FEDERAL RESERVE BANK
PHILADELPHIA

Working Papers
RESEARCH DEPARTMENT

Structural Estimation with Unstructured Data

Sara Casella
LUISS University, EIEF

Jesús Fernández-Villaverde
University of Pennsylvania, NBER, CEPR, and Visiting Scholar, Federal Reserve Bank of Philadelphia Research Department

Stephen Hansen
University College London, Federal Reserve Bank of Dallas, IFS, CEPR

Ryohei Oishi
University College London

Minchul Shin
Federal Reserve Bank of Philadelphia

WP 26-40

PUBLISHED
August 2026

ISSN: 1962-5361

DOI: https://doi.org/10.21799/frbp.wp.2026.40

Structural Estimation with Unstructured Data*

Sara Casella
LUISS University, EIEF

Jesús Fernández-Villaverde
University of Pennsylvania, NBER, CEPR

Stephen Hansen
UCL, FRB Dallas, IFS, CEPR

Ryohei Oishi
University College London

Minchul Shin
Federal Reserve Bank of Philadelphia

August 3, 2026

Abstract

Standard macroeconomic data do not cleanly separate the systematic and non-systematic components of monetary policy. We show that incorporating unstructured text data into the structural estimation of a DSGE model can sharpen this distinction. 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 new model estimates a lower response of the policy rate to inflation, higher price stickiness and lower price indexation, implying a flatter and less backward-looking price 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 Dallas, 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. We thank Sila Ordu for research assistance. Hansen gratefully acknowledges financial support from ERC Consolidator Grant 864863.

1 Introduction

Dynamic equilibrium models are designed to answer policy questions, yet the data used to estimate them record outcomes that reflect both the rules agents follow and the shocks they face, without separating the two.[1] For example, we observe the federal funds rate set by the Federal Open Market Committee (FOMC), but not how much of its movement reflects the systematic component of the Taylor rule and how much reflects departures from it.

The cost of this identification problem is highest for dynamic stochastic general equilibrium (DSGE) models. Central banks use these models to evaluate policy, yet the parameters that govern their prescriptions (e.g., the parameters related to nominal rigidities) are among the hardest to estimate.

Unstructured data contain information that could help resolve this identification problem.[2] Text, images, audio, sensor data, and high-dimensional firm or product characteristics are now routine in applied economics, used to forecast (Bybee et al. 2024), to measure latent variables such as policy uncertainty (Baker et al. 2016) and sentiment (Shapiro et al. 2022), to characterize market structure (Hoberg and Phillips 2016), and to track economic activity through satellite imagery (Henderson et al. 2012). But these applications are mostly reduced-form: text regressions, event studies, sentiment indices, and forecasting exercises that treat textual measures as exogenous regressors or as extra time series in vector autoregressions. The counterfactual experiments and welfare calculations that dynamic equilibrium models deliver have remained beyond the reach of these methods.

We show that feeding FOMC transcripts, alongside standard macroeconomic aggregates such as output and inflation, into the estimation of a DSGE model can help us achieve what reduced-form text applications cannot. The discussions recorded in these documents reveal which rate movements reflect the systematic rule and which reflect judgment, communication, or accommodation. This information shifts the parameter estimates. The estimated response of the policy rate to inflation falls, the estimated response to real activity rises, and inflation dynamics become less backward-looking because the model estimates higher price stickiness and lower price indexation.

To exploit this information, we develop a framework that links unstructured text to the latent states of a DSGE model without a priori specifying which textual signals are

[1] By dynamic equilibrium models, we refer to models that are explicit about preferences, technology, information sets, and the intertemporal equilibrium outcomes implied by those assumptions. By structural estimation, we mean estimating 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 an SQL table.

informative. We cast the DSGE model as the *core block* of a state-space system, 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, but 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. It also 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 1987Q3–2019Q4, the last quarter for which FOMC transcripts are available. The core block includes 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.

We ask 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, especially at medium horizons. The gains are concentrated in the federal funds rate, hours worked, and output. The best-performing specification combines the economic-situation and policy-discussion portions of the transcripts, suggesting that the two parts contain complementary information. Greenbook and Tealbook projections, used as an alternative non-core source, do not improve upon the no-text benchmark jointly. The spike-and-slab prior selects a small number of economically interpretable topics, related to unconventional policy, inflation dynamics, risk and policy credibility, as informative about the monetary policy shock. Many

other topics, such as trade, that may appear relevant ex ante are effectively shut down, a natural consequence of using a closed-economy model.

Without text, variation that the standard Taylor rule cannot capture (including communication, forward guidance, and deliberation) is absorbed into the systematic feedback coefficients, inflating them. With the transcripts included, this variation is reallocated to the monetary policy shock. Thus, the estimated rule shifts: the response to inflation falls and the response to real activity rises. The nominal-rigidity parameters move too: the estimation raises the Calvo price parameter and lowers price indexation. These two changes work against each other: greater price stickiness flattens the Phillips curve, while lower indexation, on its own, would steepen it. Stickiness dominates, so the curve comes out flatter; lower indexation also leaves it less backward-looking. These shifts speak to important questions: the degree of nominal rigidity, the role of indexation in inflation dynamics, and the magnitude of the Taylor-rule response to inflation, which bears directly on equilibrium determinacy (Clarida et al. 2000; Lubik and Schorfheide 2004).

The resulting posterior shifts have economically meaningful consequences. Following an adverse labor-supply shock, the smaller rise in inflation weakens the motive to tighten, while the stronger estimated response to activity gives greater weight to the contraction, producing a more muted interest-rate response than in the model without text. Together, the policy-related topic loadings, posterior shifts, and changes in impulse responses suggest that the transcript block helps discipline the decomposition of policy-rate movements into systematic rule and monetary policy shocks. Our framework does not resolve fundamental identification failures, but the shifts in posterior mass and narrower credible intervals for selected parameters are consistent with text alleviating a practical identification problem.

Our paper builds on two 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 researchers such as 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 body of work 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 (Baker et al. 2016), geopolitical risk (Caldara and Iacoviello 2022), and news sentiment (Shapiro et al. 2022), 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, Shapiro and Wilson (2022) estimate the FOMC's implicit objective function, and Aruoba and Drechsel (2025) use FOMC text to help identify monetary policy shocks. In industrial organization, recent work incorporates product text and images into structural models of demand and product differentiation (Compiani et al. 2025; Pellegrino 2025). Across these applications, unstructured data enter as observed features, regressors, external instruments, or shock series. By contrast, we treat text-derived series as noisy measurements of latent structural shocks and embed them directly in the model's measurement system, allowing the estimation to determine which series are informative jointly with the structural parameters. Although our application focuses on FOMC transcripts and a New Keynesian DSGE model, the framework extends naturally to any structural model in state-space form and to other sources of unstructured data.

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 Models 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 models

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 non-linear settings, the likelihood function can be approximated using particle filtering or other non-linear 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 ones 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 capture certain economic forces (for example, if a relevant state variable lacks 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 ensures that the non-core series can be informative without requiring them to track the structural states perfectly. 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. 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 innovation in this equation 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, conditional on the lagged state and the non-core history, and given independent measurement errors, the same-period comovement between the core and non-core innovations is driven by the structural shocks ε_t, not by v_t. This restriction is milder than it may seem. If some further disturbance moves both the non-core observables and the DSGE states, that only means the structural model is misspecified and the shock belongs in ε_t.

In practice, the state vector $S_{1:T}$ is inferred from the joint information in the observables. In the linear-Gaussian model, with fixed θ and known measurement parameters, 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. The non-core block adds noisy measurements of the same state vector, so at worst these series are uninformative and at best they sharpen the Kalman smoother. Outside the linear-Gaussian case, this holds only on average: for a non-linear or non-Gaussian filter, a given draw can raise the realized covariance, and only the expected information is guaranteed to rise.

For the structural parameters, and for a fixed, correctly specified joint likelihood under standard regularity, the score based on (Y_t^C, Y_t^{NC}) has conditional expectation given Y_t^C equal to the score based on Y_t^C alone; so the Fisher information for θ with non-core data weakly dominates the core-only information. This holds with measurement parameters fixed; once the non-core block carries its own estimated nuisance parameters, the relevant object is the information about θ after profiling them out, which need not dominate. In finite samples, these gains can be offset by overfitting when the non-core block adds many parameters. We next discuss how to mitigate this 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. Thus, we impose 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 $B_2 = \Lambda D$, where D 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

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

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 D 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 places a positive probability on each loading $\lambda_{m,i}$ being exactly zero, so the model can switch off the link between the m-th non-core series and any individual state that the data find uninformative. Selection thus operates loading by loading: within row m, some loadings are active and others are zero. A non-core series becomes effectively irrelevant when the posterior sets all of its loadings to zero or near zero, which guards against contamination from noisy measurements. The non-core block, therefore, behaves as a sparse factor model. Our approach differs from data-rich DSGE estimations such as Boivin and Giannoni (2006) and Kryshko (2011), which do not allow such sparsity and so implicitly treat all included series as informative to some degree. The hyperparameter q_m sets the prior inclusion propensity of each loading in row m: smaller q_m places more prior mass on $\lambda_{m,i} = 0$. In our empirical analysis, we follow the parameterization of Giannone et al. (2021) and also estimate the hyperparameters, including q_m; implementation details are in Appendix B.

In summary, we balance the benefits of incorporating unstructured data into DSGE es-

timation against the risk of model misspecification. When the mapping between structural model variables and observables is well understood (as it is 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 misspecification. 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:

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

where B_1 and Σ_v are diagonal, so each non-core series has its own lag coefficient and innovation variance, with no cross-series dynamics or correlated errors.

This representation involves many unknown parameters. In our empirical application, there are 30 structural parameters in the DSGE block, $M \times 2$ parameters for the intercept and own-lag coefficient 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 $M = 40$ topic-share series from the FOMC transcripts, this yields 470 parameters in total.

Posterior sampling uses a Metropolis-Hastings-within-Gibbs algorithm that exploits the model's conditional structure. We sample the DSGE structural parameters via a random-

block random-walk Metropolis-Hastings algorithm (Chib et al. 2023), draw latent states via a simulation smoother, and update the non-core measurement parameters using Gibbs steps 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 and shrinkage objects of the non-core measurement equation, $(B_0, B_1, \Lambda, \Sigma_v, \Gamma, q)$. Here $\Gamma = \{\gamma_{m,i}\}$ collects the spike-and-slab inclusion indicators, one per loading $\lambda_{m,i}$, and $q = \{q_m\}$ collects the series-specific inclusion hyperparameters. We estimate these objects jointly rather than integrate them out. Our target is the joint posterior

$$p(\theta, S_{1:T}, \varphi \mid Y_{1:T}^{C}, Y_{1:T}^{NC}). \tag{8}$$

Each MCMC iteration consists of the following steps:

1. Update $(\theta, S_{1:T})$ conditional on φ: we draw from $p(S_{1:T} \mid \theta, \varphi, Y_{1:T}^{C}, Y_{1:T}^{NC})\, p(\theta \mid \varphi, Y_{1:T}^{C}, Y_{1:T}^{NC})$ by iterating:
   - 1-1 Sample θ from $p(\theta \mid \varphi, Y_{1:T}^{C}, Y_{1:T}^{NC})$ using a random-block random-walk Metropolis-Hastings step. This step is analogous to standard MCMC estimation of a DSGE model on core observables, but with the likelihood augmented by the non-core measurement equation for Y_t^{NC}.
   - 1-2 Sample $S_{1:T}$ from $p(S_{1:T} \mid \theta, \varphi, Y_{1:T}^{C}, Y_{1:T}^{NC})$ using a simulation smoother (e.g., Durbin and Koopman 2002). This step draws 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_t^{NC} is regressed on Y_{t-1}^{NC} and S_t. The parameters φ are then updated from $p(\varphi|\theta, S_{1:T}, Y_{1:T}^{C}, Y_{1:T}^{NC})$ 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, each of length 500,000, 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 the inclusion of FOMC text improve the model's predictive performance for standard macroeconomic variables? We use pseudo-out-of-sample forecasting to compare a benchmark DSGE model estimated only on core macro variables to specifications augmented with text-based non-core measures and Greenbook/Tealbook information.
2. Which topics in FOMC policy discussions are informative about structural shocks, and which contain little or no relevant information? 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 the inclusion of 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 A New Keynesian DSGE model

We estimate a medium-scale New Keynesian DSGE model following Fernández-Villaverde (2010) and Fernández-Villaverde and Guerrón-Quintana (2021). A continuum of households choose consumption, saving, real money balances, and labor supply, and own all firms. A labor packer aggregates different types of labor into a homogeneous input. A final-goods producer aggregates a continuum of intermediate goods produced by monopolistic competitors who rent capital and labor and set prices according to a Calvo rule. Monetary policy follows an inertial Taylor rule, and government spending follows an exogenous process financed by lump-sum taxes. Both prices and wages are subject to nominal rigidities that limit how often they can be changed. Unit roots in neutral and investment-specific technology drive long-run growth.

The households. Each household j maximizes the following lifetime utility function, which is separable in consumption, c_{jt}, real money balances, M_{jt}/p_t (where p_t is the price

level), and hours worked, l_{jt}:

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

where β is the discount factor, h controls habit persistence, ϑ is the inverse of the Frisch elasticity of labor supply, 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 $\varepsilon_{\phi,t} \sim \mathcal{N}(0,1)$.

Households are monopolistic suppliers of their own type of work j. The household's budget constraint is given by:

$$\begin{aligned} c_{jt} + x_{jt} + \frac{M_{jt}}{p_t} + \frac{b_{jt+1}}{p_t} + \int q_{jt+1,t} a_{jt+1} d\omega_{j,t+1,t} \\ = w_{jt} l_{jt} + \left(r_t u_{jt} - \mu_t^{-1} \Phi\left[u_{jt}\right]\right) k_{jt-1} + \frac{M_{jt-1}}{p_t} + R_{t-1} \frac{b_{jt}}{p_t} + a_{jt} + T_t + \digamma_t, \end{aligned}$$

where w_{jt} is the real wage set by household j, r_t is the real rental price of capital, $u_{jt} > 0$ is the intensity of use of capital, $\mu_t^{-1}\Phi\left[u_{jt}\right]$ 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 $\Phi\left[1\right] = 0$, $\Phi' > 0$ and $\Phi'' > 0$.

Investment x_{jt} induces a law of motion for capital:

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

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 balanced growth path (BGP). The investment-specific technological shock follows the 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.

The labor supplied by each household j is aggregated by a labor supplier with the following production function:

$$l_t^d = \left(\int_0^1 l_{jt}^{\frac{\eta-1}{\eta}} dj\right)^{\frac{\eta}{\eta-1}},$$

where η is the elasticity of substitution among different types of labor and l_t^d is the aggregate

labor demand. The packer maximizes profits, taking wages w_{jt} and w_t as given:

$$\max_{l_{jt}} w_t l_t^d - \int_0^1 w_{jt} l_{jt} dj.$$

Optimality and the zero-profit condition deliver the demand for each differentiated labor type:

$$l_{jt} = \left(\frac{w_{jt}}{w_t}\right)^{-\eta} l_t^d, \qquad \forall j,$$

and the aggregate real wage index:

$$w_t = \left(\int_0^1 w_{jt}^{1-\eta} dj\right)^{\frac{1}{1-\eta}}.$$

We assume households face idiosyncratic wage-adjustment risk through Calvo-style staggered wage setting. In every period, a fraction $1-\theta_w$ of households can reoptimize their nominal wage. The remaining fraction θ_w partially indexes its wage to past inflation at rate $\chi_w \in [0,1]$: if a household cannot reset for τ periods, its nominal wage is scaled by $\prod_{s=1}^{\tau} \Pi_{t+s-1}^{\chi_w}$.

Given Calvo wage setting, the real-wage index evolves as a geometric average of the indexed past real wage and the new optimal real wage:

$$1 = \theta_w \left(\frac{\Pi_{t-1}^{\chi_w}}{\Pi_t}\right)^{1-\eta} \left(\frac{w_{t-1}}{w_t}\right)^{1-\eta} + (1-\theta_w)\left(\Pi_t^{w*}\right)^{1-\eta}.$$

The final-good producer. A final good is 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-good 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 $y_{it} = \left(\frac{p_{it}}{p_t}\right)^{-\varepsilon} y_t^d$ for all 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 a 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, is scaled by $z_t = A_t^{\frac{1}{1-\alpha}} \mu_t^{\frac{\alpha}{1-\alpha}}$. 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-good producers solve a two-stage problem. In the first stage, taking 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. This delivers a marginal cost, common to all producers

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

In the second stage, intermediate-good producers choose prices to maximize discounted real profits subject to Calvo pricing: each period, a fraction $1-\theta_p$ of firms reset prices, while the remainder index their prices to past inflation at a rate $\chi \in [0,1]$. Given Calvo pricing, the price index evolves

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

Monetary policy. The monetary authority sets the nominal interest rate according to a Taylor rule:

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

where Π is the inflation target (equal to inflation on the BGP), R is the BGP nominal gross return on capital, Λ_{y^d} is the BGP gross growth rate of y_t^d, and the monetary policy shock $m_t = \sigma_M \varepsilon_{m,t}$ with $\varepsilon_{m,t} \sim \mathcal{N}(0,1)$. Open market operations are financed through lump-sum transfers:

$$T_t = \frac{M_t}{p_t} - \frac{M_{t-1}}{p_t} + \frac{b_{t+1}}{p_t} - R_{t-1} \frac{b_t}{p_t}.$$

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

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

and $z_t = A_t^{\frac{1}{1-\alpha}} \mu_t^{\frac{\alpha}{1-\alpha}}$. Government consumption is also financed by lump-sum taxes, ensuring that the deficit is zero at all times.

Aggregation. Applying the definition of transfers above and the zero-profit condition for the labor packer, the aggregated budget constraint of households is equal to $c_t + x_t = w_t l_t^d + \left(r_t u_t - \mu_t^{-1}\Phi\left[u_t\right]\right) k_{t-1} + \digamma_t - g_t$. Aggregate labor is: $l_t = \left(\int_0^1 \left(\frac{w_{jt}}{w_t}\right)^{-\eta} dj\right) l_t^d = v_t^w l_t^d$, where the wage dispersion v_t^w is given by $v_t^w = \theta_w \left(\frac{w_{t-1}}{w_t}\frac{\Pi_{t-1}^{\chi_w}}{\Pi_t}\right)^{-\eta} v_{t-1}^w + (1-\theta_w)\left(\Pi_t^{w*}\right)^{-\eta}$. 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 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 Supplementary Material A. There, we also describe how we make the model stationary by rescaling the equilibrium conditions, the steady state of the rescaled equilibrium conditions, and the loglinearization of the equilibrium conditions around the steady state. After loglinearization, we finally cast the model in a state-space representation summarized in Appendix A.

Phillips curves. Finally, loglinearizing around the zero inflation steady state, we obtain the New Keynesian Phillips curve:

$$\widehat{\Pi}_t = \frac{\chi_P}{1+\beta\chi_P}\widehat{\Pi}_{t-1} + \frac{\beta}{1+\beta\chi_P}\mathbb{E}_t\widehat{\Pi}_{t+1} + \kappa_p \widehat{mc}_t$$

with slope coefficient $\kappa_p = \frac{(1-\theta_P)(1-\beta\theta_P)}{\theta_P(1+\beta\chi_P)}$; and the wage Phillips curve:

$$\widehat{\Pi}_t^w = \frac{\chi_w}{1+\beta\chi_w}\widehat{\Pi}_{t-1} + \frac{\beta}{1+\beta\chi_w}\mathbb{E}_t\widehat{\Pi}_{t+1}^w - \frac{\beta\chi_w}{1+\beta\chi_w}\widehat{\Pi}_t + \kappa_w \widehat{wgap}_t$$

with slope coefficient $\kappa_w = \frac{(1-\theta_w)(1-\beta\theta_w)}{\theta_w(1+\beta\chi_w)(1+\eta\vartheta)}$.

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 ($\triangle \log \mu_t^{-1}$),

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

As a robustness check, we also estimate the model by splicing the short-term nominal interest rate with the shadow short rates from Krippner (2015) and Wu and Xia (2016) during the zero-lower-bound period (2009Q1–2015Q4) in Supplementary Material E.

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 an 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 for comparing the incremental value of the text-based measures.

All series are sampled quarterly. 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 Material 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 near-complete account of the discussion. The transcripts are released with a substantial lag, typically about five years after the meeting. 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" portion of each meeting, in which 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 for model estimation. Below, we propose an approach for selecting the most informative part(s).

From each dataset, we extract a structured set of quarterly time series. As is typical with unstructured data, this requires a reduction in dimensionality. We consider two algorithms. The first, and the focus of our analysis, is the LDA topic model. LDA assumes each document is a mixture of a fixed number of topics, with each word generated by drawing a topic from the document's distribution and then a word from the topic's distribution. Each transcript is summarized by a vector of topic shares that uncovers the main themes in FOMC discussions and quantifies how prominently each appears.

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 counts of individual terms, 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 of our dataset, they also appear to contain less information for model estimation. For that reason, we focus on LDA in the main analysis.

LDA estimation. After pre-processing, we estimate a single LDA model on the combined corpus at the level of individual speaker utterances for $K = 20, 30, 40$ topics using the MCMC strategy of Griffiths and Steyvers (2004) implemented in the R `topicmodels` package.[4] For each topic count, we initialize three Markov chains from random seeds and retain the chain with the highest average post-burn-in log-likelihood.

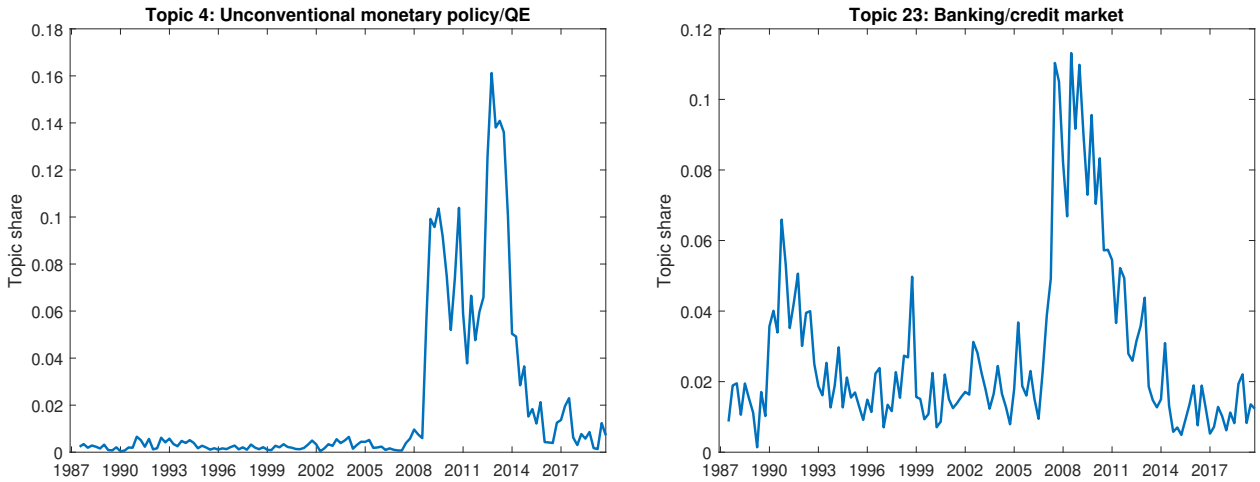
To obtain quarterly topic shares for the FOMC1, FOMC2, and combined datasets sep-

[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.

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

arately, we re-estimate LDA after aggregating the relevant individual speakers' utterances into a single document per quarter, holding the topic-term distributions fixed at their posterior means from step one. The resulting quarterly topic shares $\{\varphi_{1,t}, ..., \varphi_{K,t}\}$ are posterior means 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 15 terms with the highest probability for each topic for the combined FOMC1 and FOMC2 datasets with $K = 40$. As we explain below, this is the dataset and model specification we select for the application, so all references to specific topics should hereafter be understood to reference this setting. In some cases, the top-probability words within a topic convey a coherent theme that we manually use to assign a label to the topic. For instance, Topic 8 contains "unemploy," "labor," and "employ," so we interpret the topic as representing labor market. Topic 4 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 17 contains "panel," "left," and "chart," so we interpret this topic as presentation and explanation of figures, which is possibly irrelevant to model estimation. We expect that our econometric framework disregards conversational topics while capturing information from economically relevant topics, such as Topic 4.



Note: Topic-share series for Topics 4 and 23, estimated from the topic model with $K = 40$ using both FOMC1 and FOMC2 transcript data.

Figure 1: Time series of Topics 4 and 23

The time-series dynamics of these topics support the validity of our labels. Figure 1 presents the time-series behavior of Topics 4 and 23, estimated using the topic model with

$K = 40$ and both FOMC1 and FOMC2 datasets. Topic 4 (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. Topic 23 (Banking/Credit Market) exhibits a sharp increase following the global financial crisis. These movements align closely with standard narratives of U.S. economic and policy conditions over the sample period.

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 conduct a pseudo-out-of-sample forecasting exercise to select the combination of corpus and K that yields 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 real-time data archive.[5] The dataset contains 15 macroeconomic series, including real GDP, the GDP price index, and unemployment.[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 ($T-1, T, T+1, T+2, T+$

[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.

