx	addit	threshold
Trade	Topic22	dollar	unit	state	import	export	countri	growth	foreign	trade	global	china	japan	declin	abroad	economi
Greenbook forecast	Topic23	forecast	greenbook	project	growth	staff	assum	chang	assumpt	period	baselin	lower	path	end	higher	next
Inflation / Cost pressures	Topic24	price	increas	product	inflat	cost	rise	growth	higher	wage	pressur	energi	oil	suppli	demand	slow
??	Topic25	look	realli	thing	side	littl	lot	peopl	much	certain	term	sens	pretti	number	situat	come
ZLB / Forward guidance	Topic26	polici	can	need	monetarypolicyxx	economi	action	today	zero	monetari	import	futur	credibl	believ	public	view
SEP /guidance language	Topic27	chair	madam	path	continu	gradual	polici	economi	increas	data	move	sep	support	tealbook	outlook	appropri
Projection/Data explanation	Topic28	panel	project	shown	line	right	percent	left	show	expect	chart	next	middl	exhibit	bottom	lower
Economic growth outlook	Topic29	continu	remain	month	growth	recent	expect	improv	indic	survey	firm	outlook	increas	pace	report	posit
??	Topic30	can	signific	may	look	number	data	economi	move	clear	evid	know	inde	quit	probabl	issu

Note: For each topic, we rank terms by their probability estimated by LDA. This Table presents the top 15 words according to the ranking, along with a topic lable assigned by the authors.

real-time data archive.[5] The dataset contains 15 macroeconomic series, including real GDP, the GDP price index, unemployment, and so forth.[6]

Each variable is reported with its historical values up to four quarters back, its current estimate (nowcast), and forecasts up to eight quarters ahead, although missing values become more frequent at longer horizons. We construct a panel of 90 series by combining all 15 variables at six horizons $(T-1, T, T+1, T+2, T+3, T+4)$. We then estimate a small number of factors from these 90 series. For details of factor estimation, see Supplementary Appendix B.

3.3 Model specification

Depending on the data used in the non-core block, we consider several different model specifications. The common component across all specifications is the DSGE state-transition block

$$S_t = T(\psi)S_{t-1} + R(\psi)\varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, I) \tag{9}$$

where the state vector S_t collects the model's endogenous states, exogenous processes, and a small set of auxiliary variables (such as innovations and lags needed to write the system in first-order form):

$$S_t = [state_t', exo_t', \widehat{\widetilde{w}}_t, \widehat{\widetilde{w}}_{t-1}, \widehat{\widetilde{y}}^d_{t-1}, \widehat{\widetilde{z}}_t, \widehat{l}_t, \varepsilon_{d,t}, \varepsilon_{\phi,t}, \varepsilon_{g,t}]'. \tag{10}$$

The vector $state_t$ contains the main endogenous state variables,

$$state_t = \left(\widehat{\Pi}_t, \widehat{g}^1_t, \widehat{g}^2_t, \widehat{\widetilde{k}}_t, \widehat{R}_t, \widehat{\widetilde{y}}^d_t, \widehat{\widetilde{c}}_t, \widehat{v}^p_t, \widehat{\widetilde{q}}_t, \widehat{\widetilde{x}}_t\right)',$$

and exo_t collects the exogenous processes,

$$exo_t = \left(\widehat{\widetilde{\mu}}_t, \widehat{d}_t, \widehat{\phi}_t, \widehat{\widetilde{A}}_t, m_t, \widehat{\widetilde{g}}_t\right)'.$$

[5] https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/philadelphia-data-set

[6] Specifically, the dataset records real gross domestic product, gross domestic product price index, unemployment, headline consumer price index, core consumer price index, headline personal consumption expenditures price index, core personal consumption expenditures price index, real personal consumption expenditures, real business fixed investment, real residential investment, real federal government consumption and gross investment, real state and local government consumption and gross investment, nominal gross domestic product, housing starts, and the industrial production index.

In total, S_t is a 24×1 vector. The transition matrices $F(\theta)$ and $G(\theta)$ are known functions of the structural parameter vector θ, and we estimate 22 structural parameters in the DSGE block.

Prior distribution for structural parameters. The system matrices in the state transition are known functions of the structural parameters. We impose standard prior distributions on these parameters, as summarized in Table 2. The priors are chosen to be similar to those in Fernández-Villaverde and Guerrón-Quintana (2021), and following their specification we calibrate a subset of parameters to fixed values $\vartheta = 1, \delta = 0.025, \varepsilon = 10, f = 0$, and $\Phi_2 = 0.01$.

Benchmark model. In our benchmark specification, we use only the core variables that are the six macroeconomic time series described above. The core measurement equation takes the form

$$Y_t^C = H_0^C(\theta) + H_1^C(\theta) S_t + \Sigma_u^{1/2} u_t \tag{11}$$

where $u_t \sim \mathcal{N}(0, I)$ and Σ_u is a diagonal matrix of measurement-error variances. For example, the first element of Y_t^C is observed inflation, $\log \Pi_t^{obs}$, which we link to the model-implied inflation $\widehat{\Pi}_t$. The construction of the full core measurement equation, including the mapping from each observable to the underlying model variables, is standard and is described in detail in Appendix A.

Model with text. Our main specification augments the benchmark DSGE model with non-core variables constructed from the FOMC meeting transcripts. As described in the previous subsection, we convert the transcripts into numerical form using the LDA topic model and obtain K transformed topic-share series $a_{k,t}$. We collect these in the non-core observation vector $Y_t^{NC} = [a_{1,t}, a_{2,t}, ..., a_{K,t}]'$. As discussed earlier, we preselect a subset of the state variables that can enter the non-core measurement equation by specifying a selection matrix R such that $RS_t = \varepsilon_t = [\varepsilon_{D,t}, \varepsilon_{\varphi,t}, \varepsilon_{G,t}, \varepsilon_{\mu,t}, \varepsilon_{A,t}, \varepsilon_{M,t}]'$. That is, the selection matrix picks out the six structural shocks of the DSGE model. We then link the topic-based

Table 2: Prior distributions for DSGE structural parameters

Parameter	Description	Domain	Density	Mean	SD	LB	UB
Steady-state-related parameters							
$100(1/\beta - 1)$	β is discount factor	$(0, 1)$	Gamma	0.25	0.1	1e-7	100
g_{SS}	SS government expenditure/GDP	$(0, 1)$	Beta	0.3	0.05	0.01	0.9999
$100(\Pi^* - 1)$	Target inflation rate	$\mathbb{R}$	Gamma	0.95	0.1	0	10
Endogenous propagation parameters							
h	Habit persistence	$(0, 1)$	Beta	0.7	0.1	0.3	0.9999
α	Cobb–Douglas labor	$(0, 1)$	Normal	0.3	0.025	0.01	0.9999
κ	Investment adjustment cost	$\mathbb{R}$	Normal	4	1.5	0.01	40
θ_P	Fraction of firms with fixed prices	$(0, 1)$	Beta	0.5	0.1	0.01	0.9999
χ_P	Price indexation	$(0, 1)$	Beta	0.5	0.15	0.01	0.9999
γ_R	Taylor rule coefficient past rates	$(0, 1)$	Beta	0.75	0.1	0.01	0.9999
γ_π	Taylor rule coefficient inflation	$\mathbb{R}^+$	Normal	1.5	0.125	1.01	5
γ_Y	Taylor rule coefficient demand	$\mathbb{R}^+$	Normal	0.12	0.05	0.01	1.5
Exogenous shock parameters							
ρ_D	Persistence demand shock	$(0, 1)$	Beta	0.5	0.2	0.01	0.9999
ρ_φ	Persistence labor supply shock	$(0, 1)$	Beta	0.5	0.2	0.01	0.9999
ρ_G	Persistence government consumption shock	$(0, 1)$	Beta	0.5	0.2	0.01	0.9999
$100\lambda_\mu$	Long-run growth investment specific productivity	$\mathbb{R}$	Gamma	0.34	0.1	-5	5
$100\lambda_A$	Long-run growth productivity	$\mathbb{R}$	Gamma	0.178	0.075	-5	5
σ_D	SD demand shock innovation	$\mathbb{R}^+$	Inverse gamma	0.1	1	1e-5	100
σ_φ	SD labor supply shock innovation	$\mathbb{R}^+$	Inverse gamma	0.1	1	1e-5	100
σ_G	SD government consumption shock innovation	$\mathbb{R}^+$	Inverse gamma	0.1	1	1e-5	100
σ_μ	SD investment-specific shock	$\mathbb{R}^+$	Inverse gamma	0.1	1	1e-5	100
σ_A	SD long-run productivity	$\mathbb{R}^+$	Inverse gamma	0.1	1	1e-5	100
σ_M	SD monetary policy shock	$\mathbb{R}^+$	Inverse gamma	0.1	1	1e-5	100

Note: This table summarizes the prior distributions on the parameters of our DSGE model. SD denotes the standard deviation of the prior distribution, and LB and UB are the lower bound and upper bound constraints of the parameters, respectively, so that the parameter draw outside of this bound is discarded.

non-core variables directly to these structural shocks via the measurement equation[7]

$$Y_t^{NC} = B_0 + B_1 Y_{t-1}^{NC} + \Lambda \varepsilon_t + v_t, \quad v_t \sim \mathcal{N}(0, \Sigma_v), \tag{12}$$

so that Λ is a $K \times 6$ loading matrix. To keep the specification parsimonious, we assume that B_1 and Σ_v are diagonal. As discussed in Section 2, we impose a spike-and-slab prior on Λ, reflecting the fact that many topics derived from the transcripts may be only weakly informative about structural shocks. In particular, some topics appear primarily conversational or administrative, and the prior allows the estimation to shrink their loadings toward zero when the data do not support a strong link to the structural shocks.

We estimate this text-augmented specification for three different values of $K, \in \{20, 30, 40\}$. For each value of K, we consider three transcript corpora (FOMC1, FOMC2, FOMC-both) so that, in total, we estimate nine text-based specifications and later use the pseudo-out-of-sample forecasting exercise to select the specification that performs best.

Model with Greenbook. Text data are not the only potential source of additional information for DSGE estimation. Although the FOMC transcripts may contain useful discussion about the current state of the economy, some of that information may already be summarized in numerical form in the Greenbook/Tealbook projections. To assess this, we estimate an alternative specification in which Y_t^{NC} consists of factors extracted from the Greenbook/Tealbook nowcasts and forecasts described above. Because these projections synthesize a wide range of relationships among macroeconomic variables, it is not obvious how to map them directly to specific state variables in the DSGE model. Instead, we treat them in the same way as the text-based non-core variables and use an analogous non-core measurement equation,

$$Y_t^{NC} = B_0 + B_1 Y_{t-1}^{NC} + \Lambda \varepsilon_t + v_t, \quad v_t \sim \mathcal{N}(0, \Sigma_v) \tag{13}$$

where Y_t^{NC} now collects the Greenbook/Tealbook factors. As in the text-based specification, we assume B_1 and Σ_v are diagonal and impose a spike-and-slab prior on Λ, allowing the data to determine which Greenbook-based series, if any, load meaningfully on the structural

[7]When all transformed topic-share series from the estimated LDA are included, the non-core block is linearly dependent. A fully coherent treatment would work with $K - 1$ log-ratio coordinates. We instead keep all K series and impose diagonal restrictions on B_1 and Σ_v, which amounts to a working-independence (equation-by-equation) approximation for the non-core block. We view this as a pragmatic approximation because topic shares are estimated objects and thus not exact in finite samples. Empirically, the spike-and-slab prior yields very low inclusion probabilities for many topic-shock loadings, so the effective information content of the non-core block is driven by a smaller set of informative topics.

shocks.

3.4 Results

In this subsection, we present our empirical findings along three dimensions: (i) whether incorporating text improves the DSGE model's forecasting performance, (ii) which FOMC topics are selected as informative for the structural shocks, and (iii) how the inclusion of text alters the posterior distribution of key structural parameters and the associated impulse responses.

3.4.1 Does text improve forecasting performance?

We begin by evaluating whether incorporating unstructured data improves the DSGE model's forecasting performance. To this end, we conduct a pseudo-out-of-sample forecasting exercise and compare three classes of specifications: the benchmark model without text, the text-augmented model using FOMC transcripts, and the model augmented with Greenbook/Tealbook factors. Our primary question is whether the additional non-core measurements improve predictions for the core variables, which are common across all specifications. The forecasting evaluation also allows us to assess whether the transcripts contain information beyond the Greenbook projections and to select, among the text-based models, the preferred combination of the number of topics K and the transcript corpus (FOMC1, FOMC2, or FOMC-both).

Measuring forecasting accuracy. We start the forecasting exercise with data up to 2000Q1 and then recursively expand the sample, re-estimating each model and generating out-of-sample predictions at each step. We follow standard procedures for DSGE forecast evaluation (e.g., Del Negro and Schorfheide 2013); the detailed algorithm is described in the Supplementary Appendix C.

We use three measures of forecasting accuracy. The first is the joint log score, with larger values indicating better predictive performance. This measure summarizes the one-step-ahead predictive performance for the vector of core variables and serves as an aggregate fit criterion. The second and third measures are the root mean squared error (RMSE) and the continuous ranked probability score (CRPS), respectively. RMSE assesses point forecast accuracy, while CRPS evaluates the accuracy of the entire predictive distribution; for both, smaller values indicate better performance. RMSE and CRPS are reported variable by variable and are useful for assessing how text and Greenbook information affect predictive

Table 3: Joint Log Scores Across Model Variants

# of topic	Transcript	Joint Log Score
Without text	–	1862.15
With Greenbook	–	1861.54
K = 20	FOMC 1	1863.75
	FOMC 2	1865.96
	Both	1864.61
K = 30	FOMC 1	1863.30
	FOMC 2	**1867.80**
	Both	1862.40
K = 40	FOMC 1	1860.51
	FOMC 2	1864.97
	Both	1855.19
OpenAI embeddings	FOMC2	1864.04

Note: This table presents the joint log scores for models without text, with Greenbook text, with FOMC transcripts, across different numbers of topics ($K = 20, 30, 40$) and transcript types (FOMC1, FOMC2, and both), and OpenAI embeddings built from FOMC2 and PCA.

performance for individual core variables.

Forecast evaluation results: Joint log score. Table 3 reports the joint log scores for the benchmark model without text, the Greenbook-based specification, and the text-augmented specifications with different numbers of topics K and different transcript corpora (FOMC1, FOMC2, and FOMC-both). Five main findings emerge.

First, except for the FOMC1 specification with $K = 40$, all models that incorporate FOMC transcripts outperform the benchmark model without text in terms of the joint log score, regardless of K and the transcript subset.

Second, the specification with $K = 30$ estimated on the FOMC2 corpus achieves the highest joint log score. This suggests that the policy-discussion part of the transcripts contains more useful information for forecasting than either the economic-situation part (FOMC1) or the full combined transcript (FOMC-both). This is intuitive, as the FOMC meeting is ultimately organized around deciding how to set monetary policy in response to current and expected economic conditions.

Third, the Greenbook-based specification performs worse than both the text-based models and, in fact, even the benchmark without non-core measurements. This implies that adding

numerical projections alone does not guarantee better forecasts in our setting.

Fourth, these results highlight a classic parsimony-overfitting trade-off: when weakly informative or non-informative series are loaded into the non-core measurement equation, they can primarily add noise and reduce forecast accuracy. The poorer performance of the Greenbook specification, and the fact that FOMC2 outperforms FOMC-both for a given K, both suggest that including too much irrelevant or redundant information can lower forecast accuracy, even with our spike-and-slab prior.

Finally, we compare LDA to more modern embedding models. We embed the full text of the FOMC2 section of each meeting using OpenAI's `text-embedding-3-large` model, obtain quarterly embeddings by averaging, then reduce dimensionality via PCA and select the top thirty components for to build quarterly time series.[8] We focus on FOMC2 text and 30 dimensions to align with the most informative topic share series. While embeddings increase the joint log score compared to the no-text baseline, they do not outperform topic-based series.

Forecast evaluation results: Individual variable RMSE and CRPS. Figures 2 and 3 report the relative RMSE and CRPS for models augmented with additional data (red line for FOMC transcripts and green line for Greenbook factors) relative to the benchmark model without non-core measurements (black line, plotted in levels). Values below one indicate better performance than the benchmark. For the text-based specification, we show results for the preferred model identified above (FOMC2 with $K = 30$).

The text-augmented model consistently outperforms the benchmark for forecasting the federal funds rate and hours worked, and delivers lower RMSE across most horizons for output and wages, with gains in the range of roughly 5-15 percent. It performs slightly worse for inflation (by at most about 5 percent), while the results for PIRIC are broadly comparable to the benchmark. The CRPS results closely mirror the RMSE patterns, indicating that these improvements also hold when evaluating the full predictive distribution.

Although the one-step-ahead joint log score of the Greenbook-based model is lower than that of the benchmark, it does deliver some gains for specific variables. At the one-step-ahead horizon, the Greenbook specification improves RMSE and CRPS for output and hours, and it provides better medium-run inflation forecasts than the other models. These gains, however, are more than offset by its poor performance for interest rate forecasts, which is the main reason for its weaker overall joint fit.

The model with text is particularly effective for forecasting the federal funds rate, which is

[8] We also reduced dimesionality with UMAP but found worse performance than with PCA.

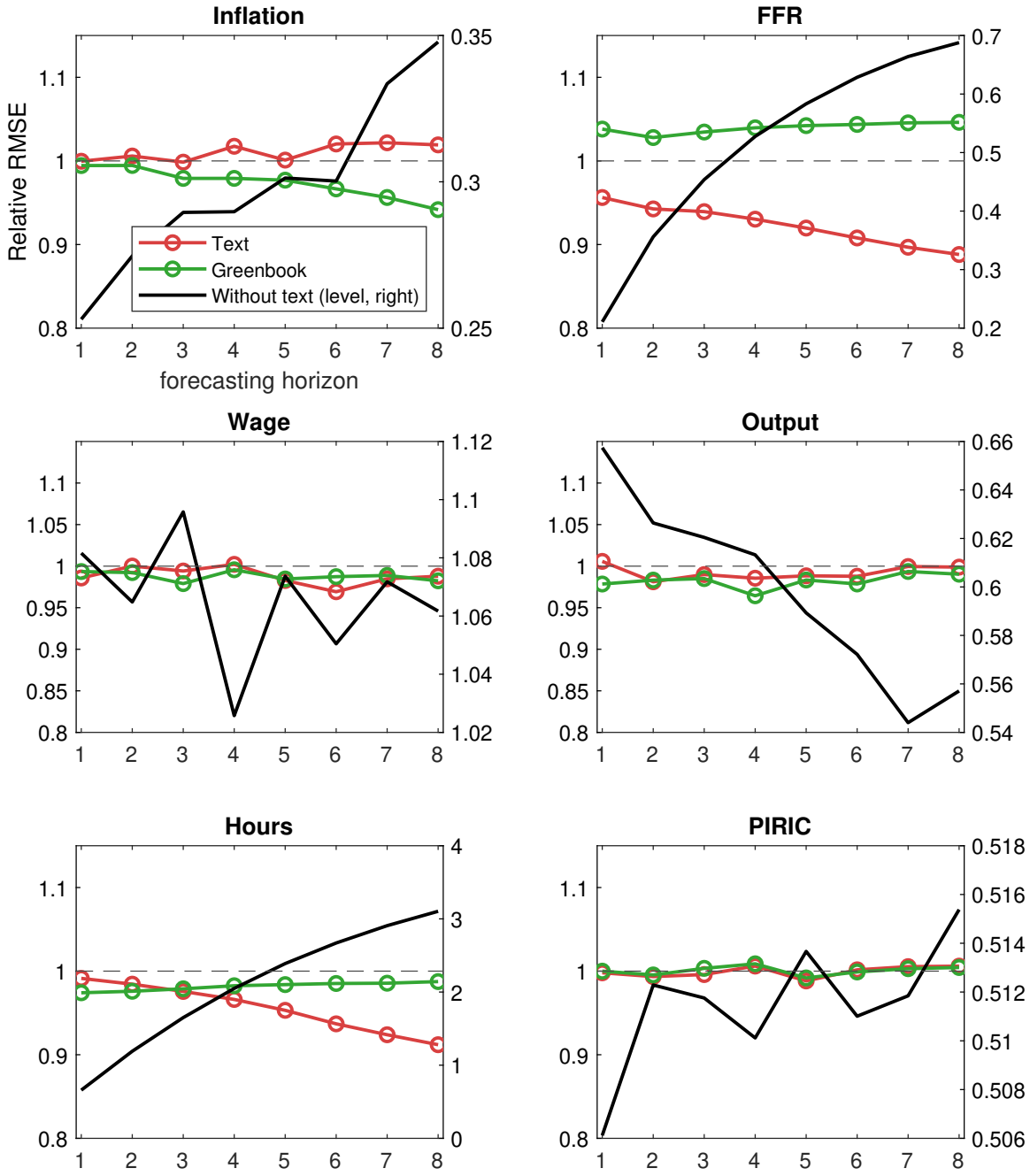
intuitive given that FOMC2 focuses on monetary policy discussions. Moreover, the improvement is not driven by a single episode but reflects persistent gains over the sample. Figure 4 shows the cumulative absolute forecast error for the federal funds rate: the gap between the text-based model and the benchmark widens gradually over time, indicating that the text-augmented specification delivers more accurate interest rate forecasts throughout the evaluation period.

In summary, the FOMC transcript data provide additional information that improves forecasting performance relative to both the benchmark model with only standard macroeconomic variables and the Greenbook-based specificaiton. The superior predictive perofrmance of the text-augmented models confirms that incorporating unstructured data into DSGE estimation yields valuable additional information for forecasting.

3.4.2 Which topics of the FOMC transcript are informative in estimation?

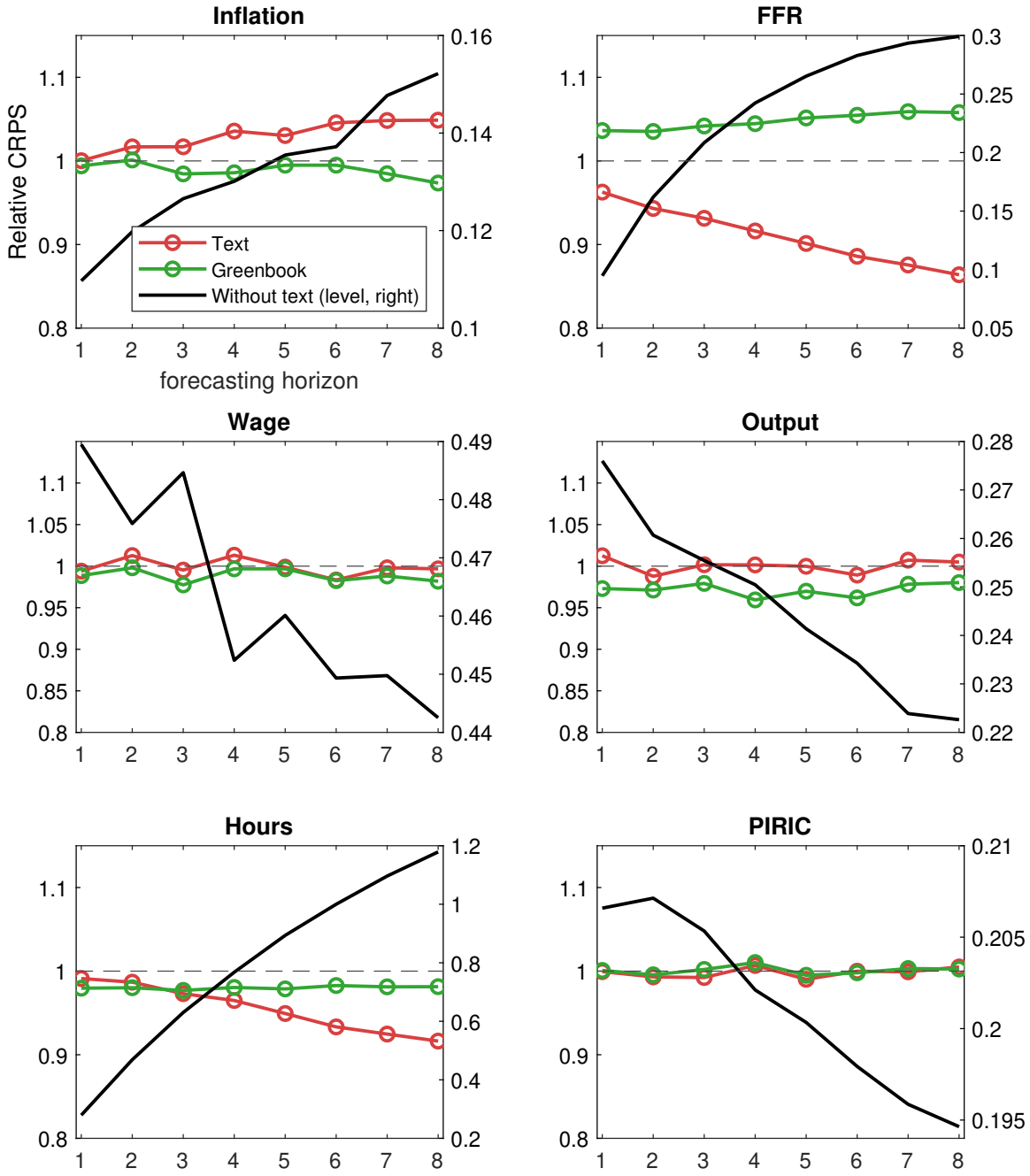
We now examine how the FOMC transcripts affect the full-sample estimation of our DSGE model, focusing on which topics are selected as informative. Throughout this subsection, we focus on our preferred specification, FOMC2 with $K = 30$. Figure 5 reports the posterior median of the factor loadings in Λ, showing how each topic-share series loads on the structural shocks. Because the spike-and-slab prior allows each loading to be exactly zero with positive probability, many entries of Λ are shrunk to zero in the posterior, yielding a sparse pattern of topic-state connections.

Several findings emerge. First, the FOMC transcripts clearly provide information that can be linked to structural shocks. The loadings on the monetary policy shock m are particularly informative and economically interpretable. The shock loads *poistively* on Topic 4 (Policy stance - tightening), Topic 5 (Business conditions / Beige Book), Topic 19 (Data / statistics), Topic 20 (Regional economy), and Topic 24 (Inflation and cost pressures), and *negatively* on and Topic 1 (Labor market), Topic 6 (Conversational / transition language), Topic 8 (Policy communication / statement drafting), Topic 21 (Unconventional monetary policy / QE), Topic 26 (Zero lower bound / forward guidance), Topic 27 (Guidance and Summary of Economic Projections (SEP) language). The positive group corresponds to a hawkish or monitoring regime: episodes of tightening are associated with greater discussion of inflation, strong activity and district conditions, and statistical justification. The negative group corresponds to a communication or accommodation regime: when policy is maintained or eased, the transcripts emphasize statement mechanics, guidance, balance-sheet tools, transitional language, and labor-market slack. Consistent with this interpretation, Figure 1 shows that Topic 21 (Unconventional Monetary Policy / QE) rises sharply after the Great Recession,



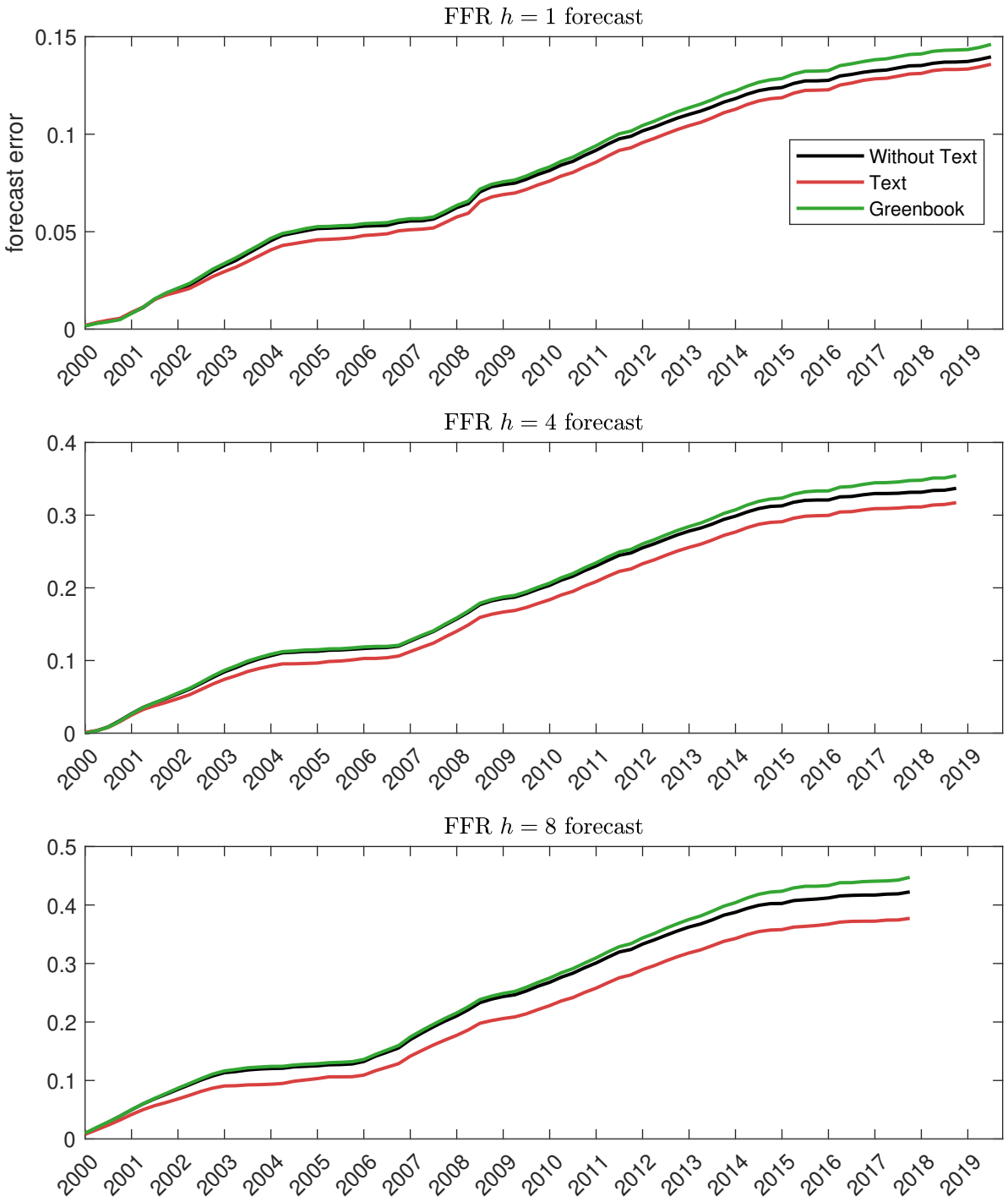
Note: This figure shows the root mean squared error (RMSE) across different horizons for each variable in the observation equation. The left axis shows the relative RMSE. The values smaller than 1 imply a better forecasting performance of the model than the model without texts. The right axis shows the RMSE of the model without additional data. The red line represents the model with FOMC transcripts, and the green line represents the model with the Greenbook data. The black line shows the level of RMSE of the model without text on the right axis.

