Ca' Foscari University of Venice

Department of Economics

Working Paper

Roberto Casarin
Antonio Peruzzi
Davide Raggi

Multiple Equilibria and the Phillips Curve: Do Agents Always Underreact?

ISSN: 1827-3580
No. 10/WP/2025

Working Papers
Department of Economics
Ca' Foscari University of Venice
No. 10/WP/2025
ISSN 1827-3580

Multiple Equilibria and the Phillips Curve: Do Agents Always Underreact?

Roberto Casarin
Ca' Foscari University of Venice

Antonio Peruzzi
Ca' Foscari University of Venice

Davide Raggi
Ca' Foscari University of Venice

Abstract
We study a New Keynesian Phillips curve in which agents deviate from the rational expectation paradigm and forecast inflation using a simple, potentially misspecified autoregressive rule. Consistency criteria à la Hommes and Zhu (2014) between perceived and actual laws of motion of inflation might allow for multiple expectational equilibria. Unfortunately, multiple equilibria models pose challenges for empirical validation. This paper proposes a latent Markov chain process to dynamically separate such equilibria. Moreover, an original Bayesian inference approach based on hierarchical priors is introduced, which naturally offers the possibility of incorporating equilibrium-identifying constraints with various degrees of prior beliefs. Finally, an inference procedure is proposed to assess a posteriori the probability that the theoretical constraints are satisfied and to estimate the equilibrium changes over time. We show that common prior assumptions regarding structural parameters favor the separation of equilibria, thereby making the Bayesian inference a natural framework for Markov–switching Phillips curve models. Empirical evidence obtained from observed inflation, output gap, and the consensus expectations from the Survey of Professional Forecasters supports multiple equilibria, and we find evidence of temporal variation in over- and under-reaction patterns, which, to the best of our knowledge, have not been previously documented. Specifically, we observe that agents tend to underreact to shocks when inflation is high and persistent, whereas they behave substantially as fully informed forecasters when the inflation level is low and stable, i.e., after the mid–nineties. We also find that the model does not suffer from the missing disinflation puzzle during the Great Recession.

Keywords
Bounded rationality; Markov Switching; Multiple equilibria; Under-reaction; Bayesian methods; Horseshoe hierarchical priors; Survey of Professional Forecasters

JEL Codes
C11, C24, E31, D84

Address for correspondence:
Davide Raggi
Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta S.Giobbe
30121 Venezia - Italy
e-mail: davide.raggi@unive.it

This Working Paper is published under the auspices of the Department of Economics of the Ca' Foscari University of Venice. Opinions expressed herein are those of the authors and not those of the Department. The Working Paper series is designed to divulge preliminary or incomplete work, circulated to favour discussion and comments. Citation of this paper should consider its provisional character.

The Working Paper Series
is available only on line
(https://www.unive.it/dec-workingpapers)
For editorial correspondence, please contact:
wp.dse@unive.it

Department of Economics
Ca' Foscari University of Venice
Cannaregio 873, Fondamenta San Giobbe
30121 Venice Italy
Fax: ++39 041 2349210

Multiple Equilibria and the Phillips Curve: Do Agents Always Underreact?

Roberto Casarin†, Antonio Peruzzi†, Davide Raggi†*
† Department of Economics, Ca' Foscari University of Venice.

Abstract

We study a New Keynesian Phillips curve in which agents deviate from the rational expectation paradigm and forecast inflation using a simple, potentially misspecified autoregressive rule. Consistency criteria à la Hommes and Zhu (2014) between perceived and actual laws of motion of inflation might allow for multiple expectational equilibria. Unfortunately, multiple equilibria models pose challenges for empirical validation. This paper proposes a latent Markov chain process to dynamically separate such equilibria. Moreover, an original Bayesian inference approach based on hierarchical priors is introduced, which naturally offers the possibility of incorporating equilibrium-identifying constraints with various degrees of prior beliefs. Finally, an inference procedure is proposed to assess a posteriori the probability that the theoretical constraints are satisfied and to estimate the equilibrium changes over time. We show that common prior assumptions regarding structural parameters favor the separation of equilibria, thereby making the Bayesian inference a natural framework for Markov–switching Phillips curve models. Empirical evidence obtained from observed inflation, output gap, and the consensus expectations from the Survey of Professional Forecasters supports multiple equilibria, and we find evidence of temporal variation in over- and under-reaction patterns, which, to the best of our knowledge, have not been previously documented. Specifically, we observe that agents tend to underreact to shocks when inflation is high and persistent, whereas they behave substantially as fully informed forecasters when the inflation level is low and stable, i.e., after the mid–nineties. We also find that the model does not suffer from the missing disinflation puzzle during the Great Recession.

JEL Codes: C11, C24, E31, D84

Keywords: Bounded rationality; Markov Switching; Multiple equilibria; Under-reaction; Bayesian methods; Horseshoe hierarchical priors; Survey of Professional Forecasters.

*Corresponding author, Email: davide.raggi@unive.it. The authors are grateful to the conference participants for helpful discussions at the 15th RCEA Bayesian Econometrics Workshop and at the Annual Conference of the International Association for Applied Econometrics (IAAE 2025). Financial support from MIUR, Grant 20228XTY79_003SH1 is gratefully acknowledged. Furthermore, the European Union funded this study – NextGenerationEU, in the framework of the GRINS – Growing Resilient, INclusive and Sustainable project (GRINS PE00000018 – CUP H73C22000930001), National Recovery and Resilience Plan (NRRP) – PE9 – Mission 4, C2, Intervention 1.3. The views and opinions expressed are solely those of the authors and do not necessarily reflect those of the European Union, nor can the European Union be held responsible for them.

1 Introduction

The New Keynesian Phillips Curve (NKPC) is arguably one of the most popular theoretical and empirical paradigms for analyzing inflation dynamics, originating from the seminal works of Fisher (1977) and Taylor (1979). In this context, inflation is mostly determined by expectations and marginal costs, rather than being driven by a sequence of past shocks, thus making the model forward-looking. A standard specification of the model reads as follows

$$\pi_t = \delta \pi^e_{t+1} + \gamma y_t + u_t, \tag{1}$$

where π_t is the observed inflation, π^e_{t+1} is its expectation, and y_t is a so-called gap variable measuring the economic slack. The *slope* parameter γ measures how the real economy affects inflation dynamics, whereas u_t is an economic shock. Because of the leading role of expectations in the model, it is of natural interest to consistently and efficiently specify the inflation expectation mechanism to be used. The rational expectation hypothesis, popularized since the contributions of Muth (1961) and Lucas (1972), remains the most widely accepted and standard paradigm in macroeconomics, largely due to its internal consistency in describing forward-looking behaviors, even though some limitations have been recently highlighted, amongst others, in Eichenbaum (2023). As stressed in Mavroeidis et al. (2014), obtaining the NKPC is easy under the rational expectation paradigm, even though it can be derived under different assumptions that allow for subjective expectations, as in Coibion et al. (2018) amongst others.

Despite its popularity, the plain NKPC under rational expectations has been criticized for its inability to adequately describe inflation data over more extended periods, particularly for its failure to account for structural breaks. For this reason, hybrid versions have been proposed. However, estimating these variants is challenging due to the parameter's weak identification, which leads to mixed empirical results (see Mavroeidis et al., 2014 for a review on these points).

One recent critique concerns the NKPC's inability to predict inflation dynamics during the Great Recession, which lasted from late 2007 to mid-2009. In particular, due to a significant reduction in the slack variable, standard NKPC models typically predict a substantial decline in inflation that has not been observed in real data. On this point, Ball and Mazumder (2011) suggested this puzzle is caused by not accounting for a time-varying slope parameter, whereas Coibion and Gorodnichenko (2015) focused on the expectation formation mechanism, for which the increase on firm's expectations during the recession, measured through the Michigan Survey of Consumers, compensated over the drop observed on the real economy variables. In general, the availability of survey data on expectations thus evidenced that sometimes agents systematically deviate from a pure rational behavior, as also stressed in Mankiw et al. (2003), Vissing-Jorgensen (2003), and Souleles (2004) and more recently in Coibion and Gorodnichenko (2012) and Bordalo et al. (2020) amongst others.

These empirical results, together with the inability of rational expectation models to match stylized facts observed in the data such as persistence (Hommes et al., 2023) led researchers to deviate from a pure version of rational expectations in favor of models with frictions, as in Mankiw and Reis (2002), Sims (2003) and Woodford (2003a), toward bounded rational alternatives (Sargent, 1999; Hommes and Zhu, 2014) or behavioural heuristics, as in Bordalo et al. (2020).

In this paper, we focus on the case in which agents are bounded rational, meaning that their forecasts of inflation are based on a misspecified econometric model. Recently, Clements (2024) provided some empirical evidence based on survey forecasts, highlighting significant heterogeneity among agents and noting that in many cases, these forecasts appear inconsistent with a hybrid version of the Phillips curve. This might suggest that agents do not use the NKPC to elicit their forecasts, but instead rely on different models.

Following Hommes and Zhu (2014), we consider a simple version of the Phillips curve in which agents form their expectations through a plain autoregressive model. Furthermore, to reach an internal consistency between the subjective prediction rule (the Perceived Law of Motion, PLM) and the NKPC under bounded rationality (the Actual Law of Motion, ALM), Hommes and Zhu (2014) propose a criterion that aims at identifying the parameters of the PLM by matching the unconditional moments of perceived and actual law of motions, in particular through their mean and the first order autocorrelation. Under suitable conditions, it has been proven that in this misspecified setup, there may be multiple equilibria that make PLM and ALM observationally equivalent. As stressed in Christiano et al. (2018), considering economic models with multiple equilibria is relevant, since each equilibrium might deliver different policy implications. Therefore, selecting the appropriate equilibrium at a given time is crucial for economic analysis. To support this intuition, Mertens and Ravn (2014) provides evidence in which the same policy might give different outcomes depending on the actual state of the economy, thus suggesting the necessity to properly recognize in which equilibrium the model is falling in.

Similarly to what happens with rational expectations, even in our setup, the number of equilibria is determined by the parameter values of the Phillips curve through a mapping that, in our case, is an analytical function that returns up to four possible solutions. To overcome the multiplicity of equilibria problem, often in the rational expectation literature a unique *relevant* equilibria, if it exists, is chosen by defining a selection criteria that helps picking the one that enjoys a given property, such as E-Stability, Forward Convergence or Minimal State Variables criteria (see for instance Driskill, 2006; Evans, 1985 or Cho and McCallum, 2015 for some definitions, results and comparisons).

This paper proposes a well-sounded inference procedure for a New Keynesian Phillips curve under bounded rationality, in which multiple expectation equilibria are jointly possible. To the best of our knowledge, it has not been done before. The main inference difficulty here is that the likelihood is not

uniquely defined due to the nonlinear parameter constraints inherent in the NKPC theory. Likelihood-based estimation methods on models displaying multiple equilibria trace back to Dagsvik and Jovanovic (1994) and Bisin et al. (2011), for instance, in which an equilibrium selection mechanism based on the maximization of the likelihood function is introduced. In line with Dagsvik and Jovanovic (1994), we assume this selection mechanism is stochastic and allows inflation to switch from one equilibrium to another over time. We contribute to the literature by proposing a Bayesian inference approach to identify the different equilibria in the NKPC jointly, estimate the parameter values within each *admissible* equilibrium, and evaluate the probability of observing each equilibrium over time. As a consequence, inflation is represented by a mixture that appears to be a suitable choice for studying the NKPC. In recent literature, it has been highlighted how the persistence and volatility of inflation have changed over time. These features have been recently explored in Lansing (2009) through a bounded rational model. Bouncing from one equilibrium to another might explain why the time-varying persistence of inflation and changing volatility are observed in real data. The original econometric contribution is threefold. First, a hidden Markov chain process is used to separate the regimes, and, unlike Dagsvik and Jovanovic (1994), the likelihood within each regime is given exactly, following the structural NKPC specification, and it is not approximated. Second, a Bayesian data-augmentation framework is introduced, which enables the estimation of latent variables and the connection between observations and equilibria. This is a crucial inferential aspect that has not been treated in previous studies (Dagsvik and Jovanovic, 1994). For the specification of the equilibrium transition model, we rely on the Markov switching literature (e.g., see Sims and Zha, 2006; Castelnuovo et al., 2014; Kaufmann, 2015), and assume a hidden Markov chain process with homogeneous transition driving the NKPC equations. Our approach not only allows us to quantify the probability of each equilibrium but also to estimate the transition probabilities and easily quantify the estimation uncertainty as a natural output of the inference procedure (e.g., see Frühwirth-Schnatter, 2006). The third relevant contribution regards the use of hierarchical prior distributions to connect the regime-specific parameters to the structural parameter values corresponding to the different equilibria. Hierarchical priors are now very popular in macroeconomic modeling (e.g., see Kadiyala and Karlsson, 1997; Canova and Ciccarelli, 2004; Canova, 2007; Canova and Ciccarelli, 2009) and can be used for various purposes, such as introducing partial pooling of information across equations and incorporating shrinkage effects in over-parameterized models. In this paper, we rely on one of the main arguments used in Bayesian inference, which considers parameter constraints as part of the extra-sample information to be included in the inference procedure through the prior distribution. Specifically, the distribution of the NKPC parameters at the second stage of the hierarchical prior is calibrated on stylized evidence, which allows for investigating the prior support assigned to different equilibria by past studies. Our prior specification assigns a positive probability to each equilibrium (a subset of the structural parameter space) and a quite

diffuse probability to the reduced-form parameter values within each equilibrium. We find that priors elicited from the literature favour equilibria separation. Regarding the first stage of the hierarchical priors which allow building a link between the NKPC equations and the bounded rational restrictions, we study different specifications, which include diffuse prior which does not assign large probability to the restrictions and the Horseshoe prior (e.g., see Carvalho et al., 2010; Cadonna et al., 2020), which favour the theoretical constraints. The main advantage of this approach is that the distribution of the parameters given the observations can be used to measure the evidence in favour of each equilibrium, to estimate the parameter values within each equilibrium, and the transition probabilities, under different levels of prior uncertainty. The quantification of posterior parameter uncertainty can be easily incorporated into forecasting exercises and policy considerations. The empirical evidence suggests that the parametric restrictions defining economic equilibria are supported by the data.

The second goal of this study is to evaluate the reaction to economic shocks at different equilibria. To this aim, we implement a test based on the impulse response functions proposed in Kučinskas and Peters (2024) and Juodis and Kučinskas (2023). In particular, given the model estimates, we investigate how the agent's expectations react to economic shocks, and we measure, for each possible equilibrium, if there is overreaction or underreaction compared to the rational expectation benchmark (Bordalo et al., 2020). Our main finding is that over/underreaction is related to the level of persistency observed in the inflation dynamics. When a highly persistent equilibrium is found, agents underreact, whereas under low persistence, the model substantially supports rationality. Finally, we test whether our empirical model accounts for the missing disinflation puzzle during the Great Recession. We find that, regardless of which equilibrium is selected, the model accurately predicts the dynamics of the observed inflation, even though the slope parameter is substantially different from zero. This feature is likely due to the nature of the bounded rational model, in which the expectation mechanism is calibrated on the observable moments of inflation.

The paper is organized as follows. Section 2 reviews the economic model proposed in Hommes and Zhu (2014), whereas Section 3 specifies the empirical strategy adopted to handle multiple equilibria. Empirical results are provided in Section 4, whereas a test for under/overreaction is provided in Section 4.4. Section 5 concludes.

2 New Keynesian Phillips Curve and Bounded Rationality

This section reviews the bounded rational NKPC as proposed in Hommes and Zhu (2014). In this setup, it is assumed agents fail to recognize the *true* sophisticated model that describes the state of the economy, either because of cognitive limitations, that is, agents might have difficulties while evaluating the impact of all the possible shocks affecting the economy in a general equilibrium framework (Eichenbaum, 2023), or because they consider too costly retrieving an extremely large set of informa-

tion necessary to implement an internally consistent fully rational model. Specifically, in this Phillips curve context, agents do not account for the role of some exogenous gap variables while eliciting their predictions on inflation. On the contrary, they prefer using a ready-to-use but misspecified econometric rule to form their expectations.

The starting point of our analysis is the standard specification of the Phillips curve as in Woodford (2003b):

$$\begin{cases} \pi_t = \delta \pi^e_{t+1} + \gamma y_t + u_t, & u_t \sim \mathcal{N}(0, \sigma^2_u), \\ y_t = a + \rho y_{t-1} + \epsilon_t, & \epsilon_t \sim \mathcal{N}(0, \sigma^2_\epsilon), \end{cases} \tag{2}$$

where π_t and π^e_{t+1} are, respectively, inflation rates and their expectations, whereas y_t is an economy slack variable that, in this paper, is the output gap. Here, y_t is assumed to be exogenous and, for simplicity, is described through a stationary AR(1) observable process. The shocks u_t and ϵ_t are uncorrelated and Gaussian with variances σ^2_u and σ^2_ϵ, respectively.

As suggested in Hommes and Zhu (2014), albeit common in the learning literature (see Fuster et al., 2010 for an extensive justification on this hypothesis), we assume the representative agent builds its expectations through a simple AR(1) process:

$$\pi_t = \tilde{\alpha} + \tilde{\beta}(\pi_{t-1} - \tilde{\alpha}) + \eta_t, \qquad \eta_t \sim \mathcal{N}(0, \sigma^2_\eta), \tag{3}$$

where η_t is a zero-mean Gaussian error with variance σ^2_η, the unconditional mean is $\tilde{\alpha}$ and the first order autocorrelation is $\tilde{\beta}$. This misspecified stochastic dynamics is also called *Perceived Law of Motion (PLM)*. Following this rule, subjective forecasts can be computed recursively, that is:

$$\pi^e_{t+1} = \tilde{\alpha} + \tilde{\beta}^2(\pi_{t-1} - \tilde{\alpha}). \tag{4}$$

If expectations are formed as in Eq. (4) and we assume the idea of self-fulfilling mistakes (Grandmont, 1998), then it must be that individual misspecified expectations affect the law of motion of inflation. Therefore, by plugging Eq. (4) into Eq. (2) we get the following *Actual Law of Motion (ALM)*, given by:

$$\begin{cases} \pi_t = \delta \left(\tilde{\alpha} + \tilde{\beta}^2(\pi_{t-1} - \tilde{\alpha}) \right) + \gamma y_t + u_t, \\ y_t = a + \rho y_{t-1} + \epsilon_t. \end{cases} \tag{5}$$

To make the subjective expectation formation mechanism coherent with the model of interest, one may impose consistency conditions that link the actual economic dynamics and the subjective expectations, ensuring that forecasts and expectations about a relevant variable are free from systematic forecasting errors.

One way to account for observable information within the forecasting problem is to assume that the individual forecasting rule, i.e., the PLM, and the true variable dynamics, that is, the ALM, share at least the same statistical moments. This requirement is called expectation consistency. A possible criterion for consistency is the *Consistent Expectation Equilibria*, proposed in Hommes and Sorger

(1998) and refined in Hommes and Zhu (2014), suggesting that PLM and ALM in Eqs. (3) and (5) must deliver the same unconditional expectation and first-order autocorrelation for the inflation rate. An extension to the multivariate case is Hommes et al. (2023).

The intuition behind this idea highlights how agents, by examining real data, can properly evaluate their unconditional moments. It is worth noting that the focus is just on unconditional average and first-order autocorrelation, which is enough to identify the parameters $\tilde{\alpha}$ and $\tilde{\beta}$ of the autoregressive model in Eq. (3).

In the Phillips Curve setup, a different forecasting rule based on a local level model has been previously considered in Lansing (2009), in which the consistency requirement is based uniquely on the first-order autocorrelation of $\Delta\pi_t$. To provide an example of how these consistency conditions are stated, note that from Eq. (3), the average inflation under PLM, namely, $E_{PLM}(\pi_t) = \tilde{\alpha}$. Therefore, the unconditional average inflation from the ALM in Eq. (5) must be equal to $\tilde{\alpha}$. The consistency criterion thus implies that:

$$E_{ALM}(\pi_t) = E_{ALM}(\delta(\tilde{\alpha} + \tilde{\beta}^2(\pi_{t-1} - \tilde{\alpha})) + \gamma y_t + u_t) \tag{6}$$

$$= \delta\left(\tilde{\alpha} + \tilde{\beta}^2(E_{ALM}(\pi_{t-1}] - \tilde{\alpha})\right) + \gamma E_{ALM}(y_t) = \tilde{\alpha} = E_{PLM}(\pi_t). \tag{7}$$

Since it must be that $E_{ALM}(\pi_t) = E_{PLM}(\pi_t) = \tilde{\alpha}$, and knowing that the expected value of y_t is $a/(1-\rho)$, for $\rho \neq 1$ and $\delta \neq 1$, we get:

$$\tilde{\alpha} = \delta\tilde{\alpha} + \gamma\frac{a}{1-\rho} \Rightarrow \tilde{\alpha} = \frac{\gamma a}{(1-\rho)(1-\delta)}. \tag{8}$$

After some algebra, it can be proved that in equilibrium, $\tilde{\beta}$ must satisfy the following condition

$$\tilde{\beta} = \delta\tilde{\beta}^2 + \frac{\gamma^2\rho(1-\delta^2\tilde{\beta}^4)}{\gamma^2(\delta\tilde{\beta}^2\rho+1) + (1-\rho^2)(1-\delta\tilde{\beta}^2\rho)\frac{\sigma_u^2}{\sigma_\epsilon^2}}. \tag{9}$$

This is a quartic function with four roots that can be computed analytically. However, not all the solutions are admissible from an economic point of view since, for instance, the admissible roots must be real and bounded in $-1 < \tilde{\beta} < 1$. Hommes and Zhu (2014) proves there is at least one admissible solution and shows how the number of admissible solutions depends on the other parameters (but not on the endogenous or the exogenous shocks).

Several special cases lie at the boundary of the parameter space. They are not considered in the analysis, as they exhibit trivial dynamics that are not helpful for most macroeconomic applications. Consider first the absence of persistence in the output gap ($\rho = 0$). In this case, the equilibrium condition reduces to $\gamma^2\tilde{\beta}(1-\tilde{\beta}) = 0$, yielding two economically meaningful solutions: a non-inflationary regime in which current inflation is independent of past inflation ($\tilde{\beta} = 0$), and a potentially explosive regime with $\tilde{\beta} = 1/\delta$, which exceeds one given that $\delta \in (0,1)$. Another instructive case arises when expected inflation does not influence actual inflation, i.e., $\delta = 0$. This leads to a unique solution for $\tilde{\beta}$:

$\tilde{\beta} = \frac{\gamma^2\rho}{\gamma^2+(1-\rho^2)\sigma_u^2/\sigma_\varepsilon^2}$, which always lies within the interval $(-1, 1)$ and is bounded in absolute value by $|\rho|$.

Finally, suppose both inflation and output gap dynamics are mutually independent, i.e., $\gamma = 0$. The equilibrium condition then simplifies to $(\rho\delta^2\tilde{\beta}^3 - \delta\rho\tilde{\beta}^2 - \delta\tilde{\beta} + 1)\tilde{\beta} = \tilde{\beta}(\delta\tilde{\beta} - 1)(\delta\rho\tilde{\beta}^2 - 1) = 0$, which admits three real solutions: $\tilde{\beta} \in \{0, 1/\delta, 1/(\delta\rho)\}$. These correspond to equilibria with increasing degrees of persistence, satisfying $0 < 1/\delta < 1/(\delta\rho)$ under the assumption that $\delta, \rho \in (0, 1)$.

3 Econometric Model and Inference

3.1 Markov-Switching Multiple Equilibria

This section proposes an econometric strategy to estimate the NKPC with possible multiple equilibria as described in Section 2. Usually, dealing with multiple equilibria is complex from an econometric perspective, as the set of constraints on the parameter space required to define each of them can make the likelihood function highly irregular. Therefore, often in the macroeconomic literature, several criteria aiming at defining the unique *interesting* equilibrium have been proposed, mainly looking at the properties of each solution and of its economic interpretability. The two most popular criteria refer to the Minimal State Variable concept and the Expectation Stability (E-Stability). McCallum (1983) suggests using the Minimum State Variable idea, where the solutions should be functions of a minimal set of variables. As an alternative, Evans (1985) proposes the concept of E-Stability, in which the *optimal solution* is the one that can be iteratively reached by an agent starting from a misspecified PLM.

Here, we chose to jointly model all the possible solutions and assess how the probability of each equilibrium changes over time through a dynamic stochastic selection criterion. As stressed in Section 2, since there are up to 4 possible consistent equilibria, we model inflation rates as a mixture of $1 \leq K \leq 4$ dynamic random processes, with K defining the number of admissible solutions. Using a mixture model to describe multiple equilibria has been proposed in Dagsvik and Jovanovic (1994), even though, in their specification, each equilibrium is chosen based on a rather ad hoc transition matrix of the Markov chain selection process, thus imposing constraints that might be too restrictive for general applicability. Moreover, in specifying their model, the multiple solution region is computed based on an approximate model. It is a function not only of parameters but also of macroeconomic shocks.