Table 1: Top-probability words by topic (40-topic model)

Label	Topic	Top1	Top2	Top3	Top4	Top5	Top6	Top7	Top8	Top9	Top10	Top11	Top12	Top13	Top14	Top15
Recession / Recovery	Topic1	economi	recess	recoveri	weak	much	econom	negat	shock	confid	turn	period	fall	begin	cours	sever
Conversation (QA)	Topic2	question	ask	know	look	thought	want	answer	differ	right	someth	tri	start	let	guess	wonder
Policy statement	Topic3	committe	altern	econom	may	indic	might	consist	federalfundsratexx	statement	object	page	note	period	maintain	paragraph
Unconventional policy / QE	Topic4	purchas	program	billion	end	expect	effect	condit	provid	treasuri	improv	market	cost	balancesheetxx	threshold	addit
Policy stance	Topic5	move	direct	point	support	tighten	recommend	agre	meet	toward	asymmetr	prefer	symmetr	eas	basispointsxx	want
Regional economy	Topic6	district	nation	area	region	continu	economi	construct	manufactur	employ	sector	state	greenbook	report	industri	strong
Business / Beige Book	Topic7	report	busi	contact	compani	firm	hear	cost	director	plan	worker	increas	heard	larg	talk	ceo
Labor market	Topic8	unemploy	labor	employ	labormarketxx	job	unemploymentratexx	forc	worker	trend	declin	past	slack	wage	level	work
Fiscal policy	Topic9	fiscal	state	tax	europ	govern	effect	european	budget	polit	problem	debt	much	fiscalpolicyxx	situat	crisi
Data release	Topic10	data	month	revis	septemb	june	march	number	octob	decemb	littl	read	indic	last	show	juli
Conversation	Topic11	seem	may	reason	might	certain	quit	point	clear	case	much	perhap	probabl	suggest	bit	though
Conversation	Topic12	realli	actual	thing	want	idontxx	someth	lot	right	much	look	term	know	tri	talk	littlebitxx
Trade	Topic13	dollar	unit	state	export	import	countri	foreign	growth	japan	project	abroad	forecast	depreci	declin	currenc
Inflation / Cost pressures	Topic14	inflat	increas	growth	product	wage	price	rise	cost	slow	pressur	trend	acceler	economi	demand	continu
Generic	Topic15	look	lot	peopl	thing	realli	side	littl	situat	number	talk	term	much	problem	sens	certain
Data / Statistics	Topic16	percent	quarter	growth	month	averag	last	end	next	around	number	third	look	half	estim	past
Presentation / Explanation	Topic17	panel	shown	line	right	left	expect	chart	show	middl	next	bottom	top	exhibit	red	can
Conversation	Topic18	know	laughter	can	peopl	world	work	cannot	great	happen	good	mean	want	call	use	come
ZLB / Forward guidance	Topic19	polici	zero	target	rule	inflat	committe	lower	bound	approach	monetarypolicyxx	can	low	today	possibl	use
Greenbook forecast	Topic20	forecast	greenbook	staff	scenario	assumpt	assum	baselin	chang	period	differ	much	end	view	somewhat	littl
Policy path	Topic21	chair	madam	move	gradual	path	support	target	continu	increas	polici	accommod	appropri	meet	neutral	economi
Household consumption / Investment	Topic22	spend	hous	invest	consum	household	incom	busi	consumpt	capit	sector	growth	demand	relat	declin	wealth
Banking / Credit Market	Topic23	bank	credit	loan	financi	fund	market	lend	mortgag	asset	spread	capit	liquid	debt	borrow	concern
Risk / Outlook	Topic24	risk	balanc	outlook	inflat	concern	downsid	growth	view	upsid	potenti	expect	signific	seem	uncertainti	bit
Manufacturing / Sales	Topic25	sale	product	price	industri	continu	good	inventori	level	manufactur	increas	auto	last	order	district	strong
Policy action	Topic26	polici	might	tighten	market	may	inflat	eas	committe	meet	staff	action	possibl	fundsratexx	potenti	seem
Conversation	Topic27	yes	governor	okay	laughter	liftoff	vice	lacker	evan	fisher	hoenig	bullard	stern	parri	break	minehan
SEP / Guidance language	Topic28	project	percent	tealbook	unemploymentratexx	forecast	inflat	end	unemploy	median	lower	assess	longer-run	sep	estim	chang
Policy credibility	Topic29	need	polici	believ	can	action	today	economi	credibl	import	act	futur	public	monetarypolicyxx	risk	view
Growth outlook	Topic30	project	forecast	expect	quarter	next	anticip	growth	recent	spend	activ	greenbook	percent	revis	percentagepointxx	consider
Positive economic indicator	Topic31	continu	remain	month	expect	improv	report	increas	firm	indic	recent	survey	pace	activ	growth	posit
Global economy	Topic32	economi	global	financi	trade	growth	uncertainti	risk	recent	remain	china	outlook	market	continu	effect	develop
Energy prices	Topic33	price	oil	energi	increas	effect	expect	suppli	higher	commod	oilpricesxx	import	product	continu	core	declin
Model analysis	Topic34	model	use	estim	gap	differ	output	measur	effect	shock	staff	analysi	actual	natur	simul	relationship
Inflation expectations	Topic35	inflat	expect	percent	measur	inflationexpectationsxx	low	-	core	term	year	coreinflationxx	past	anchor	yield	pce
Conversation	Topic36	discuss	point	comment	issu	let	talk	tabl	around	mention	want	import	agre	heard	particular	forward
Policy effect	Topic37	effect	real	economi	interestratesxx	expect	relat	nomin	aggreg	money	short-term	declin	long-term	adjust	dollar	shift
Financial market expectations	Topic38	market	come	much	expect	look	view	move	probabl	real	last	economi	surpris	suggest	seem	around
Generic	Topic39	can	signific	move	look	number	may	clear	evid	issu	inde	know	probabl	economi	stage	point
Policy statement	Topic40	statement	languag	chang	altern	word	meet	support	paragraph	communic	suggest	alternativebxx	use	sentenc	today	want

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

$3, T+4$). We then estimate a small number of factors from these 90 series. For details of factor estimation, see Supplementary Material 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):

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

The vector $state_t = \left(\widehat{\Pi}_t, \widehat{\widetilde{w}}_t, \widehat{g}^1_t, \widehat{g}^2_t, \widehat{\widetilde{k}}_t, \widehat{R}_t, \widehat{\widetilde{y}}^d_t, \widehat{\widetilde{c}}_t, \widehat{v}^p_t, \widehat{v}^w_t, \widehat{\widetilde{q}}_t, \widehat{F}_t, \widehat{\widetilde{x}}_t, \widehat{\widetilde{\lambda}}_t, \widehat{\widetilde{z}}_t\right)'$ contains the main endogenous state variables and the vector $exo_t = \left(\widehat{\widetilde{\mu}}_t, \widehat{d}_t, \widehat{\phi}_t, \widehat{\widetilde{A}}_t, m_t, \widehat{\widetilde{g}}_t\right)'$ collects the exogenous processes.

In total, S_t is a 27×1 vector. The transition matrices $T(\psi)$ and $R(\psi)$ are known functions of the structural parameter vector ψ, and we estimate 30 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, whose prior distributions are summarized in Table 2. Following Fernández-Villaverde and Guerrón-Quintana (2021), we calibrate $\vartheta = 1$, $\delta = 0.025$, $\varepsilon = 10$, $\eta = 10$, $f = 0$, 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

$$Y^C_t = H^C_0(\theta) + H^C_1(\theta)S_t + u_t, \quad u_t \sim \mathcal{N}(0, \Sigma^C_u(\theta)). \tag{11}$$

For example, the first element of Y^C_t is observed inflation, $\log \Pi^{obs}_t$, 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.

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
θ_W	Fraction of households with fixed wages	$(0, 1)$	Beta	0.5	0.1	0.01	0.9999
χ_W	Wage indexation	$(0, 1)$	Beta	0.5	0.15	0.01	0.9999
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
Measurement error parameters							
σ_{ME}	SD all measurement errors	$\mathbb{R}^+$	Inverse gamma	0.01	0.02	1e-5	100

Note: 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.

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 D such that $DS_t = \varepsilon_t = [\varepsilon_{D,t}, \varepsilon_{\phi,t}, \varepsilon_{G,t}, \varepsilon_{\mu,t}, \varepsilon_{A,t}, \varepsilon_{M,t}]'$. That is, the selection matrix picks out the six structural shocks of the DSGE model. We then link the topic-based 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 of the current state of the economy, some of that information may already be summarized numerically 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

[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 a smaller set of informative topics drives the effective information content of the non-core block.

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 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 Supplementary Material 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 h-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 square 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 performance for individual core variables.

Table 3: Joint log scores across model variants

# of topic	**Transcript**	$h = 1$	$h = 4$	$h = 8$
Without text	–	1901.2	1625.9	1406.7
With Greenbook	–	1889.0	1610.3	1384.8
$K = 20$	FOMC 1	1899.5	1631.3	1427.0
	FOMC 2	1896.6	1636.7	1450.5
	Both	1900.4	1626.7	1439.5
$K = 30$	FOMC 1	1896.7	1635.3	1432.4
	FOMC 2	1898.6	1642.6	1447.5
	Both	**1903.5**	1640.2	1440.5
$K = 40$	FOMC 1	1895.1	1619.8	1418.4
	FOMC 2	1892.2	1631.6	1452.1
	Both	1902.2	**1647.6**	**1467.2**
OpenAI embeddings	Both	1897.5	1623.8	1411.5

Note: Joint log scores for models without text, with Greenbook, and with FOMC transcripts, across different numbers of topics ($K = 20, 30, 40$) and transcript types (FOMC1, FOMC2, and both), and OpenAI embeddings using both FOMC transcripts with 40 dimensions via PCA. The forecasting horizons are $h = 1, 4, 8$. The model with the highest score is shown in bold.

Forecast evaluation results: Joint log score. Table 3 reports the joint log scores of h = 1-,4-,8-period-ahead forecasts 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, forecasting performance at the short horizon ($h = 1$) does not vary substantially across specifications. The specifications using the combined FOMC1 and FOMC2 corpus with $K = 30$ and $K = 40$ achieve higher scores than the benchmark model without text, but the differences are modest.

Second, the specifications that incorporate FOMC transcripts outperform the benchmark

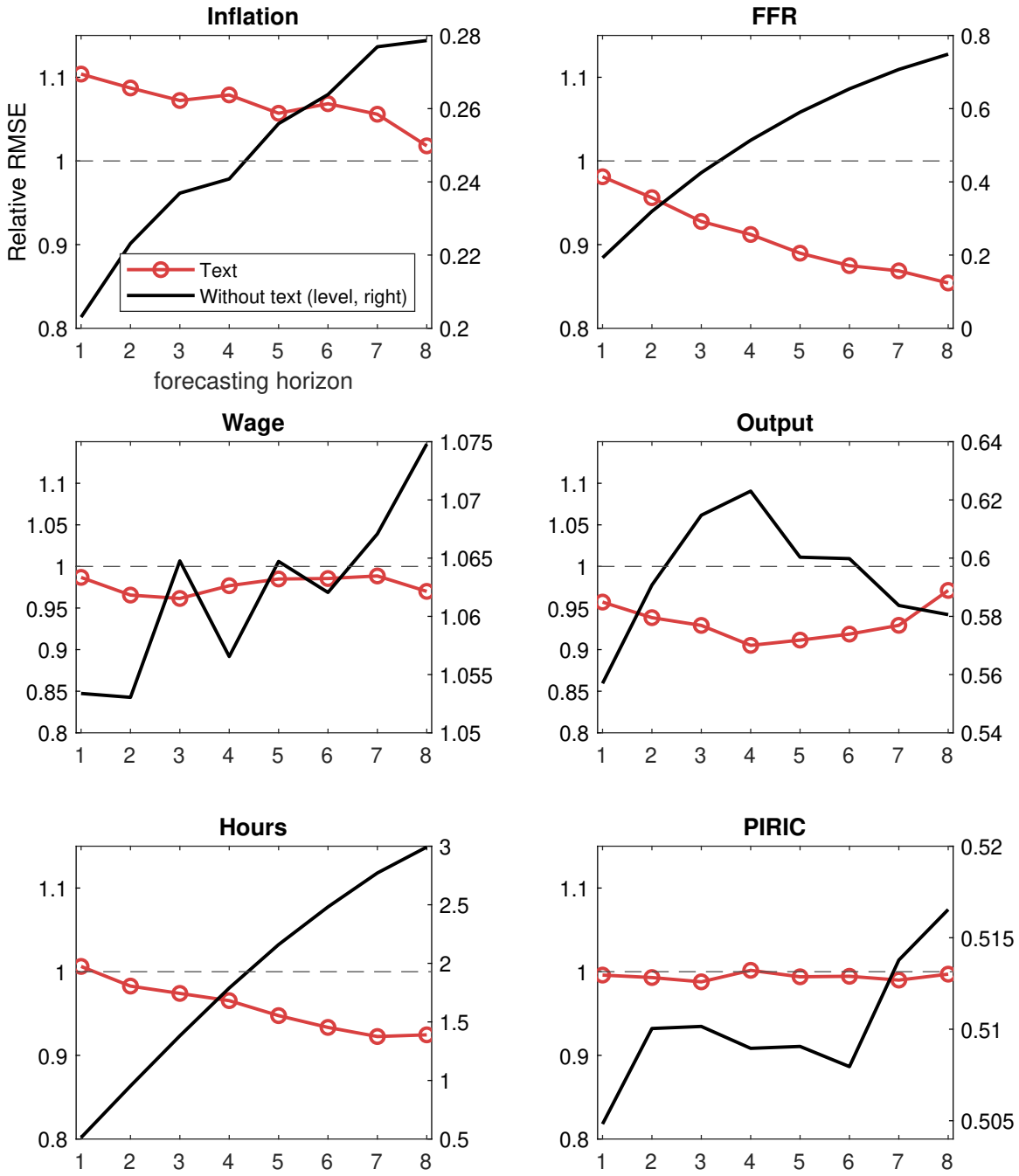
model without text at medium horizons ($h = 4, 8$). The specification with $K = 40$ estimated using the combined FOMC1 and FOMC2 corpus achieves the highest joint log score at both $h = 4$ and $h = 8$, while also remaining competitive at $h = 1$. Hence, we use this specification as our baseline model in the full-sample estimation in the next section. The finding that text improves forecasting performance mainly at medium horizons is consistent with the nature of FOMC transcripts: they contain participants' assessments of the outlook, risks, and policy path, which are likely to be more informative for forecasting over several quarters than for one-quarter-ahead nowcasting.

Third, both the economic-situation and policy-discussion portions of the transcripts contain useful forecasting information. All models that incorporate FOMC transcripts, except for the specification using FOMC1 with $K = 40$, outperform the benchmark model without text in terms of the joint log score at $h = 4$, and all text-based specifications outperform the benchmark at $h = 8$, regardless of K and transcript subset. The fact that the best-performing specification combines FOMC1 and FOMC2 further suggests that the two portions of the transcripts provide complementary information for DSGE forecasting.

Fourth, the Greenbook-based specification performs worse than all 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. 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 suggests that including too much redundant information can lower forecast accuracy, even with our spike-and-slab prior.

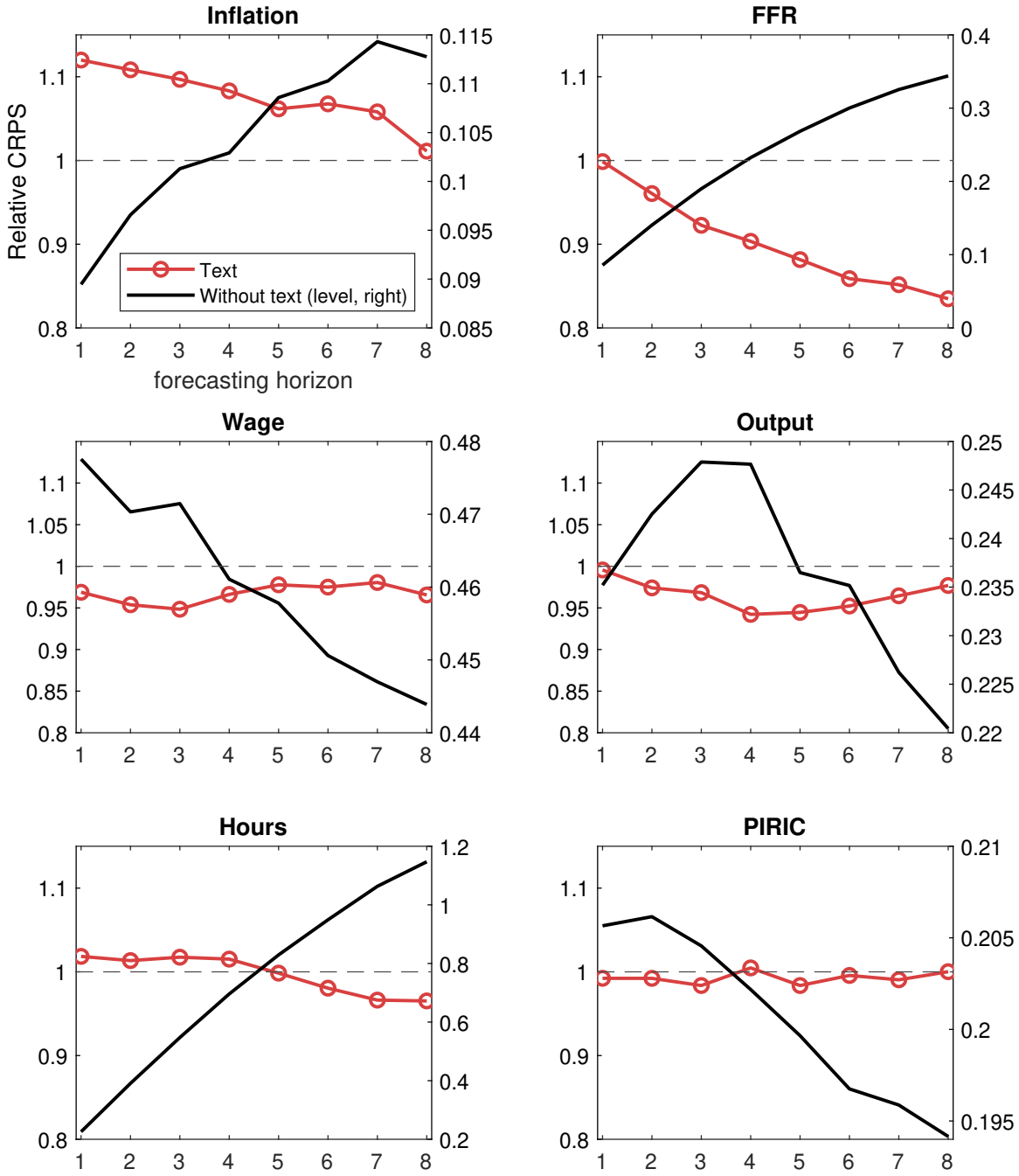
Finally, we compare LDA with modern embedding models. We embed the full text of both sections of each meeting using OpenAI's `text-embedding-3-large` model, obtain quarterly embeddings by averaging, then reduce dimensionality via PCA and retain the top forty components.[8] We focus on both FOMC 1 and 2 transcripts and 40 dimensions to align with the most informative topic-share series. The embedding-based specification performs worse than the no-text benchmark at $h = 1$, becomes comparable to the benchmark at $h = 4$, and exceeds it at $h = 8$. However, it does not outperform the topic-based specification. These results suggest that, in this application, off-the-shelf embeddings contain useful medium-horizon information, but LDA topic shares provide a more effective low-dimensional representation for DSGE estimation and forecasting.

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



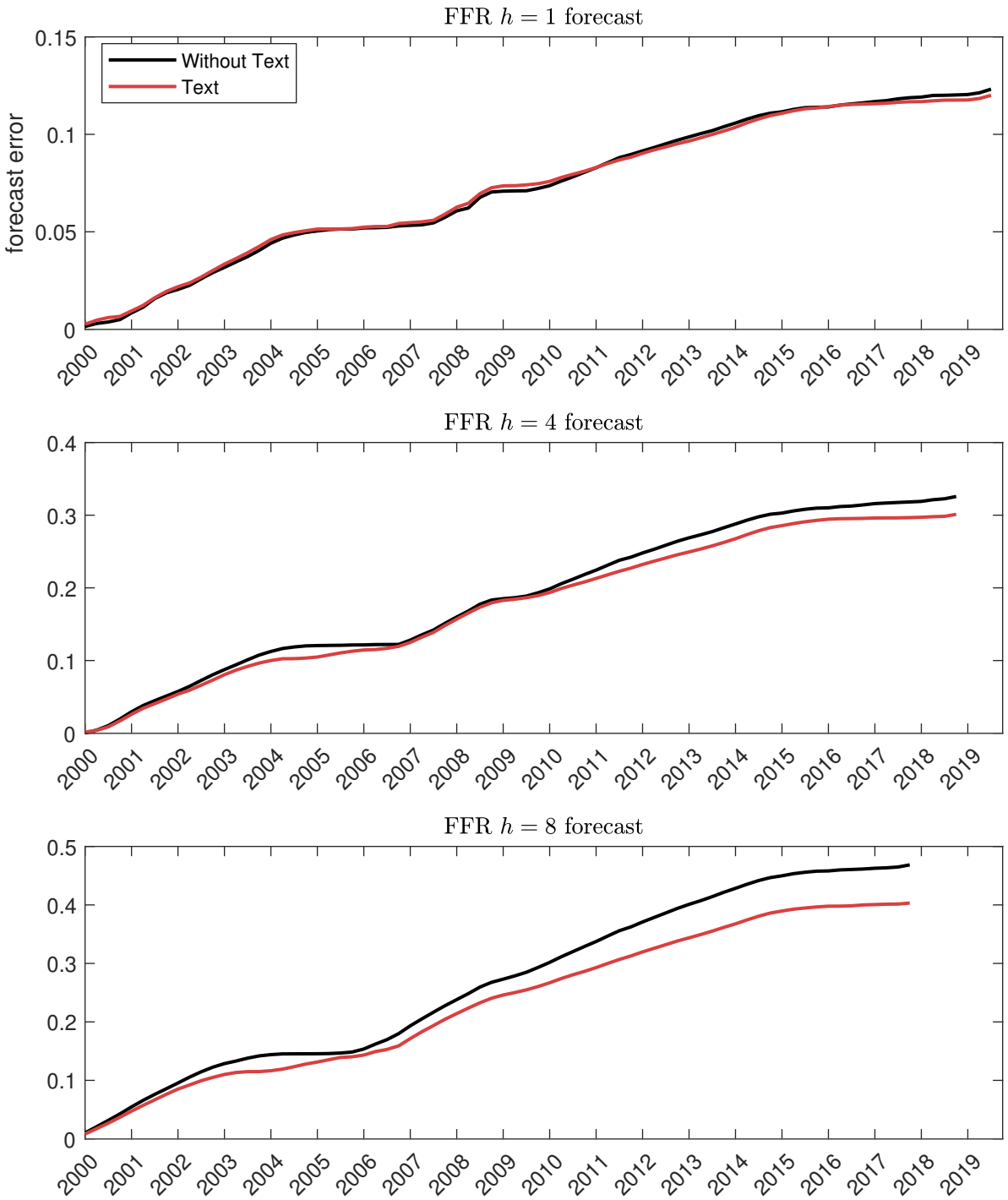
Note: RMSE across different horizons for each variable in the observation equation. The left axis shows the relative RMSE. Values smaller than 1 indicate better forecasting performance for the model than for the model without text. The right axis shows the RMSE of the model without additional data. The red line represents the model with FOMC transcripts. The black line shows the level of RMSE of the model without text on the right axis.

Figure 2: (Relative) root mean square error



Note: Continuous ranked probability score across different horizons for each variable in the observation equation. The left axis shows the relative CRPS. Values smaller than 1 indicate better forecasting performance for the model than for the model without text. The right axis shows the CRPS of the model without additional data. The red line represents the model with FOMC transcripts. The black line shows the level of CRPS of the model without text on the right axis.

Figure 3: (Relative) continuous ranked probability score



Note: Cumulative absolute error of federal funds rate forecasts from models without text and with FOMC transcripts. 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

Forecast evaluation results: Individual variable RMSE and CRPS. Figures 2 and 3 report the relative RMSE and CRPS for models augmented with the FOMC transcript data, 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, which uses the combined FOMC1 and FOMC2 corpus with $K = 40$.

The text-augmented model improves forecasting performance for most core variables, especially at medium horizons. The largest gains are for the federal funds rate and hours worked, where the relative RMSE falls steadily with the forecast horizon and reaches reductions of more than 10 percent at longer horizons. Output forecasts also improve substantially, with RMSE reductions of roughly 10 percent at medium horizons. Wage forecasts improve more modestly, while the relative price of investment goods is broadly comparable to the benchmark. By contrast, the text-augmented model performs worse for inflation, particularly at short horizons. The CRPS results closely mirror the RMSE patterns, indicating that these improvements also hold when evaluating the full predictive distribution.

These variable-level results help interpret the joint log-score findings. Although the joint log score evaluates the multivariate predictive density and is therefore not a simple aggregation of the individual RMSE or CRPS values, the pattern is consistent with the aggregate results. At the short horizon, the deterioration in inflation forecasts offsets part of the gains for the federal funds rate, output, wages, and hours, leaving only modest differences in the joint log score. At medium horizons, however, the gains for the federal funds rate, hours worked, and output become larger and more persistent, while the relative deterioration in inflation forecasts becomes smaller. As a result, the predictive information in the FOMC transcripts becomes more visible in the joint forecast evaluation. Economically, this suggests that the transcripts are more informative about the medium-horizon policy path and real activity outlook than about transitory one-quarter-ahead inflation movements.

The text-augmented model is particularly effective at forecasting the federal funds rate at medium horizons. This is intuitive because the preferred specification includes FOMC2, the policy-discussion portion of the transcripts, where participants discuss policy alternatives, risks to the outlook, forward guidance, and the appropriate policy path. These discussions provide information about the systematic component of monetary policy that is not fully captured by the standard macroeconomic observables alone. Moreover, the improvement is not driven by a single episode. Figure 4 shows the cumulative absolute forecast error for the federal funds rate. The gap between the text-augmented model and the benchmark widens gradually over time at $h = 4$ and $h = 8$, indicating that the text-based specification delivers persistent gains in interest-rate forecast accuracy throughout the evaluation period.