Figure 2: (Relative) Root mean squared error



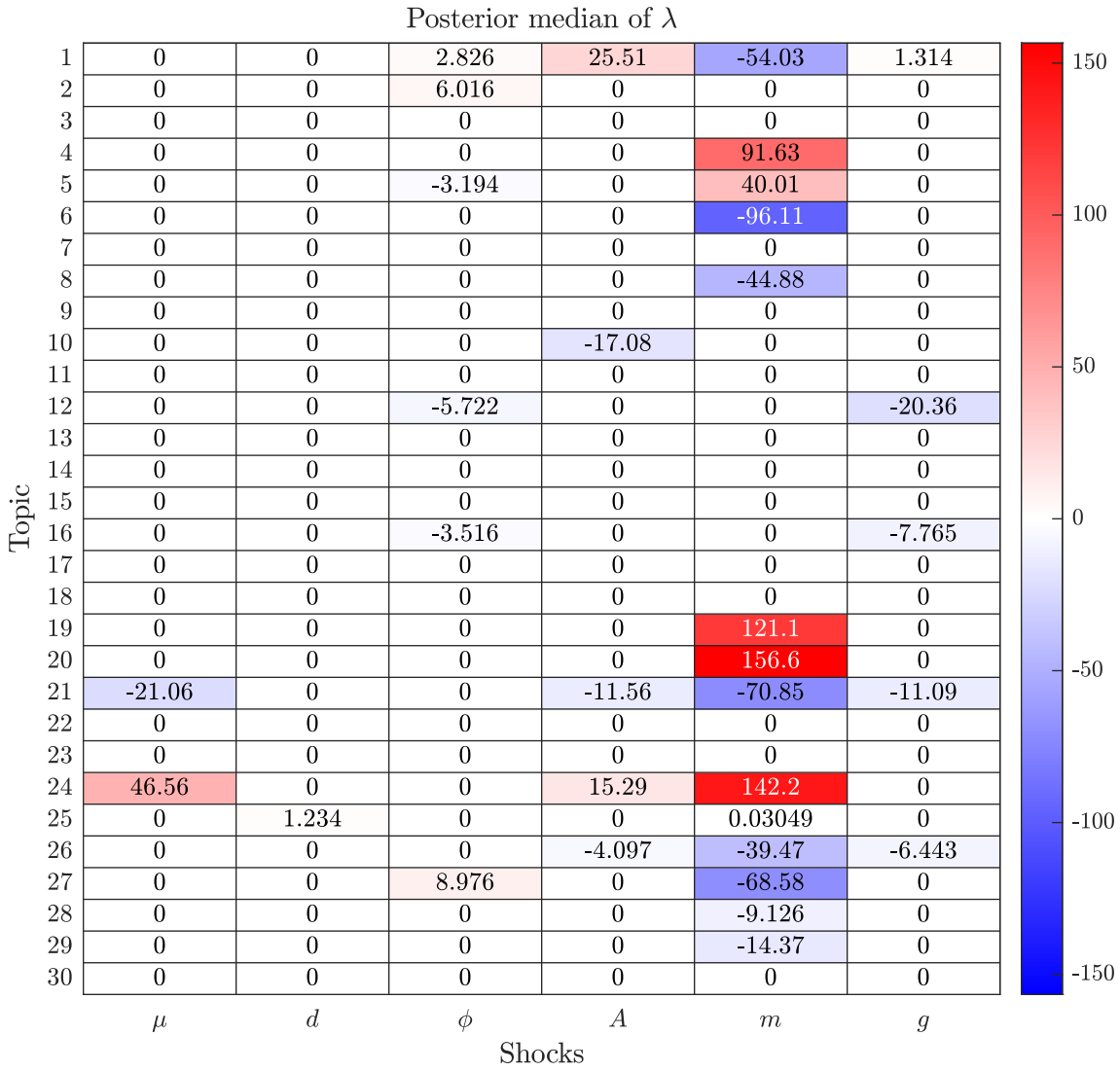
Note: This figure shows the continuous ranked probability score across different horizons for each variable in the observation equation. The left axis shows the relative RMSE. The values smaller than 1 imply a better forecasting performance of the model than the model without texts. The right axis shows the RMSE of the model without additional data. The right axis shows the RMSE of the model without additional data. The red line represents the model with FOMC transcripts, and the green line represents the model with the Greenbook data. The black line shows the level of RMSE of the model without text on the right axis.

Figure 3: (Relative) continuous ranked probability score



Note: This figure plots the cumulative absolute error of federal funds rate forecasts from models without text, with FOMC transcripts, and with Greenbook data. Forecasts are based on the mean of the predictive distribution, with forecasting horizons $h = 1, 4, 8$.

Figure 4: Cumulative absolute forecast error of the federal funds rate



Topic	μ	d	ϕ	A	m	g
1	0	0	2.826	25.51	-54.03	1.314
2	0	0	6.016	0	0	0
3	0	0	0	0	0	0
4	0	0	0	0	91.63	0
5	0	0	-3.194	0	40.01	0
6	0	0	0	0	-96.11	0
7	0	0	0	0	0	0
8	0	0	0	0	-44.88	0
9	0	0	0	0	0	0
10	0	0	0	-17.08	0	0
11	0	0	0	0	0	0
12	0	0	-5.722	0	0	-20.36
13	0	0	0	0	0	0
14	0	0	0	0	0	0
15	0	0	0	0	0	0
16	0	0	-3.516	0	0	-7.765
17	0	0	0	0	0	0
18	0	0	0	0	0	0
19	0	0	0	0	121.1	0
20	0	0	0	0	156.6	0
21	-21.06	0	0	-11.56	-70.85	-11.09
22	0	0	0	0	0	0
23	0	0	0	0	0	0
24	46.56	0	0	15.29	142.2	0
25	0	1.234	0	0	0.03049	0
26	0	0	0	-4.097	-39.47	-6.443
27	0	0	8.976	0	-68.58	0
28	0	0	0	0	-9.126	0
29	0	0	0	0	-14.37	0
30	0	0	0	0	0	0

Note: This figure shows the posterior distribution of the factor loadings of each topic share series on the exogenous process in the model. The Y-axis represents the topic share series of the FOMC transcripts, and the X-axis represents the exogenous process in our DSGE model. μ: investment specific technology shock, d: intertemporal preference shock, ϕ: labor supply shock, A: aggregate technology shock, m: monetary policy shock, g: government expenditure shock. For the labels of the topics, see Table 1.

Figure 5: Posterior median of the loadings λ

when policy was explicitly accommodative, which aligns with a negative loading on m.

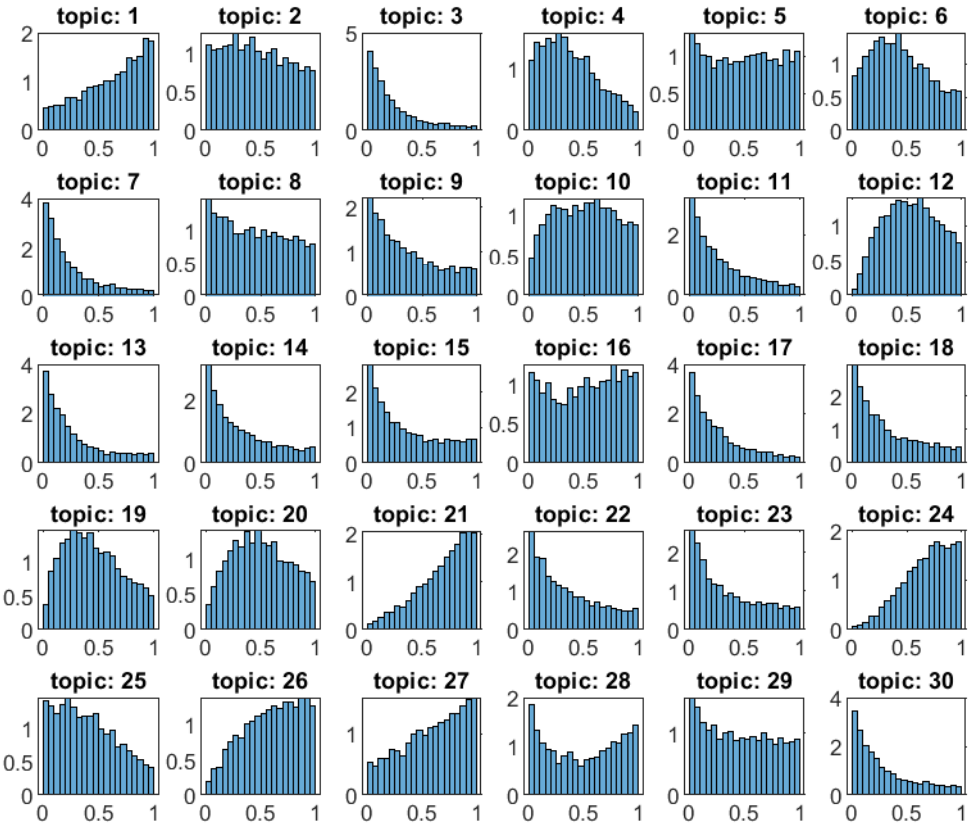
Second, some topics are loaded on several shocks simultaneously. For example, technology shocks A and μ also load on Topic 24 (Inflation and cost pressures). This does not mean that technology shocks mechanically raise inflation on impact: rather, Topic 24 captures the salience of price-cost discussions. Supply-side developments, through productivity or investment-specific technology, affect marginal costs and relative prices and thereby trigger

Committee discussion of inflationary or disinflationary pressures. This mechanism explains why Topic 24 is associated with both technology and monetary policy shocks in the estimated loadings.

Third, incorporating topic series into structural estimation also strengthens the interpretation of the topics themselves. Because topic modeling is an unsupervised technique, assigning economic labels is typically done heuristically based on vocabulary and time-series patterns. Our framework adds a structural layer: topics can be interpreted through their loadings on structural shocks. Several topics are broadly related to monetary policy in the LDA output, but the estimated Λ helps distinguish their specific roles. Topic 4, containing terms such as "tighten," "inflation," and "demand" loads positively on m confirming that it reflects tightening deliberations. Topic 21, featuring "purchase," "program," and "balance sheet," is strongly negative, capturing episodes of unconventional policy such as quantitative easing. Topic 26, with terms like "zero," "future," and "credibility," is linked to the zero lower bound and forward-guidance discussions in the post-2008 period. Together, these patterns show that joint estimation with the DSGE model delivers a clear economic taxonomy: tightening, easing, and communication topics are separated based on their structural-shock loadings, rather than solely on vocabulary.

Fourth, the results indicate that FOMC members routinely discuss issues that lie outside the scope of our benchmark DSGE model. For example, Topic 22 (Trade) and Topic 17 (Risk and Uncertainty) capture themes that are economically meaningful but exogenous to the current closed-economy, linearized New Keynesian specification. Because these dimensions are not mapped to any structural shock, the corresponding factor loadings are estimated to be essentially zero with high posterior probability. This is a desirable feature: the sparsity mechanism correctly identifies topics that are relevant for the structural model at hand and excludes those that fall outside its theoretical scope, providing a safeguard against overfitting.

The posterior distributions of the inclusion probabilities q provide complementary evidence on topic selection. Recall that q is the probability that an element of Λ is drawn from the "slab" rather than being fixed at zero. When $q = 1$, the corresponding topic is always included; when q is close to zero, the topic is effectively excluded. Figure 6 shows that the posterior distributions of q are heterogeneous across topics. Topics that are not naturally associated with our structural shocks, such as Topic 17 (Risk and Uncertainty) and Topic 22 (Trade), have posterior mass concentrated near zero. By contrast, Topic 1 (Labor), Topic 21 (Unconventional MP / QE), Topic 24 (Inflation / Cost pressures), Topic 26 (ZLB / Forward guidance), and Topic 27 (SEP /guidance language) have posterior inclusion probabilities concentrated near one, confirming that the estimation treats these topics as systematically



Note: These figures present the posterior distribution of the probability q of inclusion of each topic share series. For the labels of the topics, see Table 1.

Figure 6: Heterogeneity in the estimated probability of inclusion

informative.

Lastly, while the estimation procedure performs well in identifying meaningful associations between topics and monetary policy shocks, some linkages with other shocks are less straightforward to interpret. For example, the government spending shock ε_g loads negatively on Topic 12 (Fiscal Policy), implying that a positive shock to g_t is associated with less discussion of fiscal-policy issues. In a small- or medium-scale New Keynesian model such as ours, however, g_t is not a literal government expenditure shock but rather a residual demand shock that absorbs movements in output not explained by other factors featured in our model. An increase in g_t may therefore reflect shifts in residual demand unrelated to explicit fiscal actions, which could coincide with less emphasis on fiscal policy in the transcripts.

This example illustrates a broader limitation: some topic-shock relationships depend on how shocks are interpreted within a simplified structural framework. At the same time, such discrepancies are informative. They can serve as diagnostics for model misspecification or shock mislabeling and point towards the need for a richer model or alternative identification scheme.

Taken together, the sparse non-core measurement equation successfully selects topics from the FOMC transcripts that are informative about structural shocks and downweights topics that are irrelevant for the DSGE model. This selection improves the economic interpretation of the LDA topics, clarifies how different aspects of FOMC discussions relate to underlying shocks, and is consistent with the forecasting gains documented in the previous subsection.

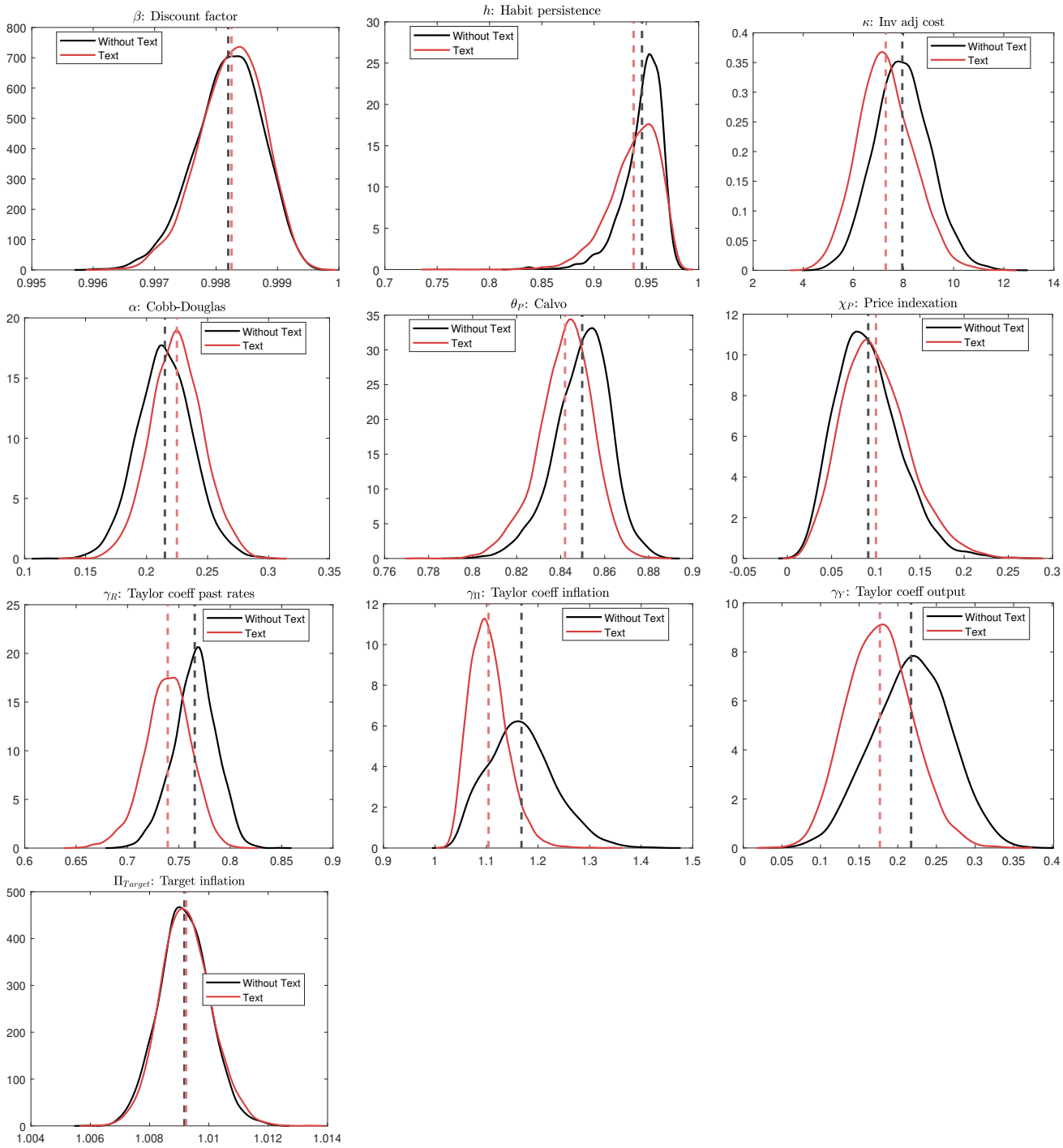
3.4.3 How does the posterior distribution of structural parameters change with FOMC transcripts?

We now turn to the structural parameter estimates and impulse response functions. The previous subsection showed that FOMC transcripts improve forecasting performance and that our framework links specific topics to structural shocks in the DSGE model. Here we ask how incorporating text changes the posterior distribution of the structural parameters, in turn, the implied dynamics of the model.[9]

Figures 7 and 8 plot the posterior distributions of the DSGE parameters for three specifications: the benchmark model without non-core variables (black), the text-augmented model using FOMC transcripts (red). These figures show that FOMC transcripts affect the posterior distribution of structural parameters in several ways. As expected, the additional information from the transcripts reduces posterior uncertainty for a subset of parameters, tightening their credible intervals relative to the benchmark.

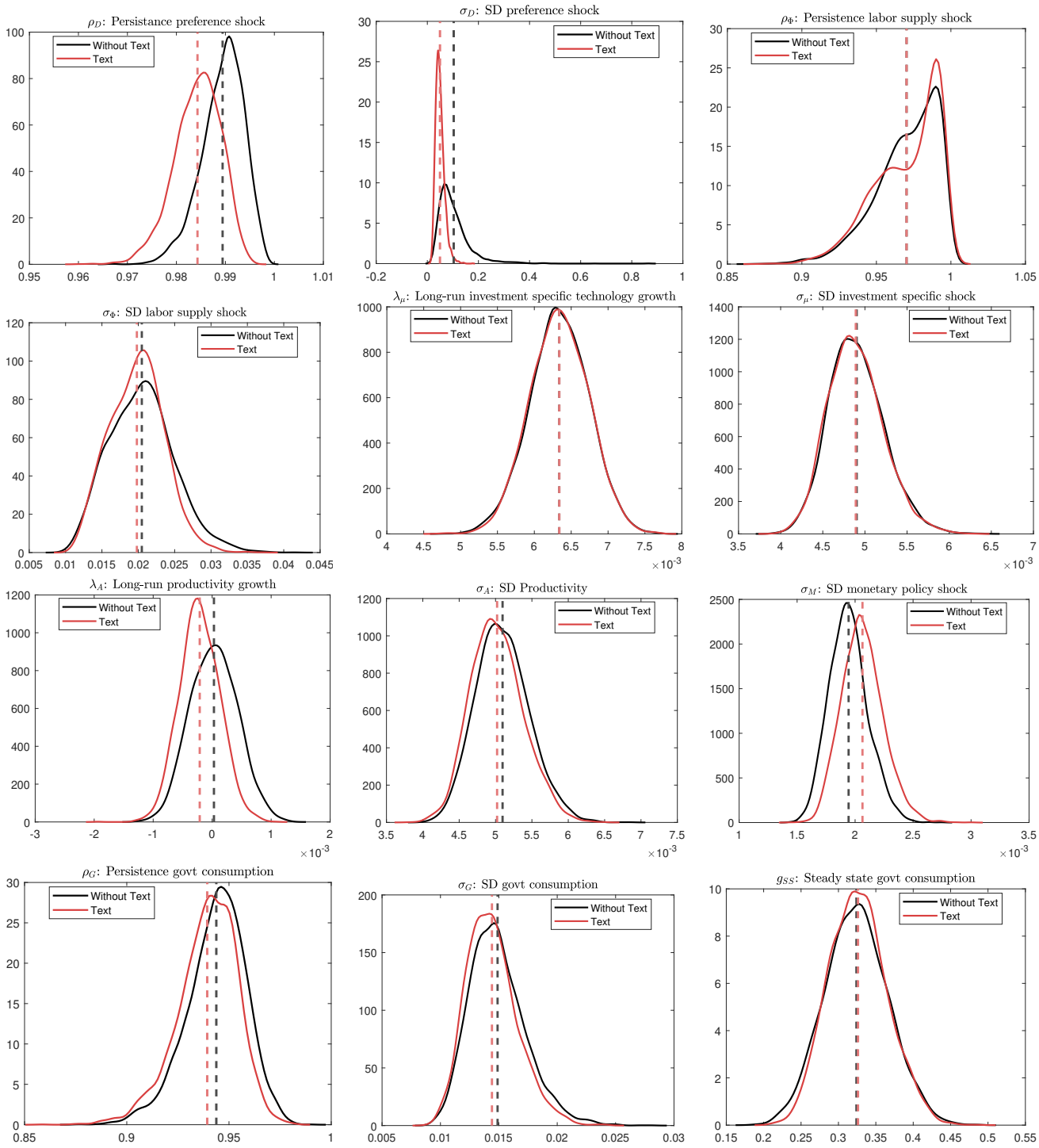
More interestingly, the posterior means of some parameters shift as well. In particular, the FOMC transcript data has a pronounced effect on the monetary-policy parameters in the Taylor rule. Including the transcripts lowers the estimated coefficients on inflation (γ_Π) and output (γ_y), suggesting that the benchmark estimation without text overstates the strength of the Fed's systematic response to these variables. The loading estimates in Λ, discussed in the previous section, help explain this pattern: the monetary policy shock absorbs variation associated with communication- and forward-looking-related topics that the standard Taylor rule cannot capture. When text is omitted, the estimation attributes this variation to the observed feedback rule, mechanically inflating the estimated γ_Π and γ_y. Once the transcripts

[9]The results for the model augmented with Greenbook factors can be found in Supplementary Appendix D.



Note: These figures show the posterior distributions of the structural parameters. The dashed vertical line is the posterior mean. The black line represents the model without text, and the red line represents the model with FOMC transcripts.

Figure 7: Posterior distribution of parameters in our DSGE model 1



Note: These figures show the posterior distributions of the structural parameters. The dashed vertical line is the posterior mean. The black line represents the model without text, and the red line represents the model with FOMC transcripts.

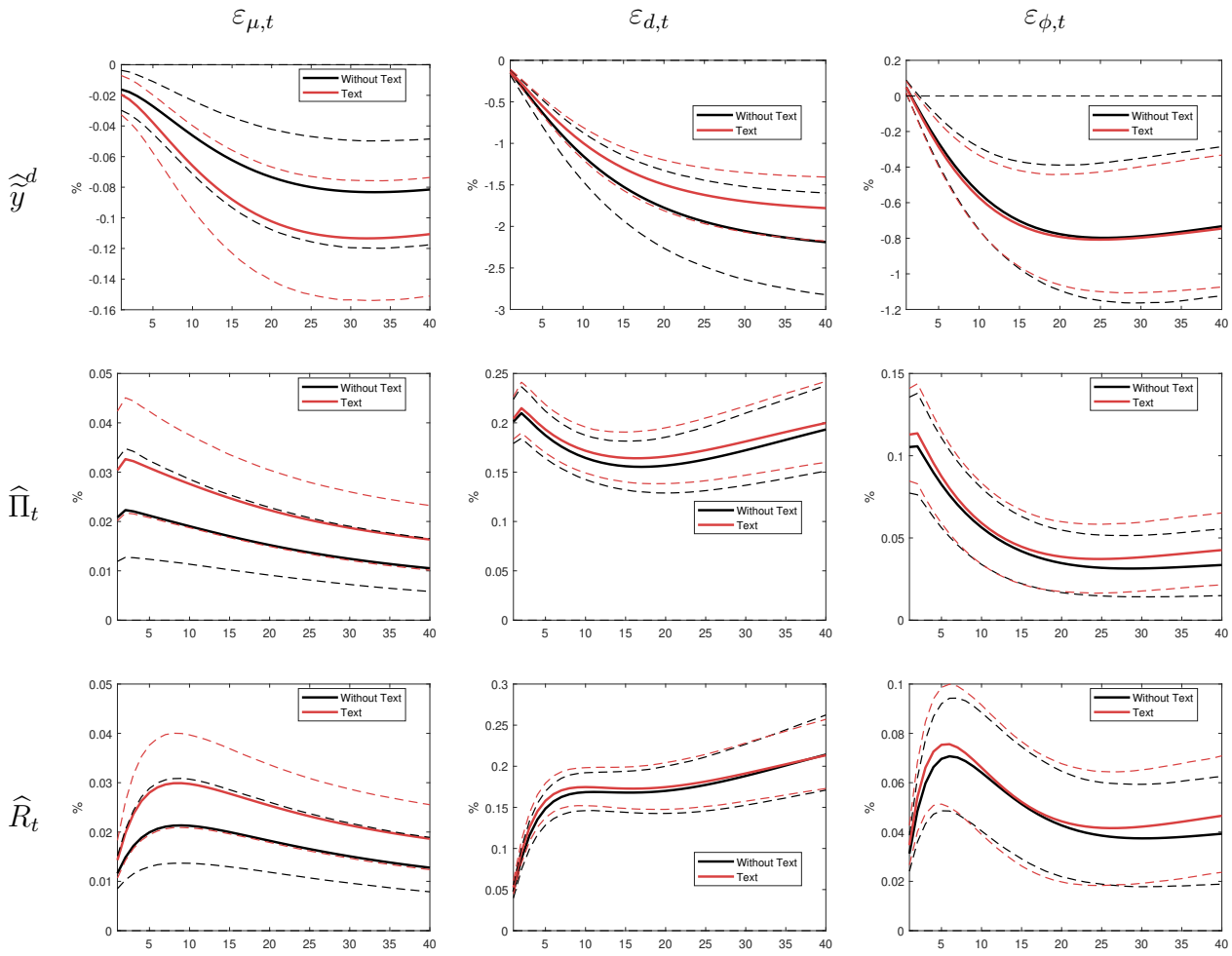
Figure 8: Posterior distribution of parameters in our DSGE model 2

are included, part of this variation is reallocated to the monetary policy shock, and the systematic coefficients adjust downward.