From now on, the equilibria parameters are denoted by $\tilde{\alpha}(\boldsymbol{\theta})$ and $\tilde{\boldsymbol{\beta}}(\boldsymbol{\theta})$ to stress their dependence from the other structural parameters $\boldsymbol{\theta} = \{\delta, \gamma, a, \rho, \sigma_u^2, \sigma_\epsilon^2\}$. Note that $\tilde{\boldsymbol{\beta}}(\boldsymbol{\theta}) = (\tilde{\beta}_1(\boldsymbol{\theta}), \ldots, \tilde{\beta}_K(\boldsymbol{\theta}))$ is a K-dimensional vector collecting the set of feasible solutions of Eq. (9). Our solution to account for multiple equilibria is based on a hierarchical procedure, where firstly inflation is defined by a standard

Markov switching process, in which persistences and the transition between them are driven by an unobservable K-states Markov chain s_t, $t = 1, \dots, T$. In this *preliminary* econometric specification, we assume the parameters $\boldsymbol{\beta} = (\beta_1, \dots, \beta_K)'$ identifying the *statistical* regimes are free and therefore not necessarily related to the equilibria defined in Eq. (9). Secondly, we impose the restrictions required to guarantee each statistical regime $\boldsymbol{\beta}$ is pinned down to the economic equilibria $\tilde{\boldsymbol{\beta}}(\boldsymbol{\theta})$.

More specifically, at the first stage, the mixture of inflation dynamics is

$$\pi_t = \begin{cases} \delta\alpha(1-\beta_1^2) + \gamma a + \delta\beta_1^2\pi_{t-1} + \gamma\rho y_{t-1} + u_t + \gamma\epsilon_t, & s_t = 1 \\ \vdots & \\ \delta\alpha(1-\beta_K^2) + \gamma a + \delta\beta_K^2\pi_{t-1} + \gamma\rho y_{t-1} + u_t + \gamma\epsilon_t, & s_t = K. \end{cases} \tag{10}$$

The dynamics of the latent Markov chain process is driven by the transition matrix $\mathbb{P}(s_t = k | s_{t-1} = l) = p_{lk}$, $l, k = 1, \dots, K$ which also provides a measure of the duration of the equilibria. It is worth noting that this specification, based on a Markov switching dynamics for $\boldsymbol{\beta}$ appears appropriate as many empirical studies provided support for a time varying persistence on the U.S. inflation (see, for instance, Barsky, 1987; Ball, 2000 or Lansing, 2009, amongst others).

Following the recent literature on the use of Surveys of Professional Forecasters (SPF) in forecasting inflation (e.g., see Clements, 2024; Mertens and Nason, 2020), to help identifying the parameters of the PLM, we also enrich the NKPC by explicitly reference Eq. (4), in which survey expectations are linked to the misspecified PLM through the following measurement equation:

$$\pi_{t+1}^{SPF} = \alpha + \nu_t(\pi_{t-1} - \alpha) + w_t, \tag{11}$$

where $\nu_t = \boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta}$, π_{t+1}^{SPF} is the observed consensus-level prediction made at time t for the inflation at $t+1$ obtained from the Survey on Professional Forecasters (Federal Reserve of Philadelphia, 2024) and w_t is a Gaussian shock with zero mean that acts as a measurement error.

The SPF-enriched Phillips curve, denoted by $\mathcal{M}^{PC}(K)$, is specified in matrix form as follows:

$$\begin{pmatrix} \pi_t \\ y_t \\ \pi_{t+1}^{SPF} \end{pmatrix} = \begin{pmatrix} \delta\alpha(1-\nu_t) + \gamma a \\ a \\ \alpha(1-\nu_t) \end{pmatrix} + \begin{pmatrix} \delta\nu_t & \gamma\rho & 0 \\ 0 & \rho & 0 \\ \delta\nu_t & 0 & 0 \end{pmatrix} \begin{pmatrix} \pi_{t-1} \\ y_{t-1} \\ \pi_t^{SPF} \end{pmatrix} + \begin{pmatrix} 1 & \gamma & 0 \\ 0 & 1 & 0 \\ 0 & 0 & 1 \end{pmatrix} \begin{pmatrix} u_t \\ \epsilon_t \\ w_t \end{pmatrix}, \tag{12}$$

where $\epsilon_t \sim \mathcal{N}(0, \sigma_\epsilon^2)$ and $u_t \sim \mathcal{N}(0, \sigma_u^2)$ are independent Gaussian error terms, while $\boldsymbol{\xi}_t = (\mathbb{I}(s_t = 1), \dots, \mathbb{I}(s_t = K))'$ is a vector of indicator variables selecting the active regime at each time t.

At this point, it is then necessary to define the link between the statistical regimes $\boldsymbol{\beta}$ and the economic equilibria $\tilde{\boldsymbol{\beta}}(\boldsymbol{\theta})$. Compared to a standard Structural VAR model, estimation of Eq. (12) is challenging if one tries to directly plug in $\tilde{\alpha}$ and $\tilde{\boldsymbol{\beta}}$, since these parameters are nonlinear combinations of the others. See Gallant (1987, 2020) for a detailed treatment of this problem. In addition, our set of restrictions leads to untractable likelihood functions, and in turn, numerical estimation algorithms

become unstable in practice (Chernozhukov and Hong, 2004). In the next section, we shall make explicit the dependence between the statistical regimes $\boldsymbol{\beta}$ and the equilibria $\tilde{\boldsymbol{\beta}}(\boldsymbol{\theta})$ solutions of Eq. (9).

3.2 Bayesian Inference

The Bayesian approach offers the possibility of including constraints on the parameter space through the prior distribution, also allowing a certain degree of belief to be assigned to the constraints themselves. Accordingly, the posterior distributions provide a measure of how the data support the prior constraints. This strategy is particularly appealing in econometrics, where prior beliefs often stem from theoretical or empirical regularities. It is thus a well-established practice to impose linear or nonlinear restrictions via the prior (in the VAR literature, notable examples include Korobilis, 2022 and Chan, 2022). Furthermore, Bayesian inference has proved to be effective and widely documented when considering nonlinear models (e.g., see Kim and Nelson, 1999; Giordani et al., 2007). An additional complication of the model arises from the number of admissible equilibria, which depends on the values assigned to the structural parameters. To overcome the problem, we set a priori the number of equilibria by constraining the parameter space to guarantee exactly K admissible solutions for Eq. (9). This further constraint is implemented numerically. As will be discussed later, under our prior specification, the number of admissible solutions for Eq. (9) is either 1 or 3. However, we can also consider the case $K = 2$ if we impose the E-stability requirement. Extensive simulations never delivered exactly 2 or 4 admissible and distinct solutions. Then each single specification, i.e., $K = 1, 2$ and 3, is evaluated separately and ex-post compared with the others.

This paper assumes a hierarchical prior distribution to smoothly include the prior constraints and connect them to the structural model parameters. At the first prior stage, the parameters α and $\boldsymbol{\beta}$ are assumed to fluctuate around the equilibrium $\tilde{\alpha}(\boldsymbol{\theta})$ and $\tilde{\boldsymbol{\beta}}(\boldsymbol{\theta})$, where $\tilde{\boldsymbol{\beta}}(\boldsymbol{\theta})$ is the vector of K real-valued and unique solutions of Eq. (9) such that $\tilde{\beta}_1(\boldsymbol{\theta}) < \ldots < \tilde{\beta}_K(\boldsymbol{\theta})$, $K \leq 4$. In particular, the prior for the unconditional average α is based on a Gaussian prior centered at $\tilde{\alpha}(\boldsymbol{\theta})$. For the persistence parameters $\boldsymbol{\beta}$, we assume a horseshoe prior (Carvalho et al., 2010; Bhadra et al., 2017). This latter choice is specifically designed to induce shrinking effects in regression modeling, and in our setup, is introduced to favor the theoretical restrictions given by the bounded rational NKPC. The Horseshoe hyperparameter estimates provide a natural way to measure how much theoretical restrictions are supported by the data[1]. It is defined as

$$\alpha|\boldsymbol{\theta} \sim \mathcal{N}(\tilde{\alpha}(\boldsymbol{\theta}), \sigma_\alpha^2), \quad \beta_k|\boldsymbol{\theta}, \tau^2, \lambda_k^2 \sim \mathcal{N}(\tilde{\beta}_k(\boldsymbol{\theta}), \tau^2\lambda_k^2), k = 1, \ldots, K, \tag{13}$$

$$\tau \sim \mathcal{C}^+(0,1), \quad \lambda_k \sim \mathcal{C}^+(0,1), k = 1, \ldots, K, \tag{14}$$

where $\mathcal{C}^+(0,1)$ is the half-Cauchy distribution with a probability density function $f(x;0,1) = 2/\pi\left(1+x^2\right)$

[1]Robustness checks based on a standard prior choice, namely a Gaussian hierarchical prior, are also available in Appendix D. In this respect, the qualitative results are similar.

for $x > 0$, σ_α^2 is an hyper parameter set to have a vague prior on α and $\mathcal{N}(m, s^2)$ denotes the Normal distribution with location m, scale s and support (a, b). Compared to the horseshoe prior of Carvalho et al. (2010) our specification assumes the prior is not centered at 0 but at $\tilde{\beta}_k(\boldsymbol{\theta})$. Integrating out λ_k from the conditional distribution of $p(\beta_k|\boldsymbol{\theta}, \lambda_k, \tau)$, one obtains a distribution

$$p(\beta_k|\boldsymbol{\theta}) = 1/\sqrt{(2\pi^3)}\exp((\beta_k - \tilde{\beta}_k(\boldsymbol{\theta})^2/2)E_1((\beta_k - \tilde{\beta}_k(\boldsymbol{\theta})^2/2) \tag{15}$$

where $E_1(\cdot)$ is the exponential integral function. This distribution has a spike at $\tilde{\beta}k(\boldsymbol{\theta})$ and tails heavier than those of the normal distribution, as stated in the following remark.

Remark 1. *Assuming $\tau = 1$, the univariate horseshoe density $p(\beta_k|\boldsymbol{\theta})$ satisfies the following: a) $\lim_{\beta_k - \tilde{\beta}_k(\boldsymbol{\theta}) \to 0} = \infty$; b) for $\beta_k - \tilde{\beta}_k(\boldsymbol{\theta}) \neq 0$*

$$\frac{C}{2}\log\left(1 + \frac{4}{(\beta_k - \tilde{\beta}_k(\boldsymbol{\theta}))^2}\right) < p(\beta_k|\boldsymbol{\theta}) < C\log\left(1 + \frac{2}{(\beta_k - \tilde{\beta}_k(\boldsymbol{\theta}))^2}\right),$$

where $C = 1/\sqrt{(2\pi^3)}$.

Proof. This remark is a straightforward extension of the result in Carvalho et al. (2010), Theorem 1. See also Appendix B.1. □

The spike and tails of the modified Horseshoe prior favor either the prior constraints or strong signals against them. The hyperparameters λ_k and τ can be used to measure a posteriori the evidence in favour of the constraints. The shrinkage weight $\kappa_k = \frac{1}{1+\lambda_k^2\tau^2} \in [0, 1]$ indicates strong adherence of β_k toward the theoretical parameter $\tilde{\beta}_k(\boldsymbol{\theta})$ when κ_k is close to 1, thus implying the economic restrictions defining the equilibria are supported by the data. A value of κ_k close to 0 points toward deviations from the economic restrictions. This property represents a main advantage, since κ_k can be used to validate the economic restrictions implied by the bounded rational model.

To guarantee the stationary requirement necessary to derive the equilibria, we further assume $0 \leq \beta_k < 1$. In principle, this prior choice also enables extensions of the NKPC that incorporate a larger number of exogenous variables and enrich the VAR specification by adding policy rules. In these extended specifications, the horseshoe prior for the NKPC core variables is centered at their theoretical values, whereas the coefficient prior for the other variables is centered at zero as in the standard horseshoe prior model.

Regarding the NKPC structural parameters $\boldsymbol{\theta}$ we assume informative prior distributions for the coefficients and an improper prior for the scale parameters:

$$\delta \sim \mathcal{U}(0, 1),\ \gamma \sim \mathcal{N}(\mu_\gamma, \sigma_\gamma^2),\ \rho \sim \mathcal{U}(0, 1) \tag{16}$$

$$a \sim \mathcal{N}(\mu_a, \sigma_a^2),\ \sigma_u^2 \sim \mathcal{U}(0, \infty),\ \sigma_\epsilon^2 \sim \mathcal{U}(0, \infty),\ \sigma_w^2 \sim \mathcal{U}(0, \infty), \tag{17}$$

such that $\boldsymbol{\theta} \in \mathcal{A}$, where $\mathcal{A} \subset [0,1]^2 \times \mathbb{R}^2 \times \mathbb{R}_+^3$ denotes the set of parameter values such that Equation 9 has K solutions. The hyperparameters μ_γ and μ_a are set based on the results from previous studies in the NKPC literature, following the calibration exercise of Hommes and Zhu (2014).

As standard in the Markov Switching literature (Frühwirth-Schnatter, 2006), for the transition matrix parameter $\mathbf{p}_l$ we assume a Dirichlet prior for each row of the transition matrix, that is:

$$\boldsymbol{p}_l \sim \mathcal{D}ir(\omega_1, \ldots, \omega_K),\ l = 1, \ldots, K, \tag{18}$$

where $\mathcal{D}ir(\omega_1, \ldots, \omega_K)$ denotes the Dirichlet distribution with concentration parameters $\omega_l > 0$.

We examine, through a simulation study, the relationship between the prior distribution on the structural parameters $\boldsymbol{\theta}$ and the number of feasible equilibria implied by the model. A detailed account of these results is provided in Appendix A. Notably, under the prior specification aligned with the calibration exercise in Hommes and Zhu (2014), the model yields a unique feasible equilibrium in the vast majority of simulations (99.86%), while scenarios with three distinct equilibria occur only rarely (0.14%). Configurations involving two or four distinct feasible equilibria do not arise under this prior.

Several alternative prior specifications for α and $\boldsymbol{\beta}$ have also been considered to demonstrate the robustness of our empirical results to the choice of the prior distribution. Specifically, a diffuse prior based on a Gaussian distribution is proposed, and some further details are reported in Appendix D.

3.3 Posterior Approximation

We approximate the joint posterior distribution by Markov-chain Monte Carlo (MCMC) sampling. Our Gibbs sampling algorithm iterates the following three blocks. In the first block, the structural parameters $\boldsymbol{\theta}$ are drawn from the conditional posterior distribution $p(\boldsymbol{\theta}|\ldots)$ by Adaptive Metropolis-Hastings (Andrieu and Thoms, 2008). In the second block, the intercept α samples are obtained using exact Monte Carlo, as the full conditional distribution $p(\alpha|\ldots)$ is a normal distribution.

The coefficients $\boldsymbol{\beta}$ are drawn from a multivariate normal, and then the candidates are accepted if they belong to the $(0,1)$ interval to impose stationarity. We adopt a Metropolis-Hastings step to obtain samples for the parameter τ and λ_k for $k = 1, \ldots, K$, as we do not have a closed-form expression for their posterior distribution under the truncated-normal assumption.

The Markov-Switching parameters and the latent variables are sampled in the third and last block. The samples for the parameters $\boldsymbol{p}_l$ are obtained by exact Monte Carlo, since the full conditional distribution $p(\boldsymbol{p}_l|\ldots)$ for $l = 1, \ldots, K$ is a Dirichlet distribution. The state trajectory $\mathbf{s}$ is sampled via a Forward-Filtering Backward-Sampling algorithm (see Frühwirth-Schnatter, 2006). Details of the algorithm together with a Directed Acyclic Graph of the complete model are described in Appendix B.3.

We assessed the efficiency of the proposed sampler through some simulation experiments. The results are summarized in Appendix C. We simulate co-movements of π_t and y_t following the DGP of Eq.

(12)[2] and in which the structural parameters have been selected to guarantee three distinct equilibria. We choose a state trajectory with increasing persistence. Estimates have been obtained by running our MCMC algorithm for 30,000 iterations with a burn-in of 15,000. The posterior distributions in the structural parameter space correctly concentrate around the ground-truth parameter values. Standard MCMC diagnostic techniques such as the MCMC trace plots reported in Fig. C.4 indicate a good mixing of the chains, confirmed by the convergence diagnostics in Tab. C.1 as suggested in Geweke (1992).

4 Empirical Results

4.1 Data

The dataset comprises quarterly U.S. macroeconomic time series spanning from 1960Q1 to 2023Q2, sourced from the FRED database (Federal Reserve Bank of St. Louis, 2025). Inflation is measured as the annualized rate of growth of the Consumer Price Index (CPI). The output gap is constructed by detrending the logarithm of the Industrial Production Index using the Hodrick-Prescott filter, following the approach of Lansing and Ma (2017).

To help identify the expectation-related parameters of the model, we supplement these data with inflation expectations from the Survey of Professional Forecasters (SPF), provided by the Federal Reserve Bank of Philadelphia (Federal Reserve Bank of Philadelphia, 2025). Specifically, we use the consensus (mean) forecast for one-quarter-ahead inflation as a proxy for the perceived expectation π^e_{t+1} in Eq. (4). The SPF expectations are available quarterly, starting from 1981 Q1.

Our likelihood consists of three main components: one related to inflation, one to the output gap, and one to SPF inflation expectations. While the inflation and output gap series span the entire sample period (1960Q1–2023Q2), the SPF inflation data are incorporated starting from 1981Q1. This approach allows us to exploit the full temporal coverage of the macroeconomic series while leveraging survey-based measures to identify subjective expectations in the latter part of the sample.

4.2 Multiple Equilibra

Estimates have been obtained by running our MCMC algorithm for 30,000 iterations with a burn-in of 15,000. We consider separately 3 cases:

- $K = 3$: the model's parameters $\boldsymbol{\theta}$ are constrained to obtain 3 distinct expectational equilibria.
- $K = 2$: we allow for 2 distinct equilibria. In this case, the focus is on the two equilibria satisfying the E-stability criterion.

[2]For simplicity, in this simulation exercise we ruled out the SPF equation.

	Multiple Equilibria Phillips Curve Model $\mathcal{M}^{PC}(K)$					
	$K=1$		$K=2$		$K=3$	
Parameters	Post. Mean	95% Cred. Interval	Post. Mean	95% Cred. Interval	Post. Mean	95% Cred. Interval
δ	0.99	[0.971, 0.999]	0.99	[0.979, 0.999]	0.99	[0.977, 0.999]
γ	0.03	[-0.065, 0.161]	0.22	[0.131, 0.303]	0.21	[0.133, 0.296]
ρ	0.91	[0.857, 0.953]	0.95	[0.914, 0.981]	0.95	[0.915, 0.981]
a	-0.04	[-0.177, 0.049]	0.00	[-0.003, 0.015]	0.00	[-0.004, 0.016]
σ_u^2	7.18	[6.045, 8.409]	5.11	[4.359, 5.992]	4.74	[3.975, 5.702]
σ_e^2	1.02	[0.877, 1.189]	1.05	[0.905, 1.229]	1.06	[0.892, 1.244]
σ_w^2	1.18	[0.943, 1.457]	0.61	[0.497, 0.752]	0.51	[0.406, 0.634]
α	3.08	[2.848, 3.329]	2.50	[2.322, 2.696]	2.43	[2.257, 2.606]
$\tilde{\alpha}$	1.17	[-2.708, 6.185]	2.38	[-2.411, 7]	2.18	[-2.543, 6.634]
β_1	0.61	[0.559, 0.655]	0.35	[0.28, 0.42]	0.33	[0.233, 0.396]
$\tilde{\beta}_1$	0.00	[0, 0.023]	0.10	[0.051, 0.217]	0.11	[0.051, 0.272]
β_2			0.92	[0.878, 0.956]	0.74	[0.662, 0.815]
$\tilde{\beta}_2$			0.99	[0.983, 0.999]	0.93	[0.739, 0.987]
β_3					0.98	[0.957, 0.999]
$\tilde{\beta}_3$					0.99	[0.984, 0.999]
$\tilde{\lambda}_1$	6.29	[0.187, 34.328]	4.70	[0.154, 28.346]	4.16	[0.139, 17.988]
$\tilde{\lambda}_2$			1.96	[0.041, 9.623]	4.82	[0.125, 27.568]
$\tilde{\lambda}_3$					7.13	[0.004, 8.165]
τ	7.94	[0.203, 30.243]	4.69	[0.088, 19.888]	1.77	[0.052, 7.889]
DIC	2443.08		2266.99		2420.94	

Table 1: Parameter estimates from the MCMC algorithm for the three models with Horseshoe Prior. The Deviance Information Criterion (DIC) is also reported.

- $K=1$: the model's parameters are constrained to obtain a unique expectational equilibrium.

The three models are compared based on the Deviance Information Criteria (DIC) of Spiegelhalter et al. (2002) and the estimates are summarized in Tab. 1, which reports posterior averages and 95% credibility intervals of the parameters. It is worth noting that, despite a prior choice that assigns a very high probability to a unique equilibrium, and a goodness-of-fit statistic that penalizes the model's complexity, the data provide strong support for multiple equilibria, as suggested by the smaller DIC statistics. In terms of fitting, the models with two stable or three equilibria are superior to the other.

In the following, we will comment on the results of the 3-regimes model, which is the most general and encompasses the others. However, qualitatively similar results are obtained with the 2-regimes specification that provides the best fit to the data and for which the results are reported in Appendix D. Figure 1 compares the parameter prior and the posterior distributions (red and black lines, respectively). The observed difference between the two suggests that the data are highly informative for the parameters of interest, as the posterior distributions move away from the priors and become more

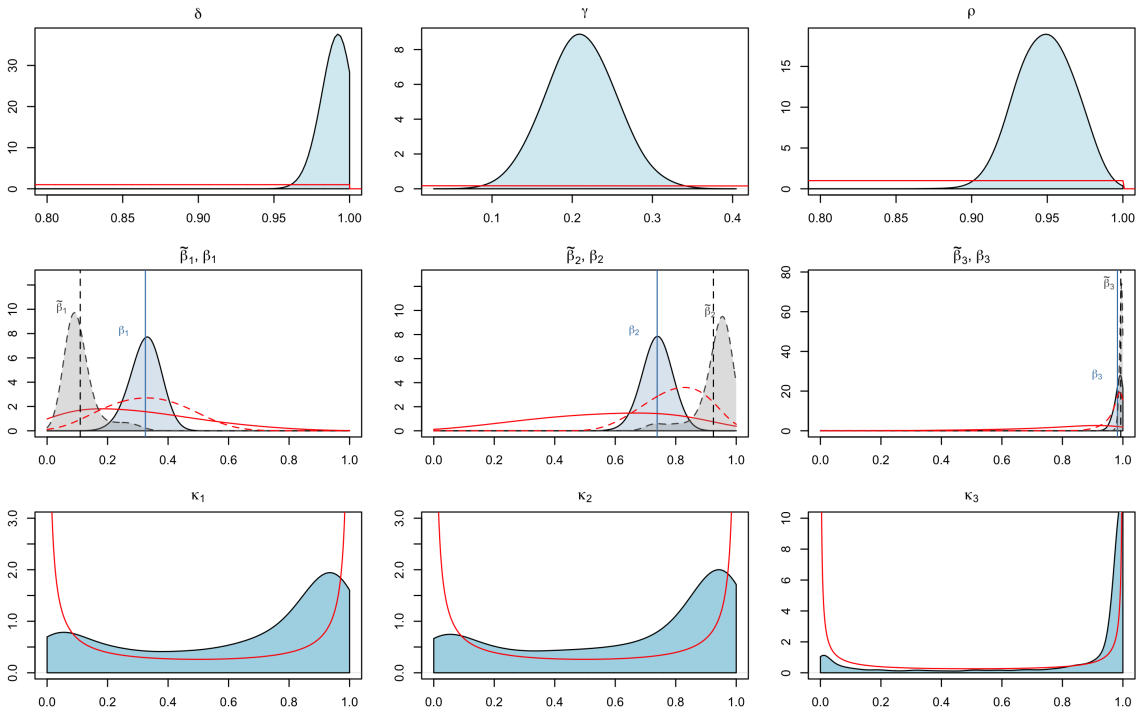


Figure 1: Top panel, posterior and prior distributions (respectively denoted by black and red lines) of the structural parameters δ (left), γ (center) and ρ (right) of the NKPC. Middle panel: posterior densities of β_k, solid black lines, $\tilde{\beta}_k$, dashed black lines, and their prior distributions, respectively solid and dashed red lines. Posterior means of β_k and $\tilde{\beta}_k$ are denoted by solid and dashed vertical lines, respectively. Bottom panel: posterior density of the implied Horseshoe shrinkage coefficient $\kappa_k = \frac{1}{1+\lambda_k^2\tau^2}$ in black, compared with its prior in red. Prior samples are obtained as described in Appendix B.2

concentrated. Figure D.1 in Appendix D shows good mixing performances of the MCMC algorithm.