In summary, FOMC transcripts improve forecasting performance relative to both the benchmark model with only standard macroeconomic variables and the Greenbook-based specification. These results indicate that unstructured policy and economic discussions contain information relevant for DSGE estimation and forecasting that is not fully captured by conventional macroeconomic observables or numerical staff projections.

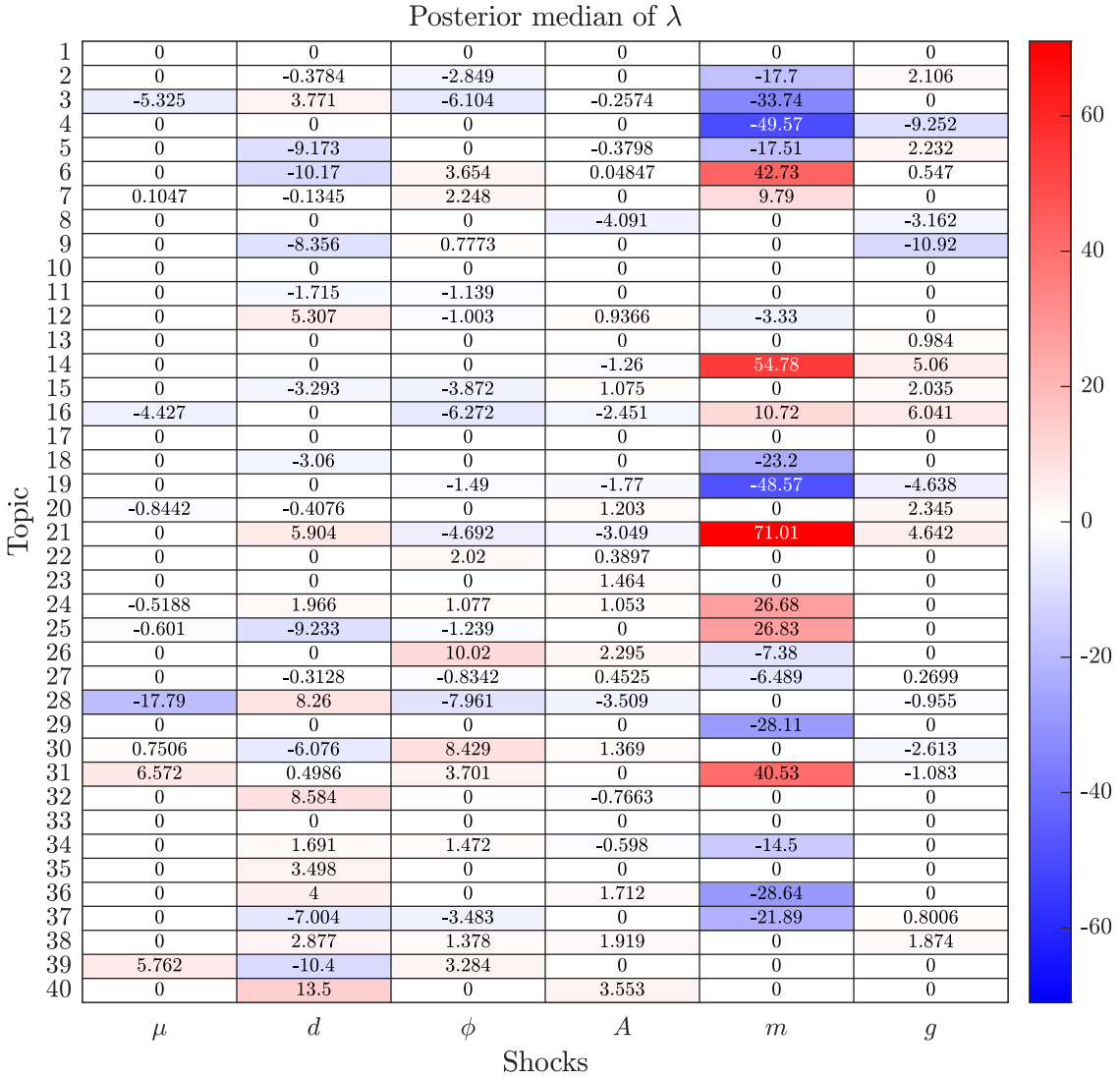
3.4.2 Which topics in the FOMC transcript are informative?

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, which uses the combined FOMC1 and FOMC2 corpus with $K = 40$. Figure 5 reports the posterior median of the factor loadings in Λ, showing how each topic-share series loads on the six structural shocks. Because the spike-and-slab prior assigns positive probability to a zero loading, many entries of Λ are shrunk to zero in the posterior, yielding a sparse pattern of topic-shock connections.

Several findings emerge. First, the FOMC transcripts provide information clearly linked to structural shocks. The loadings on the monetary policy shock m are particularly informative. The shock loads *positively* on Topic 6 (Regional economy), Topic 14 (Inflation and cost pressures), Topic 21 (Policy path), Topic 24 (Risk / Outlook), Topic 25 (Manufacturing / Sales) and Topic 31 (Positive economic indicator), and *negatively* on Topic 3 (Policy statement), Topic 4 (Unconventional policy / QE), Topic 5 (Policy stance), Topic 19 (ZLB / Forward guidance), Topic 29 (Policy credibility), and Topic 37 (Policy effect). The positive group corresponds to a hawkish regime: tightening episodes are associated with greater discussion of inflation, strong activity, district conditions, and statistical justification. The negative group corresponds to an accommodation regime: when policy is maintained or eased, transcripts emphasize statement mechanics, guidance, credibility, and balance-sheet tools. Figure 1 confirms this interpretation: Topic 4 (Unconventional policy / QE) rises sharply after the Great Recession, when policy was explicitly accommodative, consistent with a negative loading on m.

Second, some topics load on several shocks simultaneously. Topic 14 (Inflation and cost pressures), for instance, loads on both the aggregate technology shock A and the government spending shock g, which also captures an exogenous demand shift, as well as the monetary policy shock. This reflects that both demand- and supply-side developments affect marginal costs and relative prices, prompting the Committee to discuss inflationary pressures regardless of the shock's origin.

Third, incorporating topics into structural estimation strengthens the interpretation of the topics themselves. Topic modeling is unsupervised, so economic labels are typically assigned



Topic	μ	d	ϕ	A	m	g
1	0	0	0	0	0	0
2	0	-0.3784	-2.849	0	-17.7	2.106
3	-5.325	3.771	-6.104	-0.2574	-33.74	0
4	0	0	0	0	-49.57	-9.252
5	0	-9.173	0	-0.3798	-17.51	2.232
6	0	-10.17	3.654	0.04847	42.73	0.547
7	0.1047	-0.1345	2.248	0	9.79	0
8	0	0	0	-4.091	0	-3.162
9	0	-8.356	0.7773	0	0	-10.92
10	0	0	0	0	0	0
11	0	-1.715	-1.139	0	0	0
12	0	5.307	-1.003	0.9366	-3.33	0
13	0	0	0	0	0	0.984
14	0	0	0	-1.26	54.78	5.06
15	0	-3.293	-3.872	1.075	0	2.035
16	-4.427	0	-6.272	-2.451	10.72	6.041
17	0	0	0	0	0	0
18	0	-3.06	0	0	-23.2	0
19	0	0	-1.49	-1.77	-48.57	-4.638
20	-0.8442	-0.4076	0	1.203	0	2.345
21	0	5.904	-4.692	-3.049	71.01	4.642
22	0	0	2.02	0.3897	0	0
23	0	0	0	1.464	0	0
24	-0.5188	1.966	1.077	1.053	26.68	0
25	-0.601	-9.233	-1.239	0	26.83	0
26	0	0	10.02	2.295	-7.38	0
27	0	-0.3128	-0.8342	0.4525	-6.489	0.2699
28	-17.79	8.26	-7.961	-3.509	0	-0.955
29	0	0	0	0	-28.11	0
30	0.7506	-6.076	8.429	1.369	0	-2.613
31	6.572	0.4986	3.701	0	40.53	-1.083
32	0	8.584	0	-0.7663	0	0
33	0	0	0	0	0	0
34	0	1.691	1.472	-0.598	-14.5	0
35	0	3.498	0	0	0	0
36	0	4	0	1.712	-28.64	0
37	0	-7.004	-3.483	0	-21.89	0.8006
38	0	2.877	1.378	1.919	0	1.874
39	5.762	-10.4	3.284	0	0	0
40	0	13.5	0	3.553	0	0

Note: The figure reports the posterior median of the factor loading of each transformed topic-share series on each structural shock in the DSGE model. Rows correspond to FOMC transcript topics, and columns correspond to structural shocks. μ: investment-specific technology shock, d: intertemporal preference shock, ϕ: labor supply shock, A: aggregate technology shock, m: monetary policy shock, g: government spending shock. For the labels of the topics, see Table 1.

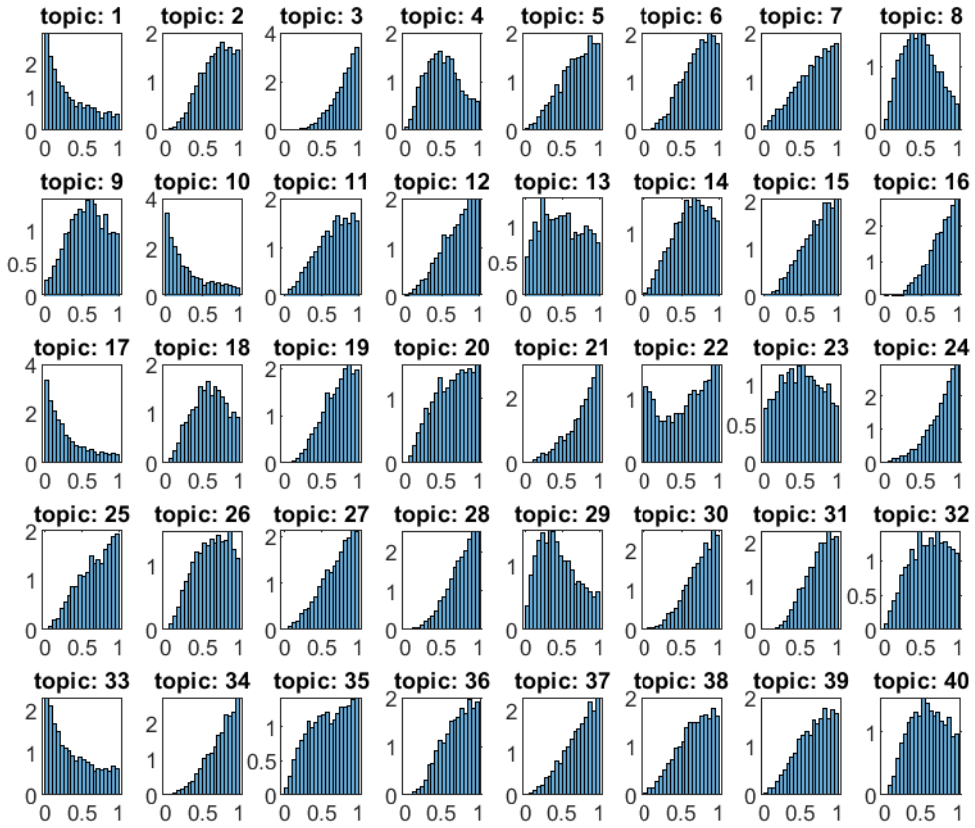
Figure 5: Posterior median of topic-shock loadings (λ)

heuristically from vocabulary and time-series patterns. The estimated loading matrix Λ adds a structural layer by showing which topics help measure each structural shock. This is particularly useful for topics related to monetary policy. Topic 14 ("inflat", "wage", "price", "cost", "pressur") and Topic 21 ("gradual", "path", "increas", "polici")[9] load positively on m, consistent with contractionary policy episodes: Topic 14 captures inflation and cost pressures associated with tightening, while Topic 21 captures policy normalization and the gradual path of rate increases. By contrast, Topic 4 ("purchas", "program", "balance sheet") loads negatively on m, capturing unconventional easing, and Topic 19 ("zero", "target", "polici") captures forward guidance at the zero lower bound. Joint estimation with the DSGE model yields a structural taxonomy that distinguishes inflation-driven tightening, policy normalization, unconventional easing, and forward guidance based on shock loadings rather than on vocabulary alone.

Fourth, the results indicate that FOMC members routinely discuss issues that lie outside the scope of our benchmark DSGE model. For example, Topic 13 (Trade) and Topic 33 (Energy prices) capture themes that are economically meaningful but exogenous to the current closed economy without explicit energy sectors. Therefore, the corresponding factor loadings of Topic 33 are estimated to be essentially zero with high posterior probability, and Topic 13 loads only slightly on g. This is a desirable feature: the sparsity mechanism correctly identifies topics relevant to the structural model at hand and excludes those that fall outside its theoretical scope, providing a safeguard against overfitting.

While the estimation procedure identifies interpretable associations between topics and monetary policy shocks, some linkages with other shocks are less straightforward. For example, Topic 9 (Fiscal policy) loads negatively on the government spending shock ε_g, implying that a positive innovation to the g_t process is associated with a lower relative share of fiscal-policy discussion in the transcripts. This is not necessarily inconsistent with the model, because g_t in a medium-scale New Keynesian model should not be interpreted only as literal government expenditure. In practice, it also acts as a residual demand shifter, absorbing movements in output that are not explained by the other structural shocks. A positive g_t innovation may therefore reflect non-fiscal demand shifts, which need not coincide with greater

[9] Although some of Topic 21's highest-probability words, such as "chair" and "madam" are institutional rather than economic, other high-probability words are associated with policy normalization and rate increases. The prominence of "madam" is consistent with the Yellen chairmanship, and the topic becomes more prominent around the beginning of the policy-normalization period. Associated transcript passages support this interpretation. For example, associated transcript passages describe policy as requiring a "gradual shift from accommodative to neutral, a path that is consistent with considering another rate increase in December."(George), note that "Financial conditions remain accommodative despite the gradual path of rate increases the Committee has put in place." (Mester), and refer to continuing on a "gradual tightening path next year" after raising the target range (Yellen).



Note: 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

discussion of fiscal policy. Such discrepancies are informative: they can serve as diagnostics for model misspecification or shock mislabeling and point toward richer model specifications.

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 is close to one, the corresponding topic is likely to have non-zero loadings on the structural shocks; when q is close to zero, the topic is effectively excluded from the non-core measurement equation. 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 33 (Energy prices), have posterior mass concentrated near zero. By contrast, Topic 4 (Unconventional MP / QE), Topic 14 (Inflation / Cost pressures), Topic 19 (ZLB / Forward guidance), and Topic 21 (Policy path) have

posterior inclusion probabilities concentrated near one, confirming that the estimation treats these topics as systematically informative.

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. These results are robust to using the shadow short rate during the zero-lower-bound period, as shown in Supplementary Material E.1.

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

We now turn to structural parameter estimates and impulse responses. The previous subsections show that FOMC transcripts improve forecasting performance and that selected topics load systematically on the structural shocks. We now ask whether this additional information changes the posterior distribution of the DSGE parameters and, in turn, the model-implied propagation of shocks. In the main text, we focus on the parameters that govern monetary policy, nominal rigidities, and real-side propagation; the remaining posterior distributions are reported in Supplementary Material D.

Figure 7 plots the posterior distributions of these DSGE parameters for two specifications: the benchmark model without non-core variables (black), and the text-augmented model using FOMC transcripts (red). This figure shows that the inclusion of FOMC transcripts alters the posterior distributions of several economically important structural parameters, particularly in the monetary-policy rule and nominal-rigidity blocks. In addition, the additional information from the transcripts reduces posterior uncertainty for a subset of parameters, tightening their credible intervals relative to the benchmark.

Relative to the benchmark model without text, the text-augmented model estimates a lower response of the policy rate to inflation, a higher response to output growth, and somewhat less interest-rate smoothing. It also estimates higher price stickiness and lower price indexation. These two changes work against each other on the slope: higher stickiness flattens the price Phillips curve, while lower indexation, on its own, would steepen it. Stickiness dominates, so the curve comes out flatter; lower indexation also leaves it less backward-looking. Wage-setting parameters move in the opposite direction along the two nominal-adjustment margins: the model estimates lower wage stickiness but higher wage indexation, suggesting that the transcript data alter how the model allocates nominal adjustment between prices and wages.

These shifts are consistent with the way transcript topics enter the non-core measurement equation. In the benchmark model, policy-rate movements must be explained using only core macroeconomic observables and the Taylor-rule residual. Once transcripts are included, topics related to the policy path, forward guidance, unconventional policy, inflation pressures, and the outlook provide additional information about policy-related shocks and communication. Some variation that would otherwise be absorbed by the systematic Taylor-rule coefficients is therefore reallocated to the monetary-policy shock, lowering the estimated response to inflation while increasing the estimated response to real activity. Similarly, the higher θ_P and lower χ_P imply that inflation dynamics in the text-augmented model are less mechanically driven by lagged inflation and less responsive on impact to marginal-cost movements.

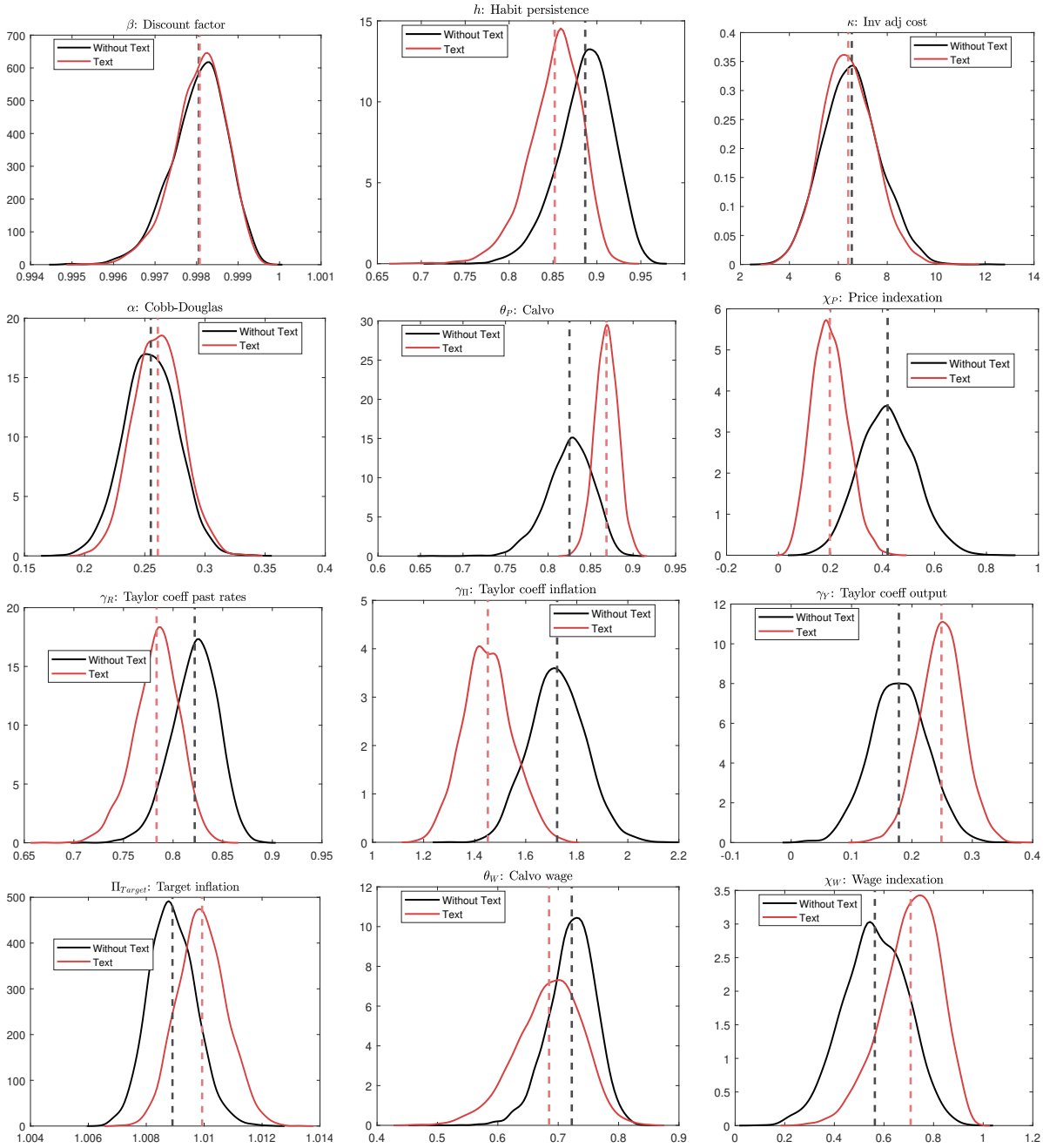
We next examine whether these posterior shifts translate into different impulse responses. We focus on three shocks that are most directly connected to the parameter changes discussed above: the labor-supply shock ε_ϕ, the neutral technology shock ε_A, and the monetary-policy shock ε_m, illustrated in Figure 8.

The responses to ε_ϕ and ε_A show that the text-augmented model implies weaker inflation responses to real-side disturbances. A labor-supply shock raises the marginal disutility of labor and generates inflationary pressure through wages and marginal costs, while a positive technology shock lowers marginal cost and is therefore disinflationary. In both cases, the inflation response is more muted when transcripts are included. This is consistent with the posterior shift toward higher price stickiness and lower price indexation: marginal-cost movements pass through less strongly to inflation, and inflation is less mechanically tied to its own lag.

The interest-rate responses to these shocks are also muted in the text-augmented model, but for slightly different reasons. For the labor-supply shock, the policy rule faces offsetting forces: the increase in inflation calls for a higher policy rate, while the decline in real activity calls for a lower one. Because the text-augmented model estimates a flatter and less backward-looking Phillips curve, the inflation response is smaller, weakening the inflation-driven motive to raise rates. As a result, the policy response is more muted. For the technology shock, the same inflation channel works in the opposite direction: lower marginal cost reduces inflation and would call for a lower policy rate. However, the shock also raises the stochastic growth component that enters measured output growth in the Taylor rule. Since the text-augmented model estimates a lower response to inflation and a higher response to output growth, the trend-growth effect can offset the disinflationary force, generating a muted or slightly positive interest-rate response.

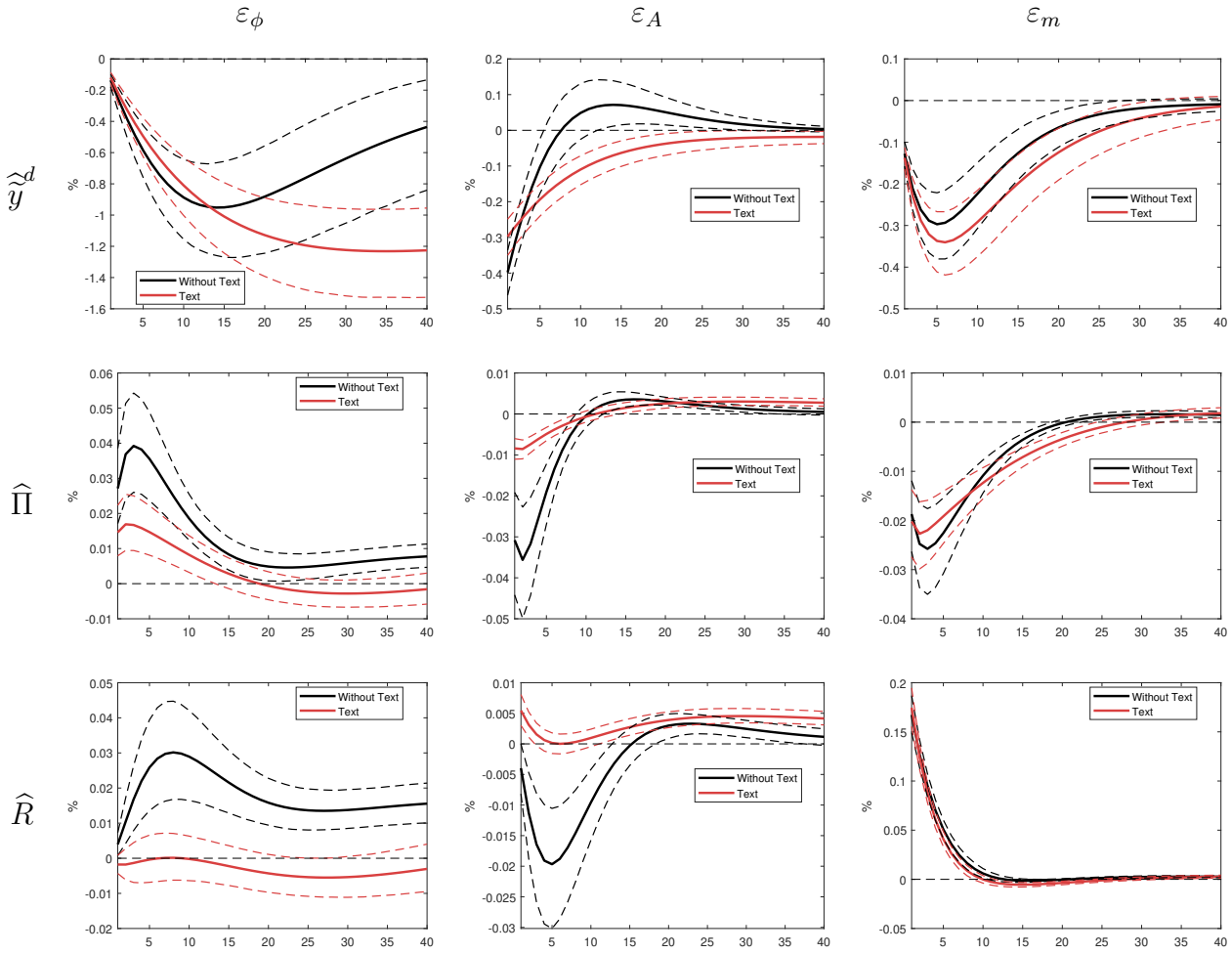
The response to the monetary-policy shock shows that the text-augmented model pre-

serves the standard qualitative transmission mechanism, even though the estimated policy-rule parameters change. In both specifications, a contractionary policy shock raises the nominal interest rate on impact and lowers aggregate demand and inflation. The main difference is in propagation rather than sign: with transcripts included, the decline in inflation is somewhat more muted and decays more slowly, while the response of aggregate demand is somewhat larger and more persistent. Thus, the transcript data alter the estimated dynamics of policy transmission, but they do not overturn the basic mechanism. Taken together, the impulse responses show that FOMC transcripts affect structural inference, not only forecasting performance.