This reallocation has implications for the nominal rigidity parameters as well. When the transcripts are included, the estimated Taylor-rule response to inflation (γ_Π) falls. Holding other parameters fixed, a weaker systematic response of the interest rate to inflation would make inflation more persistent than what we observe in the data. To offset this, the estimation lowers the Calvo parameter θ_P, steepening the Phillips curve so that inflation becomes more responsive on impact but less structurally persistent, thereby bringing the model-implied inflation dynamics back in line with the data.

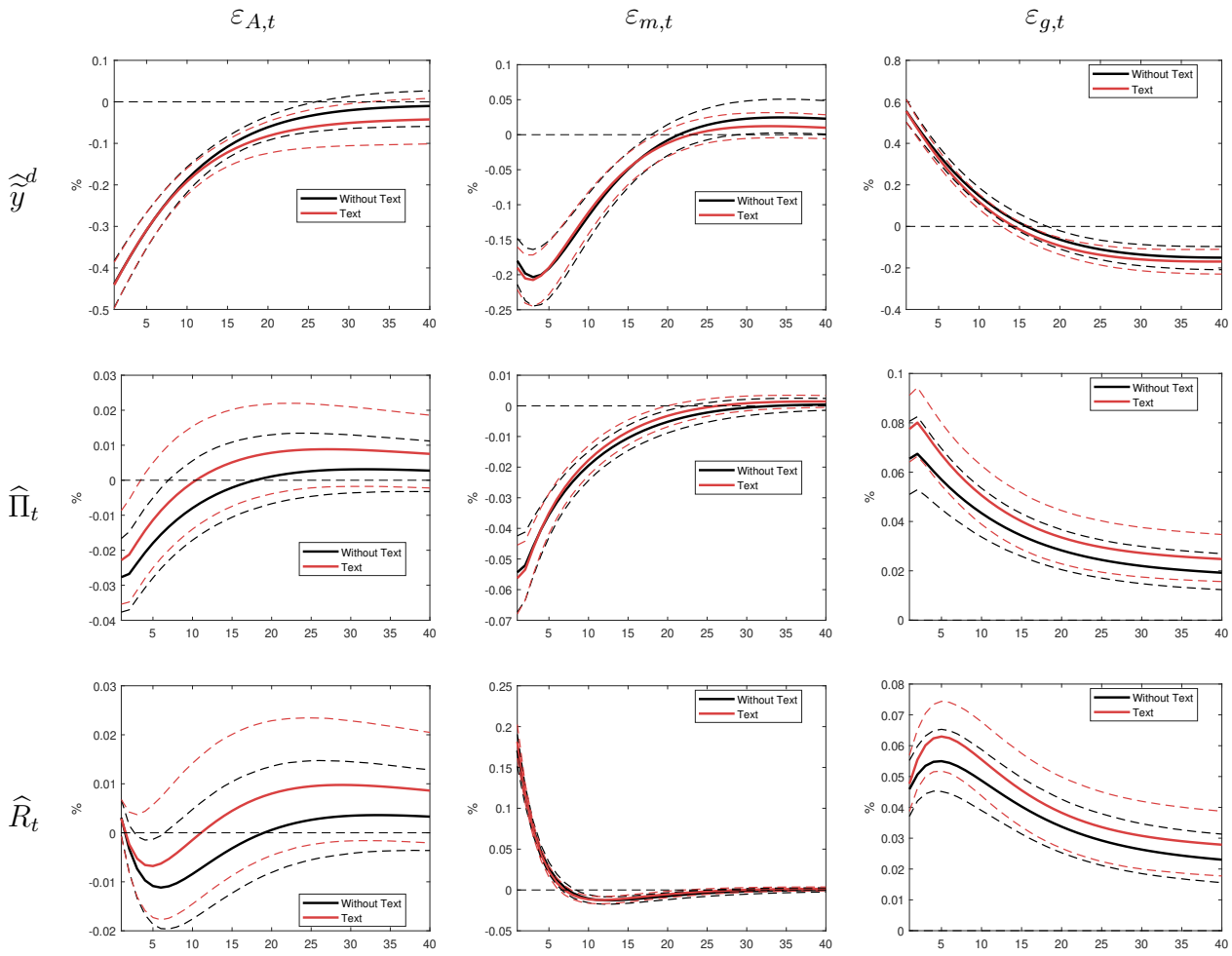
Figures 9 and 10 plot the impulse responses of aggregate demand, inflation, and the nominal interest rate to the six structural shocks in our economy. Overall, the signs and broad shapes of the responses are similar across specifications, but there are somes differences between the benchmark model and the text-augmented model.

The most pronounced divergence appears in the responses to the investment-specific technology shock. In both models, a positive investment-specific shock reallocates spending from consumption to investment, raises marginal costs, and increases prices while reducing aggregate demand. However, with FOMC transcripts included, the responses of inflation and the interest rate to this shock are noticeably stronger. This is consistent with the lower estimated Calvo parameter θ_P in the text-augmented specification: with smaller θ_P, firms adjust prices more aggressively in response to marginal-cost movements, amplifying the inflation response and, through the Taylor rule, the interest rate response. At the same time, the lower Taylor-rule coefficients on inflation and output tend to dampen the short-run response of inflation and the interest rate to technology shocks. For the remaining shocks, the impulse responses of key variables are broadly similar across the benchmark and text-augmented specifications.



Note: The figures present the impulse responses of our estimated DSGE model to a one standard deviation shock. The black line represents the model without text, and the red line represents the model with FOMC transcripts. The solid line is the posterior mean, and the dashed/dotted line is the pointwise 90 % credible interval.

Figure 9: Impulse responses of aggregate demand, inflation, and nominal interest rate to investment-specific technology, intertemporal preference, and labor supply shocks



Note: The figures present the impulse responses of our estimated DSGE model to a one standard deviation shock. The black line represents the model without text, and the red line represents the model with FOMC transcripts. The solid line is the posterior mean, and the dashed/dotted line is the pointwise 90 % credible interval.

Figure 10: Impulse responses of aggregate demand, inflation, and nominal interest rate to technology, monetary policy, and government spending shocks

4 Conclusion

Standard macroeconomic data record the joint outcome of the rules agents follow and the non-systematic departures they make from those rules. This paper shows that unstructured data can help separate the two.

We develop a non-core measurement block that links text-derived time series to the structural shocks of a DSGE model, using a spike-and-slab prior to let the data determine which series are informative. Applied to a medium-scale New Keynesian model augmented with FOMC transcripts, the framework improves predictive performance and selects a small number of policy-relevant topics as informative about the monetary policy shock. It also shifts the estimated Taylor-rule coefficients and the Calvo parameter. The estimated Phillips curve is steeper than specifications based on standard macroeconomic data alone. Without text, the econometrician cannot distinguish the systematic component of monetary policy from the non-systematic one.

The framework requires no a priori mapping from text to model objects. It extends to any structural model in state-space form and to any source of unstructured data. Promising applications include real-time implementations using publicly available policy communications such as minutes, speeches, or press conferences; alternative representations of text beyond LDA; and richer structural environments that incorporate channels the unstructured data themselves suggest.

References

Aruoba, S. B. and Drechsel, T. (2025). Identifying Monetary Policy Shocks: A Natural Language Approach. *American Economic Journal: Macroeconomics*, forthcoming.

Ash, E. and Hansen, S. (2023). Text algorithms in economics. *Annual Review of Economics*, 15(1):659–688.

Baker, S. R., Bloom, N., and Davis, S. J. (2016). Measuring Economic Policy Uncertainty. *The Quarterly Journal of Economics*, 131(4):1593–1636.

Battaglia, L., Christensen, T., Hansen, S., and Sacher, S. (2025). Inference for regression with variables generated by AI or machine learning. arXiv:2402.15585v5 (posted Apr 30, 2025; paper dated Apr 29, 2025).

Bils, M. and Klenow, P. J. (2004). Some evidence on the importance of sticky prices. *Journal of Political Economy*, 112(5):947–985.

Boivin, J. and Giannoni, M. P. (2006). Dsge models in a data-rich environment.

Bryzgalova, S., Huang, J., and Julliard, C. (2023). Bayesian solutions for the factor zoo: We just ran two quadrillion models. *The Journal of Finance*, 78(1):487–557.

Bybee, L., Kelly, B. T., Manela, A., and Xiu, D. (2024). Business news and business cycles. *Journal of Finance*, 79(5):3105–3147.

Caldara, D. and Iacoviello, M. (2022). Measuring Geopolitical Risk. *American Economic Review*, 112(4):1194–1225.

Chib, S., Shin, M., and Tan, F. (2023). DSGE-SVt: An econometric toolkit for high-dimensional dsge models with sv and t errors. *Computational Economics*, 61(1):69–111.

Cieslak, A., Hansen, S., McMahon, M., and Xiao, S. (2023). Policymakers' uncertainty. Working Paper 31849, National Bureau of Economic Research.

Clarida, R., Galí, J., and Gertler, M. (2000). Monetary policy rules and macroeconomic stability: Evidence and some theory. *Quarterly Journal of Economics*, 115(1):147–180.

Compiani, G., Morozov, I., and Seiler, S. (2025). Demand Estimation with Text and Image Data. *arXiv:2503.20711 [econ.GN]*.

Del Negro, M. and Schorfheide, F. (2013). Dsge model-based forecasting. In *Handbook of economic forecasting*, volume 2, pages 57–140. Elsevier.

Deming, D. J. and Kahn, L. B. (2018). Skill requirements across firms and labor markets: Evidence from job postings for professionals. *Journal of Labor Economics*, 36(S1):S337–S369.

Diebold, F. X., Schorfheide, F., and Shin, M. (2017). Real-time forecast evaluation of dsge

models with stochastic volatility. *Journal of Econometrics*, 201(2):322–332.

Durbin, J. and Koopman, S. J. (2002). A simple and efficient simulation smoother for state space time series analysis. *Biometrika*, 89(3):603–616.

Fernández-Villaverde, J. and Guerrón-Quintana, P. A. (2021). Estimating dsge models: Recent advances and future challenges. *Annual Review of Economics*, 13(1):229–252.

Fernández-Villaverde, J. and Rubio-Ramírez, J. F. (2006). A benchmark dsge model.

Fernández-Villaverde, J., Rubio-Ramírez, J., and Schorfheide, F. (2016). Chapter 9 - solution and estimation methods for dsge models. volume 2 of *Handbook of Macroeconomics*, pages 527–724. Elsevier.

Gelfer, S. (2019). Data-rich dsge model forecasts of the great recession and its recovery. *Review of Economic Dynamics*, 32:18–41.

Gelfer, S. (2021). Evaluating the forecasting power of an open-economy dsge model when estimated in a data-rich environment. *Journal of Economic Dynamics and Control*, 129:104177.

Gentzkow, M., Kelly, B., and Taddy, M. (2019). Text as data. *Journal of Economic Literature*, 57(3):535–74.

Giannone, D., Lenza, M., and Primiceri, G. E. (2021). Economic predictions with big data: The illusion of sparsity. *Econometrica*, 89(5):2409–2437.

Griffiths, T. L. and Steyvers, M. (2004). Finding scientific topics. *Proceedings of the National Academy of Sciences*, 101(suppl 1):5228–5235.

Gürkaynak, R. S., Sack, B., and Swanson, E. T. (2005). Do actions speak louder than words? the response of asset prices to monetary policy actions and statements. *International Journal of Central Banking*, 1(1):55–93.

Hansen, S., McMahon, M., and Prat, A. (2018). Transparency and deliberation within the FOMC: A computational linguistics approach. *The Quarterly Journal of Economics*, 133(2):801–870.

Henderson, J. V., Storeygard, A., and Weil, D. N. (2012). Measuring economic growth from outer space. *American Economic Review*, 102(2):994–1028.

Hoberg, G. and Phillips, G. (2016). Text-Based Network Industries and Endogenous Product Differentiation. *Journal of Political Economy*, 124(5):1423–1465.

Hurwicz, L. (1962). On the structural form of interdependent systems. In Nagel, E., Suppes, P., and Tarski, A., editors, *Logic, Methodology and Philosophy of Science: Proceedings of the 1960 International Congress*, pages 232–239. Stanford University Press.

Ishwaran, H. and Rao, J. S. (2005). Spike-and-slab variable selection: Frequentist and

bayesian strategies. *Annals of Statistics*, 33:730–773.

Jarociński, M. (2015). A note on implementing the durbin and koopman simulation smoother. *Computational Statistics & Data Analysis*, 91:1–3.

Kaufmann, S. and Schumacher, C. (2017). Identifying relevant and irrelevant variables in sparse factor models. *Journal of Applied Econometrics*, 32(6):1123–1144.

Kaufmann, S. and Schumacher, C. (2019). Bayesian estimation of sparse dynamic factor models with order-independent and ex-post mode identification. *Journal of Econometrics*, 210(1):116–134.

Kryshko, M. M. (2011). *Data-rich DSGE and dynamic factor models*. International Monetary Fund.

Lubik, T. A. and Schorfheide, F. (2004). Testing for indeterminacy: An application to U.S. monetary policy. *American Economic Review*, 94(1):190–217.

Ludwig, J., Mullainathan, S., and Rambachan, A. (2026). Large language models: An applied econometric framework. *Annual Review of Economics*. Forthcoming.

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.

McCracken, M. W. and Ng, S. S. (2021). Fred-qd: A quarterly database for macroeconomic research. *Federal Reserve Bank of St. Louis Review*, pages 1–44.

Nakamura, E. and Steinsson, J. (2008). Five facts about prices: A reevaluation of menu cost models. *Quarterly Journal of Economics*, 123(4):1415–1464.

Pellegrino, B. (2025). Product Differentiation and Oligopoly: A Network Approach. *American Economic Review*, 115(4):1170–1225.

Romer, C. D. and Romer, D. H. (2004). A new measure of monetary shocks: Derivation and implications. *American Economic Review*, 94(4):1055–1084.

Schorfheide, F., Sill, K., and Kryshko, M. (2010). Dsge model-based forecasting of non-modelled variables. *International Journal of Forecasting*, 26:348–373.

Shapiro, A. H., Sudhof, M., and Wilson, D. J. (2022). Measuring news sentiment. *Journal of Econometrics*, 228(2):221–243.

Appendix

A State space representation

This section explains the state-space representation (7) of our empirical model, given the log-linearized equilibrium conditions of our NK-DSGE model. First, we show how we construct our state-transition equation in detail, focusing on the state vector S_t that tracks all the model variables needed to set up our measurement equation. Second, we explain how we connect the observed macroeconomic data to the state variables of our model, which provides the core-measurement equation. Finally, we discuss how we connect unstructured data to our model, which provides the non-core measurement equation.

A.1 State transition equation

Recall that the solution of a DSGE model provides a state transition equation:

$$S_t = T(\psi)S_{t-1} + R(\psi)\varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, I) \tag{9}$$

where S_t is a vector of state variables one needs to keep track of, $T(\psi)$ and $R(\psi)$ are matrices that govern the transition of the state variables, and each of their elements is a function of the structural parameters ψ of the DSGE model.

More specifically, in our model, the solution of our DSGE model can be represented as follows:[10]

$$state_t = PP * state_{t-1} + QQ * exo_t$$

$$nstate_t = RR * state_{t-1} + SS * exo_t$$

$$exo_{t+1} = NN * exo_t + \Sigma_{\varepsilon}^{1/2} * \varepsilon_{t+1}$$

where

$$state_t = \left(\widehat{\Pi}_t, \widehat{g}_t^1, \widehat{g}_t^2, \widehat{\widetilde{k}}_t, \widehat{R}_t, \widehat{\widetilde{y}}_t^d, \widehat{\widetilde{c}}_t, \widehat{v}_t^p, \widehat{\widetilde{q}}_t, \widehat{\widetilde{x}}_t\right)',$$

$$nstate_t = \left(\widehat{\widetilde{\lambda}}_t, \widehat{\widetilde{z}}_t, \widehat{\widetilde{r}}_t, \widehat{u}_t, \widehat{\Pi}_t^*, \widehat{mc}_t, \widehat{l}_t, \widehat{\widetilde{w}}_t\right)',$$

[10] We follow Uhlig's notation, where $state_t$ is the endogenous state variable x_t, $nstate_t$ is the control (jump) variable y_t, and exo_t is the exogenous state variable z_t in their notation. The Supplemental Appendix A provides Uhlig's form.

$$exo_t = \left(\widehat{\widetilde{\mu}}_t, \widehat{d}_t, \widehat{\phi}_t, \widehat{\widetilde{A}}_t, m_t, \widehat{\widetilde{g}}_t\right)',$$

$\Sigma_\varepsilon = diag(\sigma_\mu^2, \sigma_D^2, \sigma_\phi^2, \sigma_A^2, \sigma_M^2, \sigma_G^2)$, and $PP,\ QQ,\ RR,\ SS,\ NN$ are defined by the structural parameters of the DSGE model.

Let us define a state vector as

$$S_t = [state_t', exo_t', \widehat{\widetilde{w}}_t, \widehat{\widetilde{w}}_{t-1}, \widehat{\widetilde{y}}^d_{t-1}, \widehat{\widetilde{z}}_t, \widehat{l}_t, \varepsilon_{D,t}, \varepsilon_{\phi,t}, \varepsilon_{G,t}]'. \tag{10}$$

Here, instead of tracking all the $nstate_t$ variables, we just keep track of $\widehat{\widetilde{w}}_t, \widehat{\widetilde{z}}_t$ and $\widetilde{l}_t$, as these are the only variables in $nstate_t$ needed in the core measurement equation. This is not necessary, but a smaller state vector helps to reduce computational costs. We also add $\widehat{\widetilde{w}}_{t-1}$ and $\widehat{\widetilde{y}}^d_{t-1}$ as they are needed in the core measurement equation.

Then, the state transition equation is given by setting the matrices $T(\psi)$ and $R(\psi)$ as

$$T(\psi) = \begin{bmatrix} PP & QQ * NN & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{n_e \times n_s} & NN & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ RR_{(8,\cdot)} & SS_{(8,\cdot)} * NN & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1 \times n_s} & \mathbf{0}_{1 \times n_e} & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ v_y & \mathbf{0}_{1 \times n_e} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ RR_{(2,\cdot)} & SS_{(2,\cdot)} * NN & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ RR_{(7,\cdot)} & SS_{(7,\cdot)} * NN & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1 \times n_s} & \mathbf{0}_{1 \times n_e} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1 \times n_s} & \mathbf{0}_{1 \times n_e} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1 \times n_s} & \mathbf{0}_{1 \times n_e} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$$R(\theta) = \begin{bmatrix} QQ * \Sigma_\varepsilon^{1/2} \\ \Sigma_\varepsilon^{1/2} \\ SS_{(8,:)} * \Sigma_\varepsilon^{1/2} \\ 0 \\ 0 \\ SS_{(2,:)} * \Sigma_\varepsilon^{1/2} \\ SS_{(7,:)} * \Sigma_\varepsilon^{1/2} \\ \Sigma_\varepsilon^{1/2}([d, \phi, g]') \end{bmatrix},$$

where n_s and n_e are the length of $state_t$ and exo_t, respectively, v_y is a $1 \times n_s$ vector with zeros except the 6th element being 1 (so that it selects $\widehat{\widetilde{y}}^d_t$ from S_t), $\mathbf{0}$ is a vector or matrix

of zeros, and $\Sigma([d,\phi,g])$ is the rows of covariance matrix corresponding to $\varepsilon_D, \varepsilon_\phi$ and ε_G.

Putting all together, the state transition equation is:

$$\begin{bmatrix} state_t \\ exo_t \\ \widehat{\widetilde{w}}_t \\ \widehat{\widetilde{w}}_{t-1} \\ \widehat{\widetilde{y}}^d_{t-1} \\ \widehat{\widetilde{z}}_t \\ \widehat{l}_t \\ \varepsilon_{D,t} \\ \varepsilon_{\phi,t} \\ \varepsilon_{G,t} \end{bmatrix} = \begin{bmatrix} PP & QQ*NN & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{n_e\times n_s} & NN & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ RR_{(8,\cdot)} & SS_{(8,\cdot)}*NN & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1\times n_s} & \mathbf{0}_{1\times n_e} & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ v_y & \mathbf{0}_{1\times n_e} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ RR_{(2,\cdot)} & SS_{(2,\cdot)}*NN & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ RR_{(7,\cdot)} & SS_{(7,\cdot)}*NN & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1\times n_s} & \mathbf{0}_{1\times n_e} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1\times n_s} & \mathbf{0}_{1\times n_e} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1\times n_s} & \mathbf{0}_{1\times n_e} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} state_{t-1} \\ exo_{t-1} \\ \widehat{\widetilde{w}}_{t-1} \\ \widehat{\widetilde{w}}_{t-2} \\ \widehat{\widetilde{y}}^d_{t-2} \\ \widehat{\widetilde{z}}_{t-1} \\ \widehat{l}_{t-1} \\ \varepsilon_{D,t-1} \\ \varepsilon_{\phi,t-1} \\ \varepsilon_{G,t-1} \end{bmatrix} + \begin{bmatrix} QQ*\Sigma_\varepsilon^{1/2} \\ \Sigma_\varepsilon^{1/2} \\ SS_{(8,:)}*\Sigma_\varepsilon^{1/2} \\ 0 \\ 0 \\ SS_{(2,:)}*\Sigma_\varepsilon^{1/2} \\ SS_{(7,:)}*\Sigma_\varepsilon^{1/2} \\ \Sigma_\varepsilon^{1/2}([d,\phi,g]) \end{bmatrix} \varepsilon_t.$$

A.2 Measurement equation

The measurement equation has two blocks: core and non-core measurement equations. The core measurement equation connects some macroeconomic variables to the state variables or their linear combination, which is a standard approach in the literature of DSGE model estimation. The non-core measurement equation connects unstructured data to our model.

In the following subsection, we first give a detailed formalization of the core measurement equation in our model. Then, we formulate the dynamics of the non-core data series, which provides the non-core measurement equation. Finally, we combine the core and non-core measurement equations to obtain the full state-space representation.

A.2.1 Core measurement equation

Our core observed variables are inflation, the federal funds rate, real wage growth, output growth, hours worked, and the relative price of investment with respect to the price of consumption growth, or in our notation:

$$Y_t^C = \left(\log \Pi_t, \log R_t, \triangle \log w_t, \triangle \log y_t, \log l_t, \triangle \log \mu_t^{-1}\right)'.$$

We will connect these observed series to the model as follows:

$$\log\left(\frac{P_t^{obs}}{P_{t-1}^{obs}}\right) = \log\left(\Pi_{ss}\right) + S_t(1)$$

$$\log\left(INT_t^{obs}+1\right) = \log R_{ss} + S_t(5)$$

$$\triangle \log w_t^{obs} = \Lambda_z + S_t(n_s+n_e+4) + S_t(n_s+n_e+1) - S_t(n_s+n_e+2)$$

$$\triangle \log y_t^{obs} = \Lambda_z + S_t(n_s+n_e+4) + S_t(6) - S_t(n_s+n_e+3)$$

$$\log l_t^{obs} - mean(\log l_{1:T}^{obs}) = 0 + S_t(n_s+n_e+5)$$

$$\underbrace{\triangle \log(\mu_t^{-1})^{obs}}_{\text{Relative price of capital}} = -(\log\mu_t - \log\mu_{t-1}) = -(\Lambda_\mu + z_{\mu,t}) = -\Lambda_\mu - S_t(n_s+1)$$

where $\Lambda_z = \frac{\alpha\Lambda_\mu+\Lambda_A}{(1-\alpha)}$ and $S_t(i)$ is the i-th element of S_t vector. To derive the measurement equation for the real wage growth, note that $\log(w_t)-\log(w_{t-1}) = \log(\widetilde{w}_t z_t)-\log(\widetilde{w}_{t-1}z_{t-1}) = \log(z_t)-\log(z_{t-1})+\widehat{\widetilde{w}}_t-\widehat{\widetilde{w}}_{t-1}$ and $\log(z_t)-\log(z_{t-1}) = \Lambda_z+\widehat{\widetilde{z}}_t$. Similarly, for the measurement equation of output growth, note that $\log(y_t^d)-\log(y_{t-1}^d) = \log(\widetilde{y}_t z_t)-\log(\widetilde{y}_{t-1}^d z_{t-1}) = \log(z_t)-\log(z_{t-1}) + \widehat{\widetilde{y}}_t^d - \widehat{\widetilde{y}}_{t-1}^d = \Lambda_z + \widehat{\widetilde{z}}_t + \widehat{\widetilde{y}}_t^d - \widehat{\widetilde{y}}_{t-1}^d$.

Hence, we can formulate the core measurement equation as

$$Y_t^C = H_0^C + H_1^C S_t + \Sigma_u^{1/2} u_t \tag{11}$$

where

$$H_0^C = \begin{bmatrix} \log\Pi^{ss} \\ \log R^{ss} \\ \frac{\alpha\Lambda_\mu+\Lambda_A}{1-\alpha} \\ \frac{\alpha\Lambda_\mu+\Lambda_A}{1-\alpha} \\ 0 \\ -\Lambda_\mu \end{bmatrix},$$

H_1^C is $(6\times(n_s+n_e+8))$ matrices where $H_1^C(1,1)=1, H_1^C(2,5)=1, H_1^C(4,n_s+n_e+4)=1, H_1^C(3,n_s+n_e+1)=1, H_1^C(3,n_s+n_e+2)=-1, H_1^C(4,n_s+n_e+4)=1, H_1^C(4,6)=1, H_1^C(4,n_s+n_e+3)=-1, H_1^C(5,n_s+n_e+5)=1, H_1^C(6,n_s+1)=-1$ and other elements are zero, with $H_1^C(i,j)$ denoting the (i,j) element of the matrix H_1^C, and $\Sigma_u^{1/2}$ is a matrix of zeros as we assume there is no measurement error in the core observed variables.

A.2.2 Non-core measurement equation

Suppose we have M non-core observed variables, which we denote as

$$Y_t^{NC} = (Y_{1,t}^{NC}, Y_{1,t}^{NC}, ..., Y_{M,t}^{NC})'.$$

We assume that they are generated by the following data-generating process:

$$Y_{m,t}^{NC} = b_{0,m} + b_{1,m} Y_{m,t-1}^{NC} + \lambda_m' \varepsilon_t + \sigma_{v,m} v_{m,t}, \quad v_{m,t} \sim \mathcal{N}(0,1)$$

where λ_m is $n_e \times 1$ vector of factor loadings on the structural shocks in the model. By stacking $Y_{m,t}^{NC}$, we obtain

$$Y_t^{NC} = B_0 + B_1 Y_{t-1}^{NC} + \Lambda \varepsilon_t + v_t, \quad v_t \sim \mathcal{N}(0, \Sigma_v) \tag{12}$$

where B_0 is $M \times 1$ vector, $B_1 = diag([b_{1,1}, b_{1,2}, b_{1,3}, ..., b_{1,M}])$, Λ is $M \times n_e$ matrix stacking λ_m' for $m = 1, 2, \cdots M$, and $\Sigma_v = diag([\sigma_{v,1}^2, \sigma_{v,2}^2, ..., \sigma_{v,M}^2])$.

Note that, given the observation Y_t^{NC} and the set of parameters B_0, B_1, Λ and Σ_v, we can rewrite equation (12) as

$$\tilde{Y}^{NC} \equiv Y_t^{NC} - B_0 - B_1 Y_{t-1}^{NC} = \Lambda(RS_t) + v_t, \tag{A.1}$$

where R is a deterministic selection matrix that picks structural shocks from S_t. This expression is useful in formulating the state-space representation by combining the core and non-core measurement equations.

A.2.3 Measurement equation combining core and non-core variables

By combining the core measurement equation (11) and the non-core measurement equation (A.1), we have the following measurement equation:

$$\begin{aligned} Y_t \equiv \begin{bmatrix} Y_t^C \\ \tilde{Y}_t^{NC} \end{bmatrix} &= \begin{bmatrix} H_0^C \\ \mathbf{0}_{M\times 1} \end{bmatrix} + \begin{bmatrix} H_1^C \\ \Lambda R \end{bmatrix} S_t + \begin{bmatrix} \Sigma_u^{1/2} \\ \Sigma_v^{1/2} \end{bmatrix} e_t \\ &= H_0 + H_1 S_t + e_t, \quad e_t \sim \mathcal{N}(0, \Sigma_e). \end{aligned} \tag{A.2}$$

The state transition equation (9) and the measurement equation (A.2) provide the linear-Gaussian state-space representation of our model, which enables us to compute the likelihood by the Kalman filter.

B Markov chain Monte Carlo sampling algorithm

In this section, we explain our algorithm for drawing samples from the posterior distribution in more detail. Our sampling strategy is the Metropolis-Hastings-within-Gibbs approach,

where we draw the structural parameter of the DSGE model by the random block random walk Metropolis-Hastings, the state variables by a simulation smoother, and parameters of the non-core measurement equation by Gibbs sampling and discrete approximation. Our algorithms consist of two blocks, as explained in Section 2.

B.1 First block: $p(\theta, S_{1:T}|\varphi, Y^C_{1:T}, Y^{NC}_{1:T})$

In this block, we first sample θ by applying the random block random walk Metropolis-Hastings and then apply the simulation smoother to sample $S_{1:T}$.