The estimate for the slack parameter γ is positive and different from 0, consistent with standard findings based on long time series data (Lansing, 2019), and suggesting inflation is sensitive to a possibly negative supply shock or to the consequences of an expansionary monetary policy by the Central Bank. This parameter is larger when multiple equilibria are considered. The estimates for $\boldsymbol{\beta}$ are consistent with the calibration exercise of Hommes and Zhu (2014), suggesting three well-distinct regimes of low, moderate, and high persistence in inflation, which indicates a different use of past information by the forecasters when eliciting their predictions. Our findings on the inflation's autocorrelation dynamics are consistent with Lansing (2009), which evidences high persistence in inflation until the mid-nineties that decays afterward. It is also worth noting that the gap between $\boldsymbol{\beta}$ and $\tilde{\boldsymbol{\beta}}(\boldsymbol{\theta})$ is small, thus suggesting statistical regimes are close enough to the economic equilibria.[3] On this point, the bottom panel of Figure 1, stresses how the posterior distributions of $\kappa_k = \frac{1}{1+\lambda_k^2\tau^2}$ always favor 1 instead of 0 for all

[3]Similar qualitative results are obtained while using more diffuse hierarchical priors

the regimes, thus indicating the economic restrictions defining the equilibria are supported by the data. Equivalent findings are also available for $K = 2$. The unconditional expectation on inflation $\tilde{\alpha}$, computed on a base of a diffuse hierarchical prior, is concentrated around α and is approx. 2.50, that is in line with the empirical evidence provided in Lansing (2009).

Figure 2 further illustrates how the posterior distributions of the parameters concentrate within both the structural parameter space and the corresponding reduced-form space. The figure displays representative samples from the prior distributions in each equilibrium regime (shown in light blue, gray, and pink), alongside their respective posterior samples (shown in blue, black, and red). Color coding is used to distinguish between the three equilibria. The structural parameters are plotted along the horizontal axis, while the induced reduced-form parameters are plotted on the vertical axis. Prior samples are obtained as described in Appendix B.2.

These results provide compelling evidence of well-separated equilibria, as indicated by the distinct clusters of posterior samples (dark blue and red) across all panels. These clusters correspond to three distinct levels of inflation persistence: $\tilde{\beta}_1 \approx 0.11$ (low persistence), $\tilde{\beta}_2 \approx 0.93$ (intermediate), and $\tilde{\beta}_3 \approx 0.99$ (high persistence). On this point, it is interesting to note that the 95% credibility intervals do not overlap, as also stressed by the middle panel of Figure 1. This finding suggests that the prior specification on the structural parameters allows for proper separation of the three equilibria. In the bottom-right plot, the pairs of (β_k, β_ℓ) and the dashed lines indicating the identification constraint $\beta_1 < \beta_2 < \beta_3$.

4.3 Detecting Equilibrium Changes

Overall, our framework is based on a Markov switching model, where regime selection is grounded in a theoretical economic foundation. The evidence on US inflation supports the existence of multiple equilibria and a simple representation of them in terms of inflation persistence. The proposed data-augmentation framework allowed us to estimate the active regimes (dominant equilibrium) in each period. Figure 3 reports the estimated equilibrium over time (blue stepwise line) and the expected equilibrium, providing information on the posterior probability distribution over the equilibria.

The distribution of the dominant regimes over time is of particular interest, as displayed in Fig. 3. We found evidence of low persistence from 1994Q3, corresponding to a period of substantially low and stable inflation. Outside this window, agents form their expectations using more persistent rules, switching from intermediate to high β. Regarding timing and magnitude, our findings are consistent with the descriptive evidence reported in Lansing (2009) that suggests time-varying dependent persistence and volatility. In particular, there is evidence of the following switching points, which we present by transition type.

- From low to high persistence: 1962Q3, after the end of the 1960-1961 recession and the beginning

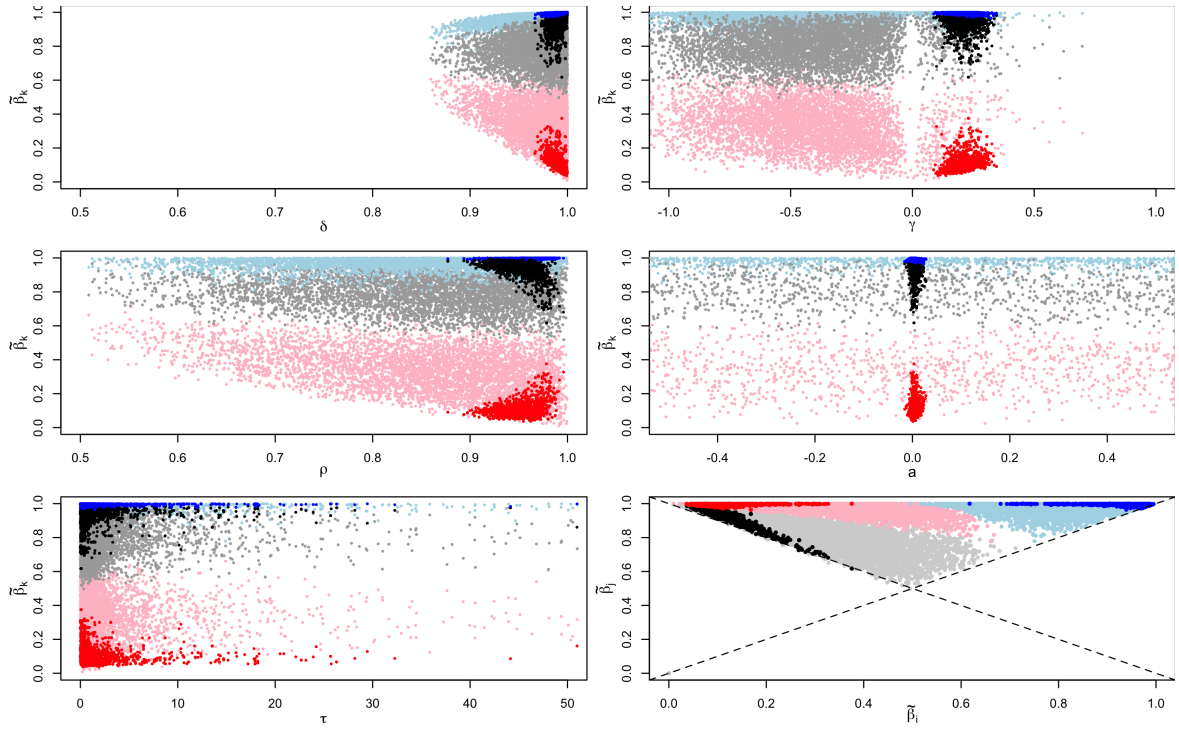


Figure 2: Draws (dots) from the prior (light blue, gray, and pink) and posterior (dark blue, black, and red) distributions of the structural (horizontal axis) and reduced form (vertical axis) parameter space. In each plot, the different colors identify different equilibria. In the bottom-right plot, the pairs of (β_k, β_ℓ) and the dashed lines indicating the identification constrain $\beta_1 < \beta_2 < \beta_3$. Prior samples are obtained as described in Appendix B.2.

of the expansion phase; 1983Q3, in the aftermath of the 1982 contraction.

- From high to low: 1985Q4, during the expansion period; 1993Q4 after the 1990 recession, a period characterized by the low FED rates, NAFTA drafted, and the Balanced Budget Act.

Different levels of persistence have implications on what type of reaction to shocks we might expect, as we will discuss extensively in Section 4.4. Note that one-step-ahead agent's inflation forecasts [4] can be written as $\pi_{t+1|t} = \beta\pi_{t|t-1} + \beta(\pi_t - \pi_{t|t-1})$, where $(\pi_t - \pi_{t|t-1})$ is the *surprise effect* indicating how agents revise their previous expectations based on their previous forecast error, and where β in this case is the steady state Kalman gain. Therefore, low persistence implies a limited reaction to previous forecasting errors, meaning that agents do not substantially revise their expectations even in the presence of large shocks. On the other hand, large persistence corresponds to substantial sensitivity to past errors while predicting inflation.

It is worth noting that in recent years, even during periods of high inflation, the model detects a low equilibrium. To explain this unexpected result, we believe that the use of SPF, as defined in Eq. (11),

[4] Here, $\pi_{t+1|t}$ is the inflation forecast based on information observed up to time t.

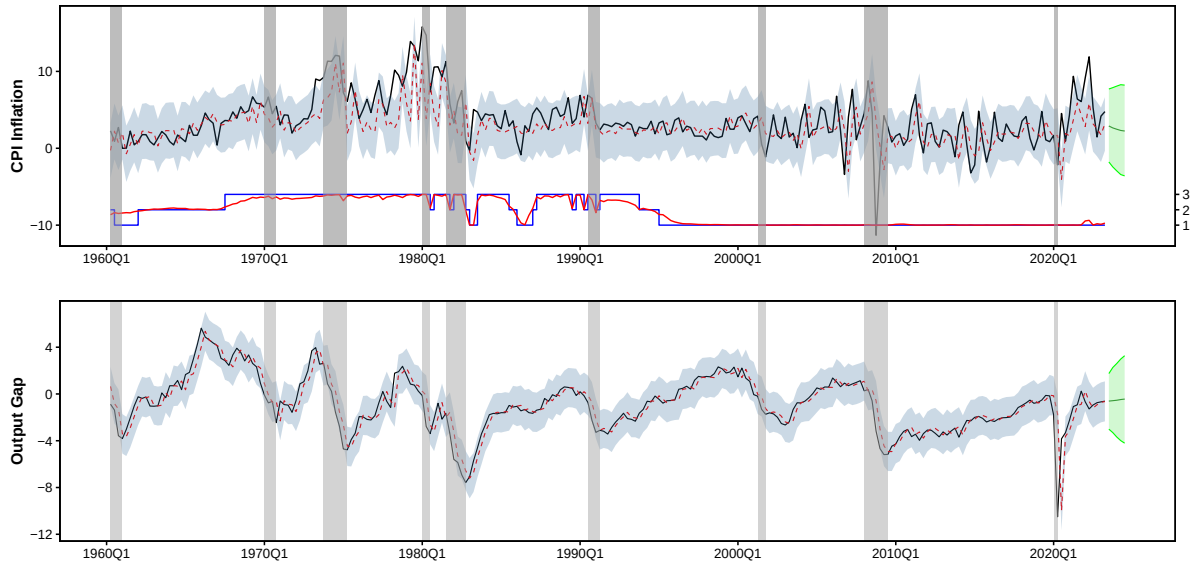


Figure 3: Top panel, observed CPI inflation (solid black line) and predicted CPI inflation (dashed red line) with their 90% credible intervals (light grey area). Below, estimated active equilibrium over time ($\hat{s}_t \in \{1, 2, 3\}$, right axis, blue line) and the regime posterior mean (right axis, red line) together with the NBER dating of the business cycle (shaded areas). Bottom panel, output gap (solid black line) and predicted output gap (dashed red line) with 90% credible interval (light grey area). The solid green line and area represent out-of-sample predictive mean and 90% credible region, respectively.

plays a crucial role. In fact, model estimates obtained by excluding the measurement equation highlight a shift towards a high or medium persistence regime from 2021Q2 to 2023Q2, in correspondence with the high level of inflation observed, as evidenced in Appendix D, Figure D.7. This result suggests that, despite the turbulent period, the information provided by subjective expectations, as measured through surveys, has enabled the identification of a period of relative calm in which agents did not substantially revise their forecasts. These survey forecasts likely predicted an increase in inflation as a consequence of COVID or geopolitical conditions. Neglecting this information would have led to the incorrect diagnosis of a different equilibrium. This evidence also stresses the importance of including data on survey expectations while analyzing the NKPC.

4.4 Did Agents Overreact or Underreact?

Although empirical evidence based on survey data mainly suggests that agents deviate from the rational expectation paradigm, the current literature still lacks a definitive answer regarding how agents react to sudden economic shocks. In particular, there is consensus in referring to overreaction in the case of financial applications, whereas underreaction is a common result in macroeconomics. Different results also depend on which data are used for testing, whether they are based on individual or consensus

surveys. Here, we aim to evaluate whether different equilibria correspond to different types of agents' responses.

Measures of reaction to shocks are typically based on forecast errors, which are usually computed as the difference between real data and forecasts collected through surveys[5] and that summarizes the gap between rational and behavioral expectations.

Here, we do not resort to surveys, but we compute the forecast errors from our model as follows:

$$e_{t+1} = \pi_{t+1} - \hat{\pi}^e_{t+1} = \pi_{t+1} - \left[\tilde{\alpha} + \tilde{\boldsymbol{\beta}}' \operatorname{diag}(\hat{\boldsymbol{\xi}}_t)\tilde{\boldsymbol{\beta}}(\pi_{t-1} - \tilde{\alpha})\right] \tag{19}$$

where $\hat{\pi}^e_{t+1}$ are the implicit expectations derived from the perceived law of motion. Recently, Kučinskas and Peters (2024) proposed to measure under/overreaction using the Impulse Response Functions (IRF) of forecast errors that, from a practical point of view, can be computed through local projections (Jordà, 2005; Olea et al., 2025), thus making their estimates model-free.

Provided that forecast errors can be represented through their Wold decomposition

$$e_t = b_0 + \sum_{j=0}^{\infty} \theta_j v_{t-j}, \tag{20}$$

where v_t are the Wold innovations, Kučinskas and Peters (2024) propose to look at the *composite bias coefficients* θ_j as a concise measure of under/overreaction[6]. Since v_t is not related to specific structural shocks, the composite bias coefficients represent a summary measure of the reaction to a generic shock unrelated to specific events.

To implement the test, we run the following set of regressions:

$$e_{t+h} = \phi_0 + \phi_h^{HIGH} e_t D_t^{(3)} + \phi_h^{MEDIUM} e_t D_t^{(2)} + \phi_h^{LOW} e_t D_t^{(1)} + \eta_t, \quad h \geq 1, \tag{21}$$

where $D_t^{(k)} = \mathbb{I}(\hat{s}_t = k)$, $k = 1, 2, 3$ is a set of indicator variables selecting the active equilibrium at time t corresponding to high, medium or low level of β.[7] As stressed in Kučinskas and Peters (2024), overreaction/underreaction effects are given by coefficients ϕ_h^j pre-multiplied by the sign of the IRF computed on inflation at the corresponding lag. Positive estimates indicate overreaction, whereas negative values suggest underreaction.

To evaluate local projections, we considered parameters and state uncertainty, for which forecast errors and the test regressions of Eq. (21) were computed at each sweep of the MCMC algorithm used to obtain the estimates of the Phillips curve. This procedure allowed to evaluate the *posterior* distribution for the parameters ϕ_h^j. The results are shown in Fig. 4, where a boxplot for the posterior distributions of the IRF is displayed.

[5] Relevant references on this topic are Coibion and Gorodnichenko (2012), and Bordalo et al. (2020) amongst others

[6] Precisely, a measure of the reaction can be derived as $-\text{sgn}(\alpha_j)\theta_j$ where α_j are the coefficients from the Wold representation of π_t.

[7] In a different setup, a similar regression has been used in Ferriere and Navarro (2025) to compute local projections for different economic regimes.

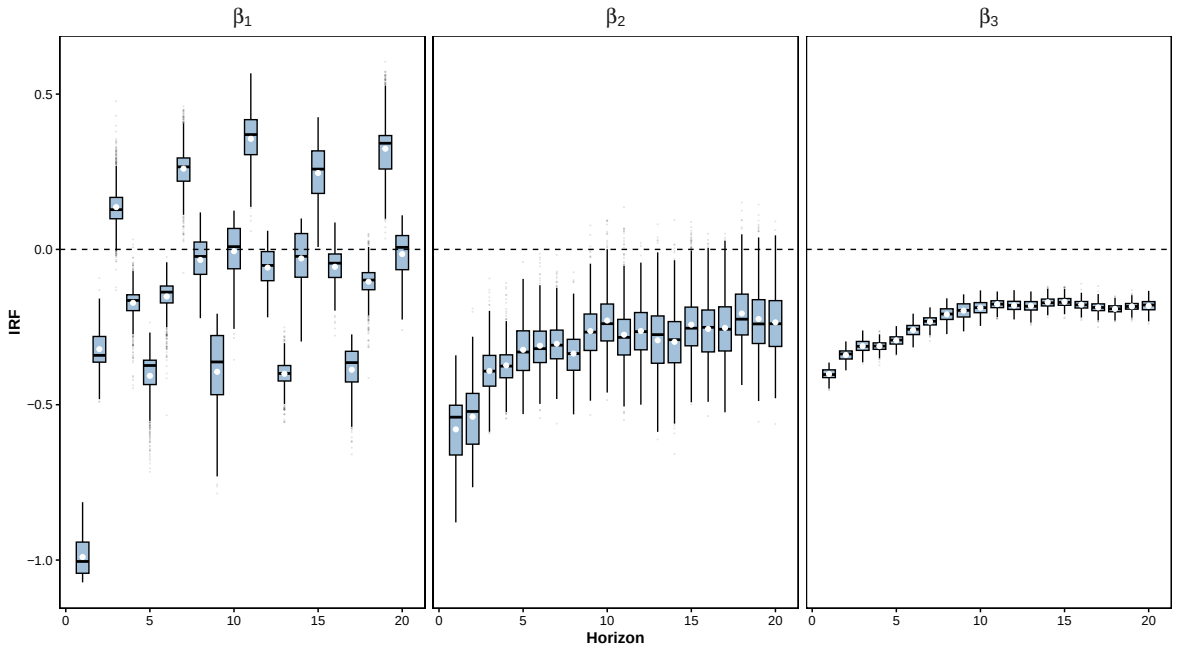


Figure 4: Composite bias coefficients via local projections at different horizons $h = 1, \ldots, 20$. The boxplots report the posterior draws of the composite bias coefficients at each horizon h.

The results in Fig. 4 highlight strong evidence in favor of underreaction when inflation is persistent or moderate, since negative bias coefficients are observed and are substantially different from zero even at large lags. For the low equilibrium, the 95% credible intervals of the coefficient often include zero and flip from positive to negative at different lags, suggesting that the average reaction over time is not systematically different from 0. These findings indicate weak and non-persistent reactions to composite shocks.

Our results, although not based on directly observable forecasting errors from the SPF, closely mimic those provided in Kučinskas and Peters (2024), indicating underreaction when inflation is perceived as persistent or moderately persistent, specifically from the 1960s to the end of 1994. This evidence is consistent with the rational inattention hypothesis of Sims (2003) in which the Impulse Response Functions are expected to display a weaker reaction compared to a model with perfect information (note that in a rational setup, the IRF on the forecast error should be null).

Nevertheless, when the system reaches the low equilibria, agents tend to be rational. This finding is consistent with the conjecture of Eichenbaum (2023), which suggests agents might be rational in quiet periods while displaying disproportionate reactions to shocks in hectic times. In this case, a low level of persistency is observed from 1995, when inflation is substantially stable and with limited variability. These results are not surprising if we note that the steady-state Kalman gain for the subjective AR(1) prediction rule corresponds to the persistence parameter that, if large, assigns a considerable weight to the previous forecast error. As a consequence, agents strongly revise their previous expectations right

after observing a surprising event. As argued in Lansing (2009), this evidence could be interpreted as a lack of confidence in the central bank's ability to maintain a stable inflation target.

Overall, if it is assumed that agents are boundedly rational, they might switch between different equilibria. When the equilibria imply high persistence, agents tend to underreact. On the other hand, when agents select a low-persistence equilibrium, average IRFs tend not to differ from zero, suggesting that the reaction to the shocks is consistent with perfect perceived information or rationality.

4.5 The Missing Disinflation Puzzle

Figure 3 also displays the in-sample inflation forecasts computed recursively from the MCMC iterations. The red dashed line is the in-sample one-step-ahead inflation forecast, whereas the gray area is the 95% credibility interval. In particular, at each time t, the one-step-ahead forecast is computed based on the parameter sampled at each iteration of the MCMC algorithm, thus controlling also for parameter uncertainty. During the Great Recession, we might have expected a drop in inflation, as evidenced by the graph for the output gap, which has not happened. It is clear that during the Great Recession, from late 2007 to mid-2009, despite a large and persistent drop in the output gap, together with a slack parameter different than 0, the predicted inflation closely follows the observed one. Note also that in this timespan, there is no evidence of parameter instability. This result is likely driven by the bounded rational hypothesis underlying our model, which anchors the expectation formation mechanism to a simple autoregressive process that well adapts to real data. This result is therefore consistent with Coibion and Gorodnichenko (2015) that assigns a dominant relevance to the expectation formation mechanism in describing the dynamics of inflation, which prevents inflation forecasts from changing drastically in correspondence of a large and systematic change of the slack variable.

5 Conclusion

In this paper, we study a New Keynesian Phillips curve in which agents are boundedly rational, i.e., form their expectations through a misspecified econometric model (Hommes and Zhu, 2014). When perceived and actual laws of motion match, multiple equilibria occur and are identified by the parameter that corresponds to the persistence of perceived inflation. We propose a strategy to jointly estimate the model's equilibria and the Phillips curve parameters through Bayesian methods. The proposed strategy allows for the easy incorporation of nonlinear constraints necessary to define different equilibria in the parameter space, which typically leads to an intractable likelihood function. Our results, based on the horseshoe prior, suggest that the data support the parameter constraints that define the economic equilibria. We evaluate whether agents underreacted or overreacted, according to the dominant equilibria observed at each given time, and find that under the high persistence equilibrium, there is evidence of underreaction, consistent with the empirical results of Kučinskas and

Peters (2024) and with the rational inattention model of Sims (2003). When the perceived persistence is low, from the mid-1990s on, the reaction to shocks does not appear systematically different from that derived from a rational expectation model.

References

Andrieu, C., Thoms, J., 2008. A Tutorial on Adaptive MCMC. Statistics and Computing 18, 343–373.

Ball, L., 2000. Near-Rationality and Inflation in Two Monetary Regimes. Technical Report 7988. National Bureau of Economic Research.

Ball, L., Mazumder, S., 2011. Inflation Dynamics and the Great Recession. Brookings Papers on Economic Activity 41, 337–405.

Barsky, R., 1987. The Fisher Hypothesis and the Forecastability and Persistence of Inflation. Journal of Monetary Economics 19, 3–24.

Bhadra, A., Datta, J., Polson, N.G., Willard, B., 2017. The Horseshoe+ Estimator of Ultra-Sparse Signals. Bayesian Analysis 12, 1105–1131.

Bisin, A., Moro, A., Topa, G., 2011. The Empirical Content of Models with Multiple Equilibria in Economies with Social Interactions. Technical Report No. 17196. National Bureau of Economic Research.

Bordalo, P., Gennaioli, N., Ma, Y., Shleifer, A., 2020. Overreaction in Macroeconomic Expectations. American Economic Review 110, 2748–2782.

Cadonna, A., Frühwirth-Schnatter, S., Knaus, P., 2020. Triple the Gamma - A Unifying Shrinkage Prior for Variance and Variable Selection in Sparse State Space and TVP Models. Econometrics 8, 20.

Canova, F., 2007. Methods for Applied Macroeconomic Research. Princeton University Press, New Jersey.

Canova, F., Ciccarelli, M., 2004. Forecasting and turning point predictions in a Bayesian panel VAR model. Journal of Econometrics 120, 327–359.

Canova, F., Ciccarelli, M., 2009. Etimating multicountry VAR models. International Economic Review 50, 929–959.

Carvalho, C.M., Polson, N.G., Scott, J.G., 2010. The Horseshoe Estimator for Sparse Signals. Biometrika , 465–480.

Castelnuovo, E., Greco, L., Raggi, D., 2014. Policy Rules, Regime Switches, and Trend Inflation: An Empirical Investigation for the United States. Macroeconomic Dynamics 18, 920–942.

Chan, J.C., 2022. Asymmetric Conjugate Priors for Large Bayesian VARs. Quantitative Economics 13, 1145–1169.

Chernozhukov, V., Hong, H., 2004. Likelihood Estimation and Inference in a Class of Nonregular Econometric Models. Econometrica 72, 1445–1480.

Cho, S., McCallum, B., 2015. Refining Linear Rational Expectation Models and Equilibria. Journal of Macroeconomics 46, 160–169.

Christiano, L.J., Eichenbaum, M., Johannsen, B.K., 2018. Does the New Keynesian Model Have a Uniqueness Problem? Technical Report No. w24612. National Bureau of Economic Research.

Clements, M.P., 2024. Do professional forecasters believe in the Phillips curve? International Journal of Forecasting 40, 1238–1254.

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

Coibion, O., Gorodnichenko, Y., 2015. Is the Phillips Curve Alive and Well after all? Inflation Expectations and the Missing Disinflation. American Economic Journal: Macroeconomics 7, 197–232.

Coibion, O., Gorodnichenko, Y., Kamdar, R., 2018. The Formation of Expectations, Inflation, and the Phillips Curve. Journal of Economic Literature 56, 1447–1491.

Dagsvik, J., Jovanovic, B., 1994. Was the Great Depression a Low-level Equilibrium? European Economic Review 38, 1711–1729.