Note: 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 the structural parameters in our DSGE model



Note: 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 lines are the posterior mean, and the dashed lines are the pointwise 90% credible interval.

Figure 8: Impulse responses of aggregate demand, inflation, and nominal interest rate to labor supply, aggregate technology, and monetary policy 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 demonstrates that unstructured data can help distinguish between 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 set of policy-relevant topics that are informative about the monetary policy shock. It also shifts the estimated Taylor-rule coefficients and the Calvo parameter. The estimated Phillips curve is flatter than the one implied by specifications based on standard macroeconomic data alone. Without text, the systematic and non-systematic components of monetary policy remain difficult to disentangle.

The framework imposes no a priori mapping from individual topics to particular model objects, though the specification does fix the class: $DS_t = \varepsilon_t$ has the text load on the structural shocks, and the spike-and-slab prior selects which shock loadings are active. It also 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. *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).

Boivin, J. and Giannoni, M. (2006). DSGE models in a data-rich environment. Working Paper 12772, National Bureau of Economic Research.

Bryzgalova, S., Huang, J., and Julliard, C. (2023). Bayesian solutions for the factor zoo: We just ran two quadrillion models. *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.

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. (2010). The econometrics of DSGE models. *SERIEs*, 1:3–49.

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. Technical report, University of Pennsylvania.

Fernández-Villaverde, J., Rubio-Ramírez, J., and Schorfheide, F. (2016). Solution and estimation methods for DSGE models. In Taylor, J. B. and Uhlig, H., editors, *Handbook of Macroeconomics, Volume 2*, 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.

Hansen, S., McMahon, M., and Prat, A. (2018). Transparency and deliberation within the FOMC: A computational linguistics approach. *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.

Krippner, L. (2015). *Zero lower bound term structure modeling: A practitioner's guide*. Springer.

Kryshko, M. M. (2011). Data-rich DSGE and dynamic factor models. Technical report, 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.

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.

Pellegrino, B. (2025). Product differentiation and oligopoly: A network approach. *American Economic Review*, 115(4):1170–1225.

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.

Shapiro, A. H. and Wilson, D. J. (2022). Taking the Fed at its word: A new approach to estimating central bank objectives using text analysis. *Review of Economic Studies*, 89(5):2768–2805.

Wu, J. C. and Xia, F. D. (2016). Measuring the macroeconomic impact of monetary policy at the zero lower bound. *Journal of Money, Credit and Banking*, 48(2-3):253–291.

Appendix

A The 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 in detail how we construct our state-transition equation, focusing on the state vector S_t, which 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, thereby providing the core measurement equation. Finally, we discuss how we connect unstructured data to our model, which provides the non-core measurement equation.

A.1 The 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.

In our model, the solution of our DSGE model can be represented as follows:

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

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

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

where

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

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

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

$$\Sigma_\varepsilon = diag(\sigma_\mu^2, \sigma_D^2, \sigma_\phi^2, \sigma_A^2, \sigma_M^2, \sigma_G^2),$$

and PP, QQ, RR, SS, NN are defined by the structural parameters of the DSGE model.

Let us define a state vector as

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

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

$$T(\psi) = \begin{bmatrix} PP & QQ * NN & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{n_e \times n_s} & NN & 0 & 0 & 0 & 0 & 0 & 0 \\ v_w & \mathbf{0}_{1\times n_e} & 0 & 0 & 0 & 0 & 0 & 0 \\ v_y & \mathbf{0}_{1\times n_e} & 0 & 0 & 0 & 0 & 0 & 0 \\ RR_{(6,\cdot)} & SS_{(6,\cdot)} * NN & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1\times n_s} & \mathbf{0}_{1\times n_e} & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1\times n_s} & \mathbf{0}_{1\times n_e} & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1\times n_s} & \mathbf{0}_{1\times n_e} & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix}$$

and

$$R(\psi) = \begin{bmatrix} QQ * \Sigma_\varepsilon^{1/2} \\ \Sigma_\varepsilon^{1/2} \\ \mathbf{0}_{1\times n_e} \\ \mathbf{0}_{1\times n_e} \\ SS_{(6,:)} * \Sigma_\varepsilon^{1/2} \\ \Sigma_\varepsilon^{1/2}([d, \phi, g]') \end{bmatrix},$$

where n_s and n_e are the length of $state_t$ and exo_t, respectively, v_y and v_w are $1 \times n_s$ vectors with zeros except the 7th and 2nd element being 1 (so that it selects $\widehat{\widetilde{y}}^d_t$ and $\widehat{\widetilde{w}}_t$ from S_t), respectively, $\mathbf{0}$ 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 of those together, the state transition equation is:

$$\begin{bmatrix} state_t \\ exo_t \\ \widehat{\widetilde{w}}_{t-1} \\ \widehat{\widetilde{y}}^{d}_{t-1} \\ \widehat{l}_t \\ \varepsilon_{D,t} \\ \varepsilon_{\phi,t} \\ \varepsilon_{G,t} \end{bmatrix} = \begin{bmatrix} PP & QQ * NN & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{n_e \times n_s} & NN & 0 & 0 & 0 & 0 & 0 & 0 \\ v_w & \mathbf{0}_{1 \times n_e} & 0 & 0 & 0 & 0 & 0 & 0 \\ v_y & \mathbf{0}_{1 \times n_e} & 0 & 0 & 0 & 0 & 0 & 0 \\ RR_{(6,\cdot)} & SS_{(6,\cdot)} * NN & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1 \times n_s} & \mathbf{0}_{1 \times n_e} & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1 \times n_s} & \mathbf{0}_{1 \times n_e} & 0 & 0 & 0 & 0 & 0 & 0 \\ \mathbf{0}_{1 \times n_s} & \mathbf{0}_{1 \times n_e} & 0 & 0 & 0 & 0 & 0 & 0 \end{bmatrix} \begin{bmatrix} state_{t-1} \\ exo_{t-1} \\ \widehat{\widetilde{w}}_{t-2} \\ \widehat{\widetilde{y}}^{d}_{t-2} \\ \widehat{l}_{t-1} \\ \varepsilon_{D,t-1} \\ \varepsilon_{\phi,t-1} \\ \varepsilon_{G,t-1} \end{bmatrix} + \begin{bmatrix} QQ * \Sigma_{\varepsilon}^{1/2} \\ \Sigma_{\varepsilon}^{1/2} \\ \mathbf{0}_{1 \times n_e} \\ \mathbf{0}_{1 \times n_e} \\ SS_{(6,:)} * \Sigma_{\varepsilon}^{1/2} \\ \Sigma_{\varepsilon}^{1/2}([d, \phi, g]') \end{bmatrix} \varepsilon_t.$$

A.2 The 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, a standard approach in the DSGE model estimation literature. The non-core measurement equation connects unstructured data to our model.

Below, 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 The 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:

$$\log\left(\frac{P_t^{obs}}{P_{t-1}^{obs}}\right) = \log(\Pi) + S_t(1) + \sigma_{ME}^{\Pi} u_t^{\Pi}, \quad u_t^{\Pi} \sim \mathcal{N}(0,1)$$

$$\log\left(INT_t^{obs} + 1\right) = \log R + S_t(6) + \sigma_{ME}^{R} u_t^{R}, \quad u_t^{R} \sim \mathcal{N}(0,1)$$

$$\triangle \log w_t^{obs} = \Lambda_z + S_t(15) + S_t(2) - S_t(n_s + n_e + 1) + \sigma_{ME}^{w} u_t^{w}, \quad u_t^{w} \sim \mathcal{N}(0,1)$$

$$\triangle \log y_t^{obs} = \Lambda_z + S_t(15) + S_t(7) - S_t(n_s + n_e + 2) + \sigma_{ME}^{y} u_t^{y}, \quad u_t^{y} \sim \mathcal{N}(0,1)$$

$$\log l_t^{obs} - mean(\log l_{1:T}^{obs}) = 0 + S_t(n_s + n_e + 3) + \sigma_{ME}^{l} u_t^{l}, \quad u_t^{l} \sim \mathcal{N}(0,1)$$

$$\underbrace{\triangle \log(\mu_t^{-1})^{obs}}_{\text{Relative price of capital}} = -(\log \mu_t - \log \mu_{t-1}) = -(\Lambda_\mu + z_{\mu,t}) = -\Lambda_\mu - S_t(n_s+1) + \sigma_{ME}^{\mu} u_t^{\mu}, \quad u_t^{\mu} \sim \mathcal{N}(0,1)$$

where $\Lambda_z = \frac{\alpha\Lambda_\mu + \Lambda_A}{(1-\alpha)}$ and $S_t(i)$ is the i-th element of the 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

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

H_1^C is a $(6 \times (n_s + n_e + 6))$ matrix, where $H_1^C(1,1) = 1, H_1^C(2,6) = 1, H_1^C(3,2) = 1, H_1^C(3, n_s + n_e + 1) = -1, H_1^C(3,15) = 1, H_1^C(4,7) = 1, H_1^C(4, n_s + n_e + 2) = -1, H_1^C(4,15) = 1, H_1^C(5, n_s + n_e + 3) = 1, H_1^C(6, n_s + 1) = -1$ and other elements are zero, with $H_1^C(i,j)$ denoting the (i,j) element of the matrix H_1^C, and Σ_u is a diagonal matrix containing the variance of the measurement errors.

A.2.2 The non-core measurement equation

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

$$Y_t^{NC} = (Y_{1,t}^{NC}, Y_{2,t}^{NC}, ..., Y_{M,t}^{NC})'.$$

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

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

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

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

A.2.3 The 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:

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

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 The Markov Chain Monte Carlo Algorithm

This section details the posterior sampling algorithm. We use a Metropolis-Hastings-within-Gibbs approach organized in two blocks, as described in Section 2: DSGE structural parameters are drawn via random-block random-walk Metropolis-Hastings, state variables via a simulation smoother, and non-core measurement parameters via Gibbs sampling with discrete updates for the inclusion indicators.

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 using the random-block random-walk Metropolis-Hastings algorithm, and then apply the simulation smoother to sample $S_{1:T}$.

In the random-block random-walk Metropolis-Hastings algorithm, first, we randomly assign structural parameters to blocks without duplication, where the total number of 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}. Given that we have 30 structural parameters, we have $1+29\times p^{block}$ blocks on average, because the total number of blocks follows a shifted binomial distribution, $1+\text{Binomial}(29, p^{block})$. We randomly assign the structural parameters to one of the blocks to ensure no duplication. We employ this randomization for each iteration of the algorithm, so the number of blocks and the assignment pattern of which parameters are placed in which block change with each draw. In the estimation, we set $p^{block} = 0.3$. Second, 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 total number of blocks and the number of parameters per block, then assign each parameter to a block at random. 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 randomly permuted vector of parameters: θ^r.

2-2 Create a block-assignment list b of the same length as θ^r that specifies which parameter belongs to which block. The total number of blocks and the number of parameters per block are randomly determined. One way of doing so is by iteratively increasing the index whenever a uniformly distributed random number is less than a specified probability p^{block}. More technically, first assign the first parameter to the first block by setting $b_1 = 1$. Then, for each subsequent parameter $i \geq 2$, draw a standard uniform random number; if it is less than a specified probability p^{block}, start a new block by setting $b_i = b_{i-1} + 1$; otherwise, assign it to the current block by setting $b_i = b_{i-1}$. 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-assignment 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 $b = 1, 2, ..., N_{block}^{(j)}$, 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

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

and where $\theta_{<b}^{(j)} = [\theta_1^{(j)}, ..., \theta_{b-1}^{(j)}]$, $\theta_{b<}^{(j-1)} = [\theta_{b+1}^{(j-1)}, ..., \theta_{N_{block}}^{(j-1)}]$, and $p(\theta)$ is the prior density. The likelihood, $p(Y_{1:T}^C, Y_{1:T}^{NC} | \theta, \varphi^{(j-1)})$, 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 $S_{1:T}^{(j)}$ from $p(S_{1:T} | Y_{1:T}^C, Y_{1:T}^{NC}, \theta^{(j)}, \varphi^{(j-1)})$ by applying the simulation smoother of Durbin and Koopman (2002) to the state-space model (7) by following the algorithm of Jarociński (2015).

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

Note that given $(Y_{1:T}^C, Y_{1:T}^{NC})$ 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, a 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

$$\lambda_{m,i}|V_m, q_m \sim \begin{cases} N(0, V_m) & \text{with probability } q_m \\ 0 & \text{with probability } 1 - q_m \end{cases} \quad \text{where } V_m = \sigma^2_{v,m}\gamma^2_m$$

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

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

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

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

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^2_m is uniform. Hence, we can directly apply the sampling algorithm developed by Giannone et al. (2021) to draw $\varphi^{(j)}$ from $p(\varphi|Y^C_{1:T}, Y^{NC}_{1:T}, \theta^{(j)}, S^{(j)}_{1:T})$, 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 their full conditional distribution, $p(S_{1:T}|\theta, Y^C_{1:T}, Y^{NC}_{1:T}, \varphi)$, so the normalization is not a straightforward issue. Thus, we implement the following steps to standardize each draw of $\varepsilon^{(j)}_t$.

First, we draw 100 samples of $S^{(i)}_{1:T}, i = 1, 2, \cdots, 100$ from $p(S_{1:T}|\theta, Y^C_{1:T})$, 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 $S^{(i)}_{1:T}$ and denote them as $\varepsilon^{(i)}_{1:T}$. Second, we compute the mean of the structural shocks across draws and obtain $\widehat{\varepsilon}_{1:T} = \frac{1}{100}\sum_{i=1}^{100} \varepsilon^{(i)}_{1:T}$. Third, we compute the standard deviations of $\widehat{\varepsilon}_{1:T}$, denoted as $\overline{\sigma}_\varepsilon$, and standardize $\varepsilon^{(j)}_t$ for each draw in the MCMC by computing $\overline{\varepsilon}^{(j)}_t = \frac{\varepsilon^{(j)}_t}{\overline{\sigma}_\varepsilon}$, and include them as regressors. 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).

Supplementary Material:

Structural Estimation with Unstructured Data

Sara Casella
LUISS University and EIEF

Jesús Fernández-Villaverde
University of Pennsylvania, NBER, CEPR

Stephen Hansen
University College London, FRB Dallas, 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, sticky-wage model as in Fernández-Villaverde (2010), to which we add government expenditure as in Fernández-Villaverde and Guerrón-Quintana (2021). The following derivations closely follow the notes in Fernández-Villaverde and Rubio-Ramírez (2006).[10] A representative household consumes, saves, holds money, supplies differentiated labor services, and sets its own wage subject to a demand curve and a Calvo friction. A final-good producer aggregates a continuum of intermediate goods from monopolistic competitors who rent capital and "packed" labor and set prices subject to their own Calvo friction. A competitive labor packer aggregates the differentiated labor inputs of households into a homogeneous input sold to intermediate producers. A consolidated public sector sets the nominal interest rate through open market operations and consumes an exogenous stream g. Long-run growth is driven by unit roots in neutral and investment-specific technology.

[10] Available at https://www.sas.upenn.edu/~jesusfv/benchmark_DSGE.pdf

A.1 Environment

A.1.1 Households

There is a continuum of households indexed by $j \in [0,1]$. Each household maximizes a lifetime utility function separable in consumption c_{jt}, real money balances M_{jt}/p_t (where p_t is the price level) and hours worked l_{jt}:

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

where β is the discount factor, h is the parameter that controls habit persistence, ϑ is the inverse of the Frisch elasticity of labor supply, 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 shifters are common to all households. The utility function is log in consumption so that the marginal rate of substitution between consumption and leisure is linear in consumption, which is necessary for a balanced growth path with constant hours.

Households trade a complete set of Arrow–Debreu securities indexed by the household j (to insure against the idiosyncratic wage-adjustment risk described below) and by time. Let a_{jt+1} denote the holdings of such securities paying one unit of consumption in event $\omega_{j,t+1,t}$, purchased at (real) price $q_{jt+1,t}$. Households also hold an amount b_{jt} of one-period government bonds that pay a nominal gross interest rate R_t.

The j-th household's budget constraint is:

$$\begin{aligned} c_{jt} + x_{jt} + \frac{M_{jt}}{p_t} + \frac{b_{jt+1}}{p_t} + \int q_{jt+1,t} a_{jt+1} d\omega_{j,t+1,t} \\ = w_{jt} l_{jt} + \left(r_t u_{jt} - \mu_t^{-1} \Phi\left[u_{jt}\right]\right) k_{jt-1} + \frac{M_{jt-1}}{p_t} + R_{t-1} \frac{b_{jt}}{p_t} + a_{jt} + T_t + \digamma_t, \end{aligned}$$

where w_{jt} is the real wage set by household j, r_t is the real rental price of capital, $u_{jt} > 0$ is the intensity of use of capital, $\mu_t^{-1}\Phi\left[u_{jt}\right]$ 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 $\Phi[1] = 0$, $\Phi' > 0$ and $\Phi'' > 0$.

Investment x_{jt} induces a law of motion for capital:

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

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 $S''[\cdot] > 0$, with Λ_x the growth rate of investment along the balanced growth path (BGP).

The investment-specific technological shock follows:

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

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

The Lagrangian associated with this household's problem is:

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

where the household maximizes 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 is 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). If $\mu_t = 1$, as in the standard neoclassical growth model, $q_t = 1$.

A.1.2 The labor packer

The labor used by intermediate-good producers as described below is supplied by a representative, competitive firm that hires the labor supplied by each household j. The labor supplier aggregates the differentiated labor of households with the following production function:

$$l_t^d = \left(\int_0^1 l_{jt}^{\frac{\eta-1}{\eta}} dj \right)^{\frac{\eta}{\eta-1}}, \tag{A.1}$$

where η is the elasticity of substitution among different types of labor and l_t^d is the aggregate labor demand. The packer maximizes profits subject to (A.1), taking wages w_{jt} and w_t as given:

$$\max_{l_{jt}} w_t l_t^d - \int_0^1 w_{jt} l_{jt} dj.$$

Optimality and the zero-profit condition deliver the demand for each differentiated labor type:

$$l_{jt} = \left(\frac{w_{jt}}{w_t} \right)^{-\eta} l_t^d, \qquad \forall j, \tag{A.2}$$

and the aggregate real wage index:

$$w_t = \left(\int_0^1 w_{jt}^{1-\eta} dj \right)^{\frac{1}{1-\eta}}.$$

A.1.3 Wage setting

Households face idiosyncratic wage-adjustment risk through Calvo-style staggered wage setting. In every period, a fraction $1-\theta_w$ of households can reoptimize their nominal wage. The remaining fraction θ_w partially indexes its wage to past inflation at rate $\chi_w \in [0,1]$: if a household cannot reset for τ periods, its nominal wage is scaled by $\prod_{s=1}^{\tau} \Pi_{t+s-1}^{\chi_w}$.

The wage-setting part of the household's problem is:

$$\max_{w_{jt}} \mathbb{E}_t \sum_{\tau=0}^{\infty} (\beta\theta_w)^\tau \left\{ -d_{t+\tau}\phi_{t+\tau}\psi \frac{l_{jt+\tau}^{1+\vartheta}}{1+\vartheta} + \lambda_{t+\tau} \prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi_w}}{\Pi_{t+s}} w_{jt} l_{jt+\tau} \right\}$$

subject to the demand curve (A.2) updated for indexation,

$$l_{jt+\tau} = \left(\prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi_w}}{\Pi_{t+s}} \frac{w_{jt}}{w_{t+\tau}} \right)^{-\eta} l_{t+\tau}^d, \qquad \forall j.$$

Substituting the demand curve:

$$\max_{w_{jt}} \mathbb{E}_t \sum_{\tau=0}^{\infty} (\beta\theta_w)^\tau \left\{ \lambda_{t+\tau} \left(\prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi_w}}{\Pi_{t+s}} \frac{w_{jt}}{w_{t+\tau}} \right)^{1-\eta} w_{t+\tau} l_{t+\tau}^d - d_{t+\tau}\phi_{t+\tau}\psi \frac{\left(\prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi_w}}{\Pi_{t+s}} \frac{w_{jt}}{w_{t+\tau}} \right)^{-\eta(1+\vartheta)} \left(l_{t+\tau}^d \right)^{1+\vartheta}}{1+\vartheta} \right\}.$$

Because markets are complete and utility is separable in labor, all reoptimizing households choose the same wage $w_{jt}^* = w_t^*$. The first-order condition is:

$$\begin{aligned} &\frac{\eta-1}{\eta} w_t^* \mathbb{E}_t \sum_{\tau=0}^{\infty} (\beta\theta_w)^\tau \lambda_{t+\tau} \left(\prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi_w}}{\Pi_{t+s}} \right)^{1-\eta} \left(\frac{w_t^*}{w_{t+\tau}} \right)^{-\eta} l_{t+\tau}^d \\ &\quad = \mathbb{E}_t \sum_{\tau=0}^{\infty} (\beta\theta_w)^\tau d_{t+\tau}\phi_{t+\tau}\psi \left(\prod_{s=1}^{\tau} \frac{\Pi_{t+s-1}^{\chi_w}}{\Pi_{t+s}} \frac{w_t^*}{w_{t+\tau}} \right)^{-\eta(1+\vartheta)} \left(l_{t+\tau}^d \right)^{1+\vartheta}. \end{aligned}$$

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

$$F_t^1 = \frac{\eta-1}{\eta} \left(w_t^* \right)^{1-\eta} \lambda_t w_t^\eta l_t^d + \beta\theta_w \mathbb{E}_t \left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}} \right)^{1-\eta} \left(\frac{w_{t+1}^*}{w_t^*} \right)^{\eta-1} F_{t+1}^1$$

and

$$F_t^2 = \psi d_t \phi_t \left(\Pi_t^{w*}\right)^{-\eta(1+\vartheta)} \left(l_t^d\right)^{1+\vartheta} + \beta\theta_w \mathbb{E}_t \left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}}\right)^{-\eta(1+\vartheta)} \left(\frac{w_{t+1}^*}{w_t^*}\right)^{\eta(1+\vartheta)} F_{t+1}^2,$$

where

$$\Pi_t^{w*} \equiv \frac{w_t^*}{w_t}.$$

The wage-setting first-order condition is $F_t^1 = F_t^2 \equiv F_t$. We need $(\beta\theta_w)^\tau \lambda_{t+\tau}$ to vanish fast enough relative to inflation for the sums to be well defined.

Given Calvo's wage setting, the real-wage index evolves as a geometric average of the indexed past real wage and the new optimal real wage:

$$w_t^{1-\eta} = \theta_w \left(\frac{\Pi_{t-1}^{\chi_w}}{\Pi_t}\right)^{1-\eta} w_{t-1}^{1-\eta} + (1-\theta_w)\, w_t^{*1-\eta},$$

or, dividing by $w_t^{1-\eta}$,

$$1 = \theta_w \left(\frac{\Pi_{t-1}^{\chi_w}}{\Pi_t}\right)^{1-\eta} \left(\frac{w_{t-1}}{w_t}\right)^{1-\eta} + (1-\theta_w) \left(\Pi_t^{w*}\right)^{1-\eta}.$$

In a symmetric equilibrium $c_{jt} = c_t$, $u_{jt} = u_t$, $k_{jt-1} = k_{t-1}$, $x_{jt} = x_t$, $\lambda_{jt} = \lambda_t$, $q_{jt} = q_t$, and $w_{jt}^* = w_t^*$. The household first-order conditions become:

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

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

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

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

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

$$F_t = \frac{\eta - 1}{\eta} \left(w_t^*\right)^{1-\eta} \lambda_t w_t^\eta l_t^d + \beta\theta_w \mathbb{E}_t \left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}}\right)^{1-\eta} \left(\frac{w_{t+1}^*}{w_t^*}\right)^{\eta-1} F_{t+1}$$

$$F_t = \psi d_t \phi_t \left(\Pi_t^{w*}\right)^{-\eta(1+\vartheta)} \left(l_t^d\right)^{1+\vartheta} + \beta\theta_w \mathbb{E}_t \left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}}\right)^{-\eta(1+\vartheta)} \left(\frac{w_{t+1}^*}{w_t^*}\right)^{\eta(1+\vartheta)} F_{t+1}.$$

A.1.4 The final-good producer

The final good is produced using intermediate goods and the technology:

$$y_t^d = \left(\int_0^1 y_{it}^{\frac{\varepsilon-1}{\varepsilon}} di \right)^{\frac{\varepsilon}{\varepsilon-1}}, \tag{A.3}$$

where ε is the elasticity of substitution. Final-good producers are perfectly competitive. Their maximization problem is:

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

delivering $y_{it} = (p_{it}/p_t)^{-\varepsilon} y_t^d$ and

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

A.1.5 Intermediate-good producers

There is a continuum of intermediate-good producers. Each intermediate-good producer i has 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 A_t follows:

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

The parameter f, which is the fixed cost of production, and $z_t = A_t^{1/(1-\alpha)} \mu_t^{\alpha/(1-\alpha)}$ guarantee that economic profits are roughly zero in the steady state. Entry and exit of intermediate producers are ruled out.