In the random block random walk Metropolis-Hastings, we first randomly assign structural parameters into blocks without duplication, where the number of total blocks and the size of each block are set randomly. To explain the sorting procedure, let us denote the probability of having a new block for each parameter as p^{block}. We have 22 structural parameters, so on average, we have $22 \times p^{block}$ blocks, as the number of total blocks follows a binomial distribution with probability of success p^{block}. When the number of blocks is determined, we randomly assign the structural parameters to one of the blocks without duplication. Note that we employ this randomization for each iteration of the algorithm, so the number of blocks and the assignment pattern regarding which parameters are placed in which block change in every draw. In the estimation, we set $p^{block} = 0.3$.

Then, given the draw of φ we run the random walk Metropolis-Hastings algorithm for each block of θ.

Finally, given the draw of (θ, φ), we draw $S_{1:T}$ by applying the simulation smoother of Durbin and Koopman (2002) to the state-space representation (7).

Below is a detailed description of the algorithm.

1 Given the initial value of all the parameters, run the following algorithm for $j = 1, 2, \cdots, N_{sim}$.

2 Randomly generate a partition of θ so that each parameter is randomly assigned to one block of parameters $\theta_1, \theta_2, ..., \theta_{N^{(j)}_{block}}$ without duplication. Essentially, the idea is to randomly select the number of total blocks and the number of parameters in each block and then assign each parameter to one block randomly. The following is one way of doing so.

2-1 Assign an independent uniform random number to each parameter. Sort the parameters by using these numbers in ascending order, which provides a new vector of parameters: θ^r.

2-2 Create a list b that specifies which parameter belongs to which block. The total number of blocks and the number of parameters in each block are determined

randomly. One way of doing so is by iteratively increasing the index whenever a uniformly distributed random number exceeds a specified probability p^{block}. As a result, you obtain a list like $b = [1, 1, 1, 2, 2, 3, ...4, 5, 5]$.

2-3 Combine the sorted parameters θ^r with the block-assinment list b. Each element of b indicates the block to which the corresponding parameter in θ^r belongs. For instance, b could be $b = [1, 1, 1, 2, 2, 3, ...4, 5, 5]$, and then the first three and second two parameters in θ^r are assigned to the first and second block, respectively, and so forth. In this example, the total number of blocks is five.

3 For each block of the parameters, we run the Metropolis-Hastings algorithm. Specifically, for $b = 1, 2, ..., N^{(j)}_{block}$, do the following:

3-1 Draw $\tilde{\theta}_b = \theta_b^{(j-1)} + \eta_b$ where $\eta_b \sim \mathcal{N}(0, c^2\hat{\Sigma}_b)$ and $\hat{\Sigma}_b$ is a matrix extracting the corresponding row and column of the negative of the inverse Hessian at the mode of log posterior density. This Hessian is computed by the model only with the core-measurement equation before running the MCMC algorithm.

3-2 Reject or accept the draw:

$$\theta_b^{(j)} = \begin{cases} \tilde{\theta}_b & \text{with probability } \alpha, \\ \theta_b^{(j-1)} & \text{with probability } 1-\alpha. \end{cases}$$

where

$$\alpha = \min\left\{ \frac{p(Y^C_{1:T}, Y^{NC}_{1:T} | \theta^{(j)}_{<b}, \tilde{\theta}_b, \theta^{(j-1)}_{b<}, \varphi^{(j-1)}) p(\theta^{(j)}_{<b}, \tilde{\theta}_b, \theta^{(j-1)}_{b<})}{p(Y^C_{1:T}, Y^{NC}_{1:T} | \theta^{(j)}_{<b}, \theta^{(j-1)}, \theta^{(j-1)}_{b<}, \varphi^{(j-1)}) p(\theta^{(j)}_{<b}, \theta^{(j-1)}_{b}, \theta^{(j-1)}_{b<})}, 1 \right\}$$

where $\theta^{(j)}_{<b} = [\theta^{(j)}_1, ..., \theta^{(j)}_{b-1}]$, $\theta^{(j-1)}_{b<} = [\theta^{(j-1)}_{b+1}, ..., \theta^{(j-1)}_{N_{block}}]$, and $p(\theta)$ is the prior density. The likelihood, $p(Y^C_{1:T}, Y^{NC}_{1:T} | \theta, \varphi^{(j-1)})$, can be computed by applying the Kalman filter to the state-space model (7). Note that we use the unconditional mean and variance of $S_{1:T}$ as the initial value in the Kalman filter.

4 Draw $S^{(j)}_{1:T}$ from $p(S_{1:T} | Y^C_{1:T}, Y^{NC}_{1:T}, \theta^{(j)}, \varphi^{(j-1)})$ by applying simulation smoother of Durbin and Koopman (2002) to the state-space model (7) by following the algorithm of Jarociński (2015).

B.2 Second block: $p(\varphi | \theta, S_{1:T}, Y^C_{1:T}, Y^{NC}_{1:T})$

Note that given $(Y^C_{1:T}, Y^{NC}_{1:T})$ and the draw of $S_{1:T}$, the measurement equation for non-core observables (12) essentially becomes a regression model with a Gaussian error where the

regressors are intercepts, 1-period lag of the observables, and the structural shocks of the model. Furthermore, the prior distribution for φ follows Giannone et al. (2021) and is formulated as

$$\lambda_m | V_m, q_m \sim \begin{cases} N(0, V_m) & \text{with probability } q_m \\ 0 & \text{with probability } 1 - q_m \end{cases} \quad \text{where } V_m = \sigma_{v,m}^2 \gamma_m^2$$

$$p(\sigma_{v,m}^2) \propto \frac{1}{\sigma_{v,m}^2}$$

$$q_m \sim \text{beta}(a, b)$$

$$R_m^2 \sim \text{beta}(A, B) \text{ where } R_m^2(\gamma_m^2, q_m) \equiv \frac{n_e q_m \gamma_m^2 \bar{v}_\varepsilon}{n_e q_m \gamma_m^2 \bar{v}_\varepsilon + 1}$$

$$B_0, B_1 \sim \text{flat}$$

for $m = 1, 2, .., M$, where $\bar{v}_\varepsilon$ is the mean of the variance of ε_t. Following Giannone et al. (2021), we assume $a = b = A = B = 1$ so that the prior for q_m and R_m^2 is uniform. Hence, we can directly apply the sampling algorithm developed by Giannone et al. (2021) to draw $\varphi^{(j)}$ from $p(\varphi | Y_{1:T}^C, Y_{1:T}^{NC}, \theta^{(j)}, S_{1:T}^{(j)})$, with a slight modification for the normalization of the regressors.

Introducing sparsity in a regression model generally requires some normalization of the regressors because the estimation result is not invariant to the volatility of the regressors. In Giannone et al. (2021), the regressors are observed data, so one can divide them by their standard deviation to have unit volatility, for instance. However, our regressors, namely ε_t, are stochastic values drawn from its full conditional distribution, $p(S_{1:T} | \theta, Y_{1:T}^C, Y_{1:T}^{NC}, \varphi)$, so the normalization is not a straightforward issue. Hence, we follow the following steps to standardize each draw of $\varepsilon_t^{(j)}$. First, we draw 100 samples of $S_{1:T}^{(i)}, i = 1, 2, \cdots, 100$ from $p(S_{1:T} | \theta, Y_{1:T}^C)$, the full conditional distribution of the state variables in the model estimated only with the core measurement equation, where θ is set to be its posterior mode. We extract only the structural shocks from the draw of $S_{1:T}^{(i)}$ and denote it as $\varepsilon_{1:T}^{(i)}$. Second, we compute the mean of the structural shocks across draws and obtain $\widehat{\varepsilon}_{1:T} = \frac{1}{100} \sum_{i=1}^{100} \varepsilon_{1:T}^{(i)}$. Third, we compute the standard deviations of $\widehat{\varepsilon}_{1:T}$, denoted as $\overline{\sigma}_\varepsilon$, and standardize $\varepsilon_t^{(j)}$ for each draw in the MCMC by computing $\overline{\varepsilon}_t^{(j)} = \frac{\varepsilon_t^{(j)}}{\overline{\sigma}_\varepsilon}$, and include them as regressors. Note that the draw of Λ in this step needs to be rescaled by $\overline{\sigma}_\varepsilon$ when used as inputs for the first block.

Otherwise, our sampling method closely follows the algorithm proposed by Giannone et al. (2021), so for the detailed algorithm, see their appendix.

Supplementary Material:

Structural Estimation of Dynamic Equilibrium Models with Unstructured Data

Sara Casella
LUISS University and EIEF

Jesús Fernández-Villaverde
University of Pennsylvania, NBER, CEPR

Stephen Hansen
University College London, IFS, CEPR

Ryohei Oishi
University College London

Minchul Shin
Federal Reserve Bank of Philadelphia

A New Keynesian DSGE model

We use a medium-scale sticky-price model that closely follows Fernández-Villaverde and Guerrón-Quintana (2021); derivations follow Fernández-Villaverde and Rubio-Ramírez (2006), modified to reflect our specification.[11] A representative household consumes, saves, holds money, and supplies labor. A final good producer aggregates a continuum of intermediate goods manufactured by monopolistic competitors. Intermediate good producers rent capital and labor and set prices subject to a Calvo friction. A consolidated public sector sets the one-period nominal interest rate through open market operations and consumes an exogenous stream g. Long-run growth is driven by two unit roots, one in neutral technology and one in investment-specific technology.

A.1 Environment

A.1.1 Households

The representative household maximizes the following lifetime utility function, which is separable in consumption, c_t, real money balances, m_t/p_t (where p_t is the price level), and

[11] The notes are available at https://www.sas.upenn.edu/~jesusfv/benchmark_DSGE.pdf

hours worked, l_t:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t d_t \left\{ \log\left(c_t - h c_{t-1}\right) + \upsilon \log\left(\frac{m_t}{p_t}\right) - \phi_t \psi \frac{l_t^{1+\vartheta}}{1+\vartheta} \right\}$$

where β is the discount factor, h is the parameter that controls habit persistence, ϑ is the inverse of Frisch labor supply elasticity, d_t is an intertemporal preference shock with law of motion:

$$\log d_t = \rho_D \log d_{t-1} + \sigma_D \varepsilon_{D,t} \text{ where } \varepsilon_{D,t} \sim \mathcal{N}(0,1),$$

and ϕ_t is a labor supply shock with law of motion:

$$\log \phi_t = \rho_\phi \log \phi_{t-1} + \sigma_\phi \varepsilon_{\phi,t} \text{ where } \varepsilon_{\phi,t} \sim \mathcal{N}(0,1).$$

The preference shock d_t will change the intertemporal first order conditions, while the preference shock ϕ_t will change the static first order conditions. Also, we have selected a utility function (log utility in consumption) whose marginal relation of substitution between consumption and leisure is linear in consumption to ensure the presence of a balanced growth path with constant hours.

Households hold an amount b_t of one-period government bonds that pay a nominal gross interest rate of R_t. For notational simplicity, we do not explicitly include markets in private securities, so all borrowing and lending is between consumers and the government. Since all consumers are identical, such claims would not be traded in equilibrium, hence their absence will not affect the equilibrium. We can interpret the current model as having complete private claims markets. The household's budget constraint is given by:

$$\begin{aligned} & c_t + x_t + \frac{m_t}{p_t} + \frac{b_{t+1}}{p_t} \\ &= w_t l_t + \left(r_t u_t - \mu_t^{-1} \Phi\left[u_t\right]\right) k_{t-1} + \frac{m_{t-1}}{p_t} + R_{t-1} \frac{b_t}{p_t} + T_t + \digamma_t \end{aligned}$$

where w_t is the real wage, r_t the real rental price of capital, $u_t > 0$ the intensity of use of capital, $\mu_t^{-1}\Phi\left[u_t\right]$ is the physical cost of use of capital in resource terms, μ_t is an investment-specific technological shock to be described momentarily, T_t is a lump-sum transfer, and $\digamma_t$ are the profits of the firms in the economy. We assume that $\Phi\left[1\right] = 0$, Φ' and $\Phi'' > 0$.

Investment x_t induces a law of motion for capital:

$$k_t = \left(1-\delta\right) k_{t-1} + \mu_t \left(1 - S\left[\frac{x_t}{x_{t-1}}\right]\right) x_t$$

where δ is the depreciation rate and $S[\cdot]$ is an adjustment cost function such that $S[\Lambda_x] = 0$, $S'[\Lambda_x] = 0$, and $S''[\cdot] > 0$ where Λ_x is the growth rate of investment along the balance growth path. We will determine that growth rate below. Note our capital timing: we index capital by the time its level is decided. The investment-specific technological shock follows an autoregressive process:

$$\mu_t = \mu_{t-1} \exp(\Lambda_\mu + z_{\mu,t}) \text{ where } z_{\mu,t} = \sigma_\mu \varepsilon_{\mu,t} \text{ and } \varepsilon_{\mu,t} \sim \mathcal{N}(0,1)$$

The value of μ_t is also the inverse of the relative price of new capital in consumption terms.

Given our description of the household's problem, the lagrangian function associated with it is:

$$\mathbb{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\begin{array}{c} d_t \left\{ \log(c_t - h c_{t-1}) + \upsilon \log\left(\frac{m_t}{p_t}\right) - \phi_t \psi \frac{l_t^{1+\vartheta}}{1+\vartheta} \right\} \\ -\lambda_t \left\{ \begin{array}{c} c_t + x_t + \frac{m_t}{p_t} + \frac{b_t}{p_t} \\ -w_t l_t - \left(r_t u_t - \mu_t^{-1} \Phi[u_t]\right) k_{t-1} - \frac{m_{t-1}}{p_t} - R_{t-1} \frac{b_{t-1}}{p_t} - T_t - \digamma_t \end{array} \right\} \\ -Q_t \left\{ k_t - (1-\delta) k_{t-1} - \mu_t \left(1 - S\left[\frac{x_t}{x_{t-1}}\right]\right) x_t \right\} \end{array} \right]$$

where they maximize over c_t, b_t, u_t, k_t, x_t, and l_t (maximization with respect to money holdings comes from the budget constraint), λ_t is the lagrangian multiplier associated with the budget constraint and Q_t the lagrangian multiplier associated with installed capital.

The first order conditions with respect to c_t, l_t, b_{t+1}, u_t, k_t and x_t are:

$$d_t (c_t - h c_{t-1})^{-1} - h\beta \mathbb{E}_t d_{t+1} (c_{t+1} - h c_t)^{-1} = \lambda_t$$

$$d_t \phi_t \psi l_t^{\vartheta} = w_t \lambda_t$$

$$\lambda_t = \beta \mathbb{E}_t \left\{ \lambda_{t+1} \frac{R_t}{\Pi_{t+1}} \right\}$$

$$r_t = \mu_t^{-1} \Phi'[u_t]$$

$$Q_t = \beta \mathbb{E}_t \left\{ (1-\delta) Q_{t+1} + \lambda_{t+1} \left(r_{t+1} u_{t+1} - \mu_{t+1}^{-1} \Phi[u_{t+1}] \right) \right\}$$

$$-\lambda_t + Q_t \mu_t \left(1 - S\left[\frac{x_t}{x_{t-1}}\right] - S'\left[\frac{x_t}{x_{t-1}}\right] \frac{x_t}{x_{t-1}}\right) + \beta \mathbb{E}_t Q_{t+1} \mu_{t+1} S'\left[\frac{x_{t+1}}{x_t}\right] \left(\frac{x_{t+1}}{x_t}\right)^2 = 0.$$

If we define the (marginal) Tobin's Q as $q_t = \frac{Q_t}{\lambda_t}$, (the ratio of the two lagrangian multipliers, or more loosely the value of installed capital in terms of its replacement cost) we

get:

$$d_t \left(c_t - hc_{t-1}\right)^{-1} - h\beta \mathbb{E}_t d_{t+1} \left(c_{t+1} - hc_t\right)^{-1} = \lambda_t$$
$$d_t \phi_t \psi l_t^{\vartheta} = w_t \lambda_t$$
$$\lambda_t = \beta \mathbb{E}_t \left\{ \lambda_{t+1} \frac{R_t}{\Pi_{t+1}} \right\}$$
$$r_t = \mu_t^{-1} \Phi' \left[u_t\right]$$
$$q_t = \beta \mathbb{E}_t \left\{ \frac{\lambda_{t+1}}{\lambda_t} \left(\left(1-\delta\right) q_{t+1} + r_{t+1} u_{t+1} - \mu_{t+1}^{-1} \Phi \left[u_{t+1}\right] \right) \right\}$$
$$1 = q_t \mu_t \left(1 - S\left[\frac{x_t}{x_{t-1}}\right] - S'\left[\frac{x_t}{x_{t-1}}\right] \frac{x_t}{x_{t-1}} \right) + \beta \mathbb{E}_t q_{t+1} \mu_{t+1} \frac{\lambda_{t+1}}{\lambda_t} S'\left[\frac{x_{t+1}}{x_t}\right] \left(\frac{x_{t+1}}{x_t}\right)^2.$$

The last equation is important. If $S\left[\cdot\right] = 0$ (i.e., there are no adjustment costs), we get:

$$q_t = \frac{1}{\mu_t}$$

i.e., the marginal Tobin's Q is equal to the replacement cost of capital (the relative price of capital). Furthermore, if $\mu_t = 1$, as in the standard neoclassical growth model, $q_t = 1$.

A.1.2 The Final Good Producer

There is one final good is produced using intermediate goods with the following production function:

$$y_t^d = \left(\int_0^1 y_{it}^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}}. \tag{A.1}$$

where ε is the elasticity of substitution.

Final good producers are perfectly competitive and maximize profits subject to the production function (A.1), taking as given all intermediate goods prices p_{ti} and the final good price p_t. As a consequence their maximization problem is:

$$\max_{y_{it}} p_t y_t^d - \int_0^1 p_{it} y_{it} di$$

The input demand functions associated with this problem are:

$$y_{it} = \left(\frac{p_{it}}{p_t}\right)^{-\varepsilon} y_t^d \qquad \forall i,$$

where y_t^d is the aggregate demand and the zero profit condition $p_t y_t^d = \int_0^1 p_{it} y_{it} di$ to deliver:

$$p_t = \left(\int_0^1 p_{it}^{1-\varepsilon} di \right)^{\frac{1}{1-\varepsilon}}.$$

A.1.3 Intermediate Good Producers

There is a continuum of intermediate goods producers. Each intermediate good producer i has access to a technology represented by a production function

$$y_{it} = A_t k_{it-1}^{\alpha} \left(l_{it}^d \right)^{1-\alpha} - f z_t$$

where k_{it-1} is the capital rented by the firm, l_{it}^d is the amount of the "packed" labor input rented by the firm, and where A_t follows the following process:

$$A_t = A_{t-1} \exp\left(\Lambda_A + z_{A,t}\right) \text{ where } z_{A,t} = \sigma_A \varepsilon_{A,t} \text{ and } \varepsilon_{A,t} \sim \mathcal{N}(0,1)$$

The parameter f, which corresponds to the fixed cost of production, and $z_t = A_t^{\frac{1}{1-\alpha}} \mu_t^{\frac{\alpha}{1-\alpha}}$ guarantee that economic profits are roughly equal to zero in the steady state. We rule out the entry and exit of intermediate good producers.

Since $z_t = A_t^{\frac{1}{1-\alpha}} \mu_t^{\frac{\alpha}{1-\alpha}}$, we have that

$$z_t = z_{t-1} \exp\left(\Lambda_z + z_{z,t}\right) \text{ where } z_{z,t} = \frac{z_{A,t} + \alpha z_{\mu,t}}{1-\alpha} \text{ and } \Lambda_z = \frac{\Lambda_A + \alpha \Lambda_\mu}{1-\alpha}.$$

Intermediate goods producers solve a two-stages problem. In the first stage, taken the input prices w_t and r_t as given, firms rent l_{it}^d and k_{it-1} in perfectly competitive factor markets in order to minimize real cost:

$$\min_{l_{it}^d, k_{it-1}} \; w_t l_{it}^d + r_t k_{it-1}$$

subject to their supply curve:

$$y_{it} = \begin{cases} A_t k_{it-1}^{\alpha} \left(l_{it}^d \right)^{1-\alpha} - f z_t & \text{if } A_t k_{it-1}^{\alpha} \left(l_{it}^d \right)^{1-\alpha} \geq f z_t \\ 0 & \text{otherwise} \end{cases}$$

Assuming an interior solution, the first order conditions for this problem are:

$$\begin{aligned} w_t &= \varrho \left(1-\alpha\right) A_t k_{it-1}^{\alpha} \left(l_{it}^d \right)^{-\alpha} \\ r_t &= \varrho \alpha A_t k_{it-1}^{\alpha-1} \left(l_{it}^d \right)^{1-\alpha} \end{aligned}$$

where ϱ is the Lagrangian multiplier or:

$$k_{it-1} = \frac{\alpha}{1-\alpha}\frac{w_t}{r_t}l_{it}^d$$

The real cost is then:

$$w_t l_{it}^d + \frac{\alpha}{1-\alpha}w_t l_{it}^d$$

or:

$$\left(\frac{1}{1-\alpha}\right)w_t l_{it}^d$$

Given that the firm has constant returns to scale, we can find the real marginal cost mc_t by setting the level of labor and capital equal to the requirements of producing one unit of good $A_t k_{it-1}^{\alpha}\left(l_{it}^d\right)^{1-\alpha} = 1$ or:

$$A_t k_{it-1}^{\alpha}\left(l_{it}^d\right)^{1-\alpha} = A_t\left(\frac{\alpha}{1-\alpha}\frac{w_t}{r_t}l_{it}^d\right)^{\alpha}\left(l_{it}^d\right)^{1-\alpha} = A_t\left(\frac{\alpha}{1-\alpha}\frac{w_t}{r_t}\right)^{\alpha}l_{it}^d = 1$$

that implies that:

$$l_{it}^d = \frac{\left(\frac{\alpha}{1-\alpha}\frac{w_t}{r_t}\right)^{-\alpha}}{A_t}$$

Then:

$$mc_t = \left(\frac{1}{1-\alpha}\right)w_t\frac{\left(\frac{\alpha}{1-\alpha}\frac{w_t}{r_t}\right)^{-\alpha}}{A_t}$$

that simplifies to:

$$mc_t = \left(\frac{1}{1-\alpha}\right)^{1-\alpha}\left(\frac{1}{\alpha}\right)^{\alpha}\frac{w_t^{1-\alpha}r_t^{\alpha}}{A_t}$$

Note that the marginal cost does not depend on i: all firms receive the same technology shocks and all firms rent inputs at the same price.

In the second stage, intermediate good producers choose the price that maximizes discounted real profits. To do so, they consider that they are under the same pricing scheme as households. In each period, a fraction $1-\theta_p$ of firms can change their prices. All other firms can only index their prices by past inflation. Indexation is controlled by the parameter $\chi \in [0,1]$, where $\chi = 0$ is no indexation and $\chi = 1$ is total indexation.

The problem of the firms is then:

$$\max_{p_{it}} \mathbb{E}_t \sum_{\tau=0}^{\infty}(\beta\theta_p)^{\tau}\frac{\lambda_{t+\tau}}{\lambda_t}\left\{\left(\prod_{s=1}^{\tau}\Pi_{t+s-1}^{\chi}\frac{p_{it}}{p_{t+\tau}} - mc_{t+\tau}\right)y_{it+\tau}\right\}$$

subject to

$$y_{it+\tau} = \left(\prod_{s=1}^{\tau} \Pi_{t+s-1}^{\chi} \frac{p_{it}}{p_{t+\tau}} \right)^{-\varepsilon} y_{t+\tau}^{d},$$

where the marginal value of a dollar to the household, is treated as exogenous by the firm. Since we have complete markets in securities and the utility function is separable in consumption, this marginal value is constant across households and, consequently, $\lambda_{t+\tau}/\lambda_t$ is the correct valuation on future profits.

Substituting the demand curve in the objective function and the previous expression, we get:

$$\max_{p_{it}} \mathbb{E}_t \sum_{\tau=0}^{\infty} (\beta\theta_p)^{\tau} \frac{\lambda_{t+\tau}}{\lambda_t} \left\{ \left(\left(\prod_{s=1}^{\tau} \Pi_{t+s-1}^{\chi} \frac{p_{it}}{p_{t+\tau}} \right)^{1-\varepsilon} - \left(\prod_{s=1}^{\tau} \Pi_{t+s-1}^{\chi} \frac{p_{it}}{p_{t+\tau}} \right)^{-\varepsilon} mc_{t+\tau} \right) y_{t+\tau}^{d} \right\}$$

or

$$\max_{p_{it}} \mathbb{E}_t \sum_{\tau=0}^{\infty} (\beta\theta_p)^{\tau} \frac{\lambda_{t+\tau}}{\lambda_t} \left\{ \left(\left(\prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi}}{\Pi_{t+s}} \frac{p_{it}}{p_t} \right)^{1-\varepsilon} - \left(\prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi}}{\Pi_{t+s}} \frac{p_{it}}{p_t} \right)^{-\varepsilon} mc_{t+\tau} \right) y_{t+\tau}^{d} \right\}$$

whose solution p_{it}^{*} implies the first order condition:

$$\mathbb{E}_t \sum_{\tau=0}^{\infty} (\beta\theta_p)^{\tau} \frac{\lambda_{t+\tau}}{\lambda_t} \left\{ \left((1-\varepsilon) \left(\prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi}}{\Pi_{t+s}} \frac{p_{it}^{*}}{p_t} \right)^{1-\varepsilon} p_{it}^{*-1} + \varepsilon \left(\prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi}}{\Pi_{t+s}} \frac{p_{it}^{*}}{p_t} \right)^{-\varepsilon} p_{it}^{*-1} mc_{t+\tau} \right) y_{t+\tau}^{d} \right\} = 0$$

or

$$\mathbb{E}_t \sum_{\tau=0}^{\infty} (\beta\theta_p)^{\tau} \lambda_{t+\tau} \left\{ \left((1-\varepsilon) \left(\prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi}}{\Pi_{t+s}} \right)^{1-\varepsilon} \frac{p_{it}^{*}}{p_t} + \varepsilon \left(\prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi}}{\Pi_{t+s}} \right)^{-\varepsilon} mc_{t+\tau} \right) y_{t+\tau}^{d} \right\} = 0$$

where, in the second step, we have dropped irrelevant constants. Note how this expression nests the usual result in the fully flexible prices case $\theta_p = 0$:

$$p_{it}^{*} = \frac{\varepsilon}{\varepsilon - 1} p_t mc_t$$

i.e., the price is equal to a mark-up over the nominal marginal cost.

Since we only consider a symmetric equilibrium, we can write that $p_{it}^* = p_t^*$ and:

$$\mathbb{E}_t \sum_{\tau=0}^{\infty} (\beta\theta_p)^\tau \lambda_{t+\tau} \left\{ \left((1-\varepsilon) \left(\prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi}}{\Pi_{t+s}} \right)^{1-\varepsilon} \frac{p_t^*}{p_t} + \varepsilon \left(\prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi}}{\Pi_{t+s}} \right)^{-\varepsilon} mc_{t+\tau} \right) y_{t+\tau}^d \right\} = 0$$

To express the previous first order condition recursively, we define:

$$g_t^1 = \mathbb{E}_t \sum_{\tau=0}^{\infty} (\beta\theta_p)^\tau \lambda_{t+\tau} \left(\prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi}}{\Pi_{t+s}} \right)^{-\varepsilon} mc_{t+\tau} y_{t+\tau}^d$$

and

$$g_t^2 = \mathbb{E}_t \sum_{\tau=0}^{\infty} (\beta\theta_p)^\tau \lambda_{t+\tau} \left(\prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi}}{\Pi_{t+s}} \right)^{1-\varepsilon} \frac{p_t^*}{p_t} y_{t+\tau}^d$$

and then the first order condition is

$$\varepsilon g_t^1 = (\varepsilon - 1) g_t^2$$

or

$$g_t^2 = \frac{\varepsilon}{\varepsilon - 1} g_t^1$$

We need $(\beta\theta_p)^\tau \lambda_{t+\tau}$ to go to zero sufficiently fast in relation with the rate of inflation for g_t^1 and g_t^2 to be well defined and stationary.