Driskill, R., 2006. Multiple Equilibria in Dynamic Rational Expectation Models: A Critical Review. European Economic Review 50, 171–210.

Eichenbaum, M., 2023. On the Limits of Rational Expectations for Policy Analysis. Canadian Journal of Economics/Revue Canadienne d'Économique 56, 1221–1237.

Evans, G.W., 1985. Expectational Stability and the Multiple Equilibrium Problem in Linear Rational Expectations Models. Quarterly Journal of Economics 100, 1217–1233.

Federal Reserve Bank of Philadelphia, 2025. Survey of Professional Forecasters. URL: https://www.philadelphiafed.org/surveys-and-data/real-time-data-research/survey-of-professional-forecasters.

Federal Reserve Bank of St. Louis, 2025. FRED (Federal Reserve Economic Data). URL: https://fred.stlouisfed.org/.

Federal Reserve of Philadelphia, 2024. Survey of Professional Forecasters: Documentation. Technical Report.

Ferriere, G., Navarro, G., 2025. The Heterogeneous Effects of Government Spending: It's All About Taxes. Review of Economic Studies 92, 1061–1125.

Fisher, S., 1977. Wage Indexation and Macroeconomic Stability. Carnegie–Rochester Conference Series on Public Policy 5, 107–147.

Frühwirth-Schnatter, S., 2006. Finite Mixture and Markov Switching models. Springer.

Fuster, A., Laibson, D., Mendel, B., 2010. Natural Expectations and Macroeconomic Fluctuations. Journal of Economic Perspectives 24, 67–84.

Gallant, R., 2020. Complementary Bayesian Method of Moments Strategies. Journal of Applied Econometrics 35, 422–439.

Gallant, R.A., 1987. Nonlinear Statistical Models. Wiley Series in Probability and Mathematical Statistics.

Geweke, J., 1992. Evaluating the Accuracy of Sampling-Based Approaches to the Calculation of Posterior Moments, in: Bayesian Statistics 4: Proceedings of the Fourth Valencia International Meeting, Dedicated to the memory of Morris H. DeGroot, 1931–1989. Oxford University Press.

Giordani, P., Kohn, R., van Dijk, D., 2007. A unified approach to nonlinearity, structural change, and outliers. Journal of Econometrics 137, 112–133.

Grandmont, J., 1998. Expectation Formation and Stability in Large Socio-economic Systems. Econometrica 66, 741–781.

Hommes, C., Mavromatis, K., Özden, T., Zha, M., 2023. Behavioral Learning Equilibria in New Keynesian Models. Quantitative Economics 14, 1401–1445.

Hommes, C., Sorger, G., 1998. Consistent Expectations Equilibria. Macroeconomic Dynamics 2, 287–321.

Hommes, C., Zhu, M., 2014. Behavioral learning equilibria. Journal of Economic Theory 150, 778–814.

Jordà, O., 2005. Estimation and Inference of Impulse Responses by Local Projections. American Economic Review 95, 161–182.

Juodis, A., Kučinskas, S., 2023. Quantifying noise in survey expectations. Quantitative Economics 14, 609–650.

Kadiyala, K.R., Karlsson, S., 1997. Numerical methods for estimation and inference in Bayesian VAR-models. Journal of Applied Econometrics 12, 99–132.

Kaufmann, S., 2015. K-state Switching Models with Time-varying Transition Distributions—Does Loan Growth Signal Stronger Effects of Variables on Inflation? Journal of Econometrics 187, 82–94.

Kim, C.J., Nelson, C.R., 1999. Has the US Economy Become More Stable? A Bayesian Approach Based on a Markov-Switching Model of the Business Cycle. Review of Economics and Statistics 81, 608–616.

Korobilis, D., 2022. A New Algorithm for Structural Restrictions in Bayesian Vector Autoregressions. European Economic Review 148, 104241.

Kučinskas, S., Peters, F.S., 2024. Measuring Under- and Overreaction in Expectation Formation. Review of Economics and Stastistics 106, 1620–1637.

Lansing, K., 2009. Time-Varying U.S. Inflation Dynamics and the New Keynesian Phillips Curve. Review of Economic Dynamics 12, 304–326.

Lansing, K., 2019. Improving the Phillips Curve with an Interaction Variable. FRBSF Economic Letter 2019–13, 1–4.

Lansing, K., Ma, J., 2017. Explaining Exchange Rate Anomalies in a Model with Taylor-rule Fundamentals and Consistent Expectations. Journal of International Money and Finance 70, 62–87.

Lucas, R., 1972. Expectations and the Neutrality of Money. Journal of Economic Theory 4, 103–124.

Makalic, E., Schmidt, D.F., 2015. A Simple Sampler for the Horseshoe Estimator. IEEE Signal Processing Letters 23, 179–182.

Mankiw, N.G., Reis, R., 2002. Sticky Information versus Sticky Prices: A Proposal to Replace the Phillips Curve. Quarterly Journal of Economics 117, 1295–1328.

Mankiw, N.G., Reis, R., Wolfers, J., 2003. Disagreement about Inflation Expectations. NBER Macroeconomics Annual 18, 209–248.

Mavroeidis, S., Plagborg-Møller, M., Stock, J., 2014. Empirical Evidence of Inflation Expectations in the New Keynesian Phillips Curve. Journal of Economic Perspectives 52, 124–188.

McCallum, B., 1983. On Non-Uniqueness in Rational Expectation models: an Attempt at Perspective. Journal of Monetary Economics 11, 139–168.

Mertens, E., Nason, J.M., 2020. Inflation and professional forecast dynamics: An evaluation of stickiness, persistence, and volatility. Quantitative Economics 11, 1485–1520.

Mertens, K.R.S.M., Ravn, M.O., 2014. Fiscal Policy in an Expectations-Driven Liquidity Trap. The Review of Economic Studies 81, 1637–1667.

Muth, J., 1961. Rational Expectations and the Theory of Price Movements. Econometrica 29, 315–335.

Olea, J., Plagborg-Møller, M., Qian, E., Wolf, C., 2025. Local Projections or VARs? A Primer for Macroeconomists. Technical Report. MIT working paper.

van der Pas, S.L., Kleijn, B.J.K., van der Vaart, A.W., 2014. The Horseshoe Estimator: Posterior Concentration around Nearly Black Vectors. Electronic Journal of Statistics 8, 2585 – 2618.

Sargent, T.J., 1999. The Conquest of American Inflation. Princeton University Press: Princeton, New Jersey.

Sims, C.A., 2003. Implications of Rational Inattention. Journal of Monetary Economics 50, 665–690.

Sims, C.A., Zha, T., 2006. Were There Regime Switches in US Monetary Policy? American Economic Review 96, 54–81.

Souleles, N., 2004. Expectations, Heterogeneous Forecast Errors, and Consumption: Micro Evidence from the Michigan Consumer Sentiment Surveys. Journal of Money, Credit and Banking 36, 39–72.

Spiegelhalter, D.J., Best, N.G., Carlin, B.P., Van Der Linde, A., 2002. Bayesian Measures of Model Complexity and Fit. Journal of the Royal Statistical Society Series B: Statistical Methodology 64, 583–639.

Taylor, J.B., 1979. Staggered Wage Setting in a Macro Model. American Economic Review 69, 108–113.

Vissing-Jorgensen, A., 2003. Perspectives on Behavioral Finance: Does 'Irrationality' Disappear with Wealth? Evidence from Expectations and Actions. NBER Macroeconomics Annual 18, 139–208.

Woodford, M., 2003a. Imperfect Common Knowledge and the Effects of Monetary Policy, in: Aghion, P., Frydman, R., Stiglitz, J., Woodford, M. (Eds.), Knowledge, Information, and Expectations in Modern Macroeconomics: In Honor of Edmund S. Phelps. Princeton: Princeton University Press, pp. 25–58.

Woodford, M., 2003b. Interest and Prices: Foundations of a Theory of Monetary Policy. Princeton University Press, Princeton, NJ,.

A Prior Distributions and Multiple Equilibria

The prior assumption given in Eq. 16-17 for the parameters of the model in Eq. 12 can be summarized as follows: $\alpha|\boldsymbol{\theta}, \sigma_\alpha^2 \sim \mathcal{N}(\tilde{\alpha}(\boldsymbol{\theta}), \sigma_\alpha^2)$ and $\beta_k|\boldsymbol{\theta}, \lambda_k^2\tau^2 \sim \mathcal{N}(\tilde{\beta}_k(\boldsymbol{\theta}), \lambda_k^2\tau^2), k = 1, \ldots, K$, where $\boldsymbol{\theta} = \{\delta, \gamma, a, \rho, \sigma_u^2, \sigma_\epsilon^2\}$ is the vector of parameters of the structural form model, $\tilde{\alpha}(\boldsymbol{\theta}) = \gamma a/((1-\rho)(1-\delta))$ is the intercept and $\tilde{\boldsymbol{\beta}}(\boldsymbol{\theta}) = (\tilde{\beta}_1(\boldsymbol{\theta}), \ldots, \tilde{\beta}_K(\boldsymbol{\theta}))$ is a vector of autoregressive coefficients.

Let $\varphi(\tilde{\beta}) = \tilde{\beta} - \delta\tilde{\beta}^2 + \gamma^2\rho(1-\delta^2\tilde{\beta}^4)/(\gamma^2(\delta\tilde{\beta}^2\rho+1) - (1-\rho^2)(1-\delta\tilde{\beta}^2\rho)\frac{\sigma_u^2}{\sigma_\varepsilon^2})$ be the quartic polynomial given by the NKPC model restrictions, we denote with $\mathcal{S}(\boldsymbol{\theta}) = \{\tilde{\beta},\, s.t.\, \varphi(\tilde{\beta}(\boldsymbol{\theta})) = 0, \tilde{\beta}(\boldsymbol{\theta}) \in \mathbb{R}\}$ the set of the real distinct solutions for a given value of $\boldsymbol{\theta}$ and with $\mathcal{S}_K(\boldsymbol{\theta}) = \{\tilde{\beta}_k(\boldsymbol{\theta}) \in \mathcal{S}(\boldsymbol{\theta}), k = 1, \ldots, K, s.t. |\mathcal{S}(\boldsymbol{\theta})| = K, \tilde{\beta}_1(\boldsymbol{\theta}) < \ldots < \tilde{\beta}_K(\boldsymbol{\theta})\}$ $K \leq 4$ the set of the ordered K real and unique solutions. The prior distribution on $\boldsymbol{\theta}$ induces a prior distribution on the coefficients $\tilde{\beta}_k(\boldsymbol{\theta})$ with full support in $\mathbb{R}$. In the following, a simulation study is conducted to evaluate the number of admissible solutions provided by Eq. (9) under our prior assumptions.

We start from the calibration exercise of Hommes and Zhu (2014), in which

- $\gamma = 0.075$ and its domain is $[0; +\infty)$;
- $\delta = 0.99$ and its domain is $[0; 1)$;
- $\rho = 0.9$ and its domain is $[0; 1)$;
- $\frac{\sigma_\epsilon^2}{\sigma_u^2} = .1$ and the ratio is strictly positive;
- $a = 0.0004$ and is real.

Therefore, we assume $\gamma \sim \mathcal{N}(0.075, 0.02^2)$, $\delta \sim \mathcal{U}(.7, .99)$, $\rho \sim \mathcal{U}(.7, .99)$, $\frac{\sigma_\epsilon^2}{\sigma_u^2} \sim \mathcal{N}(.1, 0.04^2)$ and $a \sim \mathcal{U}(0, .2)$.

We consider 100,000 independent draws from the prior distribution. For each draw, we compute the solutions of the quartic equation and we store only the real ones in the interval $(0, 1)$. For each feasible solution also the autoregressive coefficient of the Phillips curve $\phi_k = \delta\tilde{\beta}_k^2$ and intercepts $\mu_k = \delta\alpha(1-\tilde{\beta}_k^2)\tilde{\beta}_k^2 + \gamma a$ are computed.

Table A.1 reports the number of solutions obtained for different parameter values. In the great majority of cases, we observe a single feasible solution, while sometimes we observe three distinct solutions. The same table describes the occurrence of each feasible solution at its respective position. Note that the solution at position 1 is the most frequent. Solutions at positions 2,3 and 4 are less frequent.

A graphical representation of the a-priori distribution of the elements of $\tilde{\boldsymbol{\beta}}(\boldsymbol{\theta})$ in the reduced form model is displayed in case we observe three distinct and feasible solutions (left panel). The middle and right panels report the distributions of $\tilde{\mu}_k(\boldsymbol{\theta}) = \delta\alpha(\tilde{\boldsymbol{\theta}})(1-\tilde{\beta}_k^2(\boldsymbol{\theta}))$ and $\tilde{\psi}_k(\boldsymbol{\theta}) = \delta\tilde{\beta}_k^2(\boldsymbol{\theta})$ respectively.

(a) Cardinality of $\mathcal{S}(\boldsymbol{\theta})$		(b) Feasibility	
Number of solutions	sampling frequency	Solution number	sampling frequency
No solutions	–		
1 Solution	99857	Position 1	55845
2 Solutions	–	Position 2	29809
3 Solutions	143	Position 3	14422
4 Solutions	–	Position 4	210

Table A.1: Prior distribution of the number of feasible solutions (panel a) and of the position of these solutions (panel b), approximated by a set of 100,000 Monte Carlo random samples.

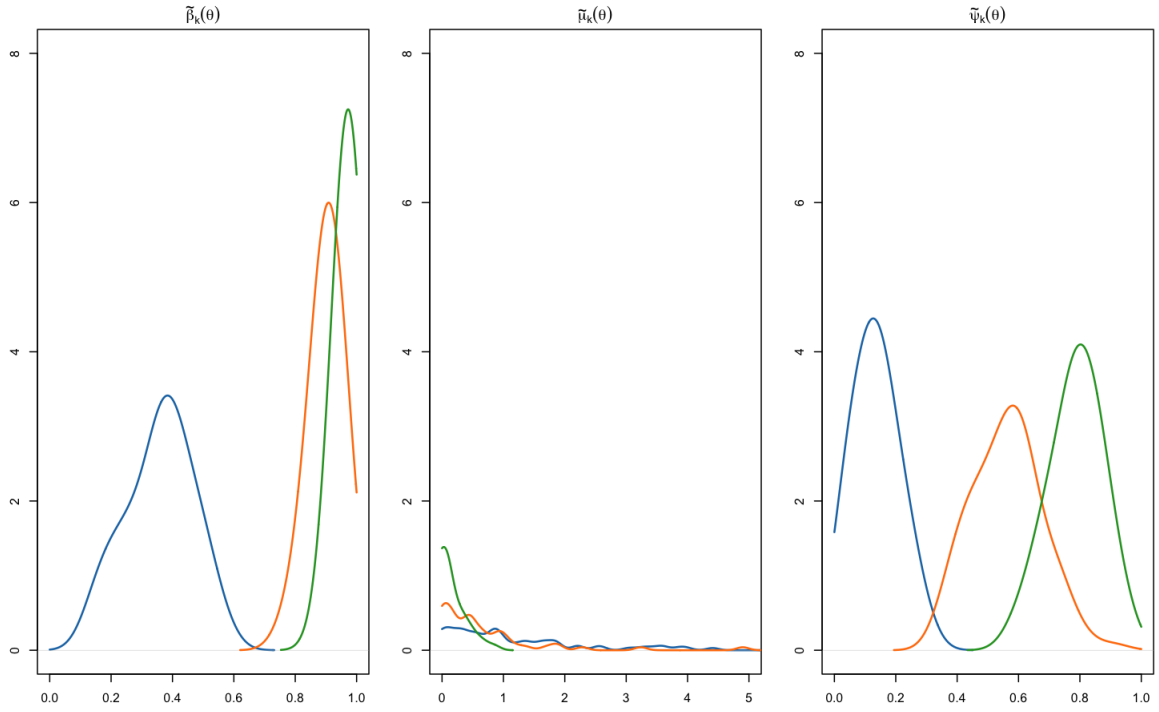


Figure A.1: Left Panel, distribution of the parameter $\tilde{\beta}_k(\boldsymbol{\theta})$ for $k = 1, \ldots, 3$. Middle Panel, distribution of the parameter $\tilde{\mu}_k(\boldsymbol{\theta}) = \delta\alpha(\tilde{\boldsymbol{\theta}})(1-\tilde{\beta}_k^2(\boldsymbol{\theta}))$. Right Panel, distribution of the parameter $\tilde{\psi}_k(\boldsymbol{\theta}) = \delta\tilde{\beta}_k^2(\boldsymbol{\theta})$. First solution in blue ($k = 1$), second in yellow ($k = 2$), and third in green ($k = 3$).

B Horseshoe Prior and Posterior Sampling Algorithm

B.1 Proof of Remark 1

This section contains an adaptation of the Proof of Theorem 1 in Carvalho et al. (2010) used for Remark 1.

For simplicity, we assume $\tau = 1$. The marginal distribution of $\beta_k - \tilde{\beta}_k(\boldsymbol{\theta})$ is given by

$$p(\beta_k|\boldsymbol{\theta}) = \int_0^\infty \frac{1}{(2\pi\lambda_k^2)^{1/2}} \exp\left(-\frac{(\beta_k - \tilde{\beta}_k(\boldsymbol{\theta}))^2}{2\lambda_k^2}\right) \frac{2}{\pi\,(1+\lambda_k^2)} d\lambda_k$$

Let $u = 1/\lambda_k^2$. Then

$$p(\beta_k|\boldsymbol{\theta}) = C \int_0^\infty \frac{1}{1+u} \exp\left(-\frac{(\beta_k - \tilde{\beta}_k(\boldsymbol{\theta}))^2 u}{2}\right) \mathrm{d}u$$

for $z = 1 + u$:

$$\begin{aligned} p(\beta_k|\boldsymbol{\theta}) &= C e^{(\beta_k - \tilde{\beta}_k(\boldsymbol{\theta}))^2/2} \int_1^\infty \frac{1}{z} \exp\left(-\frac{z(\beta_k - \tilde{\beta}_k(\boldsymbol{\theta}))^2}{2}\right) \mathrm{d}z \\ &= C \exp\left(\frac{(\beta_k - \tilde{\beta}_k(\boldsymbol{\theta}))^2}{2}\right) E_1\left(\frac{(\beta_k - \tilde{\beta}_k(\boldsymbol{\theta}))^2}{2}\right) \end{aligned}$$

where $E_1(\cdot)$ is the exponential integral function. This function satisfies tight upper and lower bounds:

$$\frac{\exp(-t)}{2} \log\left(1 + \frac{2}{t}\right) < E_1(t) < \exp(-t) \log\left(1 + \frac{1}{t}\right)$$

for all $t > 0$, which proves Part (b). Part (a) then follows from the lower bound in Equation (1), which approaches ∞ as $(\beta_k - \tilde{\beta}_k(\boldsymbol{\theta})) \to 0$.

B.2 Prior Simulation from our Model

B.2.1 Prior for $\tilde{\beta}_k(\boldsymbol{\theta})$

We obtain prior samples for $\tilde{\beta}_k(\boldsymbol{\theta})$ for $k = 1, \ldots, K$ by drawing samples from the prior of the parameters in $\boldsymbol{\theta}$ with the restriction of k solutions to the Equation 9. For the parameters u_t, ϵ_t and w_t, we draw prior samples from a Generalized Inverse Gaussian, $\mathcal{GIG}(1, 0.1, -0.5)$.

B.2.2 Prior for β_k

We obtain prior samples for $\beta_k \sim \mathcal{N}(\tilde{\beta}_k(\boldsymbol{\theta}), \tau^2\lambda_k^2)$ for $k = 1, \ldots, K$, such that $0 \leq \tilde{\beta}_k(\boldsymbol{\theta}) < 1$, by using the samples from $\tilde{\beta}_k$ previously obtained and by sampling $\tau \sim C^+(0, 1)$ and $\lambda_k \sim C^+(0, 1)$.

B.2.3 Prior for κ_k

To obtain the marginal prior distribution of κ_k, we integrate out the global shrinkage parameter τ via simulation. Specifically, we draw $\tau \sim \mathcal{C}^+(0, 1)$ and, for each draw, compute a corresponding value of κ_k from its conditional distribution given τ. This conditional density, derived in van der Pas et al. (2014), is given by

$$p_\tau(\kappa_i) = \frac{\tau}{\pi} \cdot \frac{1}{1 - (1-\tau^2)\kappa_i} \cdot \kappa_i^{-1/2}(1-\kappa_i)^{-1/2}.$$

Repeating this procedure yields an empirical distribution for κ_k that integrates over uncertainty in τ.

B.3 Posterior Sampling Algorithm

Figure B.1 provides a directed acyclic graph representation of the Bayesian model given in Eq. 12 and 16-17. The complete-data likelihood of the model given in Eq. 12 is

$$f(\boldsymbol{\pi}|\boldsymbol{\theta},\alpha,\boldsymbol{\beta},\boldsymbol{y},\boldsymbol{\xi}) = \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi\sigma_u^2}} \exp\left\{-(1/2\sigma_u^2)\left(\pi_t - \mu_{\pi,t}\right)^2\right\}$$
$$f(\boldsymbol{\pi}^{SPF}|\boldsymbol{\theta},\alpha,\boldsymbol{\beta},\boldsymbol{\pi},\boldsymbol{\xi}) = \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi\sigma_w^2}} \exp\left\{-(1/2\sigma_w^2)\left(\pi_t^{SPF} - \mu_{\pi^{SPF},t}\right)^2\right\}$$
$$f(\boldsymbol{y}|a,\rho,\sigma_\epsilon^2) = \prod_{t=2}^{T} \frac{1}{\sqrt{2\pi\sigma_\epsilon^2}} \exp\left\{-(1/2\sigma_\epsilon^2)\left(y_t - \mu_{y,t}\right)^2\right\},$$

where $\mu_{\pi,t} = \left(\delta\alpha(1-\boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta}) + \delta\boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta}\pi_{t-1} + \gamma y_t\right)$, $\mu_{\pi^{SPF},t} = \left(\alpha(1-\boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta}) + \boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta}\pi_{t-1}\right)$ and $\mu_{y,t} = (a + \rho y_{t-1})$ are the conditional means and $\boldsymbol{\xi} = (\boldsymbol{\xi}_1,\ldots,\boldsymbol{\xi}_T)$ denotes the collection of the states variables.

We approximate the joint posterior distribution by Markov-chain Monte Carlo (MCMC) sampling. Our Gibbs sampling algorithm iterates the following steps:

1. Draw $\boldsymbol{\theta}$ from $p(\boldsymbol{\theta}|\ldots)$ via Adaptive Metropolis-Hastings (AMH);
2. Draw α from $p(\alpha|\ldots)$ exactly from a normal distribution;
3. Draw $\boldsymbol{\beta}$ from $p(\boldsymbol{\beta}|\ldots)$ via MH with truncated normal proposal;
4. Draw τ from $p(\tau|\ldots)$ via MH;
5. Draw λ_k from $p(\lambda_k|\ldots)$ via MH for $k = 1,\ldots,K$;
6. Draw $\boldsymbol{p}_l$ from $p(\boldsymbol{p}_l|\boldsymbol{\xi})$ for $l = 1,\ldots,K$ from a Dirichlet distribution.
7. Draw $\mathbf{s}$ via the forward-filtering and backward-sampling algorithm.

In the following we derive the full conditional distributions used in the Gibbs sampler.

B.3.1 Posterior Sampling of $\boldsymbol{\theta}$

The posterior for $\boldsymbol{\theta}$,

$$p(\boldsymbol{\theta}|\ldots) \propto f(\boldsymbol{\pi}|\boldsymbol{\theta},\alpha,\boldsymbol{\beta},\boldsymbol{y},\boldsymbol{\xi}) f(\boldsymbol{\pi}^{SPF}|\alpha,\boldsymbol{\beta},\boldsymbol{\pi},\boldsymbol{\xi}) f(\boldsymbol{y}|a,\rho,\sigma_\epsilon^2) f(\alpha|\boldsymbol{\theta},\sigma_\alpha^2) f(\boldsymbol{\beta}|\boldsymbol{\theta},\boldsymbol{\lambda},\tau) p(\boldsymbol{\theta}),$$

where $p(\boldsymbol{\theta}) = p(\delta)p(\gamma)p(a)p(\rho)p(\sigma_u^2)p(\sigma_e^2)p(\sigma_w^2)$ and $\boldsymbol{\lambda} = (\lambda_1,\ldots,\lambda_K)$, is not analytically tractable. We sample from the full conditional of $\boldsymbol{\theta} = \{\delta,\gamma,a,\rho,\sigma_\epsilon^2,\sigma_u^2,\sigma_w^2\}$ via the Adaptive MH (AMH) algorithm with global adaptive scaling proposed in Andrieu and Thoms (2008).