Since $z_t = A_t^{1/(1-\alpha)} \mu_t^{\alpha/(1-\alpha)}$,

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

In the first stage, taking w_t and r_t as given, firms rent l_{it}^d and k_{it-1} in competitive factor markets to minimize real cost:

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

subject to $y_{it} = A_t k_{it-1}^{\alpha} \left(l_{it}^d \right)^{1-\alpha} - f z_t$ when the production side is positive. The first-order

conditions yield:

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

and since the firm has constant returns to scale, the real marginal cost is:

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

which does not depend on i.

In the second stage, intermediate-good producers choose prices. Each period, a fraction $1-\theta_p$ of firms reset prices; the remainder index to past inflation at rate $\chi \in [0,1]$.

The problem of the firm is:

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

subject to $y_{it+\tau} = \left(\prod_{s=1}^{\tau}\Pi_{t+s-1}^{\chi}\frac{p_{it}}{p_{t+\tau}}\right)^{-\varepsilon} y_{t+\tau}^d$. Following the standard derivation, the first-order condition can be written in recursive form by defining:

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

and

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

so that $\varepsilon g_t^1 = (\varepsilon - 1) g_t^2$. 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 $\Pi_t^* = p_t^*/p_t$. The price index evolves:

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

A.1.6 Monetary policy

The monetary authority sets the nominal interest rate according to a Taylor rule:

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

where Π is the inflation target (equal to inflation on the BGP), R is the BGP nominal gross return on capital, Λ_{y^d} is the BGP gross growth rate of y_t^d, and the monetary policy shock $m_t = \sigma_m \varepsilon_{m,t}$ with $\varepsilon_{m,t} \sim \mathcal{N}(0,1)$. Open market operations are financed through lump-sum transfers:

$$T_t = \frac{M_t}{p_t} - \frac{M_{t-1}}{p_t} + \frac{b_{t+1}}{p_t} - R_{t-1}\frac{b_t}{p_t}.$$

A.1.7 Fiscal policy

Government consumption is given by $g_t = \widetilde{g}_t z_t$, where $\widetilde{g}_t$ is a stationary detrended process and $z_t = A_t^{1/(1-\alpha)} \mu_t^{\alpha/(1-\alpha)}$. The mean of $\widetilde{g}_t$ is parameterized as a fraction of steady-state output: defining the steady-state government-spending-to-GDP ratio

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

we calibrate g^{ss} directly and recover the steady-state level $\widetilde{g} = g^{ss}\widetilde{y}^d$. The detrended process follows an autoregression in logs around this mean:

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

Multiplying by z_t in the level $g_t = \widetilde{g}_t z_t$ keeps real government consumption a stationary share of output along the balanced growth path. Government consumption is financed by lump-sum taxes, ensuring that the deficit is zero at all times.

A.1.8 Aggregation

Using the transfer rule and the zero-profit condition for the labor packer, the aggregate resource constraint is:

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

The demand for each intermediate good is $y_{it} = y_t^d \left(p_{it}/p_t\right)^{-\varepsilon}$, and:

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

Defining price dispersion $v_t^p = \int_0^1 \left(p_{it}/p_t\right)^{-\varepsilon} di$, Calvo's pricing gives:

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

so that:

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

Aggregating the individual labor demands (A.2) over j:

$$l_t \equiv \int_0^1 l_{jt} dj = \left(\int_0^1 \left(\frac{w_{jt}}{w_t}\right)^{-\eta} dj\right) l_t^d = v_t^w l_t^d,$$

where the wage dispersion v_t^w is given by:

$$v_t^w \equiv \int_0^1 \left(\frac{w_{jt}}{w_t}\right)^{-\eta} dj = \theta_w \left(\frac{w_{t-1}}{w_t} \frac{\Pi_{t-1}^{\chi_w}}{\Pi_t}\right)^{-\eta} v_{t-1}^w + \left(1-\theta_w\right) \left(\Pi_t^{w*}\right)^{-\eta}.$$

A.2 Equilibrium

The symmetric equilibrium is characterized by the following equations.

- The first-order conditions of the household:

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

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

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

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

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

$$F_t = \frac{\eta - 1}{\eta} \left(w_t^*\right)^{1-\eta} \lambda_t w_t^{\eta} l_t^d + \beta \theta_w \mathbb{E}_t \left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}}\right)^{1-\eta} \left(\frac{w_{t+1}^*}{w_t^*}\right)^{\eta-1} F_{t+1}$$

$$F_t = \psi d_t \phi_t \left(\Pi_t^{w*}\right)^{-\eta(1+\vartheta)} \left(l_t^d\right)^{1+\vartheta} + \beta \theta_w \mathbb{E}_t \left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}}\right)^{-\eta(1+\vartheta)} \left(\frac{w_{t+1}^*}{w_t^*}\right)^{\eta(1+\vartheta)} F_{t+1}.$$

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

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

  where they rent inputs to satisfy their static minimization problem:

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

- The wage and price indices evolve as:

$$1 = \theta_w \left(\frac{\Pi_{t-1}^{\chi_w}}{\Pi_t}\right)^{1-\eta}\left(\frac{w_{t-1}}{w_t}\right)^{1-\eta} + (1-\theta_w)\left(\Pi_t^{w*}\right)^{1-\eta}$$
$$1 = \theta_p \left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{1-\varepsilon} + (1-\theta_p)\,\Pi_t^{*1-\varepsilon}.$$

- The monetary authority 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:

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

  where the price and wage dispersions are:

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

and

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

Stationary equilibrium

Most variables are growing on average. We work with a system of stationary variables.

First, the household first-order conditions:

$$d_t \left(\frac{c_t}{z_t} - h\frac{c_{t-1}}{z_{t-1}}\frac{z_{t-1}}{z_t}\right)^{-1} - h\beta\mathbb{E}_t d_{t+1}\left(\frac{c_{t+1}}{z_{t+1}}\frac{z_{t+1}}{z_t} - h\frac{c_t}{z_t}\right)^{-1} = \lambda_t z_t$$
$$\lambda_t z_t = \beta\mathbb{E}_t\left\{\lambda_{t+1} z_{t+1}\frac{z_t}{z_{t+1}}\frac{R_t}{\Pi_{t+1}}\right\}$$
$$\mu_t r_t = \Phi'\left[u_t\right]$$
$$q_t\mu_t = \beta\mathbb{E}_t\left\{\frac{\lambda_{t+1}z_{t+1}}{\lambda_t z_t}\frac{z_t}{z_{t+1}}\frac{\mu_t}{\mu_{t+1}}\left[(1-\delta)\, q_{t+1}\mu_{t+1} + \mu_{t+1} r_{t+1} u_{t+1} - \Phi\left(u_{t+1}\right)\right]\right\}$$
$$1 = q_t\mu_t\left(1 - S\left[\frac{x_t/z_t}{x_{t-1}/z_{t-1}}\frac{z_t}{z_{t-1}}\right] - S'\left[\frac{x_t/z_t}{x_{t-1}/z_{t-1}}\frac{z_t}{z_{t-1}}\right]\frac{x_t/z_t}{x_{t-1}/z_{t-1}}\frac{z_t}{z_{t-1}}\right)$$
$$+\beta\mathbb{E}_t q_{t+1}\mu_{t+1}\frac{\lambda_{t+1}}{\lambda_t}S'\left[\frac{x_{t+1}/z_{t+1}}{x_t/z_t}\frac{z_{t+1}}{z_t}\right]\left(\frac{x_{t+1}/z_{t+1}}{x_t/z_t}\frac{z_{t+1}}{z_t}\right)^2$$
$$F_t = \frac{\eta-1}{\eta}\left(\frac{w_t^*}{z_t}\right)^{1-\eta}\lambda_t z_t\left(\frac{w_t}{z_t}\right)^{\eta} l_t^d + \beta\theta_w\mathbb{E}_t\left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}}\right)^{1-\eta}\left(\frac{w_{t+1}^*/z_{t+1}}{w_t^*/z_t}\frac{z_{t+1}}{z_t}\right)^{\eta-1} F_{t+1}$$
$$F_t = \psi d_t\phi_t\left(\Pi_t^{w*}\right)^{-\eta(1+\vartheta)}\left(l_t^d\right)^{1+\vartheta} + \beta\theta_w\mathbb{E}_t\left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}}\right)^{-\eta(1+\vartheta)}\left(\frac{w_{t+1}^*/z_{t+1}}{w_t^*/z_t}\frac{z_{t+1}}{z_t}\right)^{\eta(1+\vartheta)} F_{t+1}.$$

The firm Euler equations:

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

with

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

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

The wage and price index laws of motion become:

$$1 = \theta_w\left(\frac{\Pi_{t-1}^{\chi_w}}{\Pi_t}\right)^{1-\eta}\left(\frac{w_{t-1}/z_{t-1}}{w_t/z_t}\frac{z_{t-1}}{z_t}\right)^{1-\eta} + (1-\theta_w)\left(\Pi_t^{w*}\right)^{1-\eta}$$

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

The Taylor rule:

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

Market clearing:

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

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

and

$$l_t = v_t^w l_t^d$$

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

$$v_t^w = \theta_w\left(\frac{w_{t-1}/z_{t-1}}{w_t/z_t}\frac{z_{t-1}}{z_t}\frac{\Pi_{t-1}^{\chi_w}}{\Pi_t}\right)^{-\eta}v_{t-1}^w + (1-\theta_w)\left(\Pi_t^{w*}\right)^{-\eta}$$

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

Change of variables

We redefine variables to obtain a system in stationary variables. Define $\widetilde{c}_t = c_t/z_t$, $\widetilde{\lambda}_t = \lambda_t z_t$, $\widetilde{r}_t = r_t\mu_t$, $\widetilde{q}_t = q_t\mu_t$, $\widetilde{x}_t = x_t/z_t$, $\widetilde{w}_t = w_t/z_t$, $\widetilde{w}_t^* = w_t^*/z_t$, $\widetilde{k}_t = k_t/(z_t\mu_t)$, and $\widetilde{y}_t^d = y_t^d/z_t$. Note that $\Pi_t^{w*} = w_t^*/w_t = \widetilde{w}_t^*/\widetilde{w}_t$ is unchanged by detrending.

The equilibrium conditions become:

- Household first-order conditions:

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

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

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

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

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

$$F_t=\frac{\eta-1}{\eta}\left(\widetilde{w}_t^*\right)^{1-\eta}\widetilde{\lambda}_t\widetilde{w}_t^\eta l_t^d+\beta\theta_w\mathbb{E}_t\left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}}\right)^{1-\eta}\left(\frac{\widetilde{w}_{t+1}^*}{\widetilde{w}_t^*}\frac{z_{t+1}}{z_t}\right)^{\eta-1}F_{t+1}$$

$$F_t=\psi d_t\phi_t\left(\Pi_t^{w*}\right)^{-\eta(1+\vartheta)}\left(l_t^d\right)^{1+\vartheta}+\beta\theta_w\mathbb{E}_t\left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}}\right)^{-\eta(1+\vartheta)}\left(\frac{\widetilde{w}_{t+1}^*}{\widetilde{w}_t^*}\frac{z_{t+1}}{z_t}\right)^{\eta(1+\vartheta)}F_{t+1}.$$

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

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

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

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

with

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

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

- Wage and price index laws of motion:

$$1=\theta_w\left(\frac{\Pi_{t-1}^{\chi_w}}{\Pi_t}\right)^{1-\eta}\left(\frac{\widetilde{w}_{t-1}}{\widetilde{w}_t}\frac{z_{t-1}}{z_t}\right)^{1-\eta}+\left(1-\theta_w\right)\left(\Pi_t^{w*}\right)^{1-\eta}$$
$$1=\theta_p\left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{1-\varepsilon}+\left(1-\theta_p\right)\Pi_t^{*1-\varepsilon}.$$

- 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

$$v_t^p=\theta_p\left(\frac{\Pi_{t-1}^{\chi}}{\Pi_t}\right)^{-\varepsilon}v_{t-1}^p+\left(1-\theta_p\right)\Pi_t^{*-\varepsilon}$$
$$v_t^w=\theta_w\left(\frac{\widetilde{w}_{t-1}}{\widetilde{w}_t}\frac{z_{t-1}}{z_t}\frac{\Pi_{t-1}^{\chi_w}}{\Pi_t}\right)^{-\eta}v_{t-1}^w+\left(1-\theta_w\right)\left(\Pi_t^{w*}\right)^{-\eta}$$
$$l_t=v_t^w l_t^d$$
$$\widetilde{k}_t\frac{z_t}{z_{t-1}}\frac{\mu_t}{\mu_{t-1}}-\left(1-\delta\right)\widetilde{k}_{t-1}-\frac{z_t}{z_{t-1}}\frac{\mu_t}{\mu_{t-1}}\left(1-S\left[\frac{\widetilde{x}_t}{\widetilde{x}_{t-1}}\frac{z_t}{z_{t-1}}\right]\right)\widetilde{x}_t=0.$$

A.3 The steady state

Let $\widetilde{z}=\exp\left(\Lambda_z\right)$, $\widetilde{\mu}=\exp\left(\Lambda_\mu\right)$, $\widetilde{A}=\exp\left(\Lambda_A\right)$. The common growth rate is $\Lambda_c=\Lambda_x=\Lambda_w=\Lambda_{w^*}=\Lambda_{y^d}=\Lambda_z$.

Adopting $\Phi[u]=\Phi_1(u-1)+\frac{\Phi_2}{2}(u-1)^2$ and $S\left[x_t/x_{t-1}\right]=\frac{\kappa}{2}\left(x_t/x_{t-1}-\Lambda_x\right)^2$, steady-state

$u = 1$ gives $\widetilde{r} = \Phi_1$, $\Phi[1] = 0$, $S[\Lambda_x] = S'[\Lambda_x] = 0$. The steady-state system is:

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

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

$$\widetilde{r} = \Phi_1, \qquad \widetilde{r} = \left(1 - \tfrac{\beta}{\widetilde{z}\widetilde{\mu}}(1-\delta)\right) \Big/ \tfrac{\beta}{\widetilde{z}\widetilde{\mu}}$$

$$\left(1 - \beta\theta_w \widetilde{z}^{\eta-1}\Pi^{-(1-\chi_w)(1-\eta)}\right) F = \frac{\eta - 1}{\eta}\widetilde{w}^* \left(\Pi^{w*}\right)^{-\eta} \widetilde{\lambda} l^d$$

$$\left(1 - \beta\theta_w \widetilde{z}^{\eta(1+\vartheta)}\Pi^{\eta(1-\chi_w)(1+\vartheta)}\right) F = \psi \left(\Pi^{w*}\right)^{-\eta(1+\vartheta)} \left(l^d\right)^{1+\vartheta}$$

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

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

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

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

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

$$\frac{1 - \theta_w \Pi^{-(1-\chi_w)(1-\eta)}\widetilde{z}^{-(1-\eta)}}{1 - \theta_w} = \left(\Pi^{w*}\right)^{1-\eta}$$

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

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

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

$$l = v^w l^d$$

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

$$\frac{1 - \theta_w \widetilde{z}^{\eta}\Pi^{(1-\chi_w)\eta}}{1 - \theta_w} v^w = \left(\Pi^{w*}\right)^{-\eta}$$

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

The relationship between inflation and the optimal relative price is:

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

From the price-setting equations, the marginal cost in steady state is:

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

Likewise, the relationship between inflation and the optimal relative wage is:

$$\Pi^{w*} = \left(\frac{1 - \theta_w \Pi^{-(1-\chi_w)(1-\eta)} \widetilde{z}^{-(1-\eta)}}{1 - \theta_w} \right)^{\frac{1}{1-\eta}},$$

and

$$\widetilde{w}^* = \widetilde{w}\, \Pi^{w*}.$$

Both the optimal relative price and the optimal relative wage collapse to unity when $\Pi = 1$ and $\widetilde{z} = 1$, or under full indexation ($\chi = 1$, $\chi_w = 1$) with $\widetilde{z} = 1$.

Dividing the second wage-setting steady-state equation by the first:

$$\frac{1 - \beta\theta_w \widetilde{z}^{\eta(1+\vartheta)} \Pi^{\eta(1-\chi_w)(1+\vartheta)}}{1 - \beta\theta_w \widetilde{z}^{\eta-1} \Pi^{-(1-\chi_w)(1-\eta)}} = \frac{\psi \left(\Pi^{w*}\right)^{-\eta\vartheta} \left(l^d\right)^{\vartheta}}{\frac{\eta-1}{\eta} \widetilde{w}^* \widetilde{\lambda}},$$

which defines l^d implicitly as a function of $\widetilde{\lambda}$. Combined with the marginal utility of consumption, we obtain a second relationship between l^d and $\widetilde{\lambda}$.

The steady-state dispersions are:

$$v^p = \frac{1 - \theta_p}{1 - \theta_p \Pi^{(1-\chi)\varepsilon}} \Pi^{*-\varepsilon}, \qquad v^w = \frac{1 - \theta_w}{1 - \theta_w \widetilde{z}^{\eta} \Pi^{(1-\chi_w)\eta}} \left(\Pi^{w*}\right)^{-\eta}.$$

With ψ we calibrate the normalization $l^d = 1$ in steady state, which pins down $l = v^w$.

Once l^d is known, the marginal cost equation delivers the real wage:

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

from which $\widetilde{k} = \frac{\alpha}{1-\alpha} \frac{\widetilde{w}}{\widetilde{r}} \widetilde{z}\widetilde{\mu}$, $\widetilde{x} = \frac{\widetilde{z}\widetilde{\mu} - (1-\delta)}{\widetilde{z}\widetilde{\mu}} \widetilde{k}$, output $\widetilde{y}^d = \frac{\widetilde{A}}{\widetilde{z}} \widetilde{k}^{\alpha} (l^d)^{1-\alpha} - f$ all divided by v^p, and, using the calibrated ratio $g^{ss} \equiv \widetilde{g}/\widetilde{y}^d$ introduced above,

$$\widetilde{c} = \widetilde{y}^d - \widetilde{x} - g^{ss} \widetilde{y}^d = (1 - g^{ss}) \widetilde{y}^d - \widetilde{x}.$$

A.4 Loglinear approximations

For each variable var_t we define $\widehat{var}_t = \log var_t - \log var$, so that $var_t = var \exp^{\widehat{var}_t}$.

We start from the marginal utility of consumption. Following the same steps as in the flexible-wage version, one gets:

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

The Euler equation loglinearizes (using $R = \Pi\widetilde{z}/\beta$) to:

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

The utilization condition $\widetilde{r}_t = \Phi'[u_t]$ yields:

$$\widehat{\widetilde{r}}_t = \frac{\Phi_2}{\Phi_1}\widehat{u}_t.$$

The relationship between the shadow price of capital and the return on investment gives:

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

The investment first-order condition loglinearizes (with κ from the adjustment cost) to:

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

The wage-setting laws of motion require loglinearizing the two expressions defining F_t. For the first:

$$F_t = \frac{\eta-1}{\eta}\left(\widetilde{w}_t^*\right)^{1-\eta}\widetilde{\lambda}_t\widetilde{w}_t^{\eta}l_t^d + \beta\theta_w\mathbb{E}_t\left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}}\right)^{1-\eta}\left(\frac{\widetilde{w}_{t+1}^*}{\widetilde{w}_t^*}\frac{z_{t+1}}{z_t}\right)^{\eta-1}F_{t+1}.$$

Using the steady-state identity

$$1 - \beta\theta_w\widetilde{z}^{\eta-1}\Pi^{-(1-\eta)(1-\chi_w)} = \frac{\frac{\eta-1}{\eta}\left(\widetilde{w}^*\right)^{1-\eta}\widetilde{\lambda}\widetilde{w}^{\eta}l^d}{F},$$

one obtains:

$$\begin{aligned}\widehat{F}_t = \left(1-\beta\theta_w\widetilde{z}^{\eta-1}\Pi^{-(1-\eta)(1-\chi_w)}\right)&\left((1-\eta)\widehat{\widetilde{w}}_t^* + \widehat{\widetilde{\lambda}}_t + \eta\widehat{\widetilde{w}}_t + \widehat{l}_t^d\right)\\ &+ \beta\theta_w\widetilde{z}^{\eta-1}\Pi^{-(1-\eta)(1-\chi_w)}\mathbb{E}_t\left(\widehat{F}_{t+1} - (1-\eta)\left(\widehat{\Pi}_{t+1} - \chi_w\widehat{\Pi}_t + \Delta\widehat{\widetilde{w}}_{t+1}^*\right)\right).\end{aligned}$$

For the second expression defining F_t:

$$F_t = \psi d_t \phi_t \left(\Pi_t^{w*}\right)^{-\eta(1+\vartheta)} \left(l_t^d\right)^{1+\vartheta} + \beta\theta_w \mathbb{E}_t \left(\frac{\Pi_t^{\chi_w}}{\Pi_{t+1}}\right)^{-\eta(1+\vartheta)} \left(\frac{\widetilde{w}_{t+1}^*}{\widetilde{w}_t^*}\frac{z_{t+1}}{z_t}\right)^{\eta(1+\vartheta)} F_{t+1},$$

with steady-state identity

$$1 - \beta\theta_w \widetilde{z}^{\eta(1+\vartheta)} \Pi^{\eta(1+\vartheta)(1-\chi_w)} = \frac{\psi d \phi (l^d)^{1+\vartheta}}{F},$$

one obtains:

$$\begin{aligned}\widehat{F}_t = &\left(1 - \beta\theta_w \widetilde{z}^{\eta(1+\vartheta)} \Pi^{\eta(1+\vartheta)(1-\chi_w)}\right)\left(\widehat{d}_t + \widehat{\phi}_t + \eta(1+\vartheta)\left(\widehat{\widetilde{w}}_t - \widehat{\widetilde{w}}_t^*\right) + (1+\vartheta)\widehat{l}_t^d\right) \\ &+ \beta\theta_w \widetilde{z}^{\eta(1+\vartheta)} \Pi^{\eta(1+\vartheta)(1-\chi_w)} \mathbb{E}_t \left(\widehat{F}_{t+1} + \eta(1+\vartheta)\left(\widehat{\Pi}_{t+1} - \chi_w \widehat{\Pi}_t + \Delta\widehat{\widetilde{w}}_{t+1}^*\right)\right).\end{aligned}$$

The g^1 and g^2 equations for prices loglinearize as before:

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

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

and $\varepsilon g_t^1 = (\varepsilon - 1) g_t^2$ becomes $\widehat{g}_t^1 = \widehat{g}_t^2$.

The capital-labor ratio:

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

The marginal cost:

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

The aggregate wage law of motion:

$$1 = \theta_w \left(\frac{\Pi_{t-1}^{\chi_w}}{\Pi_t}\right)^{1-\eta} \left(\frac{\widetilde{w}_{t-1}}{\widetilde{w}_t}\frac{z_{t-1}}{z_t}\right)^{1-\eta} + (1-\theta_w)\left(\Pi_t^{w*}\right)^{1-\eta}$$

loglinearizes to:

$$\frac{\theta_w \Pi^{-(1-\chi_w)(1-\eta)} \widetilde{z}^{-(1-\eta)}}{(1-\theta_w)\left(\Pi^{w*}\right)^{1-\eta}} \left(\widehat{\Pi}_t - \chi_w \widehat{\Pi}_{t-1} + \widehat{\widetilde{w}}_t - \widehat{\widetilde{w}}_{t-1} + \widehat{\widetilde{z}}_t\right) = \widehat{\widetilde{w}}_t^* - \widehat{\widetilde{w}}_t,$$

where we have used $\widehat{\Pi}_t^{w*} = \widehat{\widetilde{w}}_t^* - \widehat{\widetilde{w}}_t$.

The aggregate price law of motion:

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

The Taylor rule:

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

The market-clearing conditions:

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

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

where $produc \equiv \frac{\widetilde{A}}{\widetilde{z}}\left(u\widetilde{k}\right)^{\alpha}\left(l^d\right)^{1-\alpha}$. Labor aggregation gives:

$$\widehat{l}_t = \widehat{v}_t^w + \widehat{l}_t^d.$$

The price-dispersion law loglinearizes to:

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

The wage-dispersion law loglinearizes to:

$$\widehat{v}_t^w = \theta_w \Pi^{\eta(1-\chi_w)} \widetilde{z}^{\eta}\left(\eta\left(\widehat{\Pi}_t - \chi_w \widehat{\Pi}_{t-1} + \widehat{\widetilde{w}}_t - \widehat{\widetilde{w}}_{t-1} + \widehat{\widetilde{z}}_t\right) + \widehat{v}_{t-1}^w\right) - \left(1 - \theta_w \Pi^{\eta(1-\chi_w)}\widetilde{z}^{\eta}\right)\eta\left(\widehat{\widetilde{w}}_t^* - \widehat{\widetilde{w}}_t\right).$$

The growth composition $\widehat{\widetilde{z}}_t = \frac{\widehat{\widetilde{A}}_t + \alpha\widehat{\widetilde{\mu}}_t}{1-\alpha}$ and the capital accumulation equation

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

complete the system.

A.5 A system of linear stochastic difference equations

We now present the 22 equations in the order used by Uhlig's algorithm.