Then, we can write the g's recursively:

$$g_t^1 = \lambda_t mc_t y_t^d + \beta\theta_p \mathbb{E}_t \left(\frac{\Pi_t^{\chi}}{\Pi_{t+1}} \right)^{-\varepsilon} g_{t+1}^1$$

and

$$g_t^2 = \lambda_t \Pi_t^* y_t^d + \beta\theta_p \mathbb{E}_t \left(\frac{\Pi_t^{\chi}}{\Pi_{t+1}} \right)^{1-\varepsilon} \left(\frac{\Pi_t^*}{\Pi_{t+1}^*} \right) g_{t+1}^2$$

where:

$$\Pi_t^* = \frac{p_t^*}{p_t}.$$

Given Calvo's pricing, the price index evolves:

$$p_t^{1-\varepsilon} = \theta_p \left(\Pi_{t-1}^{\chi} \right)^{1-\varepsilon} p_{t-1}^{1-\varepsilon} + (1-\theta_p) \, p_t^{*1-\varepsilon}$$

or, dividing by $p_t^{1-\varepsilon}$,

$$1 = \theta_p \left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t} \right)^{1-\varepsilon} + (1-\theta_p) \Pi_t^{*1-\varepsilon}$$

A.1.4 The Government Problem

The government determines monetary and fiscal policy. We assume there is no interaction between monetary and fiscal policy, that is, the results of open market operations are distributed in a lump-sum fashion to households and not transferred to the general revenue of the government.

Monetary Policy

The government sets the nominal interest rates according to the Taylor rule:

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R} \right)^{\gamma_R} \left(\left(\frac{\Pi_t}{\Pi} \right)^{\gamma_\Pi} \left(\frac{\frac{y_t^d}{y_{t-1}^d}}{\Lambda_{y^d}} \right)^{\gamma_y} \right)^{1-\gamma_R} \exp(m_t)$$

through open market operations that are financed through lump-sum transfers T_t. Those transfers insure that the deficit from monetary policy is equal to zero:

$$T_t = \frac{m_t}{p_t} - \frac{m_{t-1}}{p_t} + \frac{b_{t+1}}{p_t} - R_{t-1} \frac{b_t}{p_t}$$

The variable Π represents the target level of inflation (equal to inflation in the steady state), R steady state nominal gross return of capital, and Λ_{y^d} the steady state gross growth rate of y_t^d. The term m_t is a random shock to monetary policy that follows $m_t = \sigma_M \varepsilon_{M,t}$ where $\varepsilon_{M,t}$ is distributed according to $\mathcal{N}(0,1)$. The presence of the previous period interest rate, R_{t-1}, is justified because we want to match the smooth profile of the interest rate over time observed in U.S. data. Note that R is beyond the control of the monetary authority, since it is equal to the steady state real gross returns of capital plus the target level of inflation.

Fiscal Policy

Government consumption is given by $g_t = \widetilde{g}_t z_t$, where $\widetilde{g}_t$ follows an autoregressive process in logs:

$$\log \widetilde{g}_t = (1-\rho_g) \log \widetilde{g} + \rho_g \log \widetilde{g}_{t-1} + \sigma_g \varepsilon_{g,t} \quad \text{where } \varepsilon_{g,t} \sim \mathcal{N}(0,1)$$

and

$$z_t = A_t^{\frac{1}{1-\alpha}} \mu_t^{\frac{\alpha}{1-\alpha}}$$

Since the economy is growing, the level of real government consumption is multiplied by the term z_t defined above to keep it a stationary share of output. Government consumption is also financed by lump sum taxes that ensure that the deficit is zero at all points in time.

A.1.5 Aggregation

First, we have an expression for aggregate demand:

$$y_t^d = c_t + g_t + x_t + \mu_t^{-1}\Phi\left[u_t\right]k_{t-1}$$

With this value, the demand for each intermediate good producer is

$$y_{it} = \left(c_t + g_t + x_t + \mu_t^{-1}\Phi\left[u_t\right]k_{t-1}\right)\left(\frac{p_{it}}{p_t}\right)^{-\varepsilon} \qquad \forall i,$$

and using the production function is:

$$A_t k_{it-1}^{\alpha}\left(l_{it}^d\right)^{1-\alpha} - fz_t = \left(c_t + g_t + x_t + \mu_t^{-1}\Phi\left[u_t\right]k_{t-1}\right)\left(\frac{p_{it}}{p_t}\right)^{-\varepsilon}$$

Since all the firms have the same optimal capital-labor ratio:

$$\frac{k_{it-1}}{l_{it}^d} = \frac{\alpha}{1-\alpha}\frac{w_t}{r_t}$$

and by market clearing

$$\int_0^1 l_{it}^d di = l_t$$

and

$$\int_0^1 k_{it-1} di = u_t k_{t-1}.$$

it must be the case that:

$$\frac{k_{it-1}}{l_{it}^d} = \frac{u_t k_{t-1}}{l_t}.$$

Then:

$$A_t k_{it-1}^{\alpha}\left(l_{it}^d\right)^{1-\alpha} = A_t\left(\frac{k_{it-1}}{l_{it}^d}\right)^{\alpha}\left(l_{it}^d\right) = A_t\left(\frac{u_t k_{t-1}}{l_t^d}\right)^{\alpha} l_{it}^d$$

Integrating out

$$\int_0^1 A_t\left(\frac{u_t k_{t-1}}{l_t^d}\right)^{\alpha} l_{it}^d di = A_t\left(\frac{u_t k_{t-1}}{l_t^d}\right)^{\alpha}\int_0^1 l_{it}^d di = A_t\left(u_t k_{t-1}\right)^{\alpha}\left(l_t^d\right)^{1-\alpha}$$

and we have

$$A_t \left(u_t k_{t-1}\right)^{\alpha} \left(l_t^d\right)^{1-\alpha} - f z_t = \left(c_t + g_t + x_t + \mu_t^{-1} \Phi\left[u_t\right] k_{t-1}\right) \int_0^1 \left(\frac{p_{it}}{p_t}\right)^{-\varepsilon} di$$

Define $v_t^p = \int_0^1 \left(\frac{p_{it}}{p_t}\right)^{-\varepsilon} di$. By the properties of the index under Calvo's pricing

$$v_t^p = \theta_p \left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{-\varepsilon} v_{t-1}^p + \left(1-\theta_p\right) \Pi_t^{*-\varepsilon}.$$

we get:

$$c_t + g_t + x_t + \mu_t^{-1} \Phi\left[u_t\right] k_{t-1} = \frac{A_t \left(u_t k_{t-1}\right)^{\alpha} \left(l_t^d\right)^{1-\alpha} - f z_t}{v_t^p}$$

A.2 Equilibrium

A definition of equilibrium in this economy is standard and the symmetric equilibrium policy functions are determined by the following equations:

- The first order conditions of the household:

$$d_t \left(c_t - h c_{t-1}\right)^{-1} - h\beta \mathbb{E}_t d_{t+1} \left(c_{t+1} - h c_t\right)^{-1} = \lambda_t$$
$$d_t \phi_t \psi (l_t^s)^{\vartheta} = w_t \lambda_t$$
$$\lambda_t = \beta \mathbb{E}_t \{\lambda_{t+1} \frac{R_t}{\Pi_{t+1}}\}$$
$$r_t = \mu_t^{-1} \Phi'\left[u_t\right]$$
$$q_t = \beta \mathbb{E}_t \left\{\frac{\lambda_{t+1}}{\lambda_t} \left(\left(1-\delta\right) q_{t+1} + r_{t+1} u_{t+1} - \mu_{t+1}^{-1} \Phi\left[u_{t+1}\right]\right)\right\}$$
$$1 = q_t \mu_t \left(1 - S\left[\frac{x_t}{x_{t-1}}\right] - S'\left[\frac{x_t}{x_{t-1}}\right] \frac{x_t}{x_{t-1}}\right) + \beta \mathbb{E}_t q_{t+1} \mu_{t+1} \frac{\lambda_{t+1}}{\lambda_t} S'\left[\frac{x_{t+1}}{x_t}\right] \left(\frac{x_{t+1}}{x_t}\right)^2$$

- The firms that can change prices set them to satisfy:

$$g_t^1 = \lambda_t m c_t y_t^d + \beta \theta_p \mathbb{E}_t \left(\frac{\Pi_t^{\chi}}{\Pi_{t+1}}\right)^{-\varepsilon} g_{t+1}^1$$
$$g_t^2 = \lambda_t \Pi_t^* y_t^d + \beta \theta_p \mathbb{E}_t \left(\frac{\Pi_t^{\chi}}{\Pi_{t+1}}\right)^{1-\varepsilon} \left(\frac{\Pi_t^*}{\Pi_{t+1}^*}\right) g_{t+1}^2$$
$$\varepsilon g_t^1 = (\varepsilon - 1) g_t^2$$

where they rent inputs to satisfy their static minimization problem:

$$\frac{u_t k_{t-1}}{l_t} = \frac{\alpha}{1-\alpha}\frac{w_t}{r_t}$$
$$mc_t = \left(\frac{1}{1-\alpha}\right)^{1-\alpha}\left(\frac{1}{\alpha}\right)^{\alpha}\frac{w_t^{1-\alpha}r_t^{\alpha}}{A_t}$$

The price level evolves as:

$$1 = \theta_p\left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{1-\varepsilon} + (1-\theta_p)\,\Pi_t^{*1-\varepsilon}$$

- Government follows its Taylor rule:

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\gamma_R}\left(\left(\frac{\Pi_t}{\Pi}\right)^{\gamma_\Pi}\left(\frac{\frac{y_t^d}{y_{t-1}^d}}{\Lambda_{y^d}}\right)^{\gamma_y}\right)^{1-\gamma_R}\exp\left(m_t\right)$$

- Markets clear:

$$\begin{aligned} y_t^d &= \frac{A_t\left(u_t k_{t-1}\right)^{\alpha}\left(l_t^d\right)^{1-\alpha} - f z_t}{v_t^p} \\ y_t^d &= c_t + g_t + x_t + \mu_t^{-1}\Phi\left[u_t\right]k_{t-1} \end{aligned}$$

where price dispersion v_t^p is given by:

$$v_t^p = \theta_p\left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{-\varepsilon}v_{t-1}^p + (1-\theta_p)\,\Pi_t^{*-\varepsilon}$$

and

$$\begin{aligned} &l_t^s = l_t^d \\ &k_t - (1-\delta)\,k_{t-1} - \mu_t\left(1 - S\left[\frac{x_t}{x_{t-1}}\right]\right)x_t = 0. \end{aligned}$$

Stationary Equilibrium

Since we have growth in this model induced by technological change, most of the variables are growing in average. To solve the model, we need to make variables stationary.

Manipulating Equilibrium Conditions

First, we work on the first order conditions of the household

$$d_t\left(\frac{c_t}{z_t}-h\frac{c_{t-1}}{z_{t-1}}\frac{z_{t-1}}{z_t}\right)^{-1}-h\beta\mathbb{E}_t d_{t+1}\left(\frac{c_{t+1}}{z_{t+1}}\frac{z_{t+1}}{z_t}-h\frac{c_t}{z_t}\right)^{-1}=\lambda_t z_t$$

$$d_t\phi_t\psi(l_t)^{\vartheta}=\frac{w_t}{z_t}\lambda_t z_t$$

$$\lambda_t z_t=\beta\mathbb{E}_t\{\lambda_{t+1}z_{t+1}\frac{z_t}{z_{t+1}}\frac{R_t}{\Pi_{t+1}}\}$$

$$\mu_t r_t=\Phi'\left[u_t\right]$$

$$q_t\mu_t=\beta\mathbb{E}_t\left\{\frac{\lambda_{t+1}z_{t+1}}{\lambda_t z_t}\frac{z_t}{z_{t+1}}\frac{\mu_t}{\mu_{t+1}}\left[\left(1-\delta\right)q_{t+1}\mu_{t+1}+\mu_{t+1}r_{t+1}u_{t+1}-\Phi\left(u_{t+1}\right)\right]\right\}$$

$$1=q_t\mu_t\left(1-S\left[\frac{\frac{x_t}{z_t}}{\frac{x_{t-1}}{z_{t-1}}}\frac{z_t}{z_{t-1}}\right]-S'\left[\frac{\frac{x_t}{z_t}}{\frac{x_{t-1}}{z_{t-1}}}\frac{z_t}{z_{t-1}}\right]\frac{\frac{x_t}{z_t}}{\frac{x_{t-1}}{z_{t-1}}}\frac{z_t}{z_{t-1}}\right)$$

$$+\beta\mathbb{E}_t q_{t+1}\mu_{t+1}\frac{\lambda_{t+1}}{\lambda_t}S'\left[\frac{\frac{x_{t+1}}{z_{t+1}}}{\frac{x_t}{z_t}}\frac{z_{t+1}}{z_t}\right]\left(\frac{\frac{x_{t+1}}{z_{t+1}}}{\frac{x_t}{z_t}}\frac{z_{t+1}}{z_t}\right)^2$$

The firms that can change prices set them to satisfy:

$$g_t^1=\lambda_t z_t mc_t\frac{y_t^d}{z_t}+\beta\theta_p\mathbb{E}_t\left(\frac{\Pi_t^{\chi}}{\Pi_{t+1}}\right)^{-\varepsilon}g_{t+1}^1$$

$$g_t^2=\lambda_t z_t\Pi_t^*\frac{y_t^d}{z_t}+\beta\theta_p\mathbb{E}_t\left(\frac{\Pi_t^{\chi}}{\Pi_{t+1}}\right)^{1-\varepsilon}\left(\frac{\Pi_t^*}{\Pi_{t+1}^*}\right)g_{t+1}^2$$

$$\varepsilon g_t^1=(\varepsilon-1)g_t^2$$

where they rent inputs to satisfy their static minimization problem:

$$\frac{u_t}{l_t^d}\frac{k_{t-1}}{z_{t-1}\mu_{t-1}}=\frac{\alpha}{1-\alpha}\frac{w_t}{z_t}\frac{1}{r_t\mu_t}\frac{z_t}{z_{t-1}}\frac{\mu_t}{\mu_{t-1}}$$

$$mc_t=\left(\frac{1}{1-\alpha}\right)^{1-\alpha}\left(\frac{1}{\alpha}\right)^{\alpha}\frac{\left(\frac{w_t}{z_t}\right)^{1-\alpha}\left(r_t\mu_t\right)^{\alpha}\frac{z_t^{1-\alpha}}{\mu_t^{\alpha}}}{A_t}$$

The price level evolves as:

$$1=\theta_p\left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{1-\varepsilon}+\left(1-\theta_p\right)\Pi_t^{*1-\varepsilon}$$

Government follows its Taylor rule

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\gamma_R} \left(\left(\frac{\Pi_t}{\Pi}\right)^{\gamma_\Pi} \left(\frac{\frac{\frac{y_t^d}{z_t}}{\frac{y_{t-1}^d}{z_{t-1}}} \frac{z_t}{z_{t-1}}}{\Lambda_{y^d}} \right)^{\gamma_y} \right)^{1-\gamma_R} \exp\left(m_t\right)$$

Markets clear:

$$\frac{y_t^d}{z_t} = \frac{\mu_{t-1}^{\alpha} z_{t-1}^{\alpha} \frac{A_t}{z_t} \left(u_t \frac{k_{t-1}}{z_{t-1}\mu_{t-1}}\right)^{\alpha} \left(l_t^d\right)^{1-\alpha} - f}{v_t^p}$$

but since $\mu_{t-1}^{\alpha} z_{t-1}^{\alpha} = \frac{z_{t-1}}{A_{t-1}}$, we have

$$\frac{y_t^d}{z_t} = \frac{\frac{z_{t-1}}{A_{t-1}} \frac{A_t}{z_t} \left(u_t \frac{k_{t-1}}{z_{t-1}\mu_{t-1}}\right)^{\alpha} \left(l_t^d\right)^{1-\alpha} - f}{v_t^p}$$

$$\frac{y_t^d}{z_t} = \frac{c_t}{z_t} + \frac{x_t}{z_t} + \widetilde{g}_t + \frac{z_{t-1}}{z_t} \frac{\mu_{t-1}}{\mu_t} \Phi\left[u_t\right] \frac{k_{t-1}}{z_{t-1}\mu_{t-1}}$$

where

$$v_t^p = \theta_p \left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{-\varepsilon} v_{t-1}^p + \left(1-\theta_p\right) \Pi_t^{*-\varepsilon}$$

and

$$l_t = l_t^d$$

$$\frac{k_t}{z_t\mu_t} \frac{z_t\mu_t}{z_{t-1}\mu_{t-1}} - (1-\delta) \frac{k_{t-1}}{z_{t-1}\mu_{t-1}} - \frac{\mu_t}{\mu_{t-1}} \frac{z_t}{z_{t-1}} \left(1 - S\left[\frac{\frac{x_t}{z_t} z_t}{\frac{x_{t-1}}{z_{t-1}} z_{t-1}}\right]\right) \frac{x_t}{z_t} = 0.$$

Change of Variables

We now redefine variables to obtain a system on stationary variables that we can easily manipulate. Hence, we define $\widetilde{c}_t = \frac{c_t}{z_t}$, $\widetilde{\lambda}_t = \lambda_t z_t$, $\widetilde{r}_t = r_t \mu_t$, $\widetilde{q}_t = q_t \mu_t$, $\widetilde{x}_t = \frac{x_t}{z_t}$, $\widetilde{w}_t = \frac{w_t}{z_t}$, $\widetilde{k}_t = \frac{k_t}{z_t \mu_t}$, and $\widetilde{y}_t^d = \frac{y_t^d}{z_t}$.

Then, the set of equilibrium conditions are:

- The first order conditions of the household:

$$d_t\left(\widetilde{c}_t - h\widetilde{c}_{t-1}\frac{z_{t-1}}{z_t}\right)^{-1} - h\beta\mathbb{E}_t d_{t+1}\left(\widetilde{c}_{t+1}\frac{z_{t+1}}{z_t} - h\widetilde{c}_t\right)^{-1} = \widetilde{\lambda}_t$$

$$d_t\phi_t\psi(l_t)^{\vartheta} = \widetilde{w}_t\widetilde{\lambda}_t$$

$$\widetilde{\lambda}_t = \beta\mathbb{E}_t\{\widetilde{\lambda}_{t+1}\frac{z_t}{z_{t+1}}\frac{R_t}{\Pi_{t+1}}\}$$

$$\widetilde{r}_t = \Phi'\left[u_t\right]$$

$$\widetilde{q}_t = \beta\mathbb{E}_t\left\{\frac{\widetilde{\lambda}_{t+1}}{\widetilde{\lambda}_t}\frac{z_t}{z_{t+1}}\frac{\mu_t}{\mu_{t+1}}\left(\left(1-\delta\right)\widetilde{q}_{t+1} + \widetilde{r}_{t+1}u_{t+1} - \Phi\left(u_{t+1}\right)\right)\right\}$$

$$1 = \widetilde{q}_t\left(1 - S\left[\frac{\widetilde{x}_t}{\widetilde{x}_{t-1}}\frac{z_t}{z_{t-1}}\right] - S'\left[\frac{\widetilde{x}_t}{\widetilde{x}_{t-1}}\frac{z_t}{z_{t-1}}\right]\frac{\widetilde{x}_t}{\widetilde{x}_{t-1}}\frac{z_t}{z_{t-1}}\right)$$
$$+\beta\mathbb{E}_t\widetilde{q}_{t+1}\frac{\widetilde{\lambda}_{t+1}}{\widetilde{\lambda}_t}\frac{z_t}{z_{t+1}}S'\left[\frac{\widetilde{x}_{t+1}}{\widetilde{x}_t}\frac{z_{t+1}}{z_t}\right]\left(\frac{\widetilde{x}_{t+1}}{\widetilde{x}_t}\frac{z_{t+1}}{z_t}\right)^2$$

- The firms that can change prices set them to satisfy:

$$g_t^1 = \widetilde{\lambda}_t mc_t\widetilde{y}_t^d + \beta\theta_p\mathbb{E}_t\left(\frac{\Pi_t^{\chi}}{\Pi_{t+1}}\right)^{-\varepsilon}g_{t+1}^1$$

$$g_t^2 = \widetilde{\lambda}_t\Pi_t^*\widetilde{y}_t^d + \beta\theta_p\mathbb{E}_t\left(\frac{\Pi_t^{\chi}}{\Pi_{t+1}}\right)^{1-\varepsilon}\left(\frac{\Pi_t^*}{\Pi_{t+1}^*}\right)g_{t+1}^2$$

$$\varepsilon g_t^1 = (\varepsilon-1)g_t^2$$

where they rent inputs to satisfy their static minimization problem:

$$\frac{u_t\widetilde{k}_{t-1}}{l_t} = \frac{\alpha}{1-\alpha}\frac{\widetilde{w}_t}{\widetilde{r}_t}\frac{z_t}{z_{t-1}}\frac{\mu_t}{\mu_{t-1}}$$

$$mc_t = \left(\frac{1}{1-\alpha}\right)^{1-\alpha}\left(\frac{1}{\alpha}\right)^{\alpha}\left(\widetilde{w}_t\right)^{1-\alpha}\widetilde{r}_t^{\alpha}$$

- The price level evolves as:

$$1 = \theta_p\left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{1-\varepsilon} + \left(1-\theta_p\right)\Pi_t^{*1-\varepsilon}$$

- Government follows its Taylor rule:

$$\frac{R_t}{R} = \left(\frac{R_{t-1}}{R}\right)^{\gamma_R} \left(\left(\frac{\Pi_t}{\Pi}\right)^{\gamma_\Pi} \left(\frac{\frac{\widetilde{y}_t^d}{\widetilde{y}_{t-1}^d}\frac{z_t}{z_{t-1}}}{\Lambda_{y^d}}\right)^{\gamma_y}\right)^{1-\gamma_R} \exp\left(m_t\right)$$

- Markets clear:

$$\begin{aligned}
\widetilde{y}_t^d &= \widetilde{c}_t + \widetilde{x}_t + \widetilde{g}_t + \frac{z_{t-1}}{z_t}\frac{\mu_{t-1}}{\mu_t}\Phi\left[u_t\right]\widetilde{k}_{t-1} \\
\widetilde{y}_t^d &= \frac{\frac{A_t}{A_{t-1}}\frac{z_{t-1}}{z_t}\left(u_t\widetilde{k}_{t-1}\right)^{\alpha}\left(l_t^d\right)^{1-\alpha} - f}{v_t^p}
\end{aligned}$$

where

$$v_t^p = \theta_p\left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{-\varepsilon} v_{t-1}^p + \left(1-\theta_p\right)\Pi_t^{*-\varepsilon}$$

and

$$\begin{aligned}
l_t &= l_t^d \\
\widetilde{k}_t\frac{z_t}{z_{t-1}}\frac{\mu_t}{\mu_{t-1}} - \left(1-\delta\right)\widetilde{k}_{t-1} - \frac{z_t}{z_{t-1}}\frac{\mu_t}{\mu_{t-1}}\left(1 - S\left[\frac{\widetilde{x}_t}{\widetilde{x}_{t-1}}\frac{z_t}{z_{t-1}}\right]\right)\widetilde{x}_t &= 0.
\end{aligned}$$

A.3 The Steady-State

Now, we will find the deterministic steady-state of the model. First, let $\widetilde{z} = \exp\left(\Lambda_z\right)$, $\widetilde{\mu} = \exp\left(\Lambda_\mu\right)$, and $\widetilde{A} = \exp\left(\Lambda_A\right)$. Also, given the definition of $\widetilde{c}$, $\widetilde{x}_t$, $\widetilde{w}_t$, and $\widetilde{y}_t^d$, we have the same growth rate: $\Lambda_c = \Lambda_x = \Lambda_w = \Lambda_{y^d} = \Lambda_z$.

Then, in steady-state, the first order conditions of the household can be written as:

$$\frac{1}{\widetilde{c}-\frac{h}{\widetilde{z}}\widetilde{c}}-h\beta\frac{1}{\widetilde{c}\widetilde{z}-h\widetilde{c}}=\widetilde{\lambda}$$
$$\psi(l)^{\vartheta}=\widetilde{w}\widetilde{\lambda}$$
$$1=\beta\frac{1}{\widetilde{z}}\frac{R}{\Pi}$$
$$\widetilde{r}=\Phi'\left[1\right]$$
$$\widetilde{q}=\beta\frac{1}{\widetilde{z}\widetilde{\mu}}\left(\left(1-\delta\right)\widetilde{q}+\widetilde{r}u-\Phi\left[1\right]\right)$$
$$1=\widetilde{q}\left(1-S\left[\widetilde{z}\right]-S'\left[\widetilde{z}\right]\widetilde{z}\right)+\beta\frac{\widetilde{q}}{\widetilde{z}}S'\left[\widetilde{z}\right]\widetilde{z}^{2}$$

the first order conditions of the firm as:

$$g^{1}=\widetilde{\lambda}mc\widetilde{y}^{d}+\beta\theta_{p}\left(\frac{\Pi^{\chi}}{\Pi}\right)^{-\varepsilon}g^{1}$$
$$g^{2}=\widetilde{\lambda}\Pi^{*}\widetilde{y}^{d}+\beta\theta_{p}\left(\frac{\Pi^{\chi}}{\Pi}\right)^{1-\varepsilon}g^{2}$$
$$\varepsilon g^{1}=(\varepsilon-1)g^{2}$$
$$\frac{u\widetilde{k}}{l^{d}}=\frac{\alpha}{1-\alpha}\frac{\widetilde{w}}{\widetilde{r}}\widetilde{z}\widetilde{\mu}$$
$$mc=\left(\frac{1}{1-\alpha}\right)^{1-\alpha}\left(\frac{1}{\alpha}\right)^{\alpha}\widetilde{w}^{1-\alpha}\widetilde{r}^{\alpha}$$

the law of motion for prices as:

$$1=\theta_{p}\left(\frac{\Pi^{\chi}}{\Pi}\right)^{1-\varepsilon}+\left(1-\theta_{p}\right)\Pi^{*1-\varepsilon}$$

and the market clearing conditions as:

$$\widetilde{c}+\widetilde{x}+\widetilde{g}=\widetilde{y}^{d}$$
$$v^{p}\widetilde{y}^{d}=\frac{\widetilde{A}}{\widetilde{z}}(u\widetilde{k})^{\alpha}(l^{d})^{1-\alpha}-f$$
$$l=l^{d}$$
$$v^{p}=\theta_{p}\left(\frac{\Pi^{\chi}}{\Pi}\right)^{-\varepsilon}v^{p}+\left(1-\theta_{p}\right)\Pi^{*-\varepsilon}$$
$$\widetilde{k}\widetilde{z}\widetilde{\mu}-\left(1-\delta\right)\widetilde{k}-\widetilde{z}\widetilde{\mu}\left(1-S\left[z\right]\right)\widetilde{x}=0.$$

To find the steady-state, we need to choose functional forms for $\Phi[\cdot]$ and $S[\cdot]$. For $\Phi[u]$ we pick: $\Phi[u] = \Phi_1(u-1) + \frac{\Phi_2}{2}(u-1)^2$. Since in the steady state we have $u = 1$, then $\widetilde{r} = \Phi'[1] = \Phi_1$ and $\Phi[1] = 0$. The investment adjustment cost function is $S\left[\frac{x_t}{x_{t-1}}\right] = \frac{\kappa}{2}(\frac{x_t}{x_{t-1}} - \Lambda_x)^2$. Then, along the balanced growth path, $S[z] = S'[\Lambda_x] = 0$. Using this two expressions, we can rearrange the system of equations that determine the steady-state as:

$$(1 - h\beta/\widetilde{z})\frac{1}{1-\frac{h}{\widetilde{z}}}\frac{1}{\widetilde{c}} = \widetilde{\lambda}$$

$$\psi(l)^{\vartheta} = \widetilde{w}\widetilde{\lambda}$$

$$R = \frac{\Pi\widetilde{z}}{\beta}$$

$$\widetilde{r} = \Phi_1$$

$$\widetilde{r} = \left(1 - \frac{\beta}{\widetilde{z}\widetilde{\mu}}(1-\delta)\right)/\frac{\beta}{\widetilde{z}\widetilde{\mu}}$$

$$\left(1 - \beta\theta_p\Pi^{(1-\chi)\varepsilon}\right)g^1 = \widetilde{\lambda}mc\widetilde{y}^d$$

$$\left(1 - \beta\theta_p\Pi^{-(1-\chi)(1-\varepsilon)}\right)g^2 = \widetilde{\lambda}\Pi^*\widetilde{y}^d$$

$$\varepsilon g^1 = (\varepsilon - 1)g^2$$

$$\frac{\widetilde{k}}{l^d} = \frac{\alpha}{1-\alpha}\frac{\widetilde{w}}{\widetilde{r}}\widetilde{z}\widetilde{\mu}$$

$$mc = \left(\frac{1}{1-\alpha}\right)^{1-\alpha}\left(\frac{1}{\alpha}\right)^{\alpha}\widetilde{w}^{1-\alpha}\widetilde{r}^{\alpha}$$

$$\frac{1-\theta_p\Pi^{-(1-\chi)(1-\varepsilon)}}{1-\theta_p} = \Pi^{*1-\varepsilon}$$

$$\widetilde{c} + \widetilde{x} + \widetilde{g} = \widetilde{y}^d$$

$$v^p\widetilde{y}^d = \frac{\widetilde{A}}{\widetilde{z}}(\widetilde{k})^{\alpha}(l^d)^{1-\alpha} - f$$

$$l = l^d$$

$$\frac{1-\theta_p\Pi^{(1-\chi)\varepsilon}}{1-\theta_p}v^p = \Pi^{*-\varepsilon}$$

$$\widetilde{k} = \frac{\widetilde{z}\widetilde{\mu}}{\widetilde{z}\widetilde{\mu} - (1-\delta)}\widetilde{x}.$$

First, notice that there is some restrictions on Φ_1

$$\tilde{r} = \frac{1 - \frac{\beta}{\widetilde{z}\widetilde{\mu}}(1-\delta)}{\frac{\beta}{\widetilde{z}\widetilde{\mu}}} = \Phi_1$$

and that the nominal interest rate is:

$$R = \frac{\Pi \widetilde{z}}{\beta}$$

The relationship between inflation and optimal relative prices is:

$$\Pi^* = \left(\frac{1 - \theta_p \Pi^{-(1-\varepsilon)(1-\chi)}}{1 - \theta_p} \right)^{\frac{1}{1-\varepsilon}}$$

expression from which we obtain the following two results:

1. If there is zero price inflation ($\Pi = 1$), then $\Pi^* = 1$.
2. If there is full price indexation ($\chi = 1$), then $\Pi^* = 1$.