Denote $\Theta^* = \{\boldsymbol{\theta} : \tilde{\beta}(\boldsymbol{\theta}) \in \mathbb{R}, \quad \tilde{\beta} = F(\tilde{\beta}(\boldsymbol{\theta}) \text{ has } K \text{ real-valued solutions}, 0 \leq \delta \leq 1, 0 \leq \rho \leq 1, \sigma_u \geq 0, \sigma_e^2 \geq 0)\}$ the set of admissible values of $\boldsymbol{\theta}$. Given the starting values of the parameter $\boldsymbol{\theta}_0$, $\boldsymbol{\mu}_0$ and

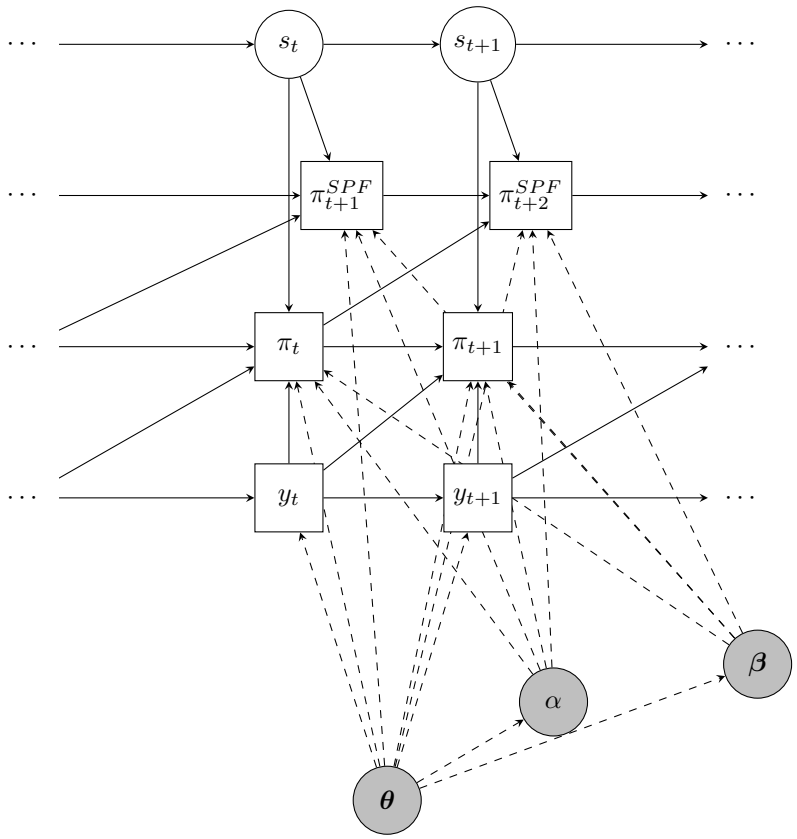


Figure B.1: Directed Acyclical Graph of the Phillips' curve model. Squares represent observed quantities (data), white circles are latent variables, and grey circles are unknown parameters. Each node in the graph is a priori conditionally independent of all other nodes other than its parental nodes (indicated by the arrows). Solid lines represent temporal DAG, and dashed lines represent relationships between parameters and the other variables. Variables conditional distributions are given in the main text.

$\boldsymbol{\Sigma}_0$, at the hth iteration, and the previous iteration values $\boldsymbol{\theta}_h$, $\boldsymbol{\mu}_h$, $\boldsymbol{\Sigma}_h$ and ℓ_h, the AMH generates samples from the posterior distribution of $\boldsymbol{\theta}$ by iterating:

1. Sample $\boldsymbol{\theta}_h^* \sim \mathcal{N}(\boldsymbol{\theta}_{h-1}, \ell_{h-1}\boldsymbol{\Sigma}_{h-1})$ such that $\boldsymbol{\theta}_h^* \in \Theta^*$
2. Set $\boldsymbol{\theta}_h = \boldsymbol{\theta}_h^*$ with probability $acc(\boldsymbol{\theta}_{h-1}, \boldsymbol{\theta}_h^*)$, otherwise $\boldsymbol{\theta}_h = \boldsymbol{\theta}_{h-1}$;
3. Update $\log(\ell_h) = \log(\ell_{h-1}) + c_h[acc(\boldsymbol{\theta}_{h-1}, \boldsymbol{\theta}_h^*) - acc^*]$;
4. Update $\boldsymbol{\mu}_h = \boldsymbol{\mu}_{h-1} + c_h(\boldsymbol{\theta}_h - \boldsymbol{\mu}_{h-1})$;
5. Update $\boldsymbol{\Sigma}_h = \boldsymbol{\Sigma}_{h-1} + c_h[(\boldsymbol{\theta}_h - \boldsymbol{\mu}_{h-1})(\boldsymbol{\theta}_h - \boldsymbol{\mu}_{h-1})' - \boldsymbol{\Sigma}_{h-1}]$;

where $c_h = h^{-\psi}$ for $\psi \in (0, 1)$, and acc^* is a target acceptance rate, here chosen to be 23.4%.

B.3.2 Posterior Sampling of α

We sample α exactly from

$$
\begin{aligned}
f(\alpha|\ldots) &\propto f(\boldsymbol{\pi}|\boldsymbol{\theta},\alpha,\boldsymbol{\beta},\boldsymbol{y},\boldsymbol{\xi}) f(\boldsymbol{\pi}^{SPF}|\boldsymbol{\theta},\alpha,\boldsymbol{\beta},\boldsymbol{\xi}) f(\alpha|\boldsymbol{\theta},\sigma_\alpha^2) \\
&\propto \prod_{t=2}^{T} \exp\left\{-\frac{1}{2\sigma_u^2}\left(\pi_t - \left(\delta\alpha(1-\boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta}) + \delta\boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta}\pi_{t-1} + \gamma y_t\right)\right)^2\right\} \\
&\quad \times \prod_{t=2}^{T} \exp\left\{-\frac{1}{2\sigma_w^2}\left(\pi_t^{SPF} - \left(\alpha(1-\boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta}) + \boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta}\pi_{t-1}\right)\right)^2\right\} \\
&\quad \times \exp\left\{-\frac{1}{2\sigma_\alpha^2}\left(\alpha - \tilde{\alpha}(\boldsymbol{\theta})\right)^2\right\} \\
&\propto \exp\left\{-\frac{1}{2\sigma_u^2}\left(\alpha^2 \sum_{t=2}^{T} \delta^2(1-\boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta})^2\right) - \frac{1}{2\sigma_w^2}\left(\alpha^2 \sum_{t=2}^{T}(1-\boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta})^2\right)\right\} \\
&\quad \times \exp\left\{-\frac{1}{2\sigma_u^2}\left(-2\alpha \sum_{t=2}^{T} \delta(1-\boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta})\left(\pi_t - \delta\boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta}\pi_{t-1} - \gamma y_t\right)\right)\right\} \\
&\quad \times \exp\left\{-\frac{1}{2\sigma_w^2}\left(-2\alpha \sum_{t=2}^{T}(1-\boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta})\left(\pi_t^{SPF} - \delta\boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta}\pi_{t-1}\right)\right)\right\} \\
&\quad \times \exp\left\{-\frac{1}{2\sigma_\alpha^2}\left(\alpha^2 - 2\alpha\tilde{\alpha}(\boldsymbol{\theta})\right)\right\} \\
&\propto \mathcal{N}(BA^{-1}, A^{-1}),
\end{aligned}
$$

where

$$
A = \frac{1}{\sigma_u^2}\sum_{t=2}^{T} \delta^2(1-\boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta})^2 + \frac{1}{\sigma_w^2}\sum_{t=2}^{T}(1-\boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta})^2 + \frac{1}{\sigma_\alpha^2} \tag{B.1}
$$

and

$$
\begin{aligned}
B = &\sum_{t=2}^{T} \frac{1}{\sigma_u^2}\delta(1-\boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta})\left(\pi_t - \delta\boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta}\pi_{t-1} - \gamma y_t\right) &\text{(B.2)} \\
&+ \sum_{t=2}^{T} \frac{1}{\sigma_w^2}(1-\boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta})\left(\pi_t^{SPF} - \boldsymbol{\beta}'\text{diag}(\boldsymbol{\xi}_t)\boldsymbol{\beta}\pi_{t-1}\right) + \frac{1}{\sigma_\alpha^2}\tilde{\alpha}(\boldsymbol{\theta}). &\text{(B.3)}
\end{aligned}
$$

B.3.3 Posterior Sampling of $\boldsymbol{\beta}$

The posterior for β_k,

$$
f(\boldsymbol{\beta}|\ldots) \propto f(\boldsymbol{\pi}|\boldsymbol{\theta},\alpha,\boldsymbol{\beta},\mathbf{y},\boldsymbol{\xi}) f(\boldsymbol{\pi}^{SPF}|\boldsymbol{\theta},\alpha,\boldsymbol{\beta},\boldsymbol{\pi},\boldsymbol{\xi}) f(\boldsymbol{\beta}|\boldsymbol{\theta},\lambda_k^2\tau^2), \tag{B.4}
$$

is not analytically tractable. Thus, we sample via an MH step.

When stationarity is assumed, then we transform the parameter to impose the stationarity constraint as follows $\beta_k = 1/(1+\exp(-b_k))$. The new target distribution is $f(\boldsymbol{\beta}|\ldots)J(\boldsymbol{\beta})$ where the Jacobian $J(\boldsymbol{\beta})$ is the product over k of $\beta_k(1-\beta_k)$. A random walk proposal is used for $\boldsymbol{b} = (b_1,\ldots,b_K)$.

B.3.4 Posterior Sampling of τ and λ_k without truncation

Under the standard horseshoe prior setup, one can adopt the parameter expansion scheme of Makalic and Schmidt (2015), which introduces auxiliary inverse-gamma variables that yield closed-form full conditionals.

One can assume the following hierarchical representation:

$$\beta_k \mid \lambda_k^2, \tau^2 \sim \mathcal{N}(\tilde{\beta}_k, \lambda_k^2 \tau^2), \tag{B.5}$$

$$\lambda_k^2 \mid \nu_k \sim \mathcal{IG}\left(\tfrac{1}{2}, \tfrac{1}{\nu_k}\right), \quad \nu_k \sim \mathcal{IG}\left(\tfrac{1}{2}, 1\right), \tag{B.6}$$

$$\tau^2 \mid \xi \sim \mathcal{IG}\left(\tfrac{1}{2}, \tfrac{1}{\xi}\right), \quad \xi \sim \mathcal{IG}\left(\tfrac{1}{2}, 1\right), \tag{B.7}$$

where $\mathcal{IG}$ denotes the inverse gamma distribution.

The full conditional distributions are:

$$f(\lambda_k^2 \mid \ldots) \propto \mathcal{IG}\left(1, \frac{(\beta_k - \tilde{\beta}_k)^2}{2\tau^2} + \frac{1}{\nu_k}\right), \tag{B.8}$$

$$f(\nu_k \mid \ldots) \propto \mathcal{IG}\left(1, \frac{1}{\lambda_k^2} + 1\right), \tag{B.9}$$

$$f(\tau^2 \mid \ldots) \propto \mathcal{IG}\left(\frac{K+1}{2}, \sum_{k=1}^{K} \frac{(\beta_k - \tilde{\beta}_k)^2}{2\lambda_k^2} + \frac{1}{\xi}\right), \tag{B.10}$$

$$f(\xi \mid \ldots) \propto \mathcal{IG}\left(1, \frac{1}{\tau^2} + 1\right). \tag{B.11}$$

This parameter expansion not only simplifies computation but also ensures that the heavy-tailed shrinkage behavior of the horseshoe prior is preserved, even when the local shrinkage is centered at $\tilde{\beta}_k$ rather than 0.

B.3.5 Posterior Sampling of τ with truncation

If the posterior distribution of τ is not available in closed form due to the truncation and hierarchical structure, and we update τ using a Metropolis–Hastings step

$$f(\tau \mid \ldots) \propto f(\boldsymbol{\beta} \mid \tau, \boldsymbol{\lambda}, \tilde{\boldsymbol{\beta}}) p(\tau), \tag{B.12}$$

where $p(\tau)$ denotes the density of the $\mathcal{C}^+(0,1)$ prior, and the likelihood term comes from the truncated Gaussian structure on the β_k's. A log-scale random walk proposal is used for τ, i.e. $\log \tau' = \log \tau + \eta, \quad \eta \sim \mathcal{N}(0, \sigma_\tau^2)$.

B.3.6 Posterior Sampling of λ_k with truncation

If the posterior distribution of each λ_k is not available in closed form due to the presence of truncation, and we update each λ_k using a Metropolis–Hastings step

$$f(\lambda_k \mid \ldots) \propto f(\beta_k \mid \lambda_k, \tau, \tilde{\beta}_k)\, p(\lambda_k), \tag{B.13}$$

where $p(\lambda_k)$ denotes the density of the $\mathcal{C}^+(0,1)$ prior. The likelihood term arises from the truncated Gaussian prior on β_k with variance $\tau^2\lambda_k^2$. A log-scale random walk proposal is used for each λ_k, i.e.

$$\log \lambda_k' = \log \lambda_k + \zeta, \quad \zeta \sim \mathcal{N}(0, \sigma_\lambda^2).$$

B.3.7 Full conditional distribution of $\boldsymbol{p}_l$

We draw $\boldsymbol{p}_l$ exactly from the posterior distribution:

$$p(\boldsymbol{p}_l|\boldsymbol{\xi}) \propto \left(\prod_{t=1}^{T}\prod_{k=1}^{K} p_{lk}^{\xi_{lt-1}\xi_{kt}}\right)\left(\prod_{k=1}^{K} p_{lk}^{\omega_k-1}\right) \propto \prod_{k=1}^{K} p_{lk}^{\sum_{t=1}^{T}(\xi_{lt-1}\xi_{kt})+\omega_k-1}$$

which is the density function of a Dirichlet distribution with parameters $\bar{\omega}_1, \ldots, \bar{\omega}_K$ where $\bar{\omega}_k = \sum_{t=1}^{T}(\xi_{lt-1}\xi_{kt}) + \omega_k$.

C Simulation study

C.1 Synthetic Data

We assess our MCMC algorithm's efficiency and the inference procedure's effectiveness through simulations. The exercise consists of simulating the time series of π_t and y_t from the data-generating process and attempting to recover the true parameter values using our estimation procedure.

We set the true parameter values as follows: $\delta = 0.95$, $\gamma = 0.05660018$, $a = 0.002$, $\rho = 0.95$, $\sigma_e^2 = 0.1/3$, $\sigma_u^2 = 0.742759/3$. These lead to the following values for $\tilde{\alpha} = 0.04528014$ and $\tilde{\boldsymbol{\beta}}$: $\tilde{\alpha} =$, $\tilde{\boldsymbol{\beta}} = (0.2535044, 0.8520700, 0.9746402)$, with $K = 3$. We further assume $g^2 = \tau^2\lambda_k^2 = 0.008$ and we sample $\alpha = 0.1374916$ and $\boldsymbol{\beta} = (0.2266622, 0.7229688, 0.9449286)$. We further assume a persistent transition matrix P and a time horizon $T = 1200$. To keep the exercise simple, we omit the $\boldsymbol{\pi}^{SPF}$ augmentation. The top plots in Fig. C.1 show the data simulated from the DGP (black solid lines) and the simulated states (red stepwise solid line).

The top plots in Fig. C.1 show the simulated series for inflation, π_t (black) the output gap, y_t (black) for $t = 1, \ldots, 1200$ and the actual hidden regime process (solid red). The left-bottom plot shows the simulated π_t (black), the state posterior probability (dashed blue) and the MAP estimate of the states (solid blue). The right-bottom plot shows the actual states (red) and estimated states (blue). In Fig. C.3, we report the estimated posterior distributions of the parameters of interest in black with prior distributions. The procedure effectively recovers the true parameter values (dashed vertical lines) since they belong to the 95%

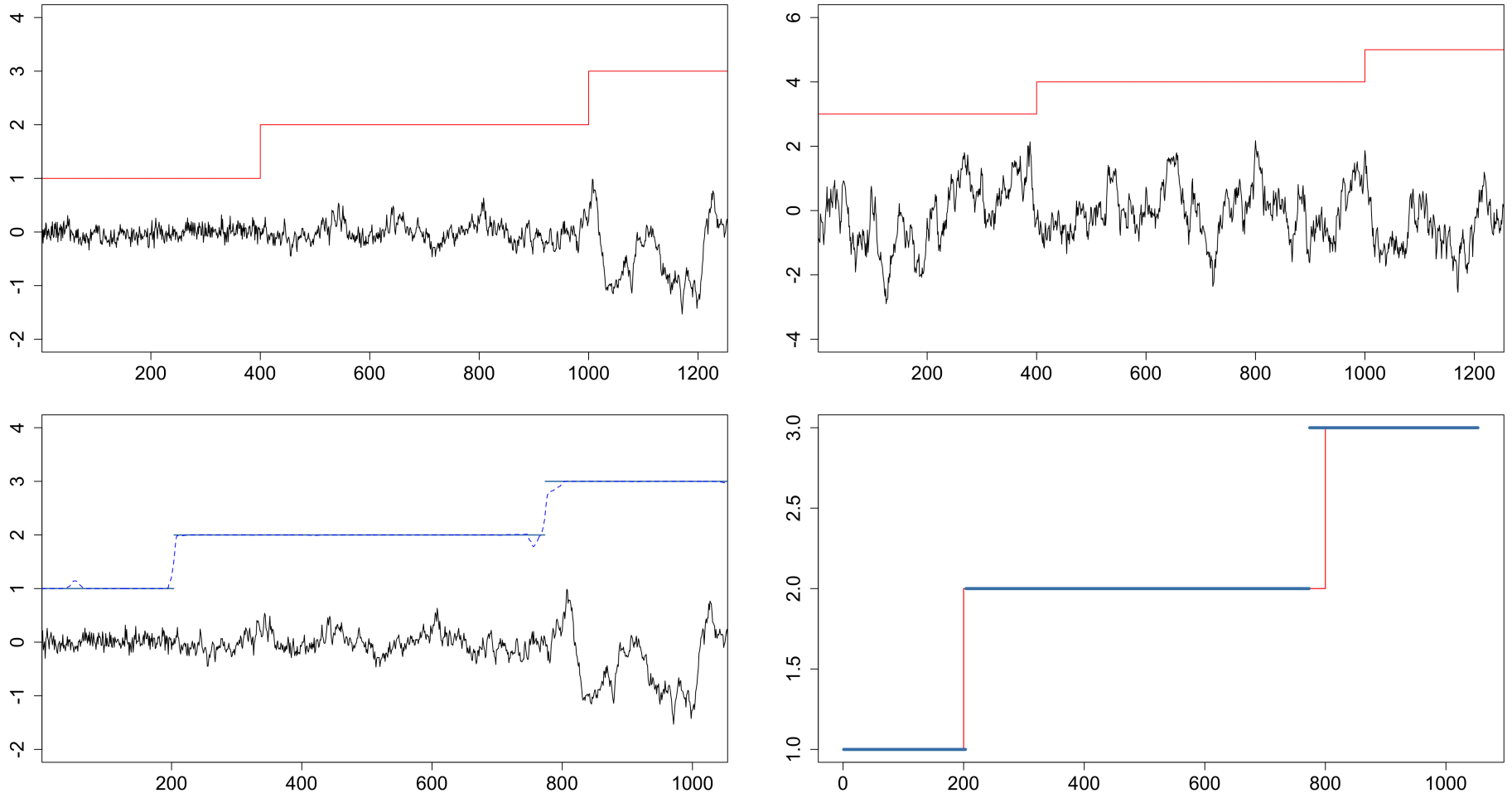


Figure C.1: Top: simulated series for inflation, π_t (left, black) and output gap, y_t (right, black) for $t = 1, \ldots, 1200$. In each plot, the simulated hidden regime process (solid red). Bottom: observed time series π_t (left, black), posterior mean of the hidden states (left, dashed blue) and state MAP estimation (left, solid red), state MAP estimation (right, solid blue) and the true simulated states (right, solid red).

Parameter	$ACF(1)$	$ACF(10)$	$ACF(30)$	%ESS	AcceptanceRate	CD p-val
(a) Before burn-in and thinning						
δ	0.96	0.77	0.61	0.63	0.25	0.00
γ	0.93	0.49	0.17	3.76	0.25	0.43
ρ	0.93	0.55	0.27	1.77	0.25	0.00
a	0.95	0.65	0.46	0.83	0.25	0.00
σ_u^2	0.91	0.38	0.07	4.77	0.25	0.04
σ_e^2	0.92	0.43	0.09	4.11	0.25	0.00
α	0.82	0.16	0.02	9.40	0.17	0.00
$\tilde{\alpha}$	0.93	0.55	0.34	1.20	0.25	0.00
β_1	0.94	0.59	0.29	2.09	0.13	0.03
β_2	0.91	0.48	0.23	3.00	0.12	0.00
β_3	0.98	0.83	0.63	0.68	0.12	0.00
$\tilde{\beta}_1$	0.94	0.63	0.36	1.82	0.25	0.23
$\tilde{\beta}_2$	0.93	0.54	0.28	2.14	0.25	0.00
$\tilde{\beta}_3$	0.95	0.70	0.56	0.76	0.25	0.07
g^2	0.21	0.08	0.04	18.16	1.00	0.00
(b) After burn-in and thinning						
δ	0.69	0.21	0.03	9.14	0.91	0.01
γ	0.48	0.02	0.01	35.05	0.91	0.20
ρ	0.51	0.07	-0.03	30.09	0.91	0.06
a	0.51	0.05	-0.00	21.56	0.91	0.08
σ_u^2	0.41	0.02	0.01	41.51	0.91	0.25
σ_e^2	0.40	0.06	0.06	42.65	0.91	0.31
α	0.19	0.02	0.01	68.23	0.83	0.41
$\tilde{\alpha}$	0.41	0.02	-0.03	25.89	0.91	0.31
β_1	0.61	-0.03	-0.00	22.73	0.70	0.03
β_2	0.46	0.06	0.01	24.20	0.69	0.33
β_3	0.81	0.24	0.03	7.35	0.69	0.02
$\tilde{\beta}_1$	0.62	0.05	0.04	17.36	0.91	0.01
$\tilde{\beta}_2$	0.52	0.01	0.04	24.62	0.91	0.18
$\tilde{\beta}_3$	0.59	0.19	0.04	10.27	0.91	0.07
g^2	0.02	-0.02	-0.01	100.00	1.00	0.08

Table C.1: MCMC Convergence Diagnostics for 30'000 iterations before (top) and after (bottom) burn-in sample removal and thinning.

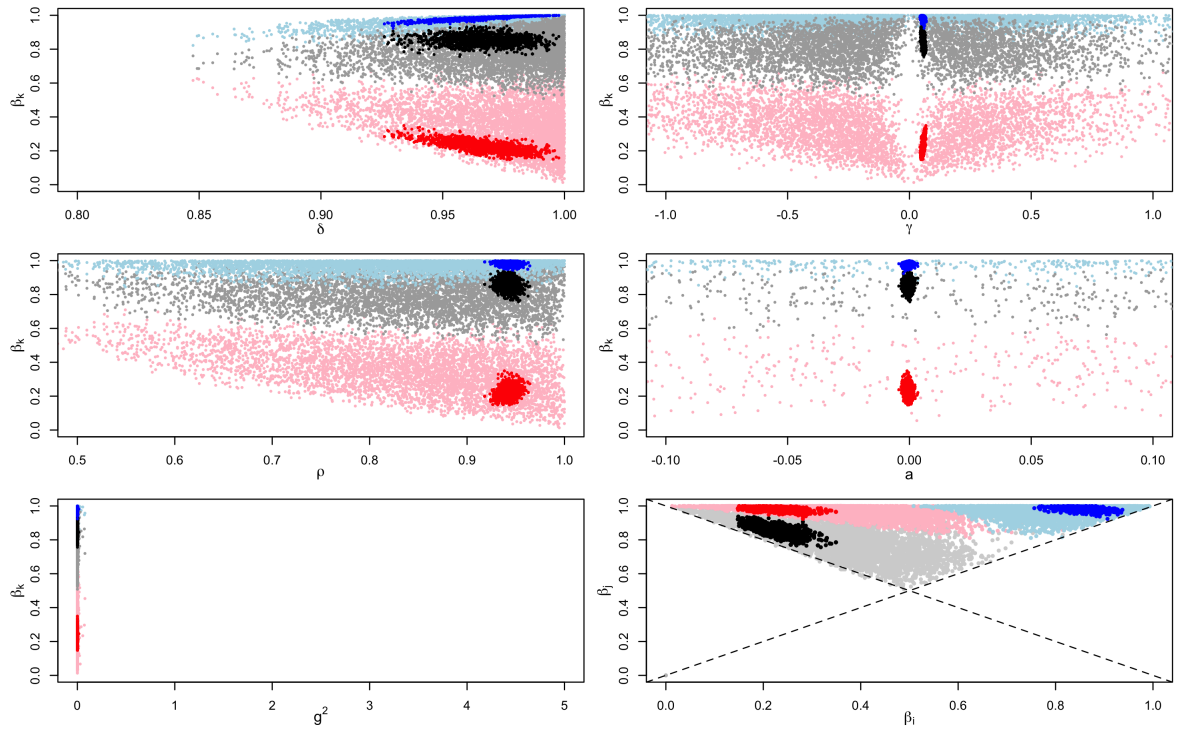


Figure C.2: Draws (dots) from the prior (light blue, grey and pink) and posterior (dark blue, black and red) distributions of the structural (horizontal axis) and reduced form (vertical axis) parameter space in simulation. In each plot, the different colours identify different equilibria. In the bottom-right plot, the pairs of (β_k, β_ℓ) and the dashed lines indicating the identification constraint $\beta_1 < \beta_2 < \beta_3$.