Equation 1 Wage law of motion. Define $a_1 = \frac{\theta_w \Pi^{-(1-\chi_w)(1-\eta)} \widetilde{z}^{-(1-\eta)}}{(1-\theta_w)(\Pi^{w*})^{1-\eta}}$. Then:

$$a_1 \widehat{\Pi}_t - \chi_w a_1 \widehat{\Pi}_{t-1} + (1+a_1)\widehat{\widetilde{w}}_t - a_1 \widehat{\widetilde{w}}_{t-1} + a_1 \widehat{\widetilde{z}}_t - \widehat{\widetilde{w}}^*_t = 0. \tag{A.4}$$

Equation 2 Price law of motion. Define $a_2 = \frac{\theta_p \Pi^{-(1-\varepsilon)(1-\chi)}}{(1-\theta_p)(\Pi^*)^{1-\varepsilon}}$. Then:

$$a_2 \widehat{\Pi}_t - a_2 \chi \widehat{\Pi}_{t-1} - \widehat{\Pi}^*_t = 0. \tag{A.5}$$

Equation 3 Utilization. With $\phi_u = \Phi_2/\Phi_1$:

$$-\widehat{\widetilde{r}}_t + \phi_u \widehat{u}_t = 0. \tag{A.6}$$

Equation 4 Price aggregator:

$$\widehat{g}^1_t - \widehat{g}^2_t = 0. \tag{A.7}$$

Equation 5 Capital–labor ratio:

$$\widehat{u}_t + \widehat{\widetilde{r}}_t + \widehat{\widetilde{k}}_{t-1} - \widehat{l}^d_t - \widehat{\widetilde{w}}_t - \widehat{\widetilde{z}}_t - \widehat{\widetilde{\mu}}_t = 0. \tag{A.8}$$

Equation 6 Marginal cost:

$$(1-\alpha)\widehat{\widetilde{w}}_t + \alpha \widehat{\widetilde{r}}_t - \widehat{mc}_t = 0. \tag{A.9}$$

Equation 7 Taylor rule:

$$-\widehat{R}_t + \gamma_R \widehat{R}_{t-1} + (1-\gamma_R)\gamma_\Pi \widehat{\Pi}_t + (1-\gamma_R)\gamma_y \widehat{\widetilde{z}}_t + (1-\gamma_R)\gamma_y \widehat{\widetilde{y}}^d_t - (1-\gamma_R)\gamma_y \widehat{\widetilde{y}}^d_{t-1} + \widehat{m}_t = 0. \tag{A.10}$$

Equation 8 Resource constraint:

$$\widetilde{c}\widehat{\widetilde{c}}_t + \widetilde{x}\widehat{\widetilde{x}}_t + g^{ss}\widetilde{y}^d \widehat{\widetilde{g}}_t + \frac{\Phi_1 \widetilde{k}}{\widetilde{z}\widetilde{\mu}} \widehat{u}_t - \widetilde{y}^d \widehat{\widetilde{y}}^d_t = 0. \tag{A.11}$$

Equation 9 Production function:

$$(\widetilde{y}^d v^p)\widehat{v}_t^p + (\widetilde{y}^d v^p)\widehat{\widetilde{y}}_t^d - (produc)\widehat{\widetilde{A}}_t + (produc)\widehat{\widetilde{z}}_t - \alpha(produc)\widehat{u}_t \\ -\alpha(produc)\widehat{\widetilde{k}}_{t-1} - (1-\alpha)(produc)\widehat{l}_t^d = 0. \tag{A.12}$$

Equation 10 Labor aggregation:

$$\widehat{l}_t - \widehat{v}_t^w - \widehat{l}_t^d = 0. \tag{A.13}$$

Equation 11 Price-dispersion law. Define $a_3 = \beta\theta_p\Pi^{\varepsilon(1-\chi)}$. Then:

$$\frac{a_3\varepsilon}{\beta}\widehat{\Pi}_t - \frac{a_3\varepsilon\chi}{\beta}\widehat{\Pi}_{t-1} + \frac{a_3}{\beta}\widehat{v}_{t-1}^p - \left(1-\frac{a_3}{\beta}\right)\varepsilon\widehat{\Pi}_t^* - \widehat{v}_t^p = 0. \tag{A.14}$$

Equation 12 Wage-dispersion law. Define $a_4 = \theta_w\Pi^{\eta(1-\chi_w)}\widetilde{z}^\eta$. Then:

$$a_4\eta\widehat{\Pi}_t - \chi_w\eta a_4\widehat{\Pi}_{t-1} + \eta\widehat{\widetilde{w}}_t - a_4\eta\widehat{\widetilde{w}}_{t-1} + a_4\eta\widehat{\widetilde{z}}_t + a_4\widehat{v}_{t-1}^w - (1-a_4)\eta\widehat{\widetilde{w}}_t^* - \widehat{v}_t^w = 0. \tag{A.15}$$

Equation 13 Capital accumulation:

$$\frac{1-\delta}{\widetilde{z}\widetilde{\mu}}\widehat{\widetilde{k}}_{t-1} + \left[1-\frac{1-\delta}{\widetilde{z}\widetilde{\mu}}\right]\widehat{\widetilde{x}}_t - \frac{1-\delta}{\widetilde{z}\widetilde{\mu}}\left(\widehat{\widetilde{z}}_t + \widehat{\widetilde{\mu}}_t\right) - \widehat{\widetilde{k}}_t = 0. \tag{A.16}$$

Equation 14 Growth composition:

$$\frac{1}{1-\alpha}\widehat{\widetilde{A}}_t + \frac{\alpha}{1-\alpha}\widehat{\widetilde{\mu}}_t - \widehat{\widetilde{z}}_t = 0. \tag{A.17}$$

Equation 15 Marginal utility of consumption:

$$\widehat{d}_t - \tfrac{h\beta}{\widetilde{z}}\mathbb{E}_t\widehat{d}_{t+1} - \tfrac{1+h^2\beta/\widetilde{z}^2}{1-h/\widetilde{z}}\widehat{\widetilde{c}}_t + \tfrac{h/\widetilde{z}}{1-h/\widetilde{z}}\widehat{\widetilde{c}}_{t-1} + \tfrac{\beta h/\widetilde{z}}{1-h/\widetilde{z}}\mathbb{E}_t\widehat{\widetilde{c}}_{t+1} - \tfrac{h/\widetilde{z}}{1-h/\widetilde{z}}\widehat{\widetilde{z}}_t - \left(1-\tfrac{h\beta}{\widetilde{z}}\right)\widehat{\widetilde{\lambda}}_t = 0. \tag{A.18}$$

Equation 16 Euler:

$$\mathbb{E}_t\left\{\widehat{\widetilde{\lambda}}_{t+1} - \widehat{\widetilde{\lambda}}_t + \widehat{R}_t - \widehat{\Pi}_{t+1}\right\} = 0. \tag{A.19}$$

Equation 17 Tobin's Q:

$$\mathbb{E}_t\widehat{\widetilde{\lambda}}_{t+1} - \widehat{\widetilde{\lambda}}_t + \tfrac{\beta(1-\delta)}{\widetilde{z}\widetilde{\mu}}\mathbb{E}_t\widehat{\widetilde{q}}_{t+1} + \left(1-\tfrac{\beta(1-\delta)}{\widetilde{z}\widetilde{\mu}}\right)\mathbb{E}_t\widehat{\widetilde{r}}_{t+1} - \widehat{\widetilde{q}}_t = 0. \tag{A.20}$$

Equation 18 Investment first-order condition:

$$\widehat{\widetilde{q}}_t + \beta\kappa\widetilde{z}^2\mathbb{E}_t\widehat{\widetilde{x}}_{t+1} - (1+\beta)\kappa\widetilde{z}^2\widehat{\widetilde{x}}_t + \kappa\widetilde{z}^2\widehat{\widetilde{x}}_{t-1} - \kappa\widetilde{z}^2\widehat{\widetilde{z}}_t = 0. \tag{A.21}$$

Equation 19 First price-setting equation (with a_3):

$$(1-a_3)\widehat{\widetilde{\lambda}}_t + (1-a_3)\widehat{mc}_t + (1-a_3)\widehat{\widetilde{y}}^d_t + \varepsilon a_3\mathbb{E}_t\widehat{\Pi}_{t+1} - \chi\varepsilon a_3\widehat{\Pi}_t + a_3\mathbb{E}_t\widehat{g}^1_{t+1} - \widehat{g}^1_t = 0. \tag{A.22}$$

Equation 20 Second price-setting equation. Define $a_7 = \beta\theta_p\Pi^{-(1-\varepsilon)(1-\chi)}$. Then:

$$(1-a_7)\widehat{\widetilde{\lambda}}_t + \widehat{\Pi}^*_t + (1-a_7)\widehat{\widetilde{y}}^d_t + (\varepsilon-1)a_7\mathbb{E}_t\widehat{\Pi}_{t+1} - \chi(\varepsilon-1)a_7\widehat{\Pi}_t - a_7\mathbb{E}_t\widehat{\Pi}^*_{t+1} + a_7\mathbb{E}_t\widehat{g}^2_{t+1} - \widehat{g}^2_t = 0. \tag{A.23}$$

Equation 21 First wage-setting equation. Define $a_5 = \beta\theta_w\widetilde{z}^{\eta-1}\Pi^{-(1-\eta)(1-\chi_w)}$. Then:

$$\begin{aligned}(1-\eta)\widehat{\widetilde{w}}^*_t + (1-a_5)\widehat{\widetilde{\lambda}}_t + (1-a_5)\eta\widehat{\widetilde{w}}_t + (1-a_5)\widehat{l}^d_t + a_5\mathbb{E}_t\widehat{F}_{t+1} \\ + (\eta-1)a_5\mathbb{E}_t\widehat{\Pi}_{t+1} - (\eta-1)a_5\chi_w\widehat{\Pi}_t - (1-\eta)a_5\mathbb{E}_t\widehat{\widetilde{w}}^*_{t+1} - \widehat{F}_t = 0.\end{aligned} \tag{A.24}$$

Equation 22 Second wage-setting equation. Define $a_6 = \beta\theta_w\widetilde{z}^{\eta(1+\vartheta)}\Pi^{\eta(1+\vartheta)(1-\chi_w)}$. Then:

$$\begin{aligned}(1-a_6)\widehat{d}_t + (1-a_6)\widehat{\phi}_t + \eta(1+\vartheta)(1-a_6)\widehat{\widetilde{w}}_t - \eta(1+\vartheta)\widehat{\widetilde{w}}^*_t + (1+\vartheta)(1-a_6)\widehat{l}^d_t - \widehat{F}_t \\ + a_6\mathbb{E}_t\widehat{F}_{t+1} + a_6\eta(1+\vartheta)\mathbb{E}_t\widehat{\Pi}_{t+1} - a_6\eta(1+\vartheta)\chi_w\widehat{\Pi}_t + a_6\eta(1+\vartheta)\mathbb{E}_t\widehat{\widetilde{w}}^*_{t+1} = 0.\end{aligned} \tag{A.25}$$

Shocks

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

The following shocks are not defined because they are i.i.d.:

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

A.6 Solving the model

Let

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

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

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

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

The system takes the form:

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

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

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

The matrices are:

$$AA = \left(\begin{array}{cccccccccccccc}
a_1 & 1+a_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_1 \\
a_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 1 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 \\
0 & 1-\alpha & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
(1-\gamma_R)\gamma_\Pi & 0 & 0 & 0 & 0 & -1 & (1-\gamma_R)\gamma_y & 0 & 0 & 0 & 0 & 0 & 0 & 0 & (1-\gamma_R)\gamma_y \\
0 & 0 & 0 & 0 & 0 & 0 & -\widetilde{y}^d & \widetilde{c} & 0 & 0 & 0 & 0 & \widetilde{x} & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & \widetilde{y}^d v^p & 0 & \widetilde{y}^d v^p & 0 & 0 & 0 & 0 & 0 & produc \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 \\
a_3\varepsilon/\beta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 \\
\eta a_4 & \eta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & \eta a_4 \\
0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1-\frac{1-\delta}{\widetilde{z}\widetilde{\mu}} & 0 & -\frac{1-\delta}{\widetilde{z}\widetilde{\mu}} \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1
\end{array} \right)$$

$$BB = \begin{pmatrix}
-\chi_w a_1 & -a_1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-\chi a_2 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & \gamma_R & -(1-\gamma_R)\gamma_y & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & -\alpha(produc) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
-\frac{a_3 \varepsilon \chi}{\beta} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{a_3}{\beta} & 0 & 0 & 0 & 0 & 0 & 0 \\
-\chi_w \eta a_4 & -\eta a_4 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_4 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & \frac{1-\delta}{\widetilde{z}\widetilde{\mu}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}$$

$$CC = \begin{pmatrix}
0 & 0 & 0 & 0 & 0 & 0 & -1 \\
0 & 0 & -1 & 0 & 0 & 0 & 0 \\
-1 & \phi_u & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
1 & 1 & 0 & -1 & 0 & 0 & 0 \\
\alpha & 0 & 0 & 0 & -1 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & \frac{\Phi_1 \widetilde{k}}{\widetilde{z}\widetilde{\mu}} & 0 & 0 & 0 & 0 & 0 \\
0 & -\alpha(produc) & 0 & -(1-\alpha)(produc) & 0 & 0 & 0 \\
0 & 0 & 0 & -1 & 0 & 1 & 0 \\
0 & 0 & -\left(1-\frac{a_3}{\beta}\right)\varepsilon & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & -(1-a_4)\eta \\
0 & 0 & 0 & 0 & 0 & 0 & 0 \\
0 & 0 & 0 & 0 & 0 & 0 & 0
\end{pmatrix}$$

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

$$FF = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\beta h/\widetilde{z}}{1-h/\widetilde{z}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ -1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{\beta(1-\delta)}{\widetilde{z}\widetilde{\mu}} & 0 & 0 & 1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \beta\kappa\widetilde{z}^{2} & 0 & 0 \\ a_3\varepsilon & 0 & a_3 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_7(\varepsilon-1) & 0 & 0 & a_7 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ a_5(\eta-1) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_5 & 0 & 0 & 0 \\ a_6\eta(1+\vartheta) & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & a_6 & 0 & 0 & 0 \end{pmatrix}$$

$$GG = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & GG_{1,8} & 0 & 0 & 0 & 0 & 0 & -\left(1-\frac{h\beta}{\widetilde{z}}\right) & GG_{1,15} \\ 0 & 0 & 0 & 0 & 0 & 1 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & -1 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 1 & 0 & GG_{4,13} & 0 & -\kappa\widetilde{z}^{2} \\ -a_3\varepsilon\chi & 0 & -1 & 0 & 0 & 0 & 1-a_3 & 0 & 0 & 0 & 0 & 0 & 0 & 1-a_3 & 0 \\ -a_7(\varepsilon-1)\chi & 0 & 0 & -1 & 0 & 0 & 1-a_7 & 0 & 0 & 0 & 0 & 0 & 0 & 1-a_7 & 0 \\ -a_5\chi_w(\eta-1) & (1-a_5)\eta & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 1-a_5 & 0 \\ -a_6\chi_w\eta(1+\vartheta) & GG_{8,2} & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & -1 & 0 & 0 & 0 \end{pmatrix}$$

where

$$\begin{aligned} GG_{1,8} &= -\frac{1+\beta h^2/\widetilde{z}^2}{1-h/\widetilde{z}}, \\ GG_{1,15} &= -\frac{h/\widetilde{z}}{1-h/\widetilde{z}}, \\ GG_{4,13} &= -(1+\beta)\kappa\widetilde{z}^2, \\ GG_{8,2} &= (1-a_6)\eta(1+\vartheta). \end{aligned}$$

$$HH = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 & \frac{h/\widetilde{z}}{1-h/\widetilde{z}} & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & \kappa\widetilde{z}^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 & 0 \end{pmatrix}$$

$$JJ = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 1-\frac{\beta(1-\delta)}{\widetilde{z}\widetilde{\mu}} & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & -a_7 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & -a_5(1-\eta) \\ 0 & 0 & 0 & 0 & 0 & 0 & a_6\eta(1+\vartheta) \end{pmatrix}$$

$$KK = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 1-a_3 & 0 & 0 \\ 0 & 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 1-a_5 & 0 & 0 & 1-\eta \\ 0 & 0 & 0 & (1-a_6)(1+\vartheta) & 0 & 0 & -\eta(1+\vartheta) \end{pmatrix}$$

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

$$MM = \begin{pmatrix} 0 & 1 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 1-a_6 & 1-a_6 & 0 & 0 & 0 \end{pmatrix}$$

$$NN = \begin{pmatrix} 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & \rho_D & 0 & 0 & 0 & 0 \\ 0 & 0 & \rho_\phi & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & 0 \\ 0 & 0 & 0 & 0 & 0 & \rho_G \end{pmatrix}, \qquad \Sigma = \begin{pmatrix} \sigma_\mu^2 & 0 & 0 & 0 & 0 & 0 \\ 0 & \sigma_D^2 & 0 & 0 & 0 & 0 \\ 0 & 0 & \sigma_\phi^2 & 0 & 0 & 0 \\ 0 & 0 & 0 & \sigma_A^2 & 0 & 0 \\ 0 & 0 & 0 & 0 & \sigma_M^2 & 0 \\ 0 & 0 & 0 & 0 & 0 & \sigma_G^2 \end{pmatrix}.$$

B Data construction

Next, 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 Greenbook 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, $\approx FFR_t/400$.
- **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 ap-

pearing 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 we extract 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 real GDP, are the quarter-over-quarter growth rate in the Greenbook, while they are transformed as $\triangle(\log(x))$ in FRED-QD. Therefore, no transformation is needed for the Greenbook GDP series.

We then estimate the factors using 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 are 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 from 1987Q3 to 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:

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

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 $\theta = \theta^{(j)}$ and $\psi = \psi^{(j)}$, which is the jth draw from the posterior distribution estimated in the sub-sample.
3. Given θ and ψ, 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 $\mathcal{N}(0, \Sigma_v)$ 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 square 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 the jth draw from the predictive distribution.

The joint log score for horizon h we report in the paper is the sum of the joint log score across forecast origins:

$$\sum_{T=2000Q1}^{2019Q4-h} \log\left(\phi(Y_{T+h}; \widehat{Y}_{T+h|T}, \widehat{\Sigma}_{\widehat{Y}_{T+h|T}})\right),$$

where $\phi(x; \mu, \Sigma) = \frac{1}{\sqrt{(2\pi)^k |\Sigma|}} \exp\left(-\frac{1}{2}(x-\mu)'\Sigma^{-1}(x-\mu)\right)$,
$\widehat{\Sigma}_{\widehat{Y}_{T+h|T}} = \frac{1}{N_{sim}-1} \sum_{j=1}^{N_{sim}} \left(Y_{T+h|T}^{(j)} - \widehat{Y}_{T+h|T}\right)\left(Y_{T+h|T}^{(j)} - \widehat{Y}_{T+h|T}\right)'$, and $k = 6$ is the number of

macroeconomic variables forecast in our application. The RMSE for variable i at forecast horizon h is

$$RMSE^i(h) = \sqrt{\frac{1}{80-h}\sum_{T=2000Q1}^{2019Q4-h}\left(Y_{i,T+h} - \widehat{Y}_{i,T+h|T}\right)^2}.$$

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

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

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 draws with CDF $F_{i,t+h|t}(u)$.[11] This measure is the discrepancy between the model's predictive CDF and the step function at the realized value $Y_{i,t+h}$. We compute it by Monte Carlo approximation:

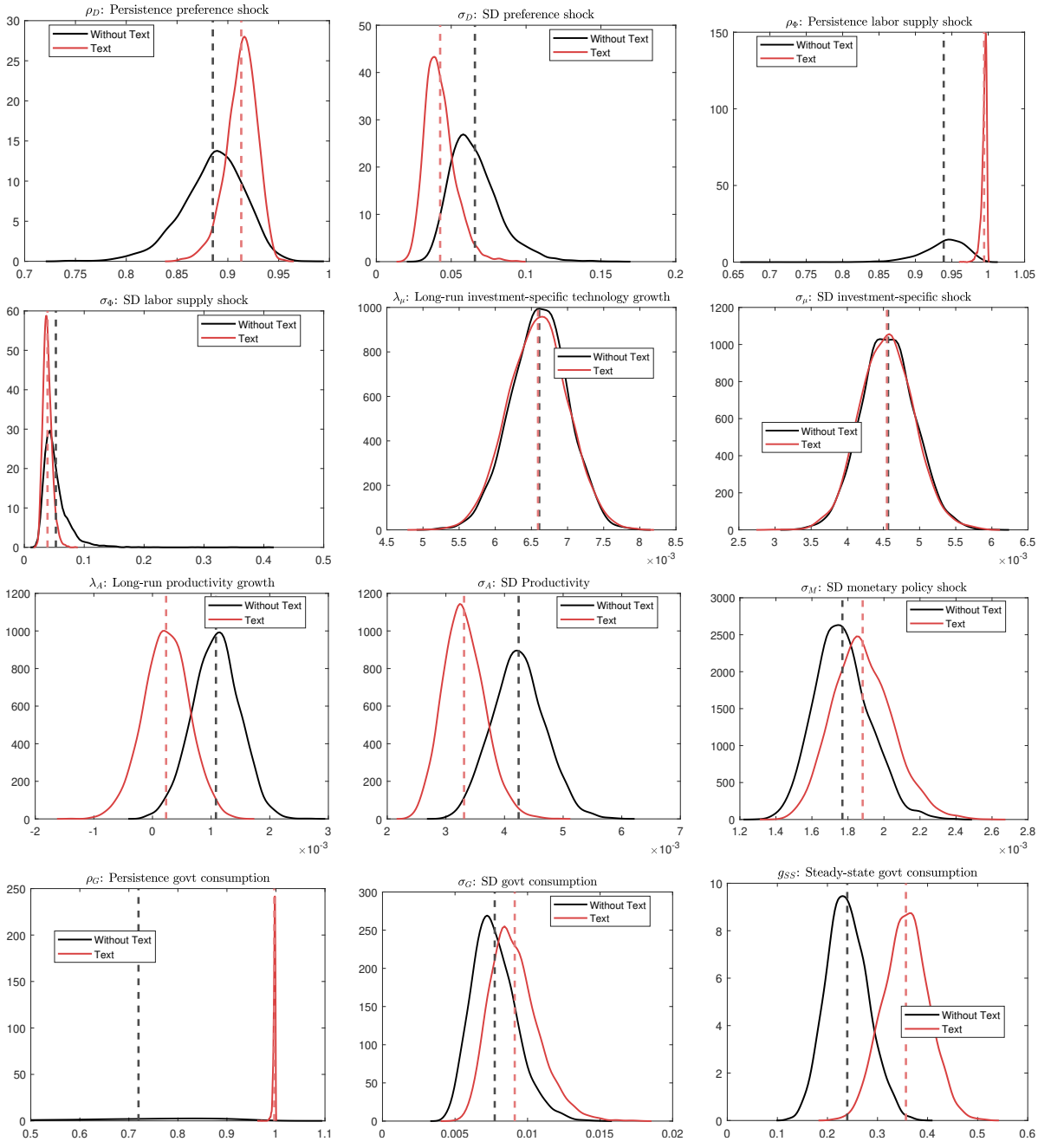
$$\widehat{CRPS}_t^i(h) = \frac{1}{N_{sim}}\sum_{j=1}^{N_{sim}}|Y_{i,t+h|t}^{(j)} - Y_{i,t+h}| - \frac{1}{2N_{sim}^2}\sum_{j=1}^{N_{sim}}\sum_{k=1}^{N_{sim}}|Y_{i,t+h|t}^{(j)} - Y_{i,t+h|t}^{(k)}|,$$

since the draws from the predictive distribution approximate $F_{i,t+h|t}$. We report $\widehat{CRPS}^i(h) = \sum_{T=2000Q1}^{2019Q4-h}\widehat{CRPS}_T^i(h)$.

[11] The second equality follows from two properties: $|X-Y| = \int_{-\infty}^{\infty} 1\{Y \leq u < X\}\,du + \int_{-\infty}^{\infty} 1\{X \leq u < Y\}\,du$, and, for independent $X \sim F$ and $Y \sim G$, $E[|X-Y|] = \int_{-\infty}^{\infty} F(u)(1-G(u))\,du + \int_{-\infty}^{\infty} G(u)(1-F(u))\,du$. Simple algebra shows the two expressions are equivalent.

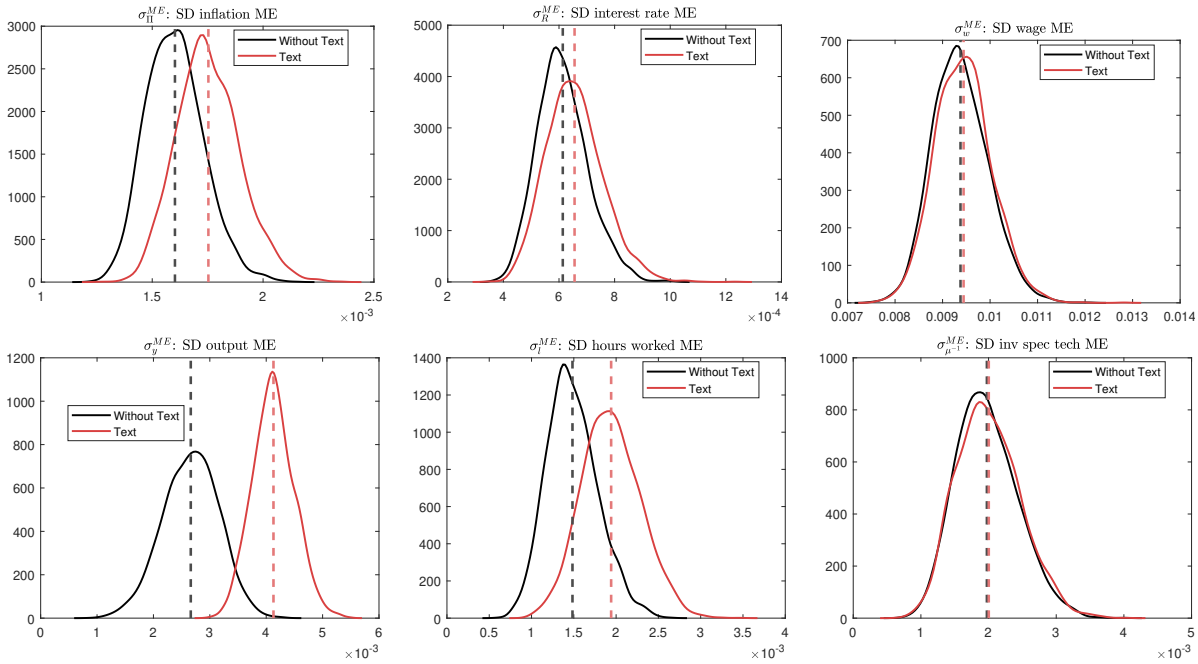
D Additional figures for the model with text

This section presents additional figures for the DSGE model estimated using FOMC transcripts (Section 3.4.3). Specifically, we report the posterior distributions of the remaining parameters and the impulse responses omitted from Figures 7 and 8.