The expression for the dispersion of prices is given by:

$$v^p = \frac{1 - \theta_p}{1 - \theta_p \Pi^{(1-\chi)\varepsilon}} \Pi^{*-\varepsilon}$$

where no price inflation or full price indexation delivers no price dispersion in steady-state.

We can calibrate the parameter ψ such that in steady state $l = 1$ (the utility function we selected does not have natural units for labor, so $l = 1$ is just a normalization that picks 1 unit of l_t to be equal to whatever number of hours we want). Labor demand must equal labor supply:

$$l^d = 1$$

Once we have l^d, we can solve for wages, capital, investment, output, and consumption. Note that wages are given by:

$$\frac{\left(1 - \beta \theta_p \Pi^{(1-\chi)\varepsilon}\right)}{\left(1 - \beta \theta_p \Pi^{-(1-\chi)(1-\varepsilon)}\right)} \frac{\varepsilon - 1}{\varepsilon} \Pi^* = mc$$

$$mc = \left(\frac{1}{1-\alpha} \right)^{1-\alpha} \left(\frac{1}{\alpha} \right)^{\alpha} \widetilde{w}^{1-\alpha} \widetilde{r}^{\alpha}$$

$$\widetilde{w} = (1-\alpha) \left(mc \left(\frac{\alpha}{\widetilde{r}} \right)^{\alpha} \right)^{\frac{1}{1-\alpha}}$$

That allows us to find $\widetilde{k}$:

$$\widetilde{k} = \frac{\alpha}{1-\alpha} \frac{\widetilde{w}}{\widetilde{r}} \widetilde{z} \widetilde{\mu}$$

We also have:

$$\widetilde{y}^d = \frac{\frac{\widetilde{A}}{\widetilde{z}} (\widetilde{k})^{\alpha} (l^d)^{1-\alpha} - f}{v^p}.$$

But, since in steady-state $\widetilde{k} = \frac{\widetilde{z}\widetilde{\mu}}{\widetilde{z}\widetilde{\mu}-(1-\delta)}\widetilde{x}$, by which we can find $\widetilde{x}$, it is the case that:

$$\widetilde{c} + \frac{\widetilde{z}\widetilde{\mu}-(1-\delta)}{\widetilde{z}\widetilde{\mu}}\widetilde{k} + g^{ss}\widetilde{y}^d = \widetilde{y}^d = \frac{\frac{\widetilde{A}}{\widetilde{z}}(\widetilde{k})^{\alpha} - f}{v^p}$$

where we conveniently define g^{ss} as the steady-state government expenditure to GDP ratio: $g^{ss} \equiv \frac{\widetilde{g}}{\widetilde{y}^d}$. Then finally:

$$\widetilde{c} \;=\; \frac{\frac{\widetilde{A}}{\widetilde{z}}\widetilde{k}^{\alpha} - f}{v^p} - \frac{\widetilde{z}\widetilde{\mu}-(1-\delta)}{\widetilde{z}\widetilde{\mu}}\widetilde{k} - g^{ss}\widetilde{y}^d$$

A.4 Loglinear approximations

For each variable var_t, we define $\widehat{var}_t = \log var_t - \log var$, where var is the steady-state value for the variable var_t. Then, we can write $var_t = var \exp^{\widehat{var}_t}$.

We start by log linearizing the marginal utility of consumption:

$$d_t \left(\widetilde{c}_t - h\widetilde{c}_{t-1}\frac{z_{t-1}}{z_t}\right)^{-1} - h\beta\mathbb{E}_t d_{t+1}\left(\widetilde{c}_{t+1}\frac{z_{t+1}}{z_t} - h\widetilde{c}_t\right)^{-1} = \widetilde{\lambda}_t. \tag{A.2}$$

It is helpful to define the auxiliary variable $aux_t = d_t\left(\widetilde{c}_t - h\widetilde{c}_{t-1}\frac{z_{t-1}}{z_t}\right)^{-1}$. Then, we have that:

$$aux_t - h\beta\mathbb{E}_t\frac{1}{\widetilde{z}_{t+1}}aux_{t+1} = \widetilde{\lambda}_t$$

where $\widetilde{z}_{t+1} = \frac{z_{t+1}}{z_t}$.

Then, (A.2) can be written as:

$$aux\exp^{\widehat{aux}_t} - \frac{h\beta}{z}aux\mathbb{E}_t\exp^{\widehat{aux}_{t+1}-\widehat{\widetilde{z}}_{t+1}} = \widetilde{\lambda}e^{\widehat{\widetilde{\lambda}}_t}$$

which can be loglinearized as:

$$aux\left(\widehat{aux}_t - \frac{h\beta}{z}\mathbb{E}_t(\widehat{aux}_{t+1} - \widehat{\widetilde{z}}_{t+1})\right) = \widetilde{\lambda}\widehat{\widetilde{\lambda}}_t. \tag{A.3}$$

Using the following two steady state relationship that $aux(1-\frac{h\beta}{\widetilde{z}}) = \widetilde{\lambda}$ and that $\mathbb{E}_t\widehat{\widetilde{z}}_{t+1} = 0$, we can write:

$$\widehat{aux}_t - \frac{h\beta}{\widetilde{z}}\mathbb{E}_t\widehat{aux}_{t+1} = \left(1-\frac{h\beta}{\widetilde{z}}\right)\widehat{\widetilde{\lambda}}_t.$$

Next, we loglinearize the auxiliary variable:

$$aux e^{\widehat{aux}_t} = d \exp^{\widehat{d}_t} \left(\widetilde{c} \exp^{\widehat{\widetilde{c}}_t} - \frac{h}{\widetilde{z}} \widetilde{c} \exp^{\widehat{\widetilde{c}}_{t-1} - \widehat{\widetilde{z}}_t} \right)^{-1}.$$

The loglinear approximation is:

$$\begin{aligned} aux\widehat{aux}_t &= d \left(\widetilde{c} - \frac{h}{\widetilde{z}} \widetilde{c} \right)^{-1} \widehat{d}_t - d \left(\widetilde{c} - \frac{h}{\widetilde{z}} \widetilde{c} \right)^{-2} \widehat{\widetilde{c}}_t \\ &\quad + d \left(\widetilde{c} - \frac{h}{\widetilde{z}} \widetilde{c} \right)^{-2} \frac{h \widetilde{c}}{\widetilde{z}} \left(\widehat{\widetilde{c}}_{t-1} - \widehat{\widetilde{z}}_t \right). \end{aligned}$$

Making use of the fact that $aux = d \left(\widetilde{c} - \frac{h}{\widetilde{z}} \widetilde{c} \right)^{-1}$, we find:

$$aux\widehat{aux}_t = aux \left(\widehat{d}_t - \left(\widetilde{c} - \frac{h}{\widetilde{z}} \widetilde{c} \right)^{-1} \widehat{\widetilde{c}}_t + \left(\widetilde{c} - \frac{h}{\widetilde{z}} \widetilde{c} \right)^{-1} \frac{h \widetilde{c}}{\widetilde{z}} \left(\widehat{\widetilde{c}}_{t-1} - \widehat{\widetilde{z}}_t \right) \right),$$

that simplifies to:

$$aux\widehat{aux}_t = aux \left(\widehat{d}_t - \left(1 - \frac{h}{\widetilde{z}} \right)^{-1} \widehat{\widetilde{c}}_t + \left(1 - \frac{h}{\widetilde{z}} \right)^{-1} \frac{h}{\widetilde{z}} \left(\widehat{\widetilde{c}}_{t-1} - \widehat{\widetilde{z}}_t \right) \right)$$

$$\widehat{aux}_t = \widehat{d}_t - \left(1 - \frac{h}{\widetilde{z}} \right)^{-1} \left(\widehat{\widetilde{c}}_t - \frac{h \widehat{\widetilde{c}}_{t-1}}{\widetilde{z}} + \frac{h \widehat{\widetilde{z}}_t}{\widetilde{z}} \right). \tag{A.4}$$

Putting the (A.3) and (A.4) together:

$$\left(1 - \frac{h\beta}{\widetilde{z}} \right) \widehat{\widetilde{\lambda}}_t =$$

$$\widehat{d}_t - \left(1 - \frac{h}{\widetilde{z}} \right)^{-1} \left(\widehat{\widetilde{c}}_t - \frac{h \widehat{\widetilde{c}}_{t-1}}{\widetilde{z}} + \frac{h \widehat{\widetilde{z}}_t}{\widetilde{z}} \right) - \frac{h\beta}{\widetilde{z}} \mathbb{E}_t \left\{ \widehat{d}_{t+1} - \left(1 - \frac{h}{\widetilde{z}} \right)^{-1} \left(\widehat{\widetilde{c}}_{t+1} - \frac{h \widehat{\widetilde{c}}_t}{\widetilde{z}} + \frac{h \widehat{\widetilde{z}}_{t+1}}{\widetilde{z}} \right) \right\}.$$

After some algebra, we arrive to the final expression:

$$(1 - \frac{h\beta}{\widetilde{z}}) \widehat{\widetilde{\lambda}}_t = \widehat{d}_t - \frac{h\beta}{\widetilde{z}} \mathbb{E}_t \widehat{d}_{t+1} - \frac{1 + h^2/z^2 \beta}{\left(1 - \frac{h}{\widetilde{z}} \right)} \widehat{\widetilde{c}}_t + \frac{h}{\widetilde{z} \left(1 - \frac{h}{\widetilde{z}} \right)} \widehat{\widetilde{c}}_{t-1} + \frac{\beta h / z}{\left(1 - \frac{h}{\widetilde{z}} \right)} \mathbb{E}_t \widehat{\widetilde{c}}_{t+1} - \frac{h}{\widetilde{z} \left(1 - \frac{h}{\widetilde{z}} \right)} \widehat{\widetilde{z}}_t$$

This expression helps to understand the role of the habit persistence parameter h. If we set

$h = 0$ (i.e., no habit), we would get:

$$\widehat{\widetilde{\lambda}}_t = \widehat{d}_t - \widehat{\widetilde{c}}_t$$

where the lags and forward terms drop.

Now, we loglinearize the Euler equation:

$$\widetilde{\lambda}_t = \beta \mathbb{E}_t \{ \widetilde{\lambda}_{t+1} \frac{z_t}{z_{t+1}} \frac{R_t}{\Pi_{t+1}} \}.$$

To do so, we write the expression as

$$\widetilde{\lambda} e^{\widehat{\widetilde{\lambda}}} = \beta \mathbb{E}_t \{ \widetilde{\lambda} e^{\widehat{\widetilde{\lambda}}_{t+1}} \frac{1}{\widetilde{z} e^{z_{z,t}}} \frac{R e^{\widehat{R}_t}}{\Pi e^{\widehat{\Pi}_{t+1}}} \}$$

By using the fact that

$$R = \frac{\Pi \widetilde{z}}{\beta}$$

we simplify to:

$$e^{\widehat{\widetilde{\lambda}}} = \mathbb{E}_t \{ e^{\widehat{\widetilde{\lambda}}_{t+1}} \frac{1}{e^{z_{z,t}}} \frac{e^{\widehat{R}_t}}{e^{\widehat{\Pi}_{t+1}}} \}$$

Now, it is easy to show that:

$$\widehat{\widetilde{\lambda}}_t = \mathbb{E}_t \{ \widehat{\widetilde{\lambda}}_{t+1} + \widehat{R}_t - \widehat{\Pi}_{t+1} \}. \tag{A.5}$$

Next, we log-linearize the static consumption/labor optimal condition

$$d_t \phi_t \psi (l_t)^{\vartheta} = \widetilde{w}_t \widetilde{\lambda}_t$$

Log-linearization delivers

$$\widehat{d}_t + \widehat{\phi}_t + \vartheta \widehat{l}_t = \widehat{\widetilde{w}}_t + \widehat{\widetilde{\lambda}}_t \tag{A.6}$$

Let us now consider:

$$\widetilde{r}_t = \Phi' \left[u_t \right].$$

First, we write:

$$\widetilde{r} e^{\widehat{\widetilde{r}}_t} = \Phi' \left[u \exp^{\widehat{u}_t} \right],$$

where the loglinear approximation delivers:

$$\widetilde{r} \widehat{\widetilde{r}}_t = \Phi'' \left[u \right] u \widehat{u}_t.$$

Since $\widetilde{r} = \Phi'[u]$, then:

$$\widehat{\widetilde{r}}_t = \frac{\Phi''[u]\,u}{\Phi'[u]}\widehat{u}_t,$$

or

$$\widehat{\widetilde{r}}_t = \frac{\Phi_2}{\Phi_1}\hat{u}_t. \tag{A.7}$$

The next equation to consider relates the shadow price of capital to the return on investment:

$$\widetilde{q}_t = \beta\mathbb{E}_t\left\{\frac{\widetilde{\lambda}_{t+1}}{\widetilde{\lambda}_t}\frac{1}{\widetilde{z}_{t+1}}\frac{1}{\widetilde{\mu}_{t+1}}\left(\left(1-\delta\right)\widetilde{q}_{t+1} + \widetilde{r}_{t+1}u_{t+1} - \Phi\left[u_{t+1}\right]\right)\right\}$$

where $\widetilde{\mu}_{t+1} = \frac{\mu_{t+1}}{\mu_t}$. We can we write this expression as

$$\widetilde{q}e^{\widehat{\widetilde{q}}_t} = \frac{\beta}{\widetilde{z}\widetilde{\mu}}\mathbb{E}_t \exp^{\Delta\widehat{\widetilde{\lambda}}_{t+1}-\widehat{\widetilde{z}}_{t+1}-\widehat{\widetilde{\mu}}_{t+1}}\left\{\begin{array}{c}(1-\delta)\,\widetilde{q}e^{\widehat{\widetilde{q}}_{t+1}}+\\ \widetilde{r}u\exp^{\widehat{\widetilde{r}}_{t+1}+\widehat{u}_{t+1}} - \Phi\left[u\exp^{\widehat{u}_{t+1}}\right]\end{array}\right\}.$$

Loglinearization delivers:

$$\begin{aligned}\widetilde{q}\widehat{\widetilde{q}}_t &= \frac{\beta}{\widetilde{z}\widetilde{\mu}}\mathbb{E}_t\left(\triangle\widehat{\widetilde{\lambda}}_{t+1} - \widehat{\widetilde{z}}_{t+1} - \widehat{\widetilde{\mu}}_{t+1}\right)\left(\left(1-\delta\right)\widetilde{q} + \widetilde{r}u - \Phi\left[u\right]\right)\\ &\quad + \frac{\beta}{\widetilde{z}\widetilde{\mu}}\left(\mathbb{E}_t\left(1-\delta\right)\widetilde{q}\widehat{\widetilde{q}}_{t+1} + \widetilde{r}u\left(\widehat{\widetilde{r}}_{t+1} + \widehat{u}_{t+1}\right) - \Phi\left[u\right]u\widehat{u}_{t+1}\right).\end{aligned}$$

Making use of the following steady-state relationships: $u = 1$, $\Phi[u] = 0$, $\widetilde{r} = \Phi'[u]$, $\widetilde{q} = 1$ and $1 = \frac{\beta}{\widetilde{z}\widetilde{\mu}}\widetilde{r} + \frac{\beta}{\widetilde{z}\widetilde{\mu}}(1-\delta)$, and $\mathbb{E}_t\left(-\widehat{\widetilde{z}}_{t+1} - \widehat{\widetilde{\mu}}_{t+1}\right) = 0$, the previous expression simplifies to:

$$\widehat{\widetilde{q}}_t = \left(\left(1-\delta\right) + \widetilde{r}\right)\frac{\beta}{\widetilde{z}\widetilde{\mu}}\mathbb{E}_t\Delta\widehat{\widetilde{\lambda}}_{t+1} + \frac{\beta\left(1-\delta\right)}{\widetilde{z}\widetilde{\mu}}\mathbb{E}_t\widehat{\widetilde{q}}_{t+1} + \frac{\beta}{\widetilde{z}\widetilde{\mu}}\widetilde{r}u\mathbb{E}_t\widehat{\widetilde{r}}_{t+1},$$

that implies:

$$\widehat{\widetilde{q}}_t = \mathbb{E}_t\Delta\widehat{\widetilde{\lambda}}_{t+1} + \frac{\beta\left(1-\delta\right)}{\widetilde{z}\widetilde{\mu}}\mathbb{E}_t\widehat{\widetilde{q}}_{t+1} + \left(1 - \frac{\beta(1-\delta)}{\widetilde{z}\widetilde{\mu}}\right)\mathbb{E}_t\widehat{\widetilde{r}}_{t+1}. \tag{A.8}$$

The next equation to loglinearize is:

$$1 = \widetilde{q}_t\left(1 - S\left[\frac{\widetilde{x}_t}{\widetilde{x}_{t-1}}\frac{z_t}{z_{t-1}}\right] - S'\left[\frac{\widetilde{x}_t}{\widetilde{x}_{t-1}}\frac{z_t}{z_{t-1}}\right]\frac{\widetilde{x}_t}{\widetilde{x}_{t-1}}\frac{z_t}{z_{t-1}}\right) + \beta\mathbb{E}_t\widetilde{q}_{t+1}\frac{\widetilde{\lambda}_{t+1}}{\widetilde{\lambda}_t}\frac{z_t}{z_{t+1}}S'\left[\frac{\widetilde{x}_{t+1}}{\widetilde{x}_t}\frac{z_{t+1}}{z_t}\right]\left(\frac{\widetilde{x}_{t+1}}{\widetilde{x}_t}\frac{z_{t+1}}{z_t}\right)^2$$

which can be rearranged as:

$$1 = \widetilde{q}\exp^{\widehat{\widetilde{q}}_t}\left(1 - S\left[\widetilde{z}\exp^{\Delta\widehat{\widetilde{x}}_t+\widehat{\widetilde{z}}_t}\right] - S'\left[\widetilde{z}\exp^{\Delta\widehat{\widetilde{x}}_t+\widehat{\widetilde{z}}_t}\right]\widetilde{z}\exp^{\Delta\widehat{\widetilde{x}}_t+\widehat{\widetilde{z}}_t}\right) +$$
$$+\beta\frac{\widetilde{q}}{\widetilde{z}}\mathbb{E}_t\exp^{\widehat{\widetilde{q}}_t-\widehat{\widetilde{z}}_{t+1}+\triangle\widehat{\widetilde{\lambda}}_{t+1}}S''\left[\widetilde{z}\exp^{\Delta\widehat{\widetilde{x}}_{t+1}+\widehat{\widetilde{z}}_{t+1}}\right]\widetilde{z}^2\exp^{2\left(\Delta\widehat{\widetilde{x}}_{t+1}+\widehat{\widetilde{z}}_{t+1}\right)}.$$

Taking the loglinear approximation (and using the fact that $\widetilde{q}=1$) we get:

$$0 = \widehat{\widetilde{q}}_t - S''\left[\widetilde{z}\right]\widetilde{z}^2\left(\Delta\widehat{\widetilde{x}}_t + \widehat{\widetilde{z}}_t\right) + \frac{\beta}{z}S''\left[\widetilde{z}\right]\widetilde{z}^3\mathbb{E}_t\left(\Delta\widehat{\widetilde{x}}_{t+1} + \widehat{\widetilde{z}}_{t+1}\right).$$

Reorganizing:

$$\kappa\widetilde{z}^2\left(\Delta\widehat{\widetilde{x}}_t + \widehat{\widetilde{z}}_t\right) = \widehat{\widetilde{q}}_t + \beta\kappa\widetilde{z}^2\mathbb{E}_t\Delta\widehat{\widetilde{x}}_{t+1} \tag{A.9}$$

where κ comes from the adjustment cost function.

Let us loglinearize the law of motion for g_t^1 and g_t^2. First consider

$$g_t^1 = \widetilde{\lambda}_t mc_t\widetilde{y}_t^d + \beta\theta_p\mathbb{E}_t\left(\frac{\Pi_t^\chi}{\Pi_{t+1}}\right)^{-\varepsilon}g_{t+1}^1$$

that can be rewritten as:

$$g^1\exp^{\widehat{g}_t^1} = \widetilde{\lambda}mc\widetilde{y}^d\exp^{\widehat{\widetilde{\lambda}}_t+\widehat{mc}_t+\widehat{\widetilde{y}}_t^d} + \beta\theta_p g^1\Pi^{\varepsilon(1-\chi)}\mathbb{E}_t\exp^{\varepsilon(\widehat{\Pi}_{t+1}-\chi\widehat{\Pi}_t)+\widehat{g}_{t+1}^1}.$$

If we loglinearize that last expression, we get:

$$g^1\widehat{g}_t^1 = \widetilde{\lambda}mc\widetilde{y}^d\left(\widehat{\widetilde{\lambda}}_t + \widehat{mc}_t + \widehat{\widetilde{y}}_t^d\right) + \beta\theta_p g^1\Pi^{\varepsilon(1-\chi)}\mathbb{E}_t\left(\varepsilon(\widehat{\Pi}_{t+1}-\chi\widehat{\Pi}_t) + \widehat{g}_{t+1}^1\right).$$

It can be shown that in steady state

$$1 - \beta\theta_p\Pi^{\varepsilon(1-\chi)} = \frac{\widetilde{\lambda}mc\widetilde{y}^d}{g^1}$$

therefore,

$$\widehat{g}_t^1 = \left(1-\beta\theta_p\Pi^{\varepsilon(1-\chi)}\right)\left(\widehat{\widetilde{\lambda}}_t + \widehat{mc}_t + \widehat{\widetilde{y}}_t^d\right) + \beta\theta_p\Pi_t^{\varepsilon(1-\chi)}\mathbb{E}_t\left(\varepsilon(\widehat{\Pi}_{t+1}-\chi\widehat{\Pi}_t) + \widehat{g}_{t+1}^1\right). \tag{A.10}$$

Let us now consider:

$$g_t^2 = \widetilde{\lambda}_t \Pi_t^* \widetilde{y}_t^d + \beta\theta_p \mathbb{E}_t \left(\frac{\Pi_t^\chi}{\Pi_{t+1}}\right)^{1-\varepsilon} \left(\frac{\Pi_t^*}{\Pi_{t+1}^*}\right) g_{t+1}^2,$$

that can rewritten as:

$$g^2 \exp^{\widehat{g}_t^2} = \widetilde{\lambda}\Pi^* \widetilde{y}^d \exp^{\widehat{\widetilde{\lambda}}_t + \widehat{\Pi}_t^* + \widehat{\widetilde{y}}_t^d} + \beta\theta_p \Pi^{-(1-\varepsilon)(1-\chi)} g^2 \mathbb{E}_t \exp^{-(1-\varepsilon)\left(\widehat{\Pi}_{t+1} - \chi\widehat{\Pi}_t\right) - \left(\widehat{\Pi}_{t+1}^* - \widehat{\Pi}_t^*\right) + \widehat{g}_{t+1}^2}.$$

If we loglinearize that last expression, we get:

$$\begin{aligned} g^2 \widehat{g}_t^2 &= \widetilde{\lambda}\Pi^* \widetilde{y}^d \left(\widehat{\widetilde{\lambda}}_t + \widehat{\Pi}_t^* + \widehat{\widetilde{y}}_t^d\right) \\ &\quad + \beta\theta_p \Pi^{-(1-\varepsilon)(1-\chi)} g^2 \mathbb{E}_t \left(-(1-\varepsilon)\left(\widehat{\Pi}_{t+1} - \chi\widehat{\Pi}_t\right) - \left(\widehat{\Pi}_{t+1}^* - \widehat{\Pi}_t^*\right) + \widehat{g}_{t+1}^2\right). \end{aligned}$$

It can be shown that:

$$1 - \beta\theta_p \Pi^{-(1-\varepsilon)(1-\chi)} = \frac{\widetilde{\lambda}\Pi^* \widetilde{y}^d}{g^2},$$

therefore:

$$\begin{aligned} \widehat{g}_t^2 &= \left(1 - \beta\theta_p \Pi^{-(1-\varepsilon)(1-\chi)}\right) \left(\widehat{\widetilde{\lambda}}_t + \widehat{\Pi}_t^* + \widehat{\widetilde{y}}_t^d\right) \\ &\quad + \beta\theta_p \Pi^{-(1-\varepsilon)(1-\chi)} \mathbb{E}_t \left(-(1-\varepsilon)\left(\widehat{\Pi}_{t+1} - \chi\widehat{\Pi}_t\right) - \left(\widehat{\Pi}_{t+1}^* - \widehat{\Pi}_t^*\right) + \widehat{g}_{t+1}^2\right). \quad \text{(A.11)} \end{aligned}$$

Note that is easy to show that $\varepsilon g_t^1 = (\varepsilon - 1) g_t^2$ loglinearizes to:

$$\widehat{g}_t^1 = \widehat{g}_t^2. \quad \text{(A.12)}$$

Now, let us loglinearize the relationship between the capital-labor ratio and the real wage-real interest rate

$$\frac{u_t \widetilde{k}_{t-1}}{l_t^d} = \frac{\alpha}{1-\alpha} \frac{\widetilde{w}_t}{\widetilde{r}_t} \frac{z_t}{z_{t-1}} \frac{\mu_t}{\mu_{t-1}}.$$

It is easy to show that

$$\widehat{u}_t + \widehat{\widetilde{k}}_{t-1} - \widehat{l}_t^d = \widehat{\widetilde{w}}_t - \widehat{\widetilde{r}}_t + \widehat{\widetilde{z}}_t + \widehat{\widetilde{\mu}}_t. \quad \text{(A.13)}$$

Let us loglinearize the marginal cost

$$mc_t = \left(\frac{1}{1-\alpha}\right)^{1-\alpha} \left(\frac{1}{\alpha}\right)^{\alpha} (\widetilde{w}_t)^{1-\alpha} \widetilde{r}_t^{\alpha}$$

to get

$$\widehat{mc}_t = (1-\alpha)\widehat{\widetilde{w}}_t + \alpha\widehat{\widetilde{r}}_t. \tag{A.14}$$

Now we concentrate on the aggregate price law of motion:

$$1 = \theta_p \left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{1-\varepsilon} + (1-\theta_p)\,\Pi_t^{*1-\varepsilon}$$

that can be rewritten as:

$$1 = \theta_p \Pi^{-(1-\varepsilon)(1-\chi)} \exp^{-(1-\varepsilon)(\widehat{\Pi}_t - \chi\widehat{\Pi}_{t-1})} + (1-\theta_p)\,(\Pi^*)^{(1-\varepsilon)} \exp^{(1-\varepsilon)\widehat{\Pi}_t^*}$$

that loglinearizes to:

$$\frac{\theta_p \Pi^{-(1-\varepsilon)(1-\chi)}}{(1-\theta_p)\,(\Pi^*)^{(1-\varepsilon)}}(\widehat{\Pi}_t - \chi\widehat{\Pi}_{t-1}) = \widehat{\Pi}_t^*. \tag{A.15}$$

The Taylor rule loglinearizes to:

$$\widehat{R}_t = \gamma_R \widehat{R}_{t-1} + (1-\gamma_R)\left(\gamma_\Pi \widehat{\Pi}_t + \gamma_y(\triangle\widehat{\widetilde{y}}_t^d + \widehat{\widetilde{z}}_t)\right) + \widehat{m}_t. \tag{A.16}$$

The market clearing conditions:

$$l_t = l_t^d,$$

$$\widetilde{c}_t + \widetilde{x}_t + \widetilde{g}_t + \frac{\Phi\,[u_t]\,\widetilde{k}_{t-1}}{\widetilde{\mu}_t \widetilde{z}_t} = \widetilde{y}_t^d,$$

and

$$v_t^p \widetilde{y}_t^d = \frac{\widetilde{A}_t}{\widetilde{z}_t}\left(u_t \widetilde{k}_{t-1}\right)^{\alpha}\left(l_t^d\right)^{1-\alpha} - f$$

can be written as:

$$l \exp^{\widehat{l}_t} = l^d \exp^{\widehat{l}_t^d},$$

$$\widetilde{c}\exp^{\widehat{\widetilde{c}}_t} + \widetilde{x}\exp^{\widehat{\widetilde{x}}_t} + \widetilde{g}\exp^{\widehat{\widetilde{g}}_t} + \frac{\Phi\left[u\exp^{\widehat{u}_t}\right]\widetilde{k}\exp^{\widehat{\widetilde{k}}_{t-1}}}{\widetilde{z}\widetilde{\mu}\exp^{\widehat{\widetilde{\mu}}_t + \widehat{\widetilde{z}}_t}} = \widetilde{y}^d \exp^{\widehat{\widetilde{y}}_t^d},$$

and

$$v^p \widetilde{y}^d \exp^{\widehat{v}_t^p + \widehat{\widetilde{y}}_t^d} = \frac{\widetilde{A}}{\widetilde{z}}\left(\widetilde{k}\right)^{\alpha}\left(\widetilde{l}^d\right)^{1-\alpha} \exp^{\widehat{\widetilde{A}}_t - \widehat{\widetilde{z}}_t + \alpha\left(\widehat{u}_t + \widehat{\widetilde{k}}_{t-1}\right) + (1-\alpha)\widehat{l}_t^d} - f$$

respectively, where $\widetilde{A}_{t+1} = \frac{A_{t+1}}{A_t}$.