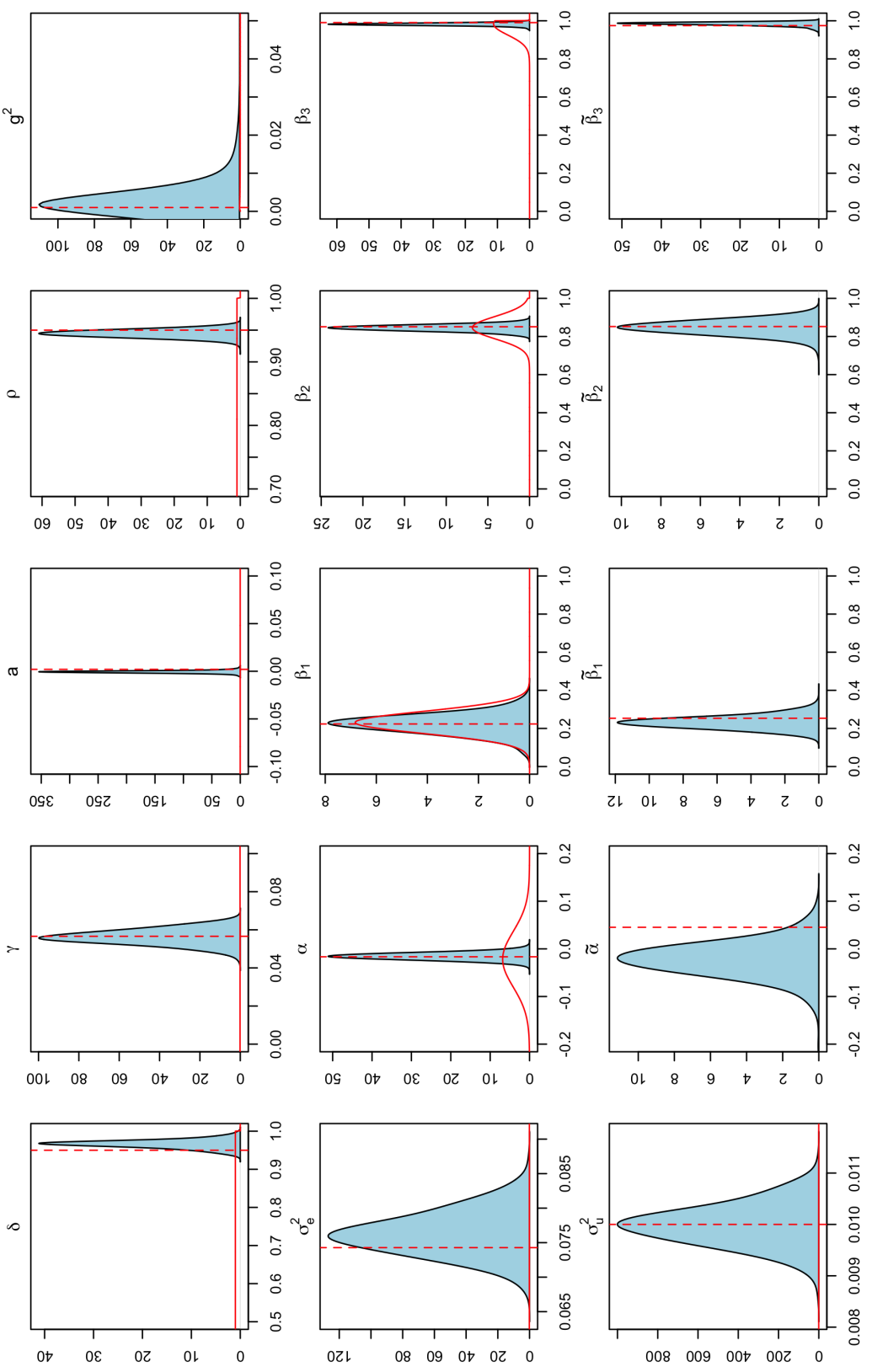


Figure C.3: Posteriors of the parameters of interest in black, prior distributions in red and true values in simulation in dashed red.

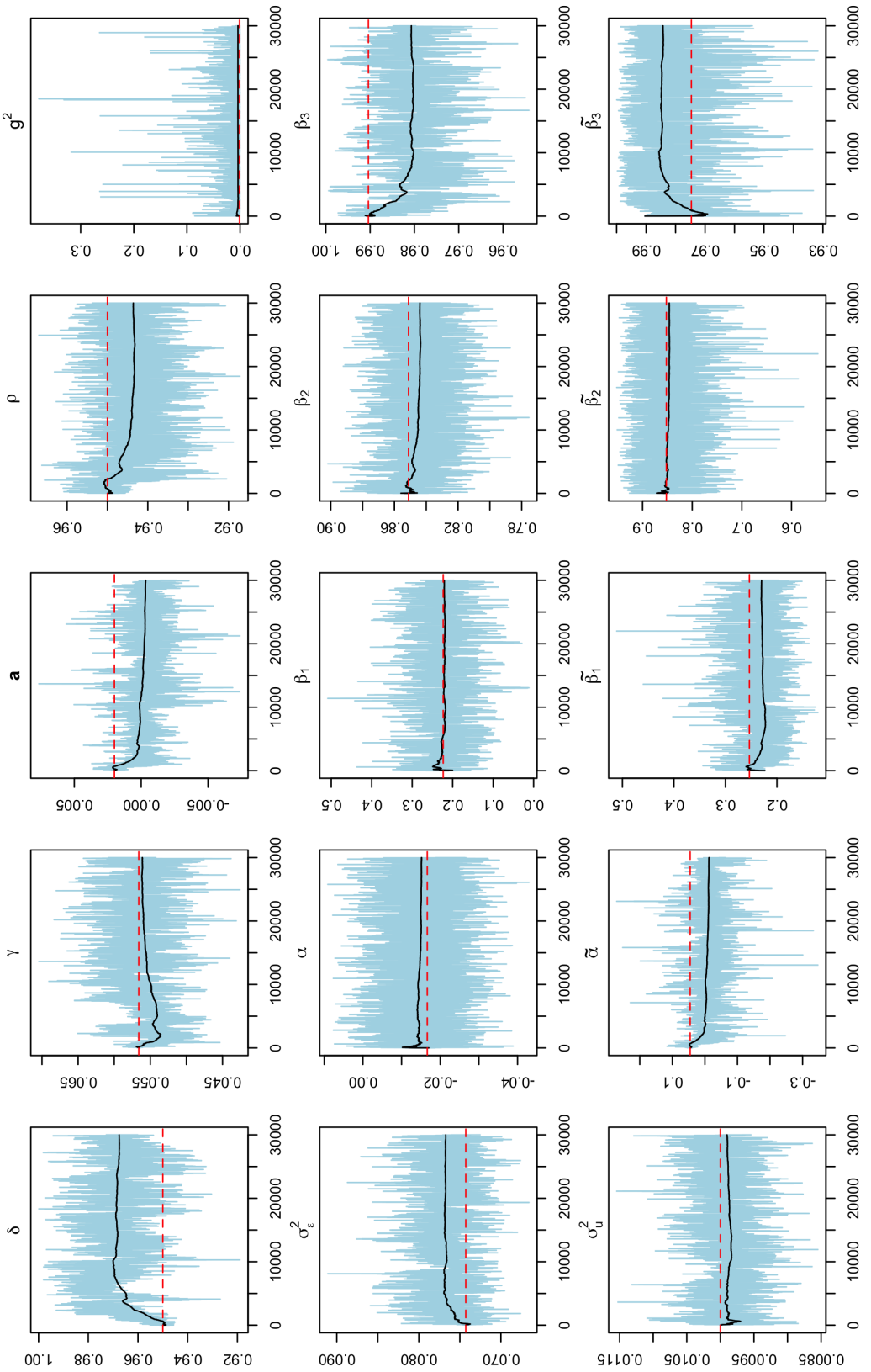


Figure C.4: Posterior MCMC draws for the parameters of interest (solid black) and the true parameter values (dashed red).

D Further Empirical Results

This section presents further results and robustness checks. Firstly, we provide further detailed results in support of the empirical evidence shown in Section 4, and secondly, we discuss the role of the prior choices for the parameters.

D.1 Further results with the Horseshoe prior

We illustrate the performance of the estimation procedure and verify the identifiability of the parameters. First, we show in Fig. D.1 the trace plot obtained from the MCMC sampler used to approximate the posterior distribution given in Fig. D.2. Empirical evidence suggests that the MCMC algorithm converges nicely to a steady state.

We report the results obtained for the case $K = 2$. Overall, the empirical estimates do not qualitatively differ from the three equilibria case, even though the DIC statistics penalize substantially the added complexity of a third regime. Figures D.3 to D.6 suggest a clear identification for the parameters and the equilibria. Furthermore, the trace plot indicates a good mixing of the MCMC.

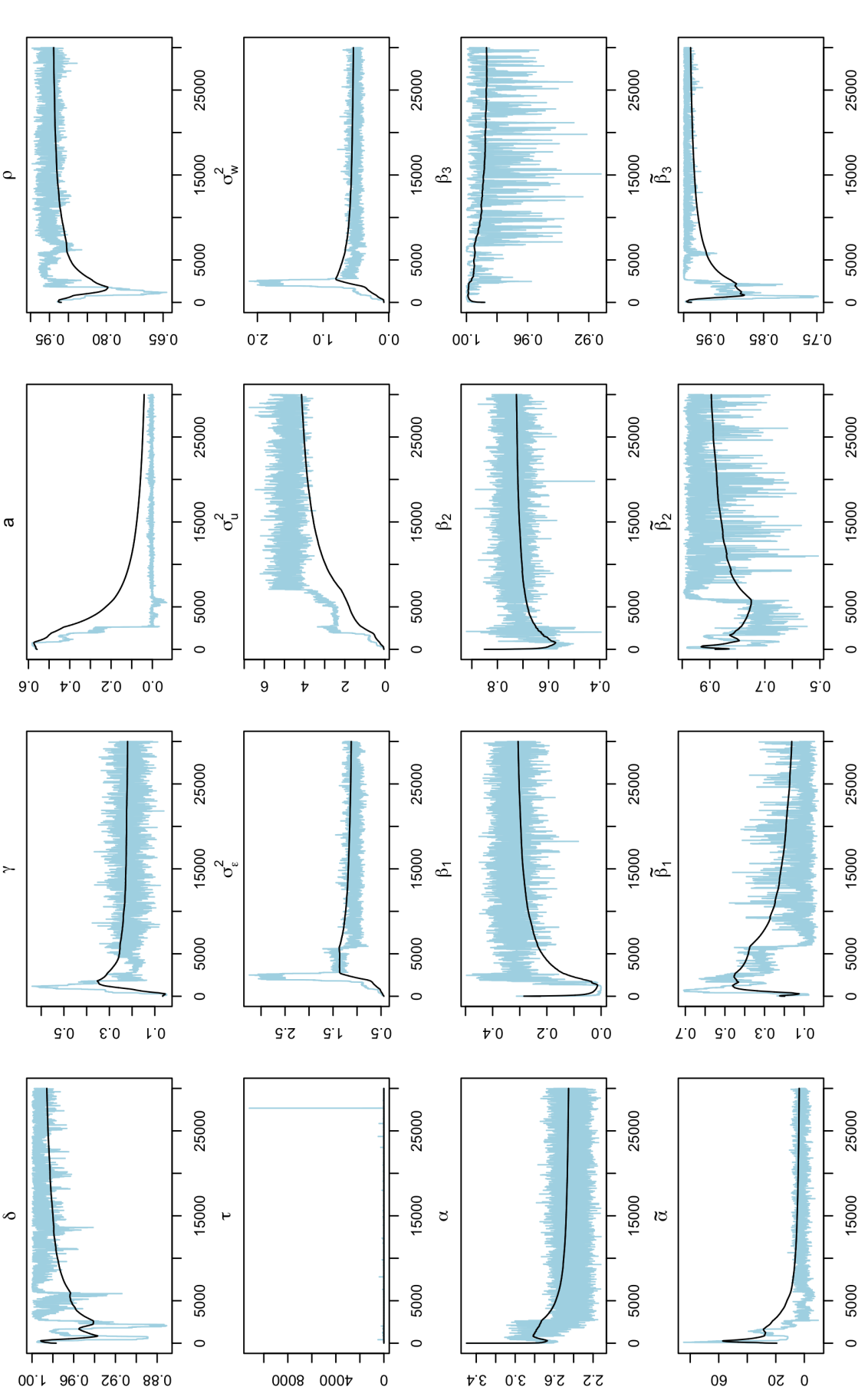


Figure D.1: Horseshoe prior: Posterior trace plots for the model with three regimes ($K = 3$). The light blue line represents the posterior draw at each iteration. The solid black line is the cumulative posterior mean

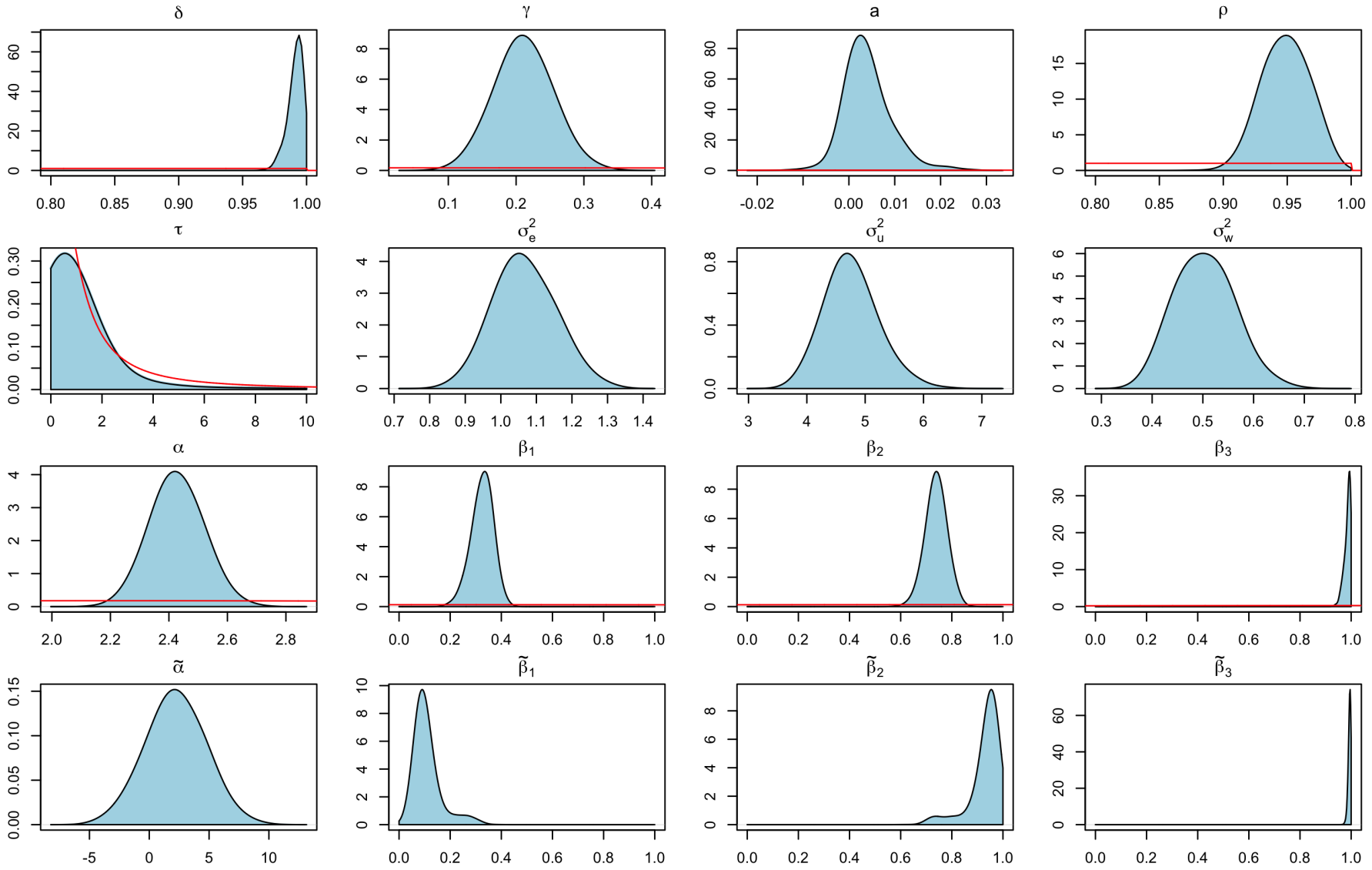


Figure D.2: Horseshoe prior: Prior vs. Posterior distributions for the model with three regimes ($K = 3$). The straight red curve is the prior distribution, whereas the black density is the estimated posterior.

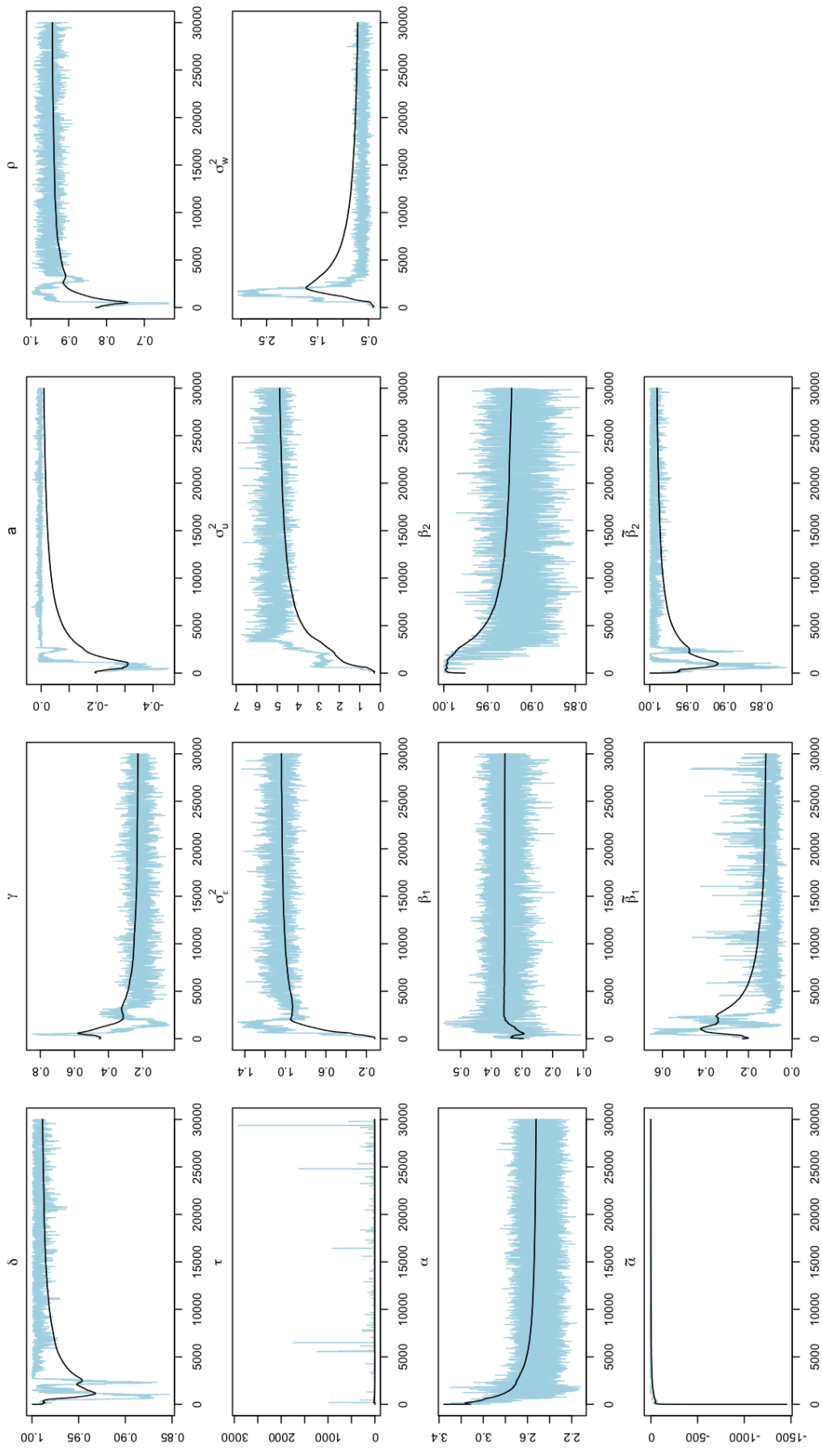


Figure D.3: Horseshoe prior: Posterior trace plots for the model with three regimes ($K = 3$). The light blue line represents the posterior draw at each iteration. The solid black line is the cumulative posterior mean

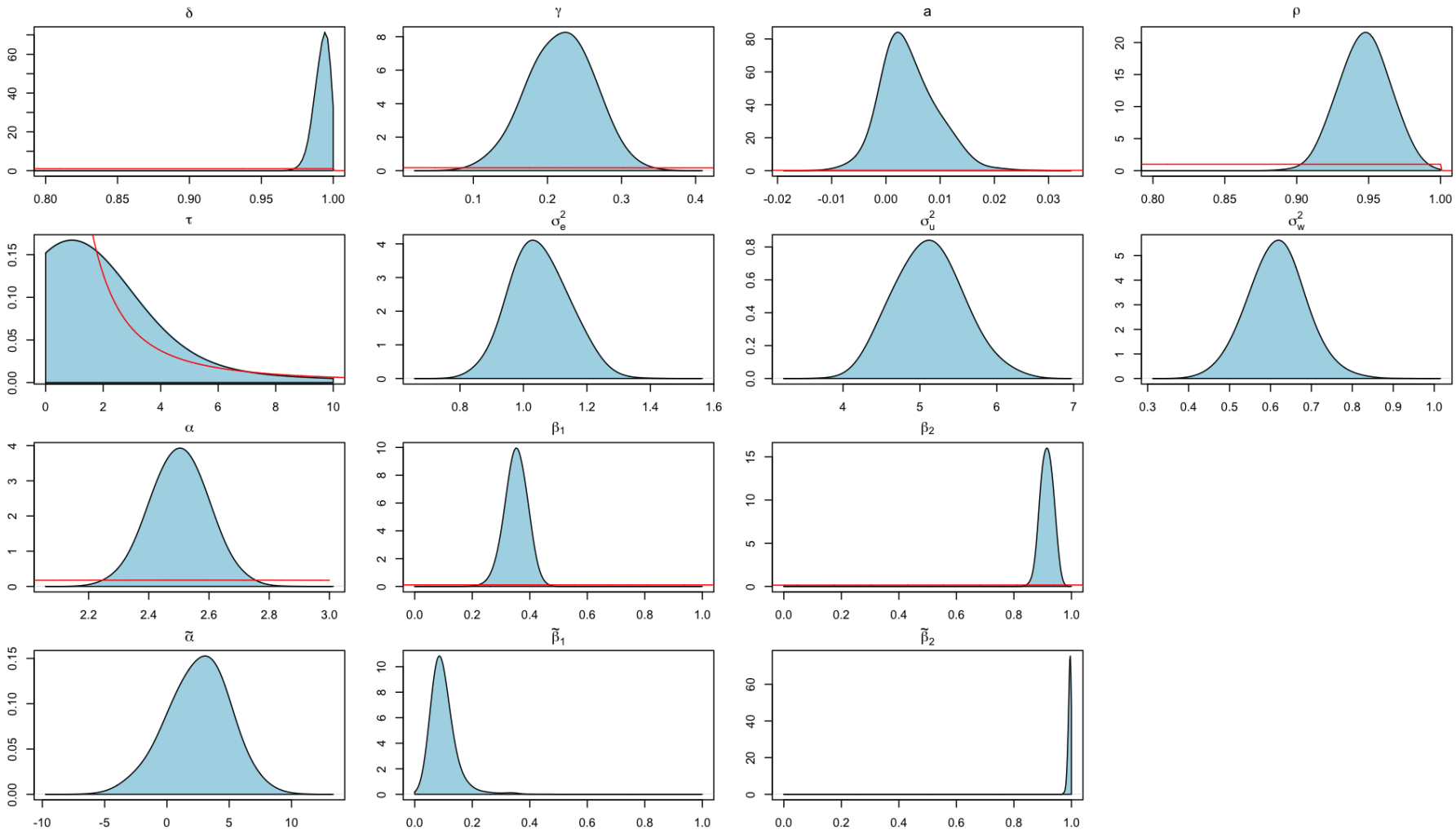


Figure D.4: Horseshoe prior: Prior vs. Posterior distributions for the model with two regimes ($K = 2$). The straight red curve is the prior distribution, whereas the black density is the estimated posterior.