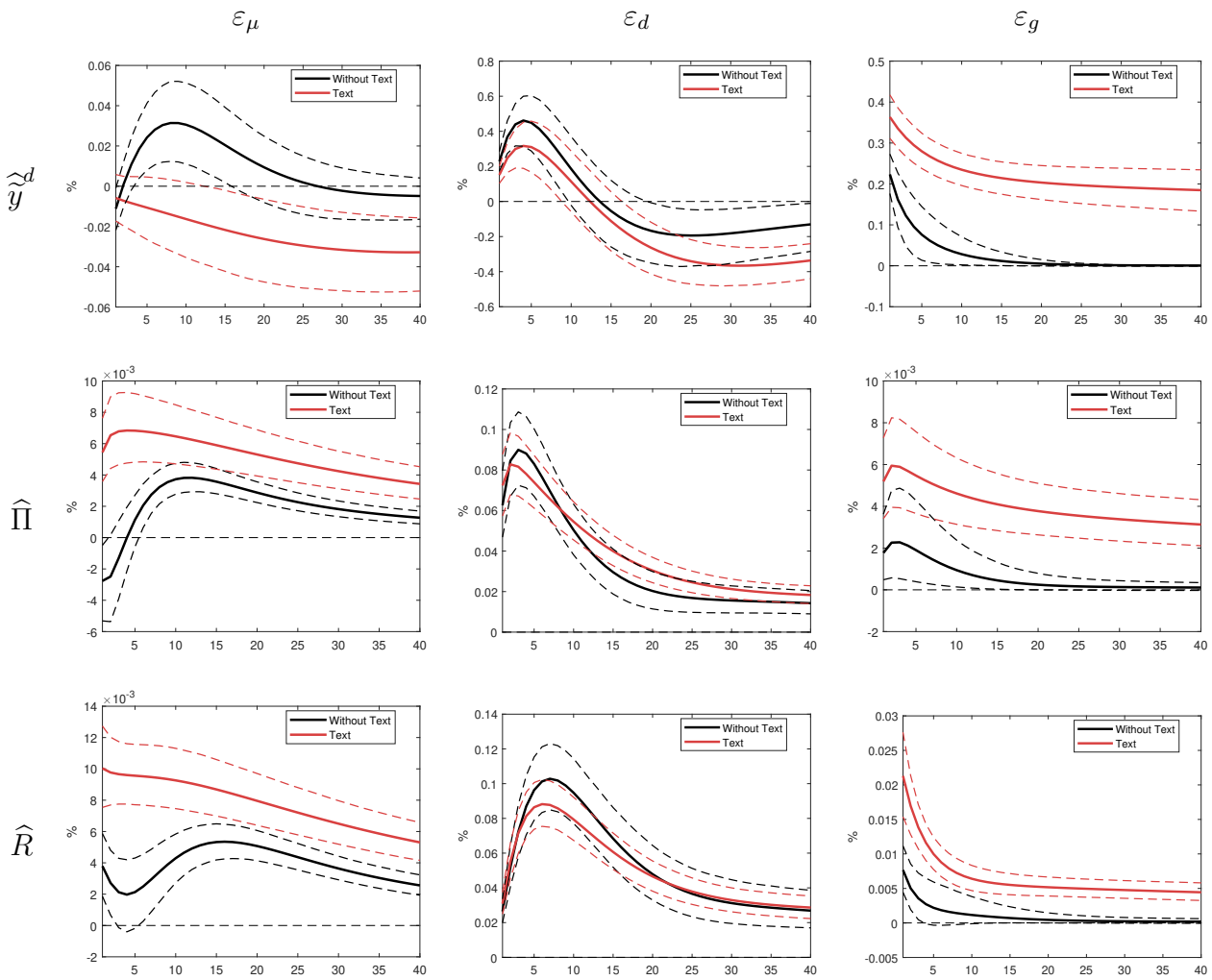
Note: The dashed vertical line is the posterior mean. The black line represents the model without text, and the red line represents the model with FOMC transcripts.

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



Note: 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 D.2: Posterior distribution of the other parameters in our DSGE model 2



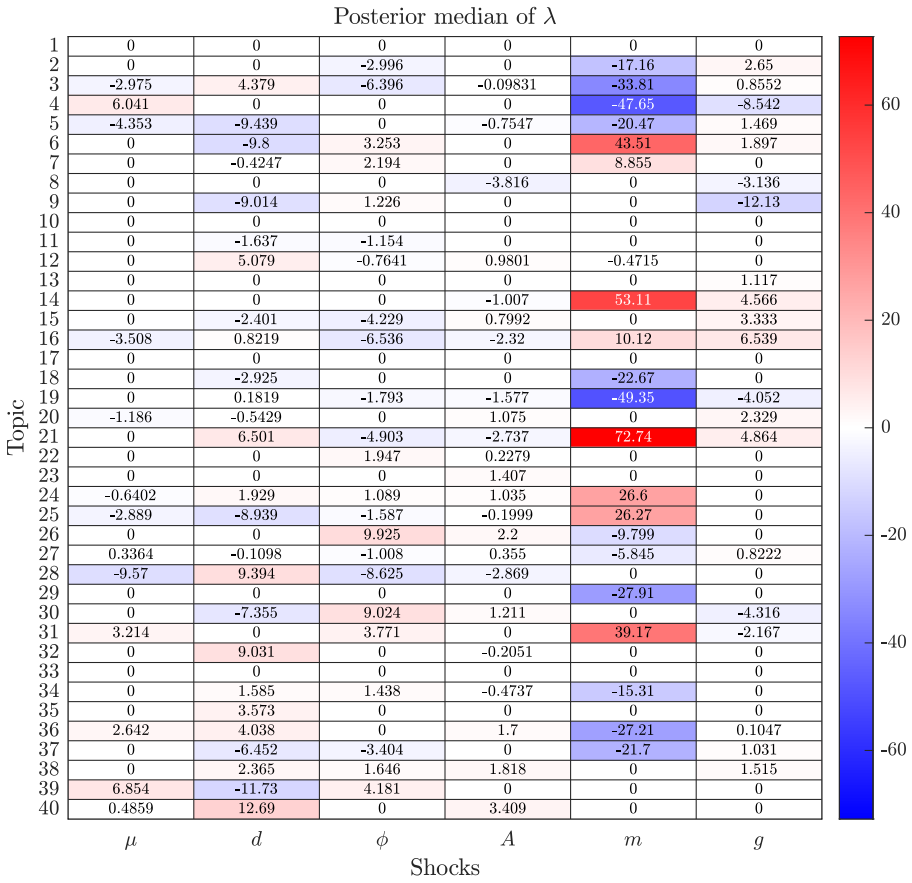
Note: Impulse responses of our estimated DSGE model to a one-standard-deviation shock. The black line represents the model without text, and the red line represents the model with FOMC transcripts. The solid lines are the posterior means, and the dashed lines are the pointwise 90% credible interval.

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

E Robustness checks using the shadow short rates

As a robustness check, we re-estimate the model by splicing the short-term nominal interest rate with the shadow short rates (SSR) from Wu and Xia (2016) and Krippner (2015) during the zero-lower-bound period (2009Q1–2015Q4). We find that the results described in the main paper remain robust to this alternative specification. The following figures present results using the shadow short rate (SSR) from Wu and Xia (2016); however, the results obtained with the SSR from Krippner (2015) are quantitatively similar.

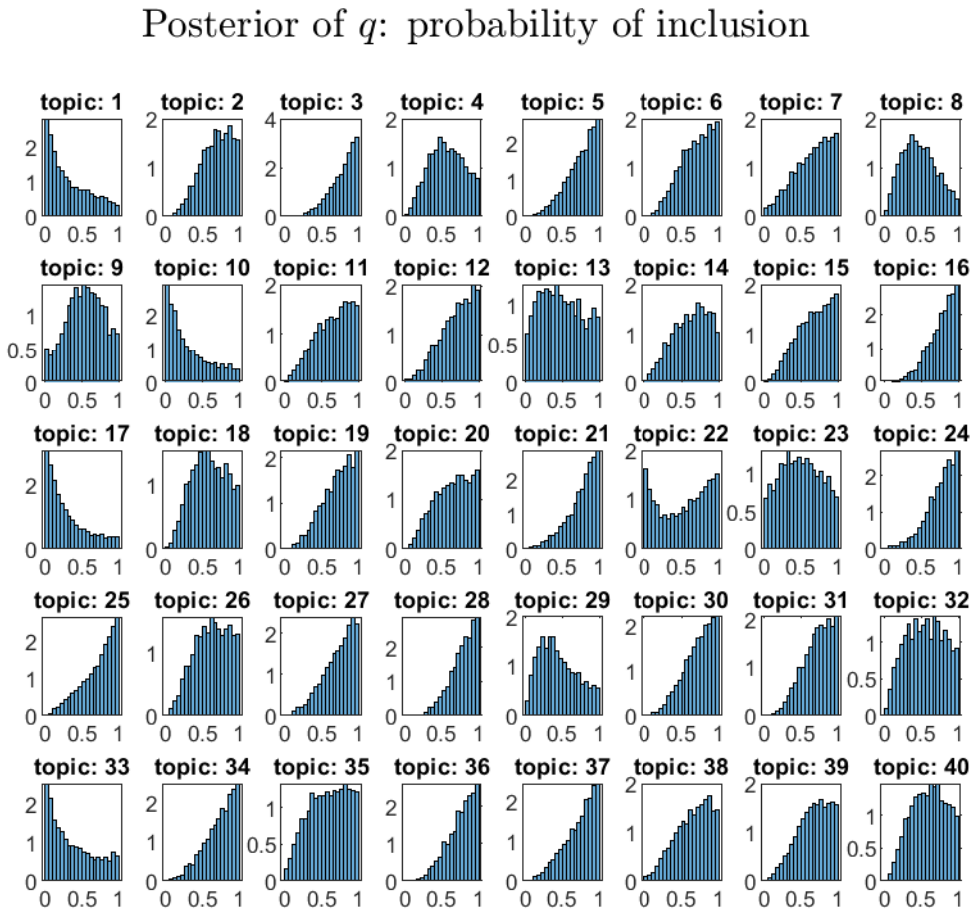
E.1 Topic selection



Topic	μ	d	ϕ	A	m	g
1	0	0	0	0	0	0
2	0	0	-2.996	0	-17.16	2.65
3	-2.975	4.379	-6.396	-0.09831	-33.81	0.8552
4	6.041	0	0	0	-47.65	-8.542
5	-4.353	-9.439	0	-0.7547	-20.47	1.469
6	0	-9.8	3.253	0	43.51	1.897
7	0	-0.4247	2.194	0	8.855	0
8	0	0	0	-3.816	0	-3.136
9	0	-9.014	1.226	0	0	-12.13
10	0	0	0	0	0	0
11	0	-1.637	-1.154	0	0	0
12	0	5.079	-0.7641	0.9801	-0.4715	0
13	0	0	0	0	0	1.117
14	0	0	0	-1.007	53.11	4.566
15	0	-2.401	-4.229	0.7992	0	3.333
16	-3.508	0.8219	-6.536	-2.32	10.12	6.539
17	0	0	0	0	0	0
18	0	-2.925	0	0	-22.67	0
19	0	0.1819	-1.793	-1.577	-49.35	-4.052
20	-1.186	-0.5429	0	1.075	0	2.329
21	0	6.501	-4.903	-2.737	72.74	4.864
22	0	0	1.947	0.2279	0	0
23	0	0	0	1.407	0	0
24	-0.6402	1.929	1.089	1.035	26.6	0
25	-2.889	-8.939	-1.587	-0.1999	26.27	0
26	0	0	9.925	2.2	-9.799	0
27	0.3364	-0.1098	-1.008	0.355	-5.845	0.8222
28	-9.57	9.394	-8.625	-2.869	0	0
29	0	0	0	0	-27.91	0
30	0	-7.355	9.024	1.211	0	-4.316
31	3.214	0	3.771	0	39.17	-2.167
32	0	9.031	0	-0.2051	0	0
33	0	0	0	0	0	0
34	0	1.585	1.438	-0.4737	-15.31	0
35	0	3.573	0	0	0	0
36	2.642	4.038	0	1.7	-27.21	0.1047
37	0	-6.452	-3.404	0	-21.7	1.031
38	0	2.365	1.646	1.818	0	1.515
39	6.854	-11.73	4.181	0	0	0
40	0.4859	12.69	0	3.409	0	0

Note: The figure reports the posterior median of the factor loading of each transformed topic-share series on each structural shock in the DSGE model. Rows correspond to FOMC transcript topics, and columns correspond to structural shocks. μ: investment-specific technology shock, d: intertemporal preference shock, ϕ: labor supply shock, A: aggregate technology shock, m: monetary policy shock, g: government spending shock. For the labels of the topics, see Table 1.

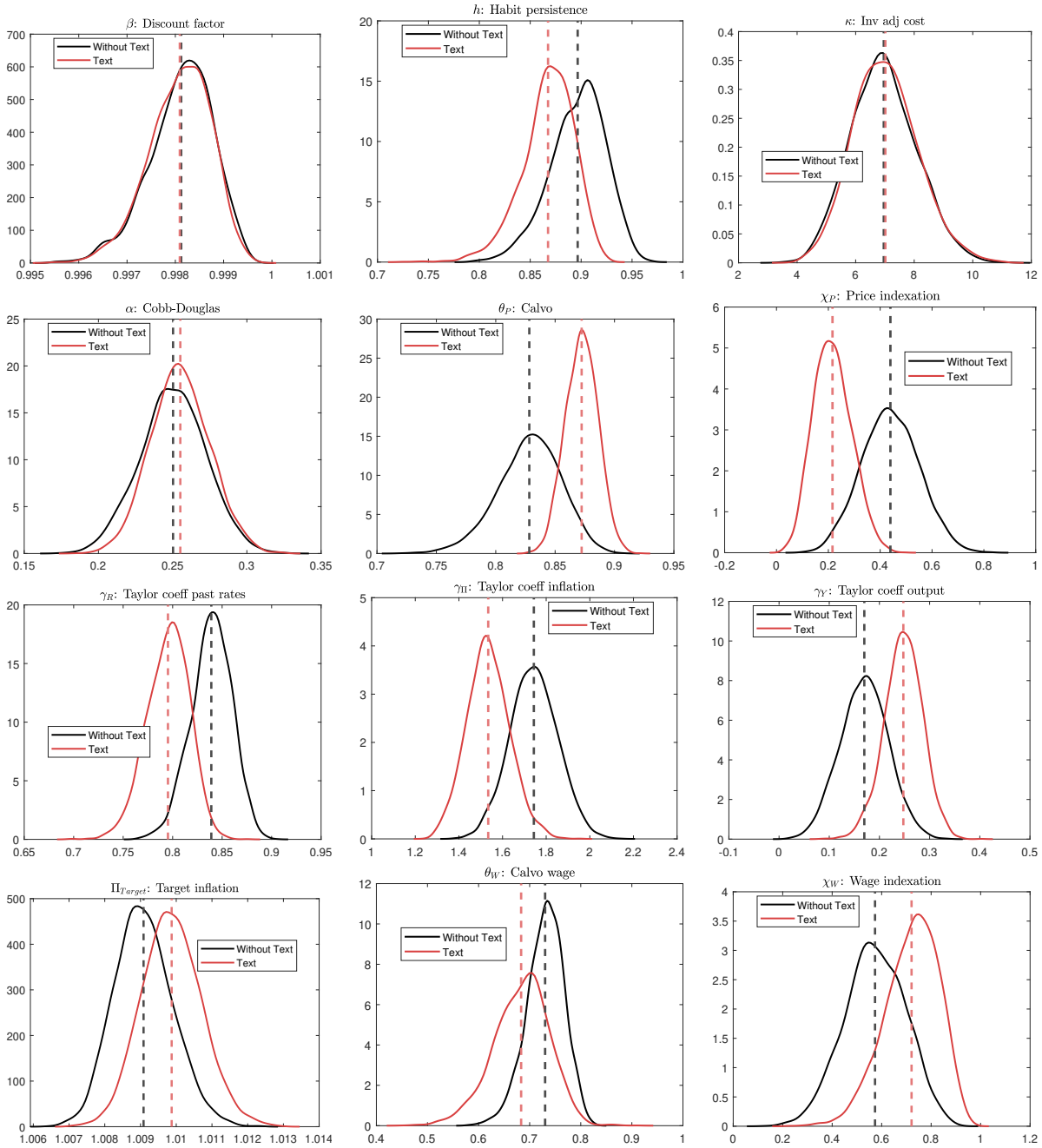
Figure E.1: Posterior median of the loadings λ



Note: Posterior distribution of the probability q of inclusion of each topic-share series. For the labels of the topics, see Table 1.

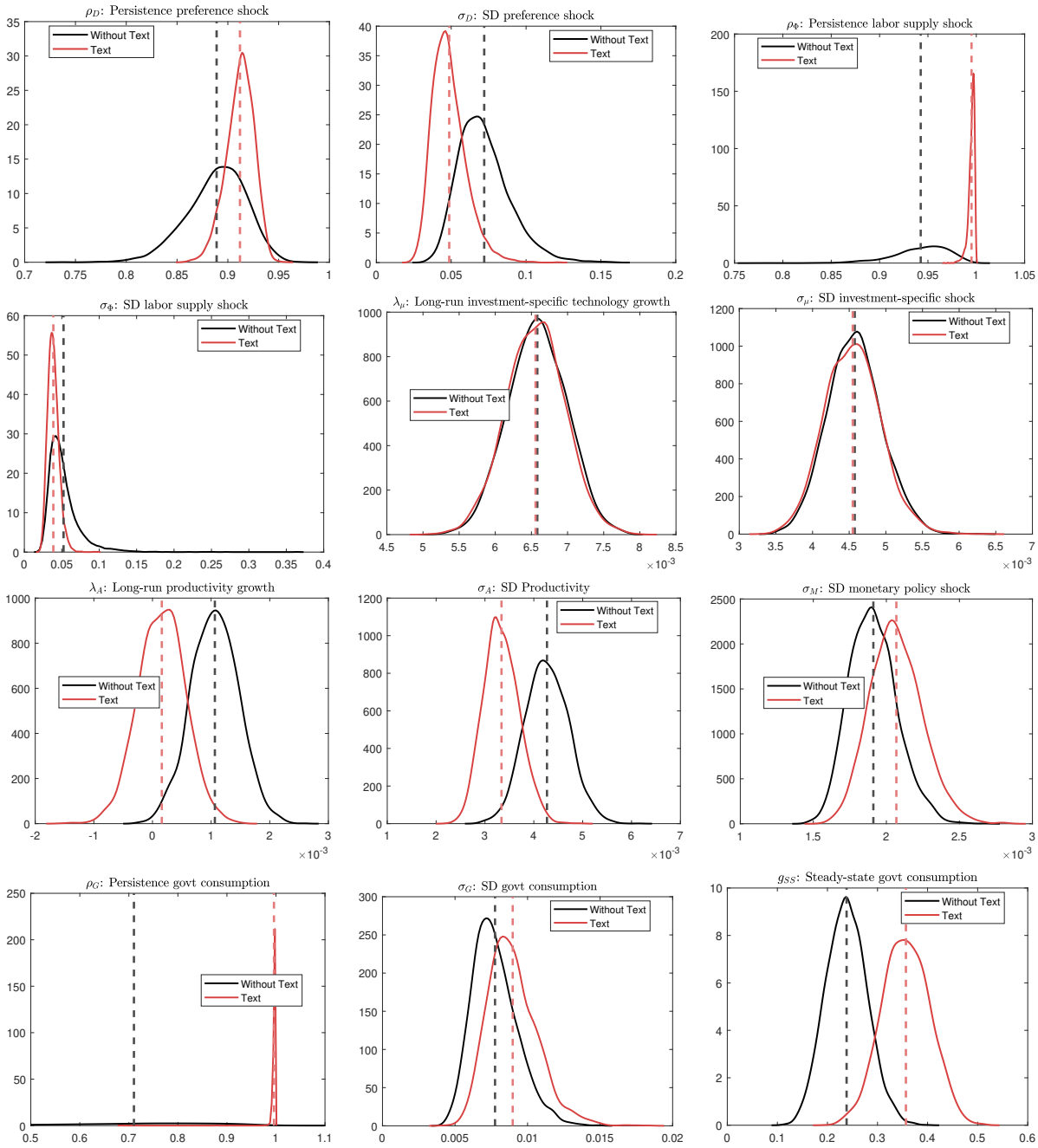
Figure E.2: Heterogeneity in the estimated probability of inclusion

E.2 Posterior distributions and the impulse responses



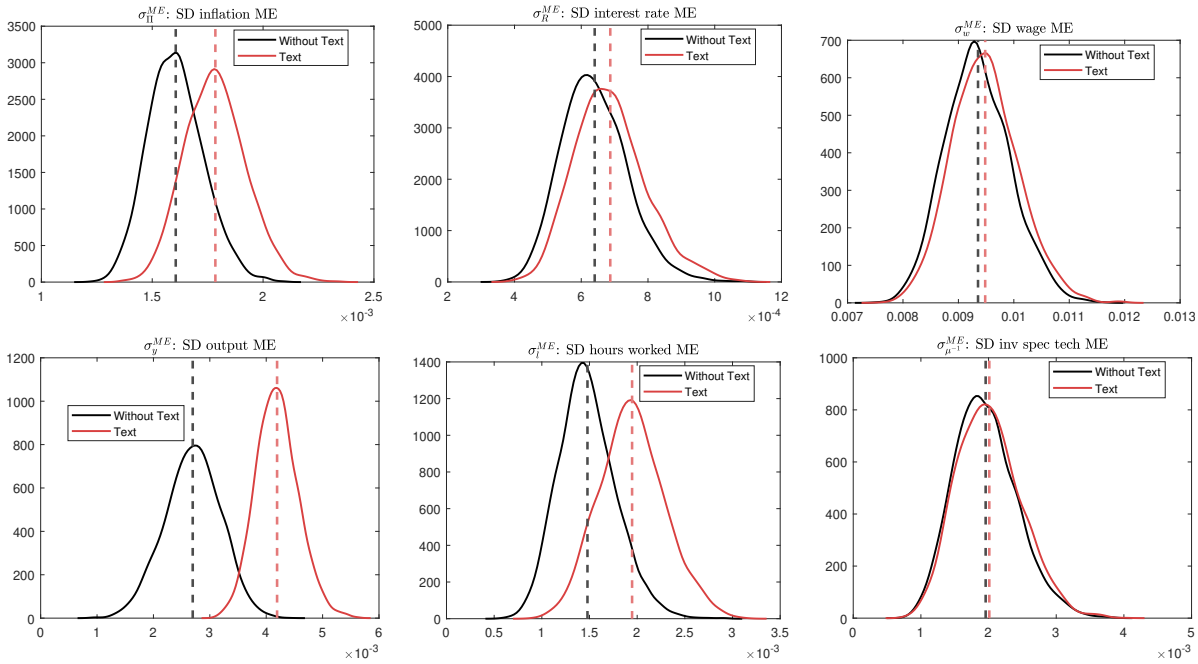
Note: The dashed vertical line is the posterior mean. The black line represents the model without text, and the red line represents the model with FOMC transcripts.