Loglinearizing we get:

$$\widehat{l}_t = \widehat{l}_t^d, \tag{A.17}$$

$$\widetilde{c}\widehat{\widetilde{c}}_t + \tilde{x}\widehat{\widetilde{x}}_t + g^{ss}\tilde{y}^d\widehat{\widetilde{g}}_t + \frac{\Phi_1\widetilde{k}}{\widetilde{z}\widetilde{\mu}}\widehat{u}_t = \tilde{y}^d\widehat{\widetilde{y}}_t^d \tag{A.18}$$

because $g^{ss} \equiv \frac{\widetilde{g}}{\widetilde{y}^d}$, and

$$(\tilde{y}^d v^p)\left(\widehat{v}_t^p + \widehat{\widetilde{y}}_t^d\right) = \frac{\widetilde{A}}{\widetilde{z}}\left(u\widetilde{k}\right)^{\alpha}\left(\widetilde{l}^d\right)^{1-\alpha}\left(\widehat{\widetilde{A}}_t - \widehat{\widetilde{z}}_t + \alpha\left(\widehat{u}_t + \widehat{\widetilde{k}}_{t-1}\right) + (1-\alpha)\,\widehat{l}_t^d\right). \tag{A.19}$$

Let us consider now

$$v_t^p = \theta_p\left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{-\varepsilon} v_{t-1}^p + (1-\theta_p)\,\Pi_t^{*-\varepsilon}$$

that can be written as

$$v^p \exp^{\widehat{v}_t^p} = \theta_p\Pi^{\varepsilon(1-\chi)}v^p \exp^{\varepsilon(\widehat{\Pi}_t - \chi\widehat{\Pi}_{t-1})+\widehat{v}_{t-1}^p} + (1-\theta_p)\,\Pi^{*-\varepsilon}\exp^{\varepsilon\widehat{\Pi}_t^*}.$$

Using

$$\left(1-\theta_p\Pi^{\varepsilon(1-\chi)}\right)v^p = (1-\theta_p)\,\Pi^{*-\varepsilon}$$

we get:

$$\widehat{v}_t^p = \theta_p\Pi^{\varepsilon(1-\chi)}\left(\varepsilon(\widehat{\Pi}_t - \chi\widehat{\Pi}_{t-1}) + \widehat{v}_{t-1}^p\right) - \left(1-\theta_p\Pi^{\varepsilon(1-\chi)}\right)\varepsilon\widehat{\Pi}_t^*. \tag{A.20}$$

Let us consider now

$$z_t = A_t^{\frac{1}{1-\alpha}}\mu_t^{\frac{\alpha}{1-\alpha}}.$$

By definition, we have

$$\widetilde{z}_t \equiv \frac{z_t}{z_{t-1}} = \left(\frac{A_t}{A_{t-1}}\right)^{\frac{1}{1-\alpha}}\left(\frac{\mu_t}{\mu_{t-1}}\right)^{\frac{\alpha}{1-\alpha}} = \widetilde{A}_t^{\frac{1}{1-\alpha}}\widetilde{\mu}_t^{\frac{\alpha}{1-\alpha}},$$

so we obtain

$$\widehat{\widetilde{z}}_t = \frac{\widehat{\widetilde{A}}_t + \alpha\widehat{\widetilde{\mu}}_t}{1-\alpha}$$

Finally, let us loglinearize the law of motion of capital

$$\tilde{k}_t\widetilde{z}_t\widetilde{\mu}_t = (1-\delta)\,\tilde{k}_{t-1} + \widetilde{\mu}_t\widetilde{z}_t\left(1 - S\left[\frac{\tilde{x}_t}{\tilde{x}_{t-1}}\widetilde{z}_t\right]\right)\tilde{x}_t.$$

If we rearrange terms, we get:

$$\tilde{k}\widetilde{z}\widetilde{\mu}\exp^{\widehat{\tilde{k}}_t+\widehat{\widetilde{z}}_t+\widehat{\widetilde{\mu}}_t} = (1-\delta)\,\tilde{k}\exp^{\widehat{\tilde{k}}_{t-1}} + \widetilde{z}\widetilde{\mu}\exp^{\widehat{\widetilde{z}}_t+\widehat{\widetilde{\mu}}_t}\left(1 - S\left[ze^{\widehat{\widetilde{z}}_t+\triangle\widehat{\widetilde{x}}_t}\right]\right)\tilde{x}\exp^{\widehat{\widetilde{x}}_t}.$$

Loglinearizing:

$$\tilde{k}\widetilde{z\mu}\left(\widehat{\tilde{k}}_t+\widehat{\tilde{z}}_t+\widehat{\tilde{\mu}}_t\right)=(1-\delta)\,\tilde{k}\widehat{\tilde{k}}_{t-1}+\widetilde{z\mu}\tilde{x}(\widehat{\tilde{z}}_t+\widehat{\tilde{\mu}}_t+\widehat{\tilde{x}}_t).$$

Note that, using $\tilde{k}=\frac{\widetilde{z\mu}}{\widetilde{z\mu}-(1-\delta)}\tilde{x}$, we can rearrange the previous expression to get:

$$\widehat{\tilde{k}}_t+\widehat{\tilde{z}}_t+\widehat{\tilde{\mu}}_t=\frac{(1-\delta)}{\widetilde{z\mu}}\widehat{\tilde{k}}_{t-1}+\frac{\widetilde{z\mu}-(1-\delta)}{\widetilde{z\mu}}(\widehat{\tilde{z}}_t+\widehat{\tilde{\mu}}_t+\widehat{\tilde{x}}_t)$$

or

$$\widehat{\tilde{k}}_t=\frac{(1-\delta)}{\widetilde{z\mu}}\widehat{\tilde{k}}_{t-1}+\frac{\widetilde{z\mu}-(1-\delta)}{\widetilde{z\mu}}\widehat{\tilde{x}}_t-\frac{1-\delta}{\widetilde{z\mu}}\left(\widehat{\tilde{z}}_t+\widehat{\tilde{\mu}}_t\right). \tag{A.21}$$

A.5 System of Linear Stochastic Difference Equations

We now present the equations in the system in an orderly fashion.

Equation 1 The first equation is:

$$\frac{\theta_p\Pi^{-(1-\varepsilon)(1-\chi)}}{(1-\theta_p)\,(\Pi^*)^{(1-\varepsilon)}}(\widehat{\Pi}_t-\chi\widehat{\Pi}_{t-1})=\widehat{\Pi}_t^*$$

In order to make notation for compact, define $a_2=\frac{\theta_p\Pi^{-(1-\varepsilon)(1-\chi)}}{(1-\theta_p)(\Pi^*)^{(1-\varepsilon)}}$. Substituting for a_2:

$$a_2(\widehat{\Pi}_t-\chi\widehat{\Pi}_{t-1})=\widehat{\Pi}_t^*$$

Rearranging:

$$a_2\widehat{\Pi}_t-a_2\chi\widehat{\Pi}_{t-1}-\widehat{\Pi}_t^*=0 \tag{A.22}$$

Equation 2 The second equation is

$$\widehat{\tilde{r}}_t=\phi_u\hat{u}_t,$$

where $\phi_u=\Phi_2/\Phi_1$. Then,

$$-\widehat{\tilde{r}}_t+\phi_u\hat{u}_t=0 \tag{A.23}$$

Equation 3 The third equation is

$$\widehat{g}_t^1=\widehat{g}_t^2$$

Rearranging:

$$\widehat{g}_t^1 - \widehat{g}_t^2 = 0 \tag{A.24}$$

Equation 4 The fourth equation is:

$$\widehat{u}_t + \widehat{\widetilde{k}}_{t-1} - \widehat{l}_t = \widehat{\widetilde{w}}_t - \widehat{\widetilde{r}}_t + \widehat{\widetilde{z}}_t + \widehat{\widetilde{\mu}}_t$$

Rearranging:

$$\widehat{u}_t + \widehat{\widetilde{r}}_t + \widehat{\widetilde{k}}_{t-1} - \widehat{l}_t - \widehat{\widetilde{w}}_t - \widehat{\widetilde{z}}_t - \widehat{\widetilde{\mu}}_t \tag{A.25}$$

Equation 5 The fifth equation is:

$$\widehat{mc}_t = (1-\alpha)\widehat{\widetilde{w}}_t + \alpha\widehat{\widetilde{r}}_t$$

Rearranging:

$$(1-\alpha)\widehat{\widetilde{w}}_t + \alpha\widehat{\widetilde{r}}_t - \widehat{mc}_t = 0 \tag{A.26}$$

Equation 6 The sixth equation is:

$$\widehat{R}_t = \gamma_R \widehat{R}_{t-1} + (1-\gamma_R)\left(\gamma_\Pi \widehat{\Pi}_t + \gamma_y(\triangle\widehat{\widetilde{y}}_t^d + \widehat{\widetilde{z}}_t)\right) + \widehat{m}_t$$

Rearranging:

$$-\widehat{R}_t + \gamma_R\widehat{R}_{t-1} + (1-\gamma_R)\gamma_\Pi\widehat{\Pi}_t + (1-\gamma_R)\gamma_y\widehat{\widetilde{z}}_t + (1-\gamma_R)\gamma_y\widehat{\widetilde{y}}_t^d - (1-\gamma_R)\gamma_y\widehat{\widetilde{y}}_{t-1}^d + \widehat{m}_t = 0 \tag{A.27}$$

Equation 7 The seventh equation is:

$$\tilde{c}\widehat{\widetilde{c}}_t + \tilde{x}\widehat{\widetilde{x}}_t + g^{ss}\tilde{y}^d\widehat{\widetilde{g}}_t + \frac{\Phi_1\widetilde{k}}{\widetilde{z}\widetilde{\mu}}\widehat{u}_t = \tilde{y}^d\widehat{\widetilde{y}}_t^d$$

Note that in the code, because we have solved for the log-steady state, constants enter as follows: $\tilde{c} = \exp\left[\log\left(\tilde{c}\right)\right]$, $\tilde{x} = \exp\left[\log\left(\tilde{x}\right)\right]$, $\tilde{k} = \exp\left[\log\left(\tilde{k}\right)\right]$, $\tilde{y}^d = \exp\left[\log\left(\tilde{y}^d\right)\right]$. Rearranging:

$$\tilde{c}\widehat{\widetilde{c}}_t + \tilde{x}\widehat{\widetilde{x}}_t + g^{ss}\tilde{y}^d\widehat{\widetilde{g}}_t + \frac{\Phi_1\widetilde{k}}{\widetilde{z}\widetilde{\mu}}\widehat{u}_t - \tilde{y}^d\widehat{\widetilde{y}}_t^d = 0 \tag{A.28}$$

Equation 8 The eighth equation is:

$$(\tilde{y}^d v^p)\left(\widehat{v}_t^p + \widehat{\tilde{y}}_t^d\right) = \frac{\widetilde{A}}{\widetilde{z}}\left(u\widetilde{k}\right)^{\alpha}\left(\widetilde{l}^d\right)^{1-\alpha}\left(\widehat{\widetilde{A}}_t - \widehat{\widetilde{z}}_t + \alpha\left(\widehat{u}_t + \widehat{\widetilde{k}}_{t-1}\right) + (1-\alpha)\,\widehat{l}_t\right)$$

Define the parameter $produc = \frac{\widetilde{A}}{\widetilde{z}}\left(u\widetilde{k}\right)^{\alpha}\left(\widetilde{l}^d\right)^{1-\alpha}$. Substituting:

$$(\tilde{y}^d v^p)\left(\widehat{v}_t^p + \widehat{\tilde{y}}_t^d\right) = produc\left(\widehat{\widetilde{A}}_t - \widehat{\widetilde{z}}_t + \alpha\left(\widehat{u}_t + \widehat{\widetilde{k}}_{t-1}\right) + (1-\alpha)\,\widehat{l}_t\right)$$

Rearranging:

$$\begin{aligned}(\tilde{y}^d v^p)\widehat{v}_t^p + (\tilde{y}^d v^p)\widehat{\tilde{y}}_t^d - (produc)\widehat{\widetilde{A}}_t + (produc)\widehat{\widetilde{z}}_t - (\alpha)(produc)\widehat{u}_t \\ -(\alpha)(produc)\widehat{\widetilde{k}}_{t-1} - (1-\alpha)\,(produc)\widehat{l}_t = 0\end{aligned} \tag{A.29}$$

Equation 9 The ninth equation is:

$$\widehat{v}_t^p = \theta_p \Pi^{\varepsilon(1-\chi)}\left(\varepsilon(\widehat{\Pi}_t - \chi\widehat{\Pi}_{t-1}) + \widehat{v}_{t-1}^p\right) - \left(1 - \theta_p\Pi^{\varepsilon(1-\chi)}\right)\varepsilon\widehat{\Pi}_t^*$$

Define $a_3 = \beta\theta_p\Pi^{\varepsilon(1-\chi)}$. Then, $\theta_p\Pi^{\varepsilon(1-\chi)} = \frac{a_3}{\beta}$. This type of parameter definition will become clearer when we analyze the price setting equations. Then, substituting:

$$\widehat{v}_t^p = \frac{a_3}{\beta}\varepsilon\widehat{\Pi}_t - \chi\varepsilon\frac{a_3}{\beta}\widehat{\Pi}_{t-1} + \frac{a_3}{\beta}\widehat{v}_{t-1}^p - \left(1 - \frac{a_3}{\beta}\right)\varepsilon\widehat{\Pi}_t^*$$

And rearranging:

$$\frac{a_3\varepsilon}{\beta}\widehat{\Pi}_t - \frac{a_3\varepsilon\chi}{\beta}\widehat{\Pi}_{t-1} + \frac{a_3}{\beta}\widehat{v}_{t-1}^p - \left(1 - \frac{a_3}{\beta}\right)\varepsilon\widehat{\Pi}_t^* - \widehat{v}_t^p = 0 \tag{A.30}$$

Equation 10 The tenth equation is:

$$\widehat{\widetilde{k}}_t = \frac{(1-\delta)}{\widetilde{z}\widetilde{\mu}}\widehat{\widetilde{k}}_{t-1} + \frac{\widetilde{z}\widetilde{\mu} - (1-\delta)}{\widetilde{z}\widetilde{\mu}}\widehat{\widetilde{x}}_t - \frac{1-\delta}{\widetilde{z}\widetilde{\mu}}\left(\widehat{\widetilde{z}}_t + \widehat{\widetilde{\mu}}_t\right)$$

which we rearrange to:

$$\frac{(1-\delta)}{\widetilde{z}\widetilde{\mu}}\widehat{\widetilde{k}}_{t-1} + \left[1 - \frac{(1-\delta)}{\widetilde{z}\widetilde{\mu}}\right]\widehat{\widetilde{x}}_t - \frac{1-\delta}{\widetilde{z}\widetilde{\mu}}\left(\widehat{\widetilde{z}}_t + \widehat{\widetilde{\mu}}_t\right) - \widehat{\widetilde{k}}_t = 0 \tag{A.31}$$

Equation 11 The eleventh equation is:

$$\widehat{\widetilde{z}}_t = \frac{\widehat{\widetilde{A}}_t + \alpha\widehat{\widetilde{\mu}}_t}{1-\alpha}$$

which we rearrange to:

$$\frac{1}{1-\alpha}\widehat{\widetilde{A}}_t + \frac{\alpha}{1-\alpha}\widehat{\widetilde{\mu}}_t - \widehat{\widetilde{z}}_t = 0 \tag{A.32}$$

Equation 12 The twelfth equation is:

$$\widehat{d}_t + \widehat{\phi}_t + \vartheta\widehat{l}_t = \widehat{\widetilde{w}}_t + \widehat{\widetilde{\lambda}}_t$$

which we rearrange to:

$$\widehat{d}_t + \widehat{\phi}_t + \vartheta\widehat{l}_t - \widehat{\widetilde{w}}_t - \widehat{\widetilde{\lambda}}_t = 0$$

Equation 13 The thirteenth equation is:

$$\left(1 - \frac{h\beta}{\widetilde{z}}\right)\widehat{\widetilde{\lambda}}_t = \widehat{d}_t - \frac{h\beta}{\widetilde{z}}\mathbb{E}_t\widehat{d}_{t+1} - \frac{1+h^2/z^2\beta}{\left(1-\frac{h}{\widetilde{z}}\right)}\widehat{\widetilde{c}}_t + \frac{h}{\widetilde{z}\left(1-\frac{h}{\widetilde{z}}\right)}\widehat{\widetilde{c}}_{t-1} + \frac{\beta h/z}{\left(1-\frac{h}{\widetilde{z}}\right)}\mathbb{E}_t\widehat{\widetilde{c}}_{t+1} - \frac{h}{\widetilde{z}\left(1-\frac{h}{\widetilde{z}}\right)}\widehat{\widetilde{z}}_t$$

Rearranging:

$$\widehat{d}_t - \frac{h\beta}{\widetilde{z}}\mathbb{E}_t\widehat{d}_{t+1} - \frac{1+h^2/\widetilde{z}^2\beta}{\left(1-\frac{h}{\widetilde{z}}\right)}\widehat{\widetilde{c}}_t + \frac{h/\widetilde{z}}{\left(1-\frac{h}{\widetilde{z}}\right)}\widehat{\widetilde{c}}_{t-1} + \frac{\beta h/\widetilde{z}}{\left(1-\frac{h}{\widetilde{z}}\right)}\mathbb{E}_t\widehat{\widetilde{c}}_{t+1} - \frac{h/\widetilde{z}}{\left(1-\frac{h}{\widetilde{z}}\right)}\widehat{\widetilde{z}}_t - \left(1 - \frac{h\beta}{\widetilde{z}}\right)\widehat{\widetilde{\lambda}}_t = 0 \tag{A.33}$$

Equation 14 The fourteenth equation is:

$$\widehat{\widetilde{\lambda}}_t = \mathbb{E}_t\{\widehat{\widetilde{\lambda}}_{t+1} + \widehat{R}_t - \widehat{\Pi}_{t+1}\}$$

which we rearrange to:

$$\mathbb{E}_t\{\widehat{\widetilde{\lambda}}_{t+1} - \widehat{\widetilde{\lambda}}_t + \widehat{R}_t - \widehat{\Pi}_{t+1}\} = 0 \tag{A.34}$$

Equation 15 The fifteenth equation is:

$$\widehat{\widetilde{q}}_t = \mathbb{E}_t\Delta\widehat{\widetilde{\lambda}}_{t+1} + \frac{\beta\left(1-\delta\right)}{\widetilde{z}\widetilde{\mu}}\mathbb{E}_t\widehat{\widetilde{q}}_{t+1} + \left(1 - \frac{\beta(1-\delta)}{\widetilde{z}\widetilde{\mu}}\right)\mathbb{E}_t\widehat{\widetilde{r}}_{t+1}$$

which we rearrange to:

$$\mathbb{E}_t\widehat{\widetilde{\lambda}}_{t+1} - \widehat{\widetilde{\lambda}}_t + \frac{\beta\,(1-\delta)}{\widetilde{z}\widetilde{\mu}}\mathbb{E}_t\widehat{\widetilde{q}}_{t+1} + \left(1 - \frac{\beta(1-\delta)}{\widetilde{z}\widetilde{\mu}}\right)\mathbb{E}_t\widehat{\widetilde{r}}_{t+1} - \widehat{\widetilde{q}}_t = 0 \tag{A.35}$$

Equation 16 The sixteenth equation is:

$$\kappa\widetilde{z}^2\left(\Delta\widehat{\widetilde{x}}_t + \widehat{\widetilde{z}}_t\right) = \widehat{\widetilde{q}}_t + \beta\kappa\widetilde{z}^2\mathbb{E}_t\Delta\widehat{\widetilde{x}}_{t+1}$$

which we rearrange by undoing the first-difference operator:

$$\widehat{\widetilde{q}}_t + \beta\kappa\widetilde{z}^2\mathbb{E}_t\widehat{\widetilde{x}}_{t+1} - (1+\beta)\kappa\widetilde{z}^2\widehat{\widetilde{x}}_t + \kappa\widetilde{z}^2\widehat{\widetilde{x}}_{t-1} - \kappa\widetilde{z}^2\widehat{\widetilde{z}}_t = 0 \tag{A.36}$$

Equation 17 The seventeenth first equation is

$$\widehat{g}_t^1 = \left(1 - \beta\theta_p\Pi^{\varepsilon(1-\chi)}\right)\left(\widehat{\widetilde{\lambda}}_t + \widehat{mc}_t + \widehat{\widetilde{y}}_t^d\right) + \beta\theta_p\Pi_t^{\varepsilon(1-\chi)}\mathbb{E}_t\left(\varepsilon(\widehat{\Pi}_{t+1} - \chi\widehat{\Pi}_t) + \widehat{g}_{t+1}^1\right)$$

As before, define $a_3 = \beta\theta_p\Pi^{\varepsilon(1-\chi)}$. Then:

$$(1-a_3)\,\widehat{\widetilde{\lambda}}_t + (1-a_3)\,\widehat{mc}_t + (1-a_3)\,\widehat{\widetilde{y}}_t^d + \varepsilon a_3\mathbb{E}_t\widehat{\Pi}_{t+1} - \chi\varepsilon a_3\widehat{\Pi}_t + a_3\mathbb{E}_t\widehat{g}_{t+1}^1 - \widehat{g}_t^1 = 0 \tag{A.37}$$

Equation 18 The eighteenth equation is

$$\begin{aligned}\widehat{g}_t^2 &= \left(1 - \beta\theta_p\Pi^{-(1-\varepsilon)(1-\chi)}\right)\left(\widehat{\widetilde{\lambda}}_t + \widehat{\Pi}_t^* + \widehat{\widetilde{y}}_t^d\right)\\ &\quad + \beta\theta_p\Pi^{-(1-\varepsilon)(1-\chi)}\mathbb{E}_t\left(-(1-\varepsilon)\left(\widehat{\Pi}_{t+1} - \chi\widehat{\Pi}_t\right) - \left(\widehat{\Pi}_{t+1}^* - \widehat{\Pi}_t^*\right) + \widehat{g}_{t+1}^2\right)\end{aligned}$$

Define: $a_7 = \beta\theta_p\Pi^{-(1-\varepsilon)(1-\chi)}$. Substituting:

$$(1-a_7)\,\widehat{\widetilde{\lambda}}_t + \widehat{\Pi}_t^* + (1-a_7)\,\widehat{\widetilde{y}}_t^d + (\varepsilon-1)a_7\mathbb{E}_t\widehat{\Pi}_{t+1} - \chi(\varepsilon-1)a_7\widehat{\Pi}_t - a_7\mathbb{E}_t\widehat{\Pi}_{t+1}^* + a_7\mathbb{E}_t\widehat{g}_{t+1}^2 - \widehat{g}_t^2 = 0$$

Shocks The preference and labor supply shocks have the following structure:

$$\begin{aligned}\widehat{d}_t &= \rho_D\widehat{d}_{t-1} + \sigma_D\varepsilon_{D,t}\\ \widehat{\phi}_t &= \rho_\phi\widehat{\phi}_{t-1} + \sigma_\phi\varepsilon_{\phi,t}\end{aligned}$$

The government consumption process is

$$\widehat{\widetilde{g}}_t = \rho_G \widehat{\widetilde{g}}_{t-1} + \sigma_G \varepsilon_{G,t}$$

Note that Uhlig does not allow to write something like $\widehat{d}_t = \rho \widehat{d}_{t-1} + \sigma_D \varepsilon_{D,t}$. We declare the variance-covariance matrix later.

The following shocks are not defined because of their i.i.d. nature.

$$\begin{aligned} \widehat{\widetilde{\mu}}_t &= z_{\mu,t} \\ \widehat{\widetilde{A}}_t &= z_{A,t} \\ m_t &= \sigma_M \varepsilon_{M,t} \end{aligned}$$

A.6 Solving the Model

Now, let

$$state_t = \left(\widehat{\Pi}_t, \widehat{g}_t^1, \widehat{g}_t^2, \widehat{\widetilde{k}}_t, \widehat{R}_t, \widehat{\widetilde{y}}_t^d, \widehat{\widetilde{c}}_t, \widehat{v}_t^p, \widehat{\widetilde{q}}_t, \widehat{\widetilde{x}}_t \right)',$$

$$nstate_t = \left(\widehat{\widetilde{\lambda}}_t, \widehat{\widetilde{z}}_t, \widehat{\widetilde{r}}_t, \widehat{u}_t, \widehat{\Pi}_t^*, \widehat{mc}_t, \widehat{l}_t, \widehat{\widetilde{w}}_t \right)',$$

$$exo_t = \left(\widehat{\widetilde{\mu}}_t, \widehat{d}_t, \widehat{\phi}_t, \widehat{\widetilde{A}}_t, m_t, \widehat{\widetilde{g}}_t \right)',$$

and

$$\varepsilon_t = \left(\varepsilon_{\mu,t}, \varepsilon_{D,t}, \varepsilon_{\phi,t}, \varepsilon_{A,t}, \varepsilon_{M,t}, \varepsilon_{G,t} \right)'.$$

We write the system defined above in Uhlig's format, i.e.:

$$0 = AA * state_t + BB * state_{t-1} + CC * nstate_t + DD * exo_t,$$

$$0 = \mathbb{E}_t \left(\begin{array}{c} FF * state_{t+1} + GG * state_t + HH * state_{t-1} \\ +JJ * nstate_{t+1} + KK * nstate_t + LL * exo_{t+1} + MM * exo_t \end{array} \right),$$

and

$$exo_{t+1} = NN * exo_t + \Sigma^{1/2} * \varepsilon_{t+1} \text{ with } \mathbb{E}_t \varepsilon_{t+1} = 0.$$

The matrices in Uhlig's notation are as following:

$$AA = \begin{pmatrix} a_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ (1-\gamma_R)\gamma_\Pi & 0 & 0 & 0 & -1 & (1-\gamma_R)\gamma_y & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & -\tilde{y}^d & \tilde{c} & 0 & 0 & \tilde{x} \\ 0 & 0 & 0 & 0 & 0 & \tilde{y}^d v^p & 0 & \tilde{y}^d v^p & 0 & 0 \\ a_3\varepsilon/\beta & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 \\ 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 1-\frac{(1-\delta)}{\tilde{z}\tilde{\mu}} \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$BB = \begin{pmatrix} -\chi a_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & \gamma_R & -(1-\gamma_R)\gamma_y & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & -\alpha(produc) & 0 & 0 & 0 & 0 & 0 & 0 \\ \frac{-a_3\varepsilon\chi}{\beta} & 0 & 0 & 0 & 0 & 0 & 0 & \frac{a_3}{\beta} & 0 & 0 \\ 0 & 0 & 0 & \frac{1-\delta}{\tilde{z}\tilde{\mu}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$CC = \begin{pmatrix} 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \\ 0 & 0 & -1 & \phi_u & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 1 & 1 & 0 & 0 & -1 & -1 \\ 0 & 0 & \alpha & 0 & 0 & -1 & 0 & (1-\alpha) \\ 0 & (1-\gamma_R)\gamma_y & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & \frac{\Phi_1 \tilde{k}}{\tilde{z}\tilde{\mu}} & 0 & 0 & 0 & 0 \\ 0 & produc & 0 & -\alpha(produc) & 0 & 0 & -(1-\alpha)(produc) & 0 \\ 0 & 0 & 0 & 0 & -(1-\frac{a_3}{\beta})\varepsilon & 0 & 0 & 0 \\ 0 & -\frac{(1-\delta)}{\tilde{z}\tilde{\mu}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & \theta & -1 \end{pmatrix}$$

$$DD = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & g^{ss}\tilde{y}^d \\ 0 & 0 & 0 & -produc & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ -\frac{1-\delta}{\tilde{z}\tilde{\mu}} & 0 & 0 & 0 & 0 & 0 \\ \frac{\alpha}{1-\alpha} & 0 & 0 & \frac{1}{1-\alpha} & 0 & 0 \\ 0 & 1 & 1 & 0 & 0 & 0 \end{pmatrix}$$

$$FF = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \frac{\beta h/\tilde{z}}{1-\frac{h}{\tilde{z}}} & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\beta(1-\delta)}{\tilde{z}\tilde{\mu}} & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \beta\kappa\tilde{z}^2 \\ a_3\varepsilon & a_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_7(\varepsilon-1) & 0 & a_7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$GG = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & GG_{1,7} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & GG_{4,10} \\ -a_3\varepsilon\chi & -1 & 0 & 0 & 0 & 1-a_3 & 0 & 0 & 0 & 0 \\ -a_7(\varepsilon-1)\chi & 0 & -1 & 0 & 0 & 1-a_7 & 0 & 0 & 0 & 0 \end{pmatrix}$$

where

$$\begin{aligned} GG_{1,7} &= -\frac{1+\beta h^2/\tilde{z}^2}{1-\frac{h}{\tilde{z}}} \\ GG_{4,10} &= -(1+\beta)\kappa\tilde{z}^2 \end{aligned}$$

$$HH = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & \frac{h}{\tilde{z}(1-\frac{h}{\tilde{z}})} & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \kappa\tilde{z}^2 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$JJ = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1 & 0 & 1-\frac{\beta(1-\delta)}{\tilde{z}\tilde{\mu}} & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & -a_7 & 0 & 0 & 0 \end{pmatrix}$$

$$KK = \begin{pmatrix} -(1-h\beta/\tilde{z}) & -\frac{h}{\tilde{z}(1-\frac{h}{\tilde{z}})} & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & -\kappa\tilde{z}^2 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1-a_3 & 0 & 0 & 0 & 0 & 1-a_3 & 0 & 0 \\ 1-a_7 & 0 & 0 & 0 & 1 & 0 & 0 & 0 \end{pmatrix}$$

$$LL = \begin{pmatrix} 0 & -\beta h/\tilde{z} & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$MM = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$NN = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \rho_D & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho_\phi & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \rho_G \end{pmatrix}$$

$$\Sigma = \begin{pmatrix} \sigma_\mu^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_D^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_\phi^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_A^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_M^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_G^2 \end{pmatrix}$$

B Data construction

In this section, we explain how we constructed our data in detail. First, we explain the macroeconomic dataset used in the core measurement equation. Then, we discuss how we processed the FOMC transcript data. Finally, we explain how we extracted factors from the Greenbok data.