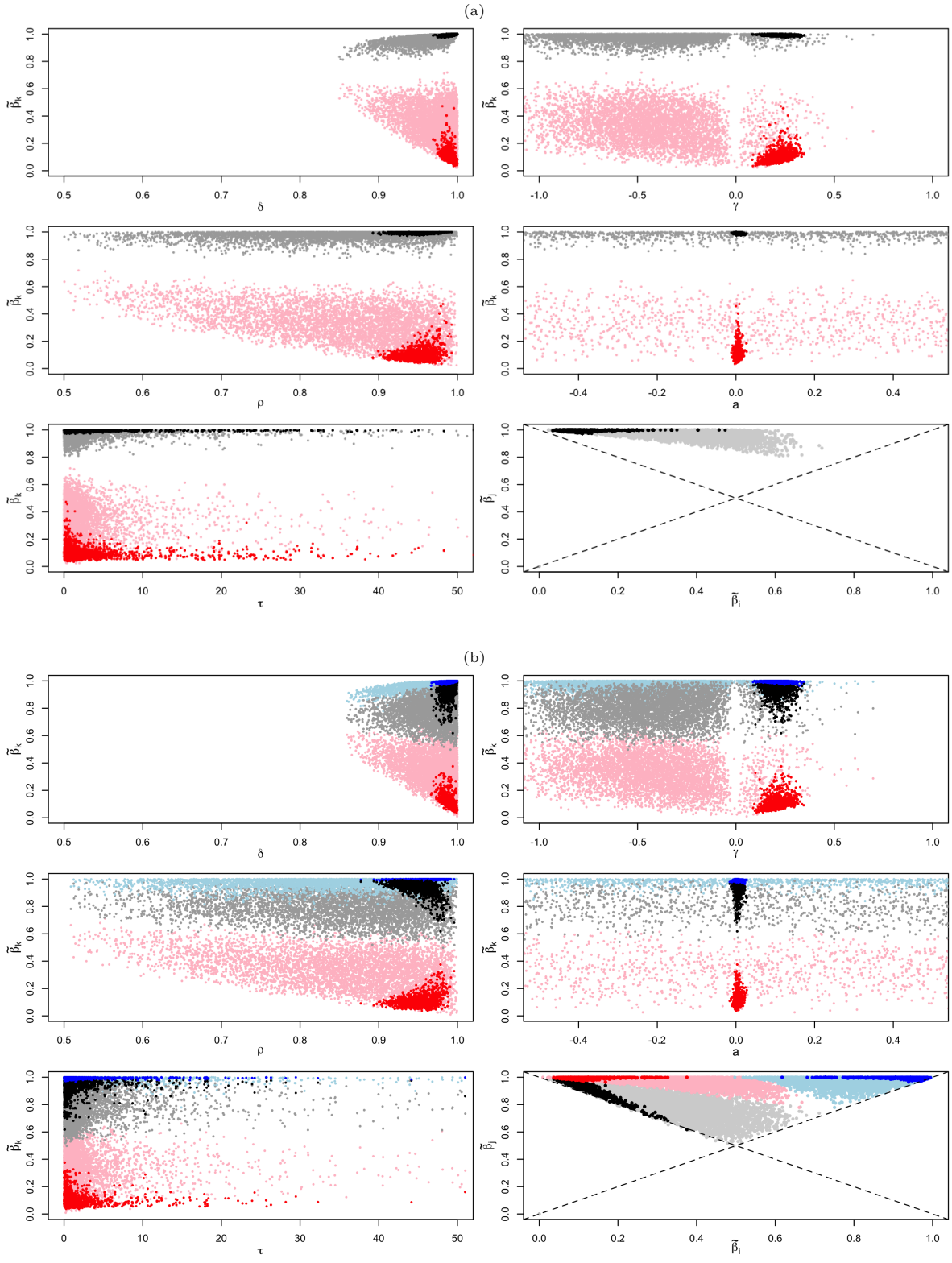


Figure D.5: Horseshoe prior: Draws (dots) from the prior (light blue, grey and pink) and posterior (dark blue, black and red) distributions of the structural (horizontal axis) and reduced form (vertical axis) parameter space for the model with $K = 2$ (top) and $K = 3$ (bottom). In each plot, the different colours identify different equilibria. In the bottom-right plot, the pairs of (β_k, β_ℓ) and the dashed lines indicating the identification constraint $\beta_1 < \beta_2$.

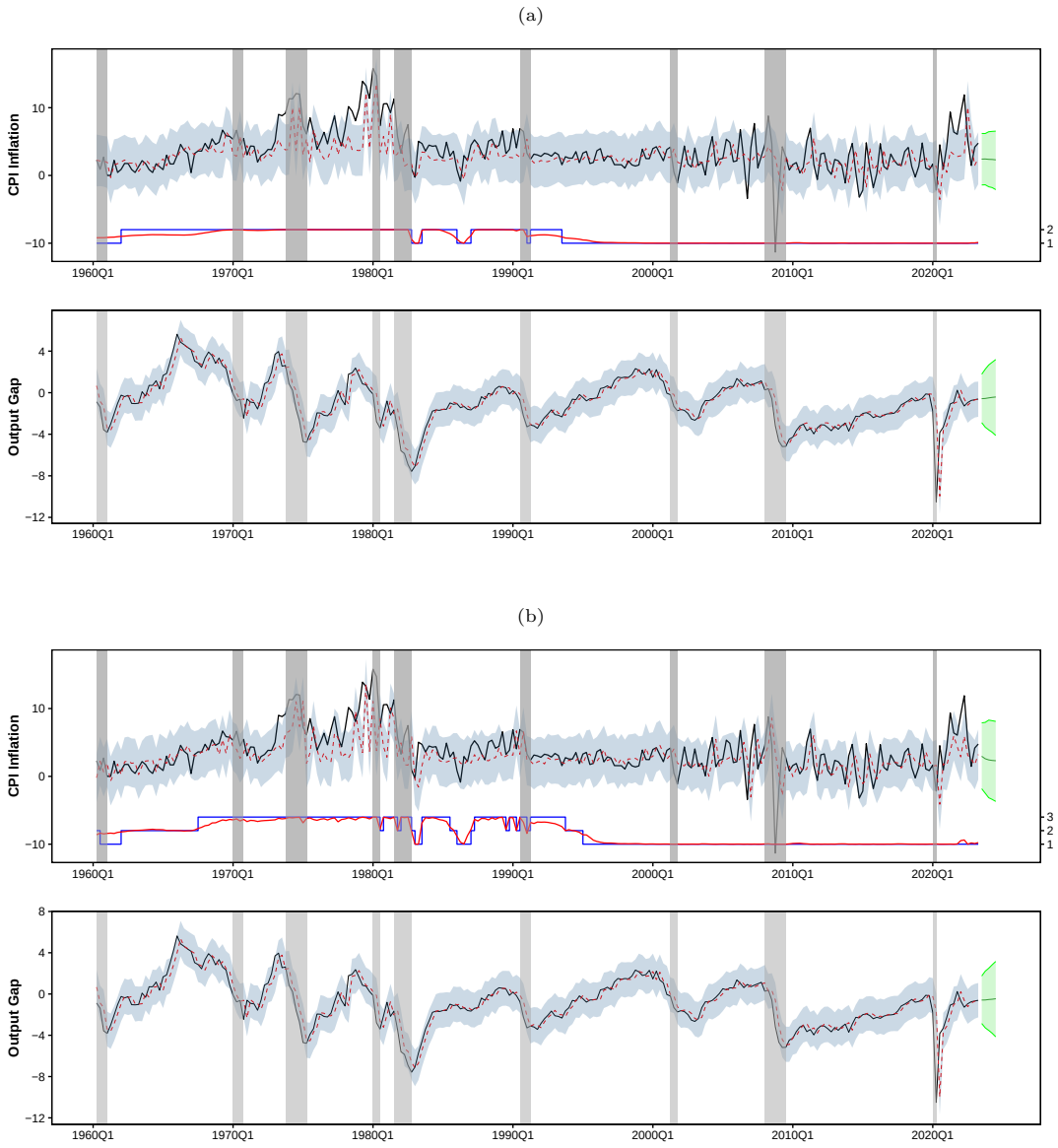


Figure D.6: Horseshoe prior: Estimated equilibria with $K = 2$ (top) and $K = 3$ (bottom). In each panel: Top plots, observed CPI inflation (solid black line) and predicted CPI inflation (dashed red line) with 90% credible interval (light grey area). Below, estimated active equilibrium over time ($\hat{s}_t \in \{1, 2\}$, right axis, blue line) and the regime posterior mean (right axis, red line) together with the NBER dating of the business cycle (shaded areas). Bottom plots, output gap (solid black line) and predicted output gap (dashed red line) with 90% credible interval (light grey area). Solid green line and area represent out-of-sample predictive mean and 90% credible region, respectively.

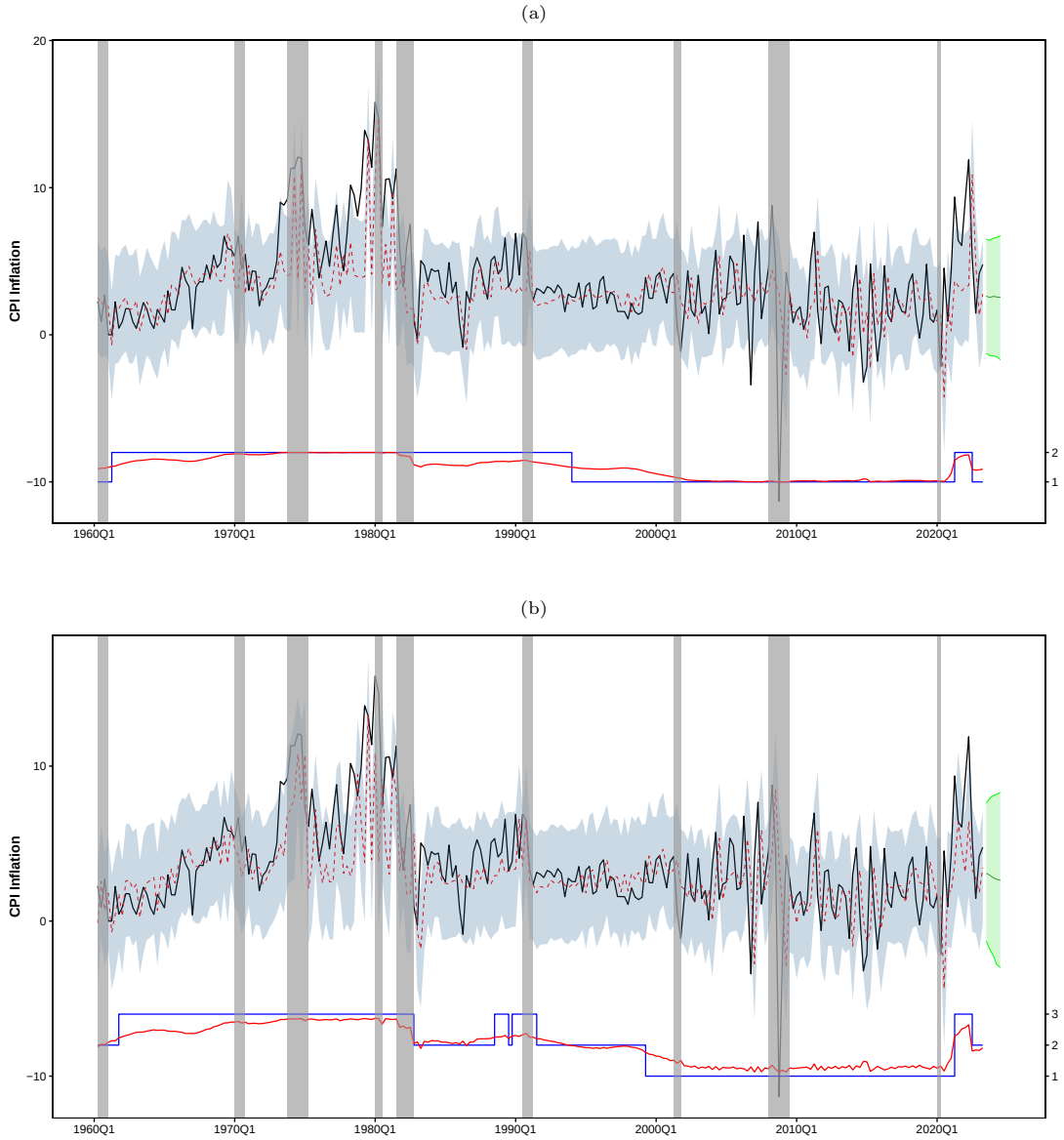
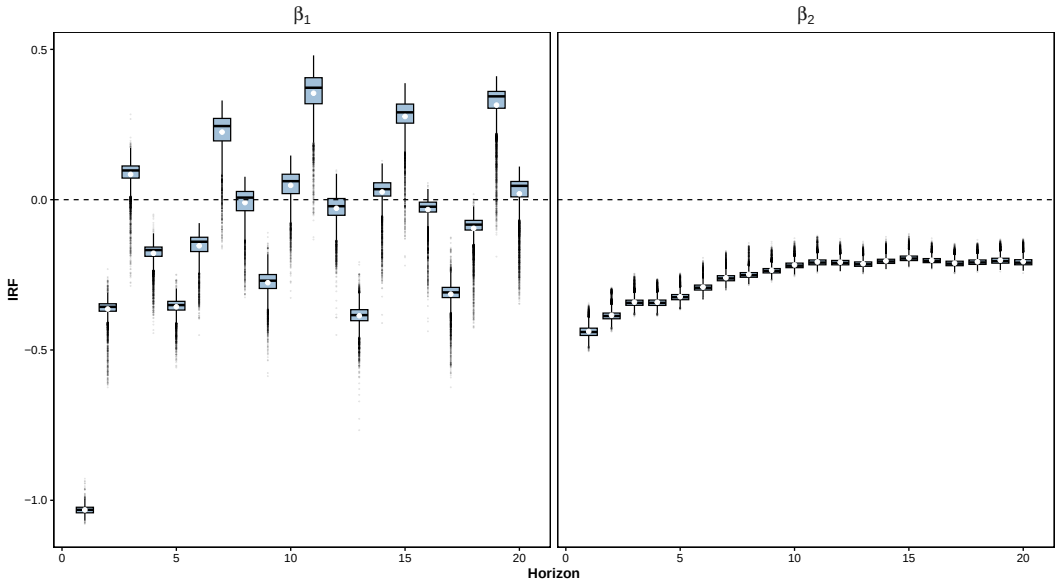


Figure D.7: Horseshoe prior: Estimated equilibria with $K = 2$ (top) and $K = 3$ (bottom) without the SPF equation. Top plot: observed CPI inflation (solid black line) and predicted CPI inflation (dashed red line) with 90% credible interval (light grey area). Below, estimated active equilibrium over time ($\hat{s}_t \in \{1, 2\}$, right axis, blue line) and the regime posterior mean (right axis, red line) together with the NBER dating of the business cycle (shaded areas). Solid green line and area represent out-of-sample predictive mean and 90% credible region, respectively.

(a)



(b)

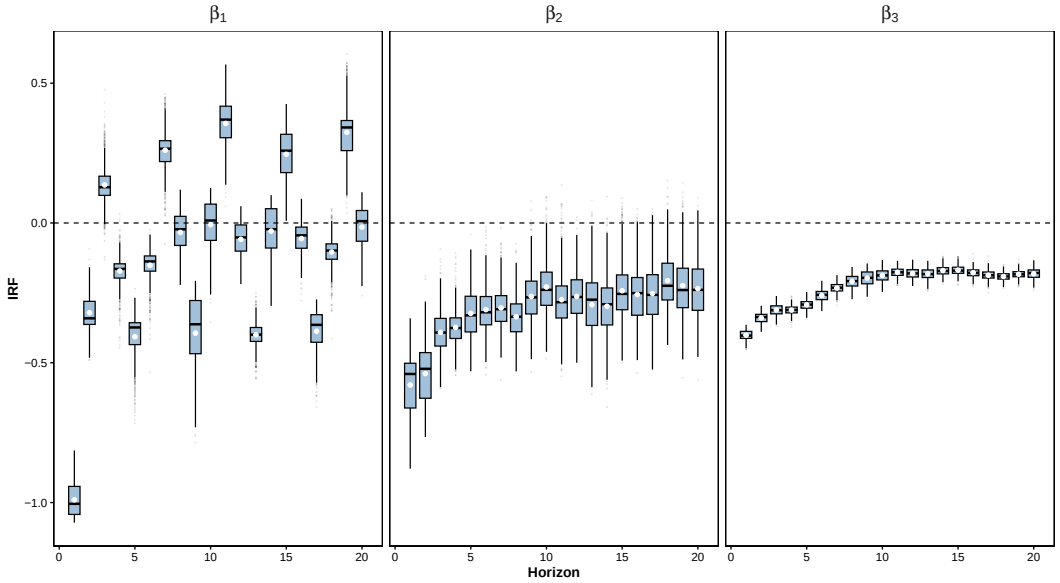


Figure D.8: Composite bias coefficients via local projections at different horizons $h = 1, \ldots, 20$ for the model with $k = 2$ (top) and $K = 3$ (bottom). The boxplots report the posterior draws of the composite bias coefficients at each horizon h.

D.2 Hierarchical diffuse prior

As an alternative to the hierarchical horseshoe prior proposed in the paper, we consider a diffuse variant based on Gaussian distributions. Two conditionally independent normal prior distributions measure the uncertainty level in the prior belief :

$$\alpha|\boldsymbol{\theta}, g^2 \sim \mathcal{N}(\tilde{\alpha}(\boldsymbol{\theta}), g^2), \quad \beta_k|\boldsymbol{\theta}, g^2 \sim \mathcal{TN}(\tilde{\beta}_k(\boldsymbol{\theta}), g^2, 0, 1),\ k = 1, \ldots, K \tag{D.1}$$

with $\beta_1 < \beta_2 < \ldots < \beta_K$, where g^2 is an hyper-parameter to be estimated and $\mathcal{TN}(m, s^2, a, b)$ denotes the Truncated normal distribution with location m, scale s and support (a, b).

At the second stage of the hierarchical prior, we choose an inverse gamma prior for the shrinking parameter g^2

$$g^2 \sim \mathcal{IG}(a_g, b_g), \tag{D.2}$$

where $\mathcal{IG}(a, b)$ denotes the Inverse Gamma distributions with shape $a > 0$ and scale $b > 0$.

	Multiple equilibria Phillips Curve Model $\mathcal{M}^{PC}(K)$					
	$K = 1$		$K = 2$		$K = 3$	
Parameters	Post. Mean	95% Cred. Interval	Post. Mean	95% Cred. Interval	Post. Mean	95% Cred. Interval
δ	0.99	[0.978 0.999]	0.99	[0.98 0.999]	0.99	[0.978 0.999]
γ	0.10	[0.018 0.198]	0.21	[0.122 0.3]	0.20	[0.121 0.296]
ρ	0.92	[0.874 0.957]	0.95	[0.919 0.979]	0.95	[0.919 0.979]
a	0.02	[0.001 0.054]	0.00	[0.001 0.011]	0.01	[0 0.012]
σ_u^2	7.09	[6.014 8.277]	5.09	[4.316 5.967]	4.70	[4.022 5.478]
σ_e^2	1.03	[0.875 1.203]	1.06	[0.903 1.24]	1.06	[0.899 1.234]
σ_w^2	1.18	[0.963 1.445]	0.61	[0.494 0.744]	0.51	[0.409 0.627]
α	3.06	[2.721 3.404]	2.51	[2.292 2.717]	2.43	[2.227 2.632]
$\tilde{\alpha}$	2.77	[0.737 4.772]	2.55	[0.83 4.357]	2.45	[0.664 4.226]
β_1	0.61	[0.559 0.653]	0.36	[0.279 0.422]	0.33	[0.239 0.4]
$\tilde{\beta}_1$	0.01	[0 0.033]	0.10	[0.05 0.186]	0.10	[0.051 0.192]
β_2			0.91	[0.873 0.955]	0.72	[0.646 0.805]
$\tilde{\beta}_2$			0.99	[0.983 0.999]	0.94	[0.838 0.984]
β_3					0.98	[0.947 1]
$\tilde{\beta}_3$					0.99	[0.983 0.999]
g^2	1.97	[0.246 8.782]	1.01	[0.198 4.014]	0.66	[0.163 2.371]
DIC	2448.07		2266.46		2408.66	

Table D.1: Parameter estimates from the MCMC algorithm for the three models and diffuse hierarchical priors. DIC values are also provided.

The results reported in Table D.1 are similar to those reported in the main text, although with this specification, there is a larger gap between equilibrium parameters and statistical regimes. Policy

implications, such as underreaction and missing disinflation, closely align with the results presented in the main paper.

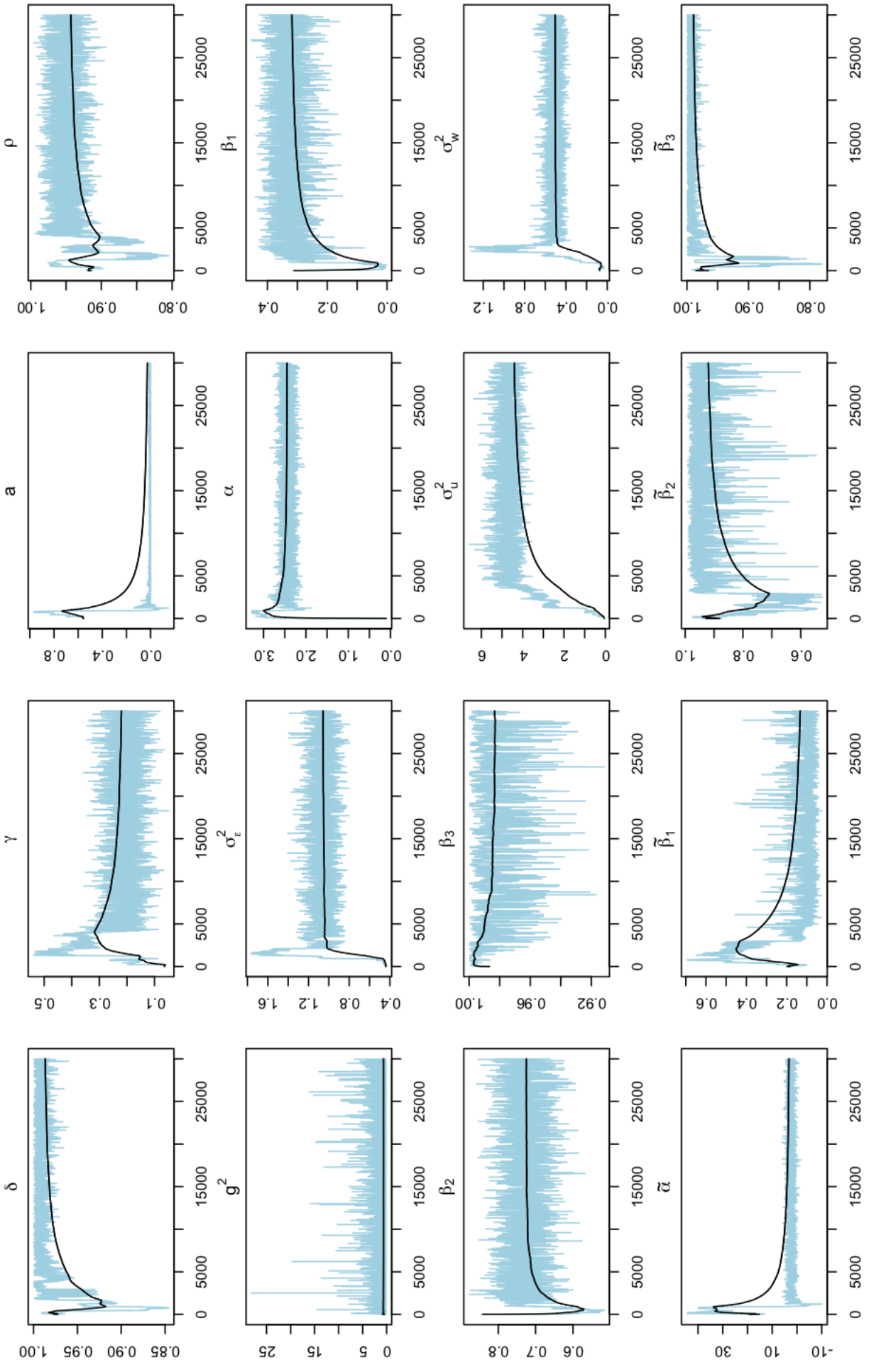


Figure D.9: Hierarchical diffuse priors: Posterior trace plots for the model and three regimes ($K = 3$). The light blue line represents the posterior draw at each iteration. The solid black line is the cumulative posterior mean