Figure E.3: Posterior distribution of parameters in our DSGE model 1



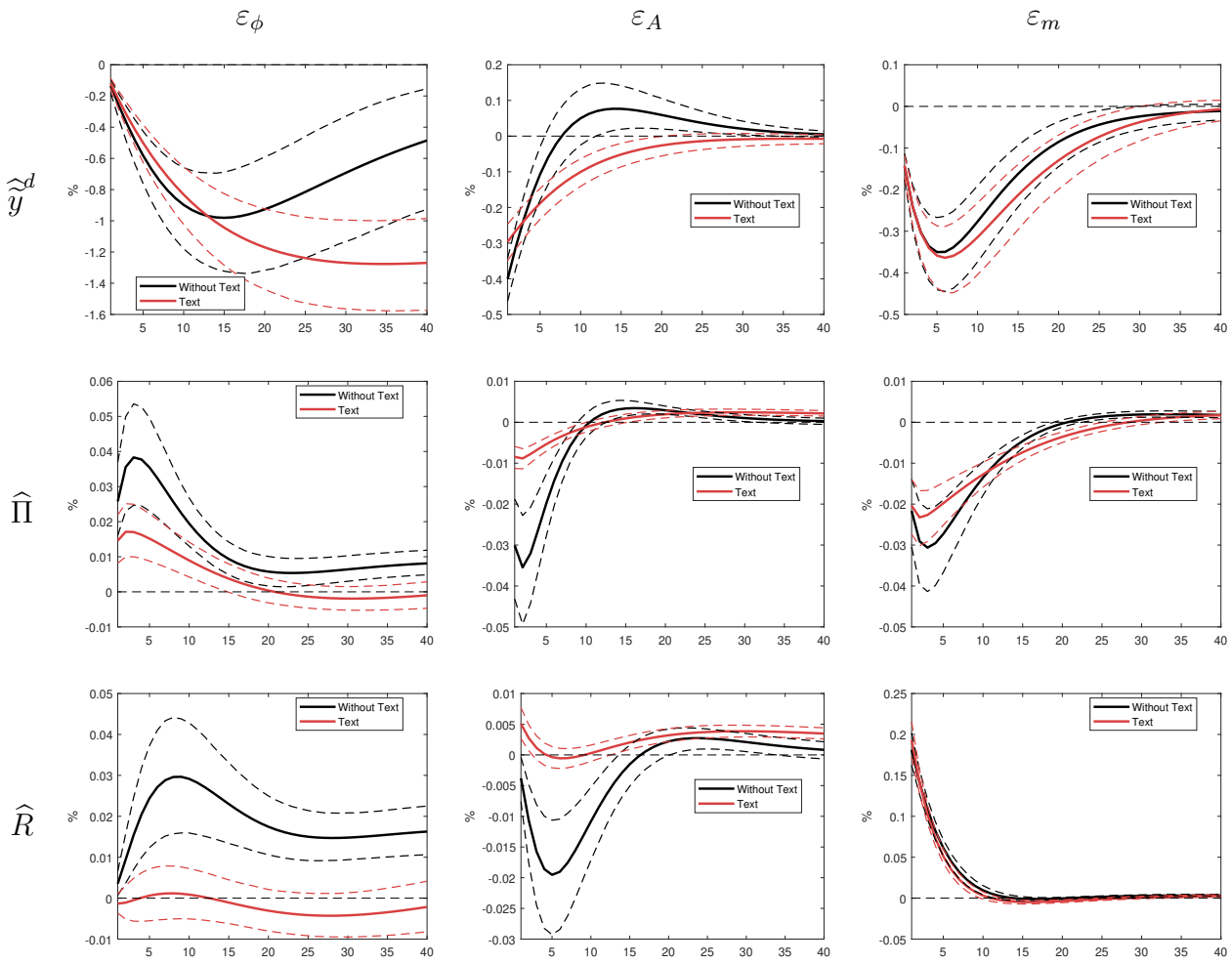
Note: The dashed vertical line is the posterior mean. The black line represents the model without text, and the red line represents the model with FOMC transcripts.

Figure E.4: Posterior distribution of parameters in our DSGE model 2



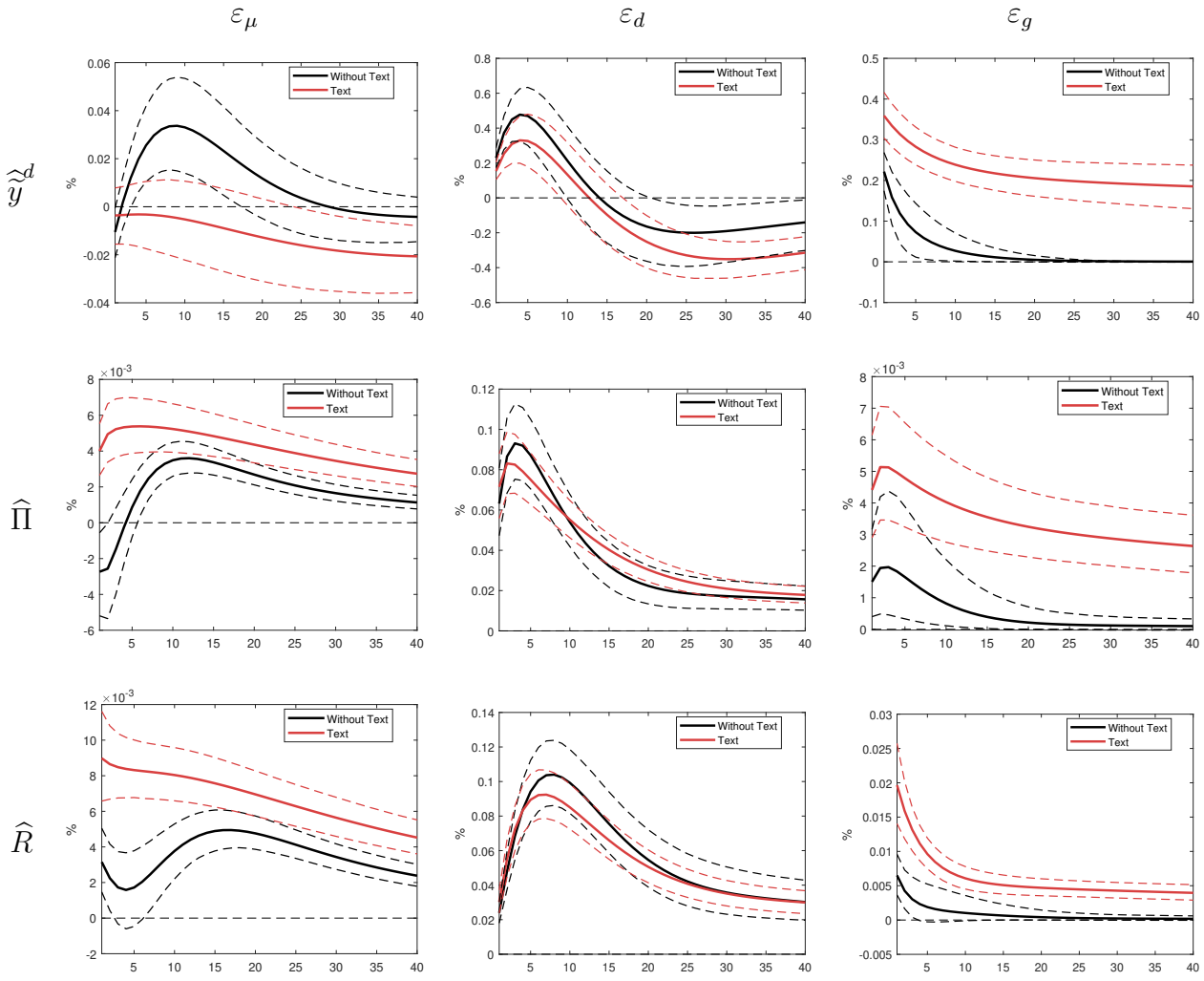
Note: The dashed vertical line is the posterior mean. The black line represents the model without text, and the red line represents the model with FOMC transcripts.

Figure E.5: Posterior distribution of parameters in our DSGE model 3



Note: 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 lines are the posterior mean, and the dashed lines are the pointwise 90% credible interval.

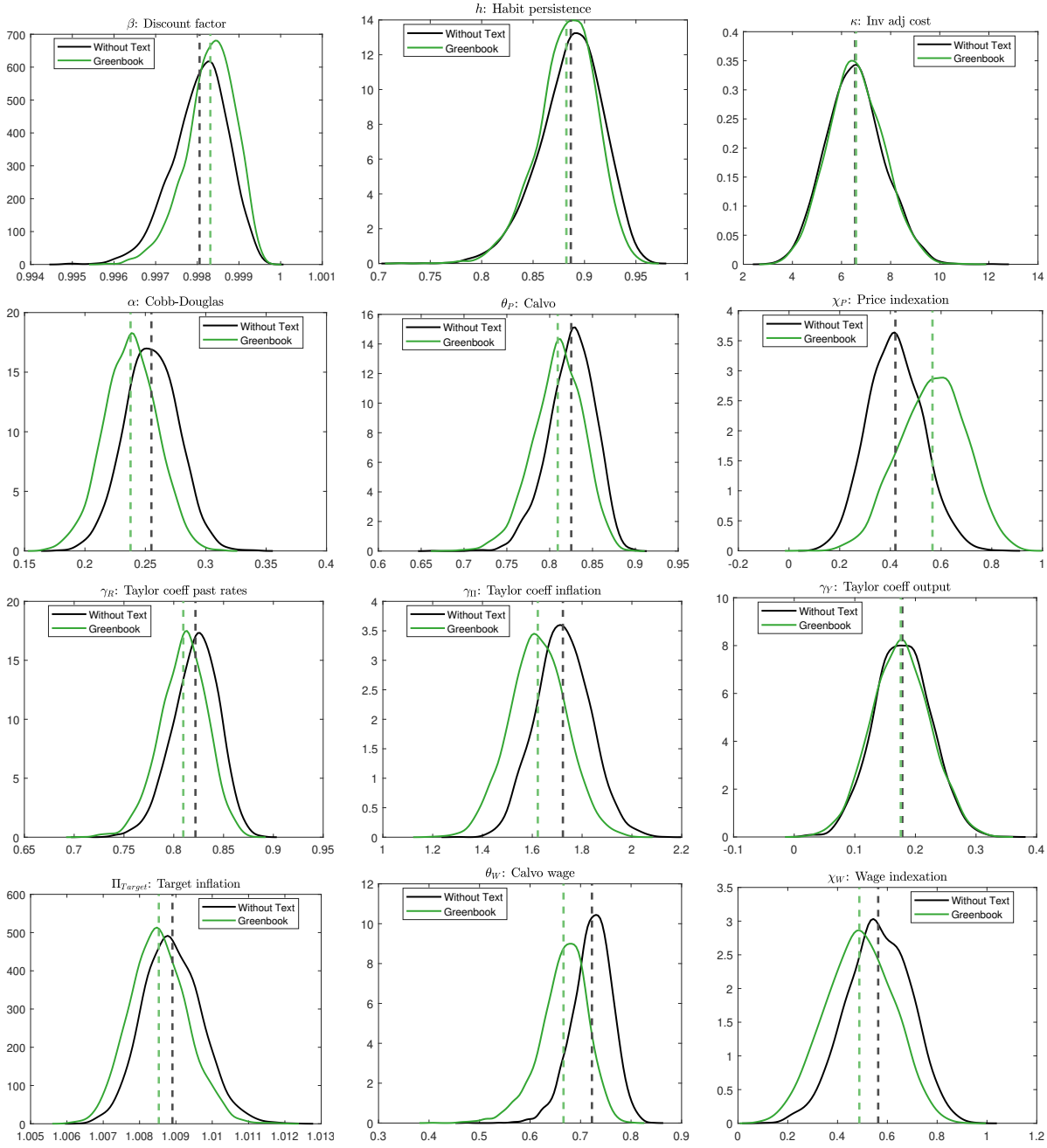
Figure E.6: Impulse responses of aggregate demand, inflation, and nominal interest rate to labor supply, aggregate technology, and monetary policy shocks



Note: 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 lines are the posterior mean, and the dashed lines are the pointwise 90% credible interval.

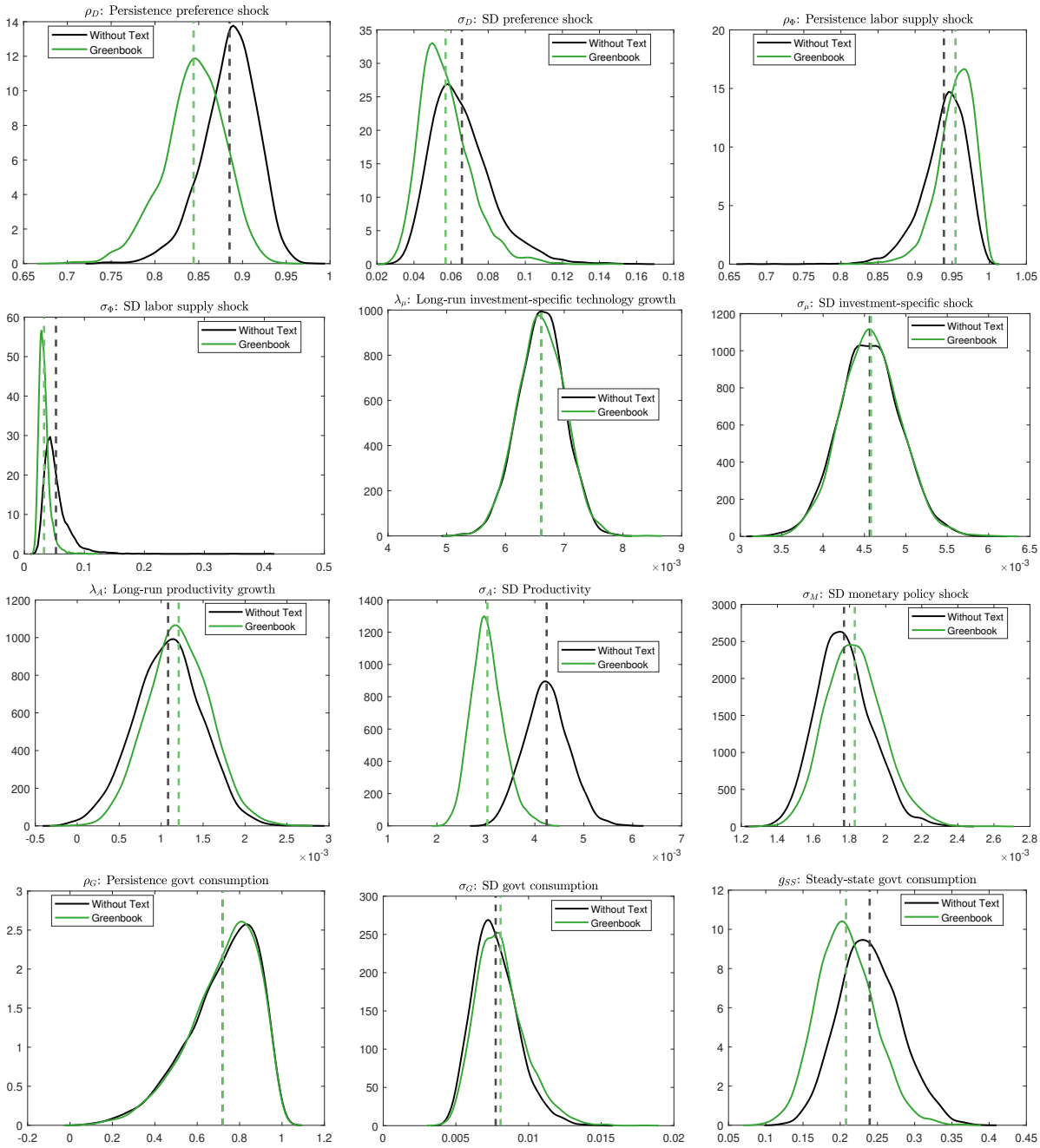
Figure E.7: Impulse responses of aggregate demand, inflation, and nominal interest rate to investment-specific technology, intertemporal preference, and government spending shocks

F Posterior distributions and impulse responses with Greenbook data



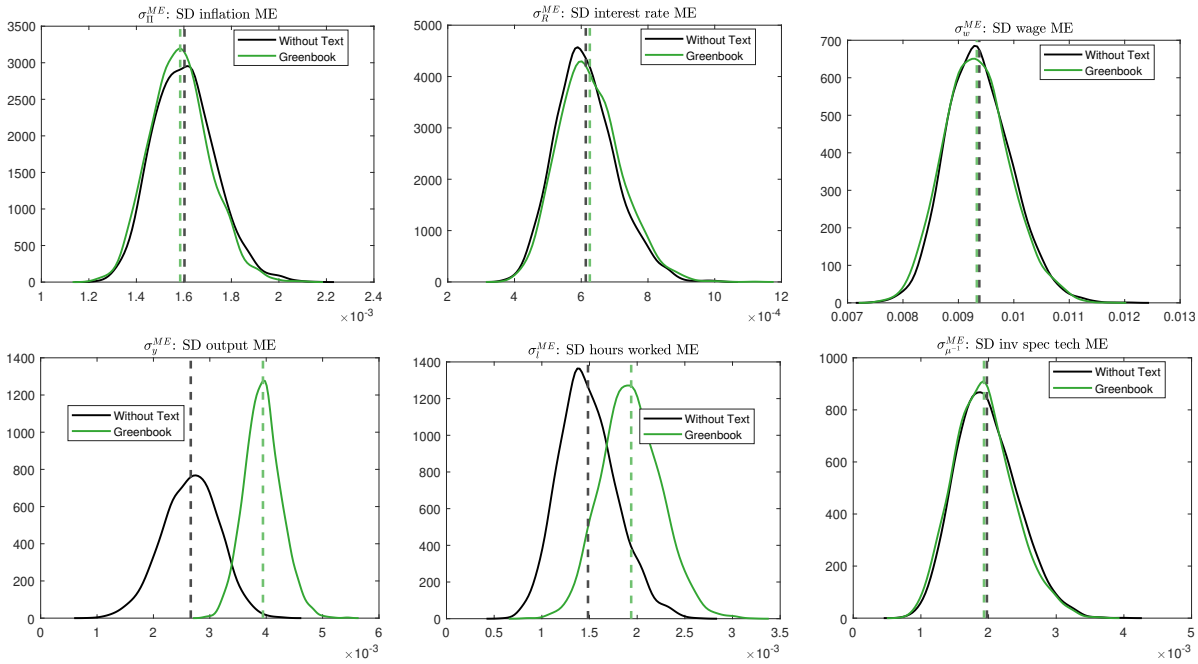
Note: The dashed vertical line is the posterior mean. The black line represents the model without text, and the green line represents the model with Greenbook data.

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



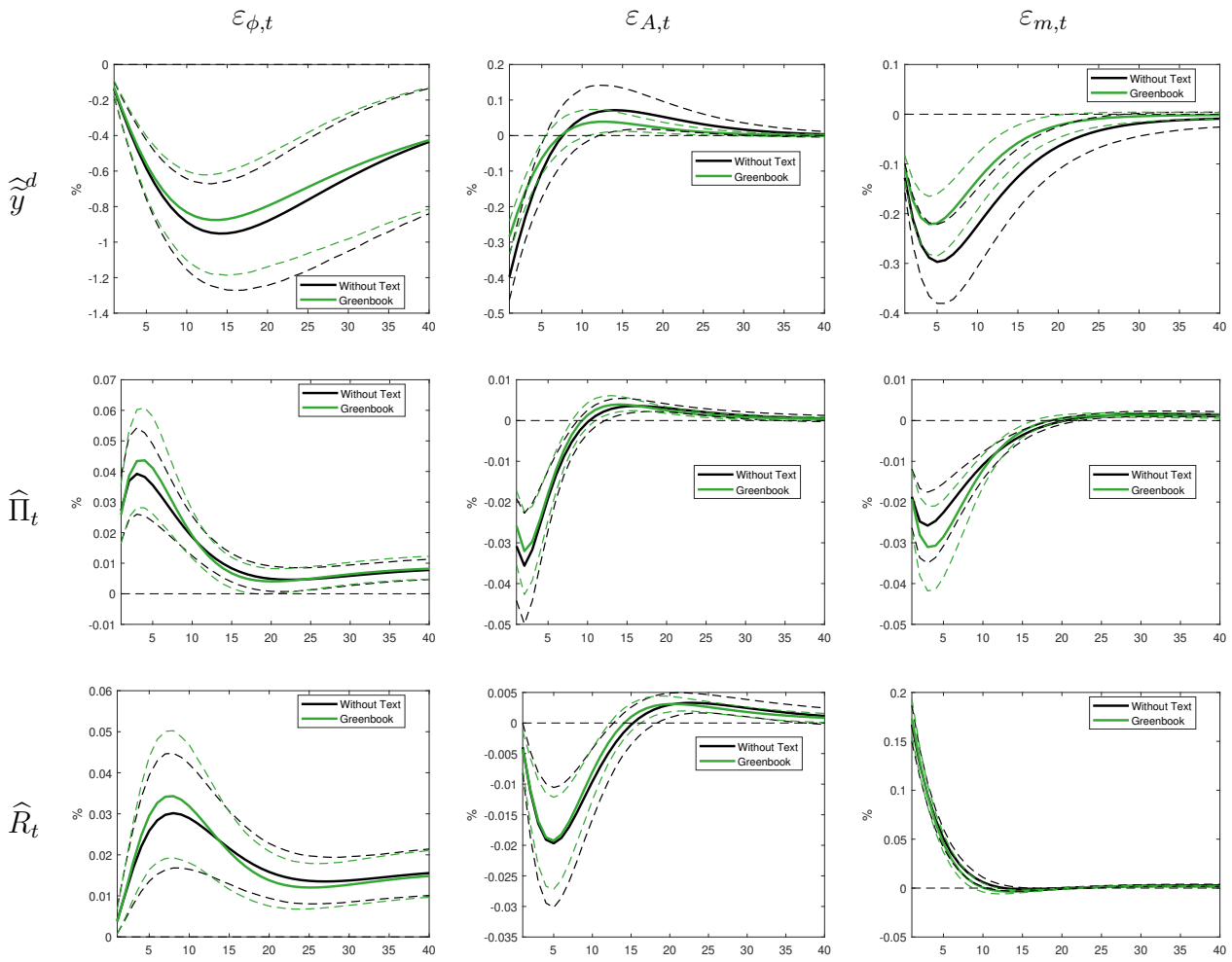
Note: The dashed vertical line is the posterior mean. The black line represents the model without text, and the green line represents the model with Greenbook data.

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



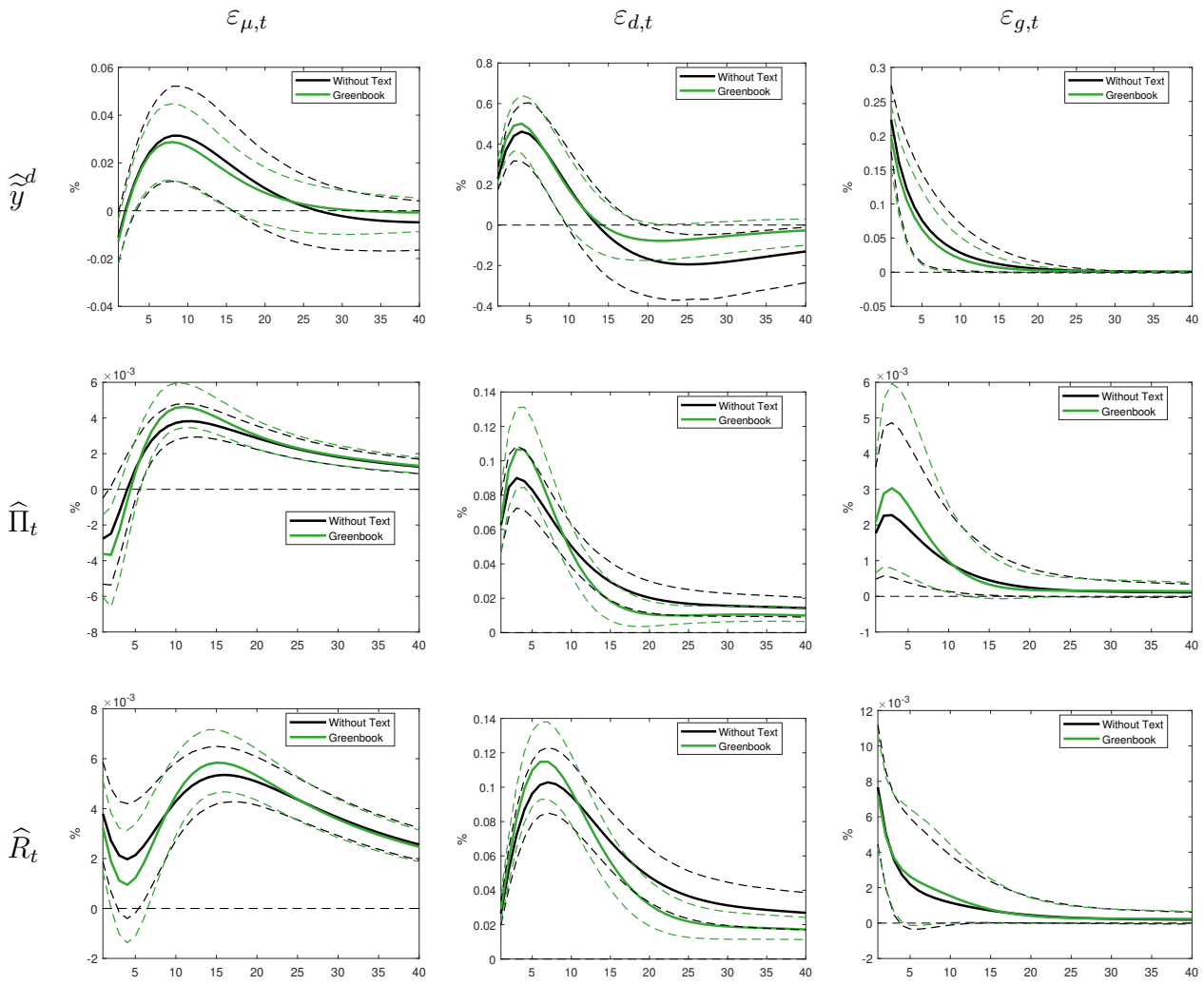
Note: The dashed vertical line is the posterior mean. The black line represents the model without text, and the green line represents the model with Greenbook data.

Figure F.3: Posterior distribution of parameters in our DSGE model 3



Note: 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 lines are the posterior mean, and the dashed lines are the pointwise 90% credible interval.

Figure F.4: Impulse responses of aggregate demand, inflation, and nominal interest rate to labor supply, aggregate technology, and monetary policy shocks



Note: 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 lines are the posterior mean, and the dashed lines are the pointwise 90% credible interval.

Figure F.5: Impulse responses of aggregate demand, inflation, and nominal interest rate to investment-specific technology, intertemporal preference, and government spending shocks