B.1 Macroeconomic data

Our macroeconomic data come from the Federal Reserve Bank of St.Louis' FRED database (https://fred.stlouisfed.org/), collected at a quarterly frequency. We construct the data as follows:

- **Inflation**. Take the series for the GDP deflator, *FRED* mnemonic "GDPDEF", call it $GDPP_t$. Calculate net quarterly inflation as $\log(GDPP_t/GDPP_{t-1})$.
- **Federal Funds Rate**. Take the series for the effective federal funds rate, *FRED* mnemonic "FEDFUNDS" (percent), call it FFR_t. Calculate the gross rate as $1 + FFR_t/400$ and take log.
- **Real wage growth**. Take the *FRED* mnemonic "COMPRNFB" (Nonfarm Business Sector: Real Compensation Per Hour). Call it $WAGE_t$. Define real wage growth as $\log(WAGE_t/WAGE_{t-1})$.
- **Real output per capita growth**. Take the *FRED* mnemonic "A939RX0Q048SBEA" (Real gross domestic product per capita). Call it $GDPPC_t$. Define real output per capita growth as $\log(GDPPC_t/GDPPC_{t-1})$.
- **Per Capita Hours Index**. Take the index of average weekly nonfarm business hours, (*FRED* mnemonic "PRS85006023"), call it $HOURS_t$. Take the quarterly average of the civilian non-institutional population, (*FRED* mnemonic "CNP16OV"), call it POP_t. Take the number of employed civilians (*FRED* mnemonic "CE16OV"), call it EMP_t. Calculate per capita hours as $HOURS_t \cdot EMP_t/POP_t$. Take logs.
- **Relative price of investment goods growth**. Take the relative price of investment goods, (*FRED* mnemonic "PIRIC"), call it $PIRIC_t$. Define the relative price of investment goods growth as $\log(PIRIC_t/PIRIC_{t-1})$.

B.2 Text data

Verbatim and fully attributed transcripts of Federal Open Market Committee meetings are available from https://www.federalreserve.gov/monetarypolicy/fomc_historical_year.htm. We take FOMC meetings from the beginning of Alan Greenspan's Chairmanship (Aug 1987) through 2019, the last year transcripts were available at the time of writing. This period comprises 259 meetings.

Starting with Greenspan, FOMC meetings had a regular structure that included a discussion of the economic situation (FOMC1) followed by a discussion of the policy stance (FOMC2). Each section is opened by staff presentations and then each FOMC member shares his or her views in sequence. The labeling of these sections is manual and follows

Cieslak et al. (2023).

We preprocess the text by applying standard cleaning operations including lowercasing, contraction splitting, and punctuation removal. We replace frequent bigrams and trigrams with single tokens, stem all words, and remove stopwords. We further exclude tokens appearing fewer than 15 times in the corpus, as well as a small set of generic high-frequency terms. We estimate LDA topic models on the combined FOMC1 and FOMC2 corpus, which includes statements from both FOMC members and staff, totaling 76,056 statements with 8,381 unique stems. We use Gibbs sampling with 15,000 burn-in iterations, 5,000 sampling iterations, and a thinning interval of 50, and consider models with 20, 30, and 40 topics.

To construct quarterly topic share time series, we aggregate all statements by FOMC members only (i.e., excluding staff statements) within each quarter and infer topic proportions holding fixed the estimated topic-word distributions from the baseline LDA model. We do this separately for FOMC1 (33,069 member statements concatenated into one text per quarter) and FOMC2 (29,447 member statements similarly concatenated).

B.3 Estimating factors for the Greenbook data

Before extracting factors from the Greenbook data, some variables must be transformed to be stationary. We do so by following the transformation code in the FRED-QD database by McCracken and Ng (2021). However, note that some data may require different treatments when applying the transformation of McCracken and Ng (2021). For instance, the Greenbook data on inflation are the quarter-over-quarter growth rate, whereas in FRED-QD they are transformed as $\triangle^2(\log(x))$. Hence, we simply need to take another difference for the Greenbook inflation rate. In addition, the real quantity variables, such as the real GDP, are the quarter-over-quarter growth rate in the Greenbook, while they are transformed as $\triangle(\log(x))$ in FREDQD. Therefore, no transformation is needed for the Greenbook GDP series.

We then estimate factors by applying the EM algorithm proposed in McCracken and Ng (2021). The method is essentially a principal component analysis, but the iterative procedure embedded in the EM algorithm enables us to handle the missing data consistently. There are two possible reasons why we encounter missing values: (a) the data is indeed missing, which is especially the case for longer forecast horizons in old data, and (b) some outliers are removed following the criteria given by McCracken and Ng (2021).

We estimate factors for the full-sample period of 1987Q1-2019Q4. We then standardize the factors and incorporate them into the non-core measurement equation.

C Details on forecasting

C.1 Sampling algorithm from a predictive distribution

In this subsection, we explain in detail how to obtain out-of-sample predictions. Particularly, we follow the standard method, such as Del Negro and Schorfheide (2013) and sample from the predictive distribution $p(Y_{T+1}, \cdots Y_{T+H}|Y_{1:T})$.

Recall our state-space representation:

$$\begin{aligned} Y_t^C &= H_0^C(\theta) + H_1^C(\theta)S_t + u_t, \quad u_t \sim \mathcal{N}(0, \Sigma_u^C(\theta)) \\ Y_t^{NC} &= B_0 + B_1 Y_{t-1}^{NC} + \Lambda(RS_t) + v_t, \quad v_t \sim \mathcal{N}(0, \Sigma_v) \\ S_t &= T(\psi)S_{t-1} + R(\psi)\varepsilon_t, \quad \varepsilon_t \sim \mathcal{N}(0, \Sigma_\varepsilon) \end{aligned} \tag{7}$$

We use the following algorithm to draw samples from the predictive distribution for each sub-sample.

1. Repeat the following algorithm for $j = 1, 2, \cdots N_{sim}$.
2. Set $\theta = \theta^{(j)}$, which is the jth draw from the posterior distribution estimated in the sub-sample.
3. Given θ, draw $S_T^{(j)}$ by the simulation smoother.
4. Generate shocks $\varepsilon_{T+1}^{(j)}, \cdots \varepsilon_{t+H}^{(j)}$ from $\mathcal{N}(0, I)$ and use the state-transition equation to simulate a path of $S_{T+1}^{(j)}, \cdots S_{T+H}^{(j)}$.
5. Generate measurement errors $u_{T+1}^{(j)}, \cdots u_{t+H}^{(j)}$ from $\mathcal{N}(0, I)$ and use the core-measurement equation to obtain the prediction: $Y_{T+1|T}^{C,(j)}, \cdots Y_{T+H|T}^{C,(j)}$.

C.2 Measures of forecasting accuracy

We use three measures of forecasting accuracy: the joint log score, the root mean squared error (RMSE), and the continuous ranked probability score (CRPS).

When we need a point-forecast $\widehat{Y}_{T+h|T}$, it is computed by the Monte Carlo integrations following Diebold et al. (2017):

$$\widehat{Y}_{T+h|T} = \int_{Y_{T+h}} Y_{t+h} p(Y_{T+h}|Y_{1:T}) dY_{T+h} \approx \frac{1}{N_{sim}} \sum_{j=1}^{N_{sim}} Y_{T+h|T}^{(j)}.$$

where $Y_{T+h|T}^{(j)}$ is a vector of jth draw from the predictive distribution.

The joint log score we report in the paper is the sum of the joint log score at each

forecasting period:

$$\sum_{T=2000Q1}^{2019Q3} \log\left(\phi(Y_{T+1}; \widehat{Y}_{T+1|T}, \widehat{\Sigma}_{\widehat{Y}_{T+1|T}})\right)$$

where $\phi(x;\mu,\Sigma) = \frac{1}{\sqrt{(2\pi)^k|\Sigma|}} \exp\left(-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\right)$,
$\widehat{\Sigma}_{\widehat{Y}_{T+1|T}} = \frac{1}{N_{sim}-1}\sum_{j=1}^{N_{sim}}\left(Y_{T+1|T}^{(j)} - \widehat{Y}_{T+1|T}\right)\left(Y_{T+1|T}^{(j)} - \widehat{Y}_{T+1|T}\right)'$, and $k = 6$ corresponding to the number of macroeconomic variables forecasted in our application.

The RMSE for variable i at forecast horizon h can be computed as

$$RMSE^i(h) = \sqrt{\frac{1}{76-h}\sum_{T=2000Q1}^{2019Q4-h}\left(Y_{i,T+h} - \widehat{Y}_{i,T+h|T}\right)^2}.$$

The CRPS for variable i at forecast horizon h at time t is defined as

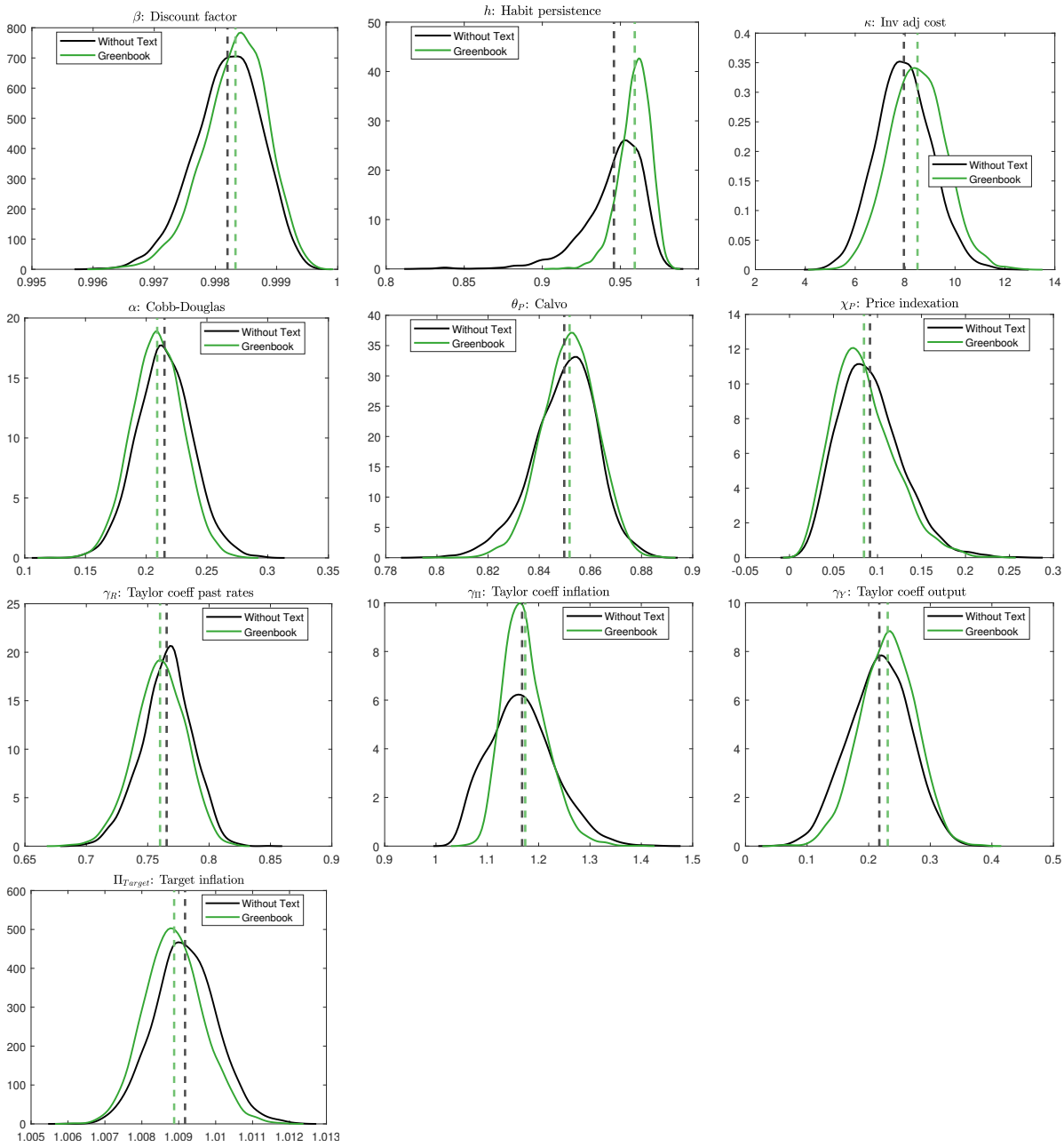
$$\begin{aligned} CRPS_t^i(h) &= \int_{-\infty}^{\infty}\left(F_{i,t+h|t}(u) - 1\{Y_{i,t+h} \le u\}\right)^2 du \\ &= E[|X_1 - Y_{i,t+h}|] - \frac{1}{2}E[|X_1 - X_2|] \end{aligned}$$

where $F_{i,t+h|t}(z)$ is the h-step-ahead cumulative predictive distribution for variable i at time t, and X_1 and X_2 are independent random variables whose CDF is given by $F_{i,t+h|t}(u)$.[12] This measure computes the discrepancy between the cumulative predictive distribution of the model and its empirical counterpart. We compute it by the following Monte Carlo approximation:

$$\widehat{CRPS}_t^i(h) = \frac{1}{N_{sim}}\sum_{j=1}^{N_{sim}}|Y_{i,t+h|t}^{(j)} - Y_{i,t+h}| - \frac{1}{2N_{sim}^2}\sum_{i=1}^{N_{sim}}\sum_{j=1}^{N_{sim}}|Y_{i,t+h|t}^{(i)} - Y_{i,t+h|t}^{(j)}|$$

because $F_{i,t+h|t}$ can be approximated by the draws from the predictive distribution. We report the sumed values in the paper: $\widehat{CRPS}^i(h) = \sum_{T=2000Q1}^{2019Q4-h}\widehat{CRPS}_T^i(h)$.

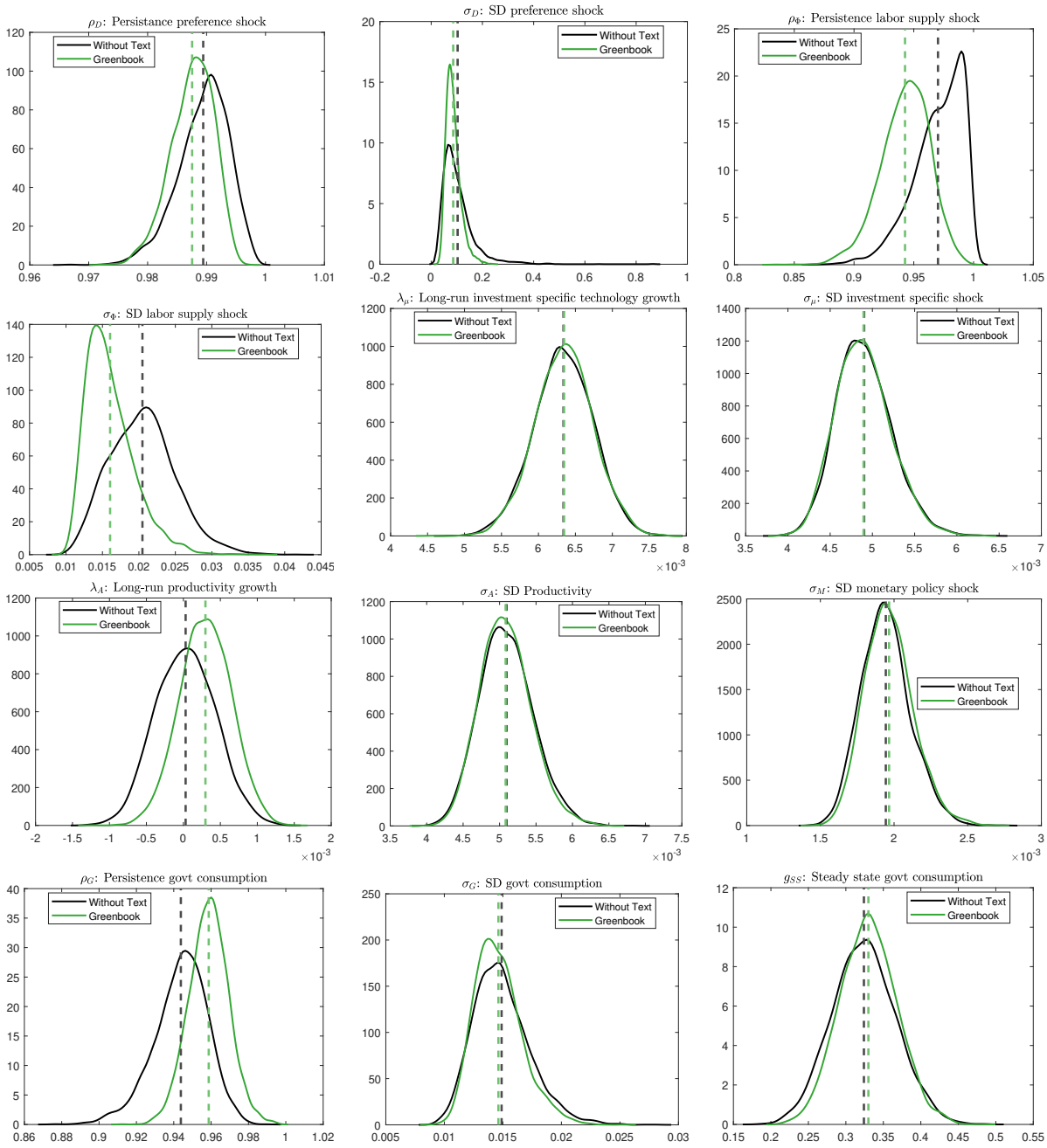
[12] The second equality follows from the following two properties: $|X-Y| = \int_{-\infty}^{\infty} 1\{Y \le u < X\}du + \int_{-\infty}^{\infty} 1\{X \le u < Y\}du$ and if X and Y are independent random variables with $X \sim F$ and $Y \sim G$, then $E[|X-Y|] = \int_{-\infty}^{\infty} F(u)(1-G(u))du + \int_{-\infty}^{\infty} G(u)(1-F(u))du$. Then, a simple algebra shows that these two expressions are equivalent.



Note: The dashed vertical line is the posterior mean. The black line represents the model without text, and the green line represents the model with Greenbook data.

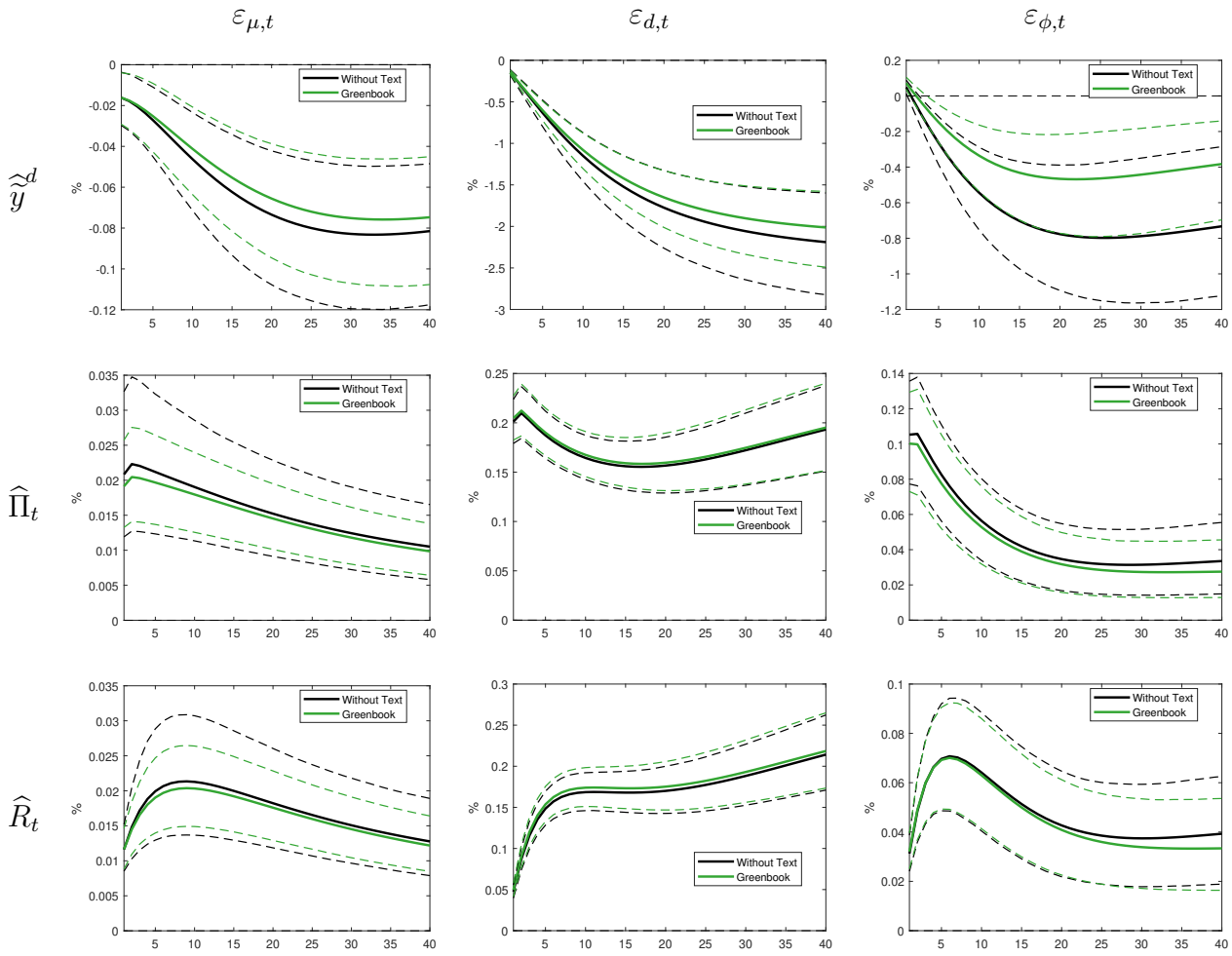
Figure D.1: Posterior distribution of parameters in our DSGE model 1

D Posterior Distributions and Impulse Responses with Greenbook data



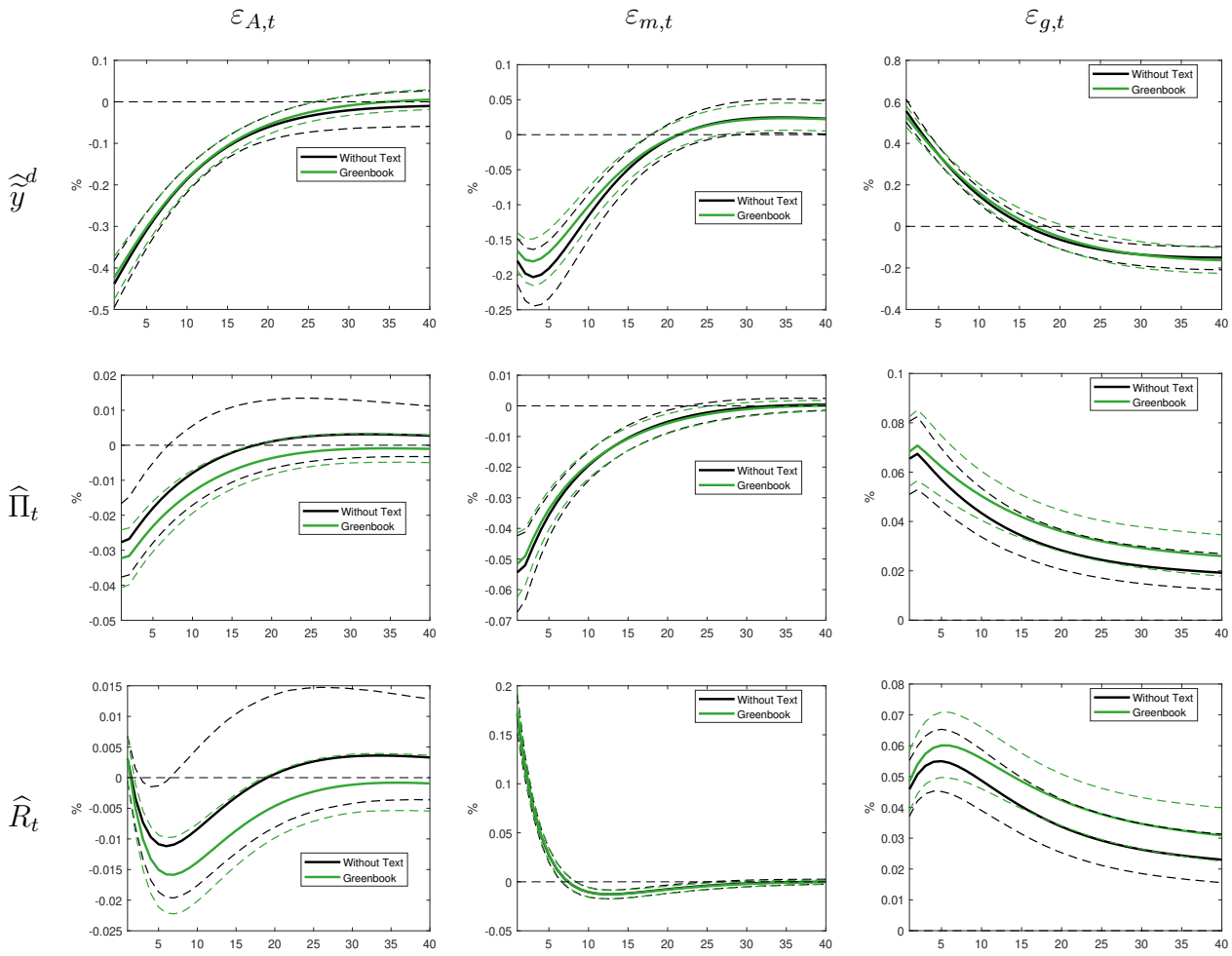
Note: The dashed vertical line is the posterior mean. The black line represents the model without text, and the green line represents the model with Greenbook data.

Figure D.2: Posterior distribution of parameters in our DSGE model 2



Note: The figures present the impulse responses of our estimated DSGE model to a one standard deviation shock. The black line represents the model without text, and the green line represents the model with Greenbook data. The solid line is the posterior mean, and the dashed line is the pointwise 90 % credible interval.

Figure D.3: Impulse responses of aggregate demand, inflation, and nominal interest rate to investment-specific technology, intertemporal preference, and labor supply shocks



Note: The figures present the impulse responses of our estimated DSGE model to a one standard deviation shock. The black line represents the model without text, and the green line represents the model with Greenbook data. The solid line is the posterior mean, and the dashed line is the pointwise 90 % credible interval.

Figure D.4: Impulse responses of aggregate demand, inflation, and nominal interest rate to technology, monetary policy, and government spending shocks