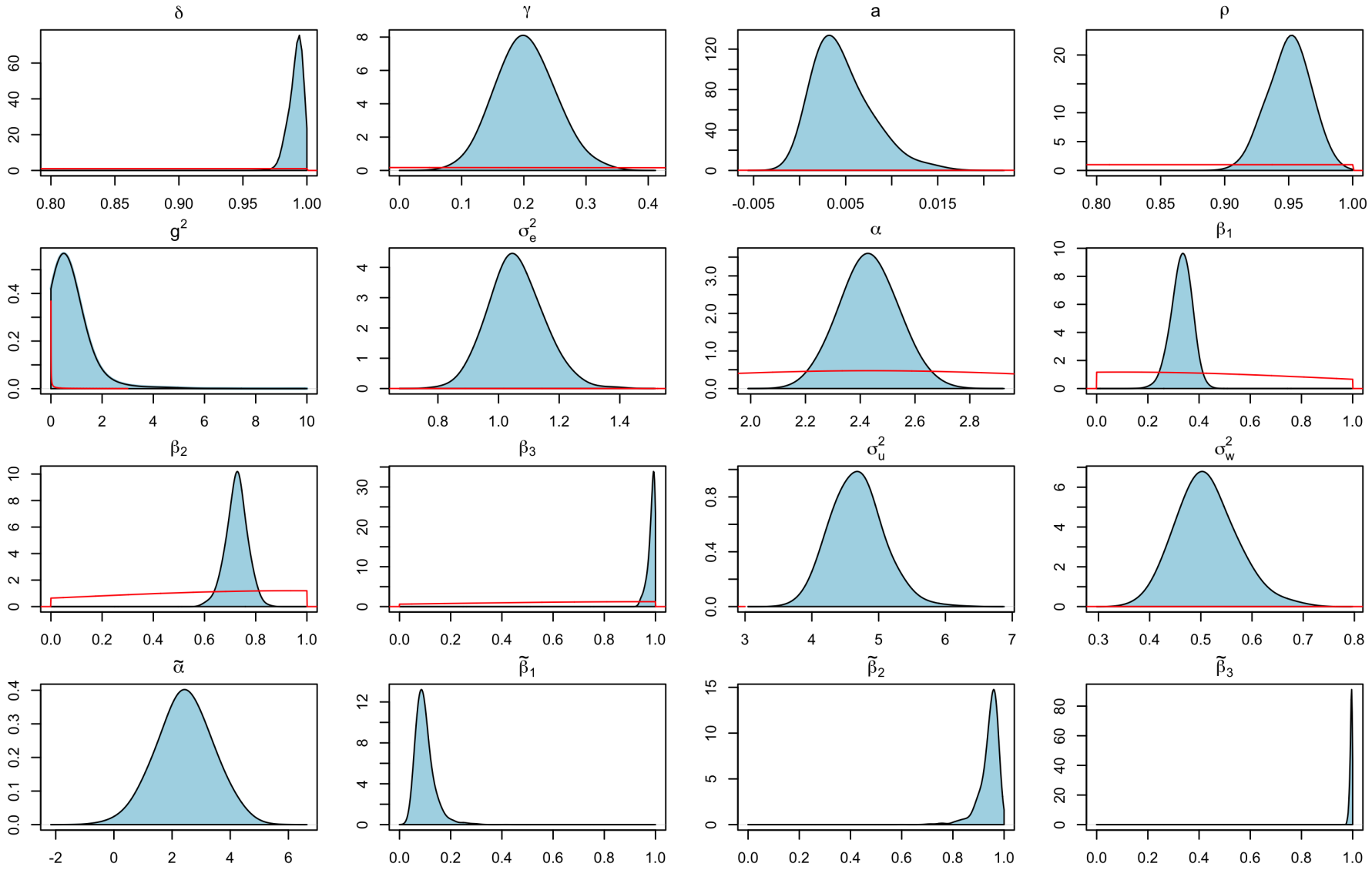


Figure D.10: Hierarchical diffuse priors: Prior vs. Posterior distributions for the model with an Horseshoe prior and three regimes ($K = 3$). The straight red curve is the prior distribution, whereas the black density is the estimated posterior.

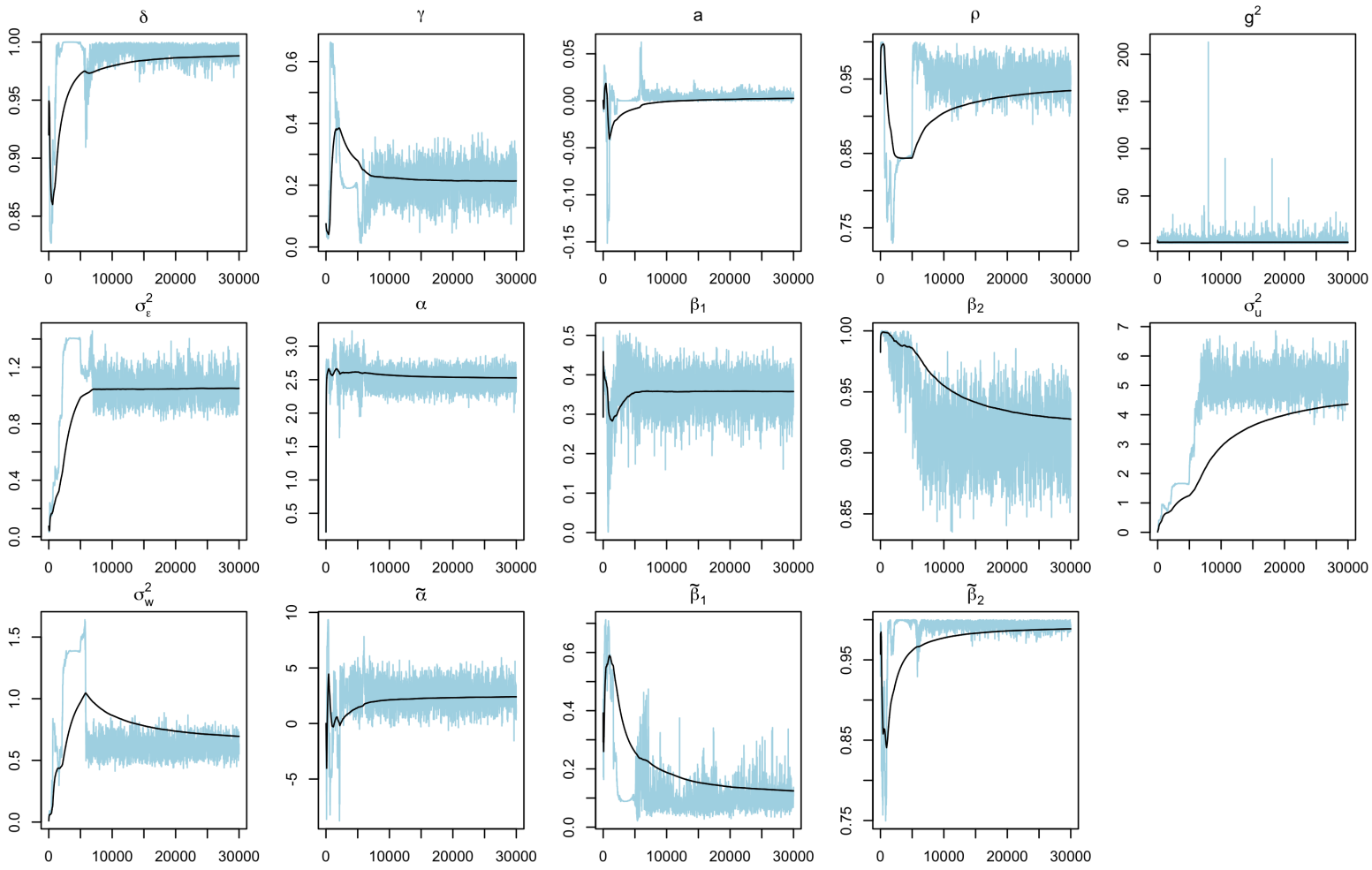


Figure D.11: Hierarchical diffuse priors: Posterior trace plots for the model with two regimes ($K = 2$). The light blue line represents the posterior draw at each iteration. The solid black line is the cumulative posterior mean

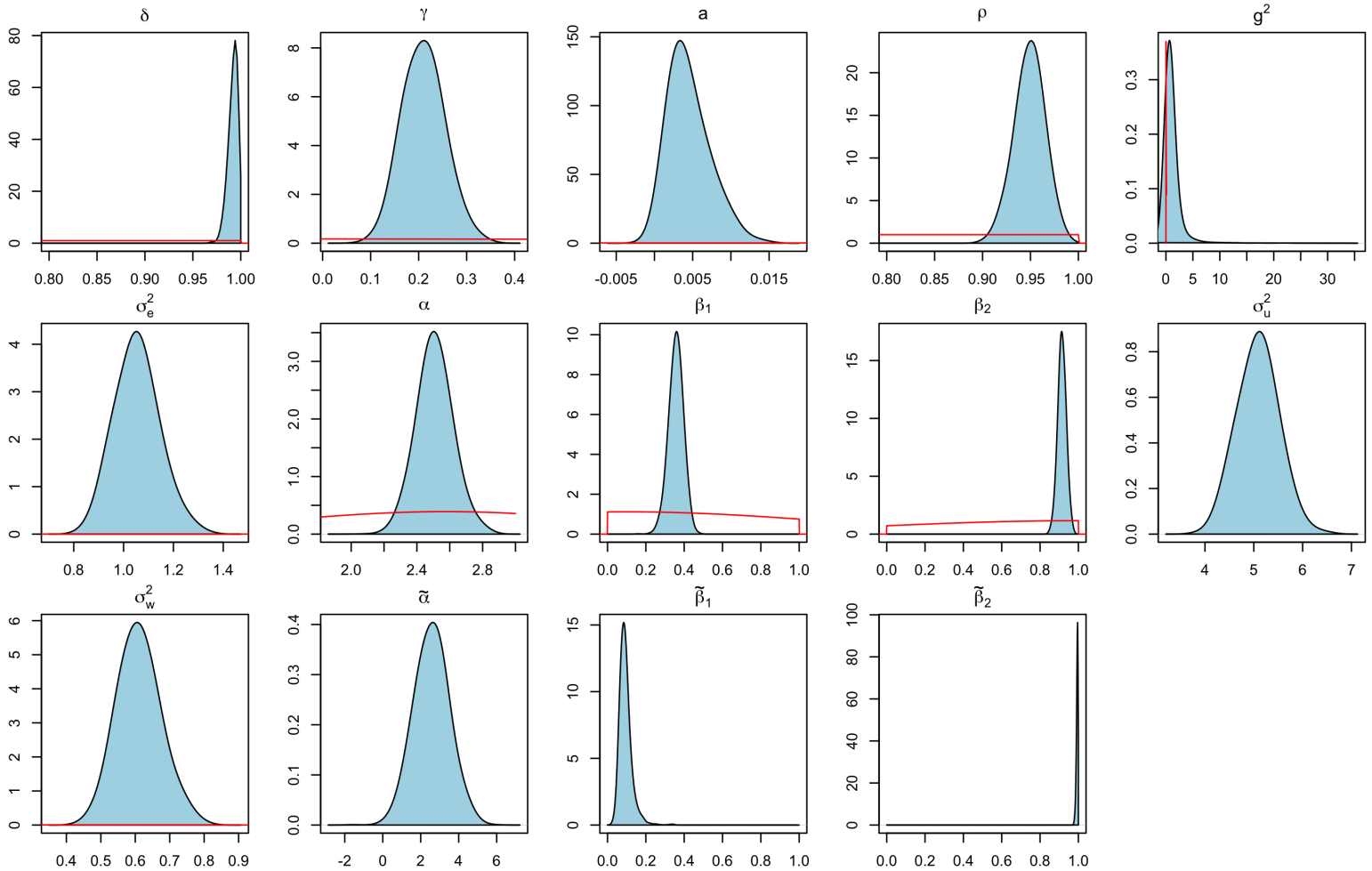


Figure D.12: Hierarchical diffuse priors: Prior vs. Posterior distributions for the model with two regimes ($K = 2$). The straight red curve is the prior distribution, whereas the black density is the estimated posterior.

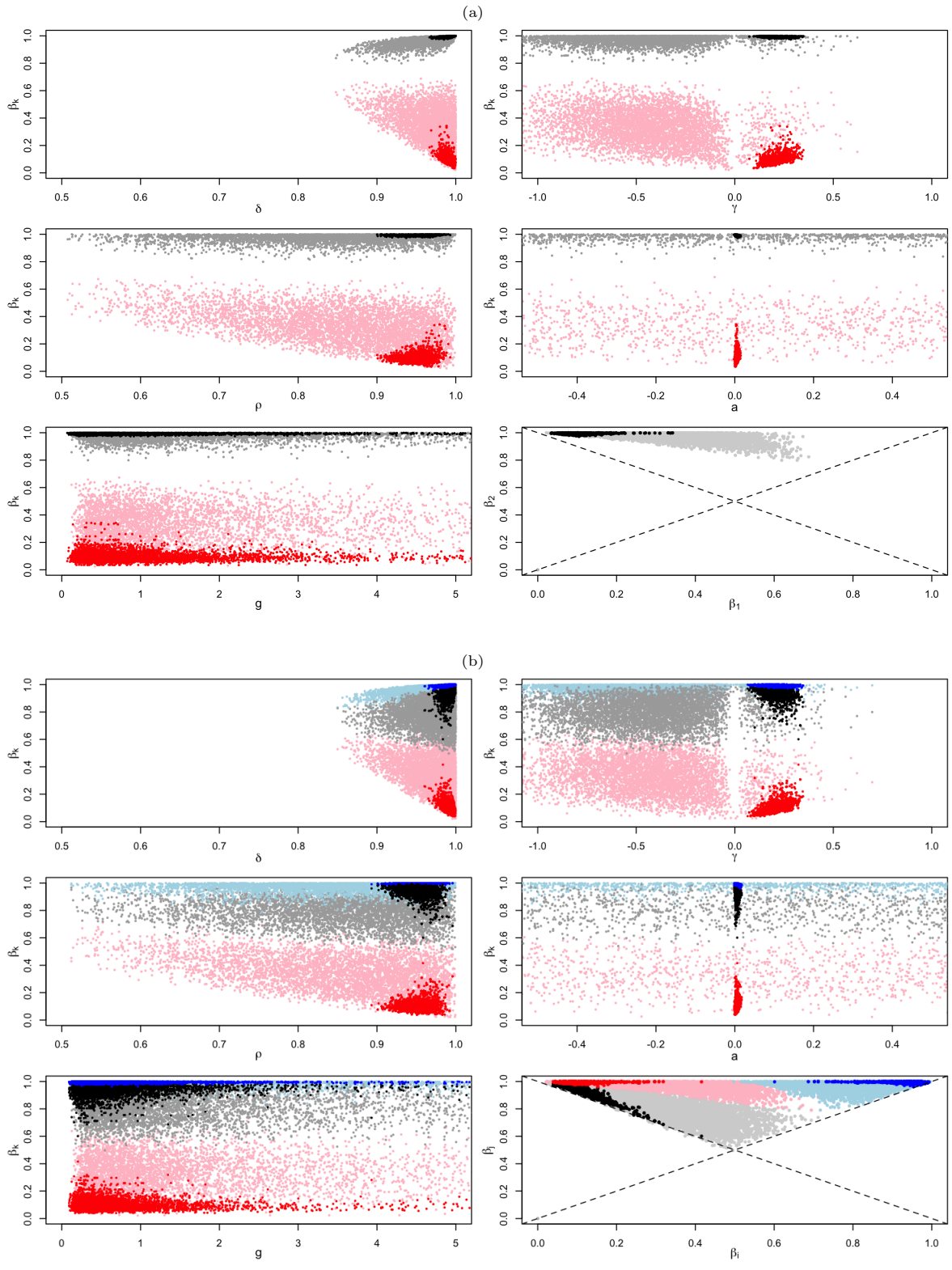


Figure D.13: Hierarchical diffuse priors: Draws (dots) from the prior (light blue, grey and pink) and posterior (dark blue, black and red) distributions of the structural (horizontal axis) and reduced form (vertical axis) parameter space for the model with $K = 2$ (top) and $K = 3$ (bottom). In each plot, the different colours identify different equilibria. In the bottom-right plot, the pairs of (β_k, β_ℓ) and the dashed lines indicating the identification constraint $\beta_1 < \beta_2$.

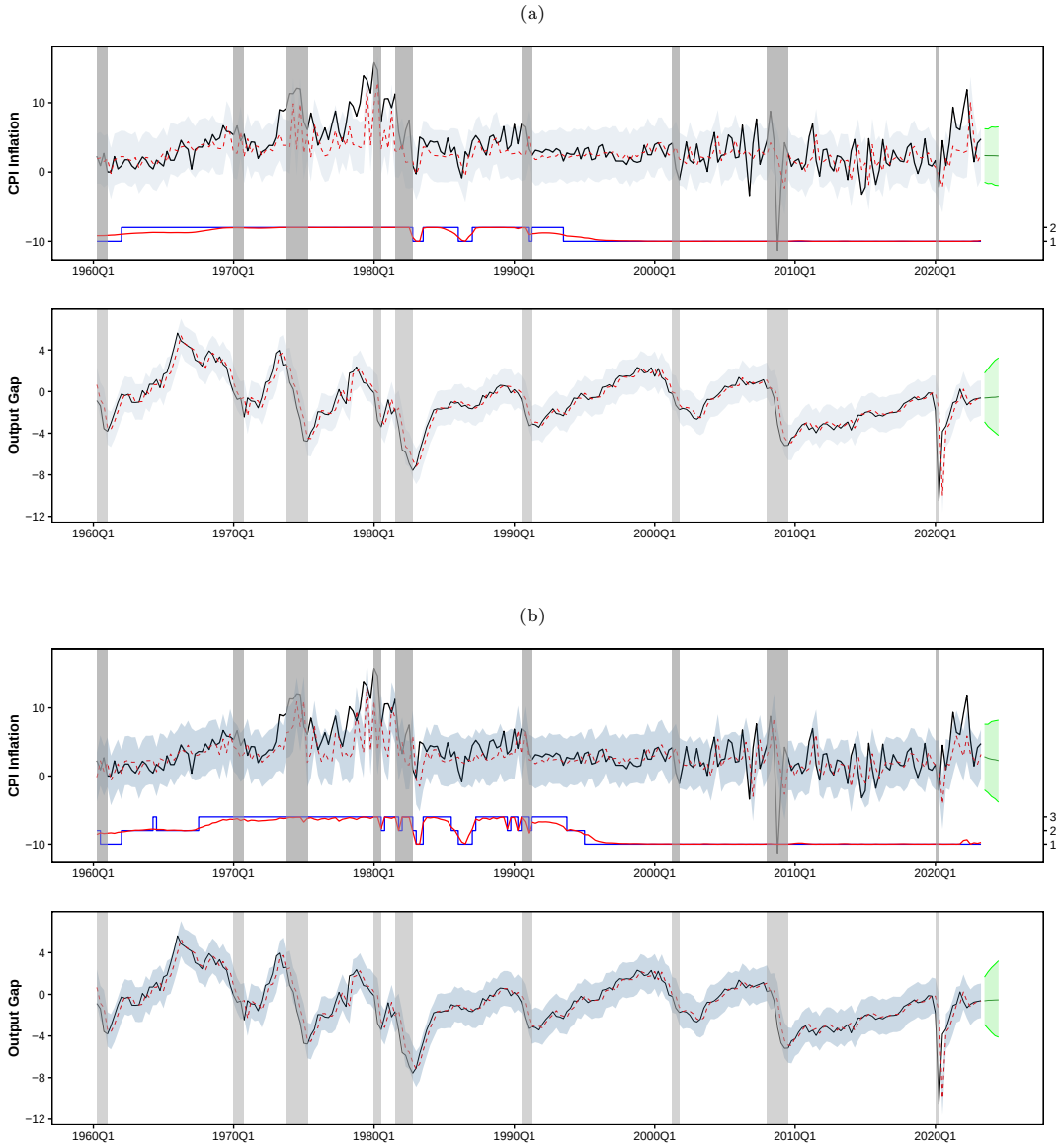
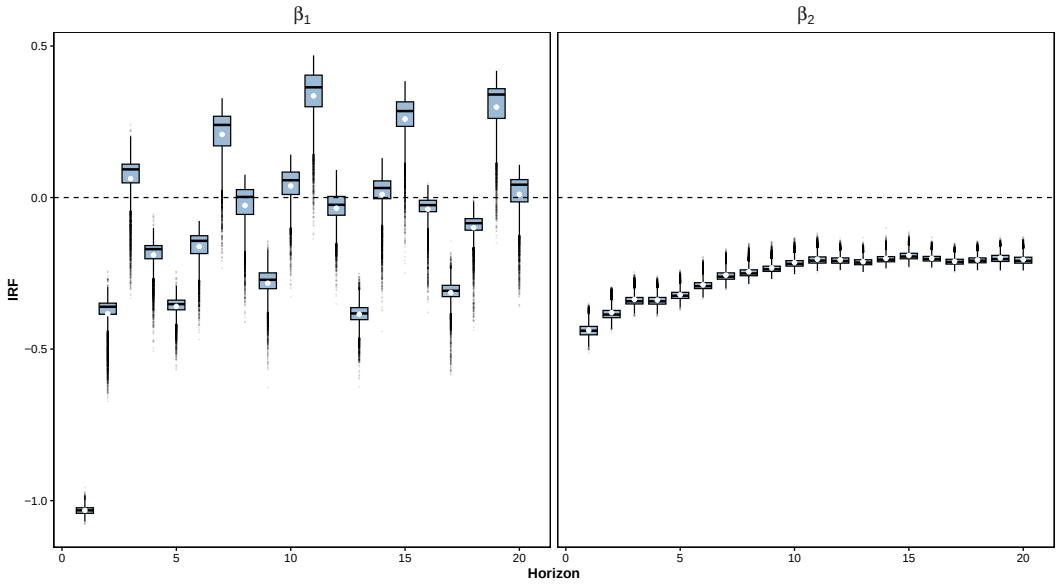


Figure D.14: Hierarchical diffuse priors: Estimated equilibria with $K = 2$ (top) and $K = 3$ (bottom). In each panel: Top plots, observed CPI inflation (solid black line) and predicted CPI inflation (dashed red line) with 90% credible interval (light grey area). Below, estimated active equilibrium over time ($\hat{s}_t \in \{1, 2\}$, right axis, blue line) and the regime posterior mean (right axis, red line) together with the NBER dating of the business cycle (shaded areas). Bottom plots, output gap (solid black line) and predicted output gap (dashed red line) with 90% credible interval (light grey area). Solid green line and area represent out-of-sample predictive mean and 90% credible region, respectively.

(a)



(b)

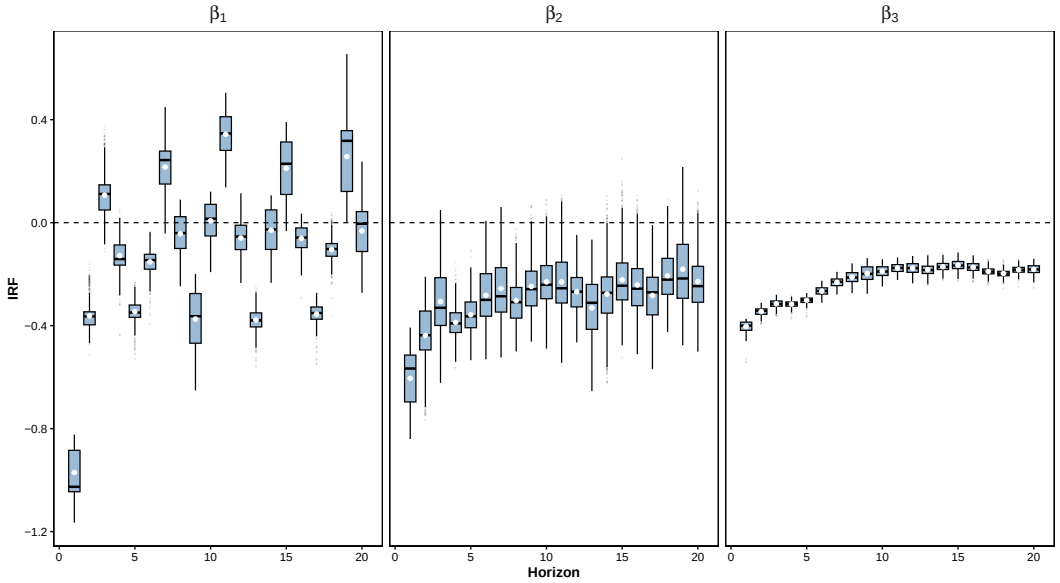


Figure D.15: Hierarchical diffuse priors: Composite bias coefficients via local projections at different horizons $h = 1, \ldots, 20$ for the model with $K = 2$ (top) and $K = 3$ (bottom). The boxplots report the posterior draws of the composite bias coefficients at each horizon h.