Price Gaps and Inflation Dynamics*

Miguel Bandeira† Laura Castillo-Martínez‡ Shiyuan Wang§

September 2026

Abstract

The inflationary effect of a shock depends not only on its size, but also on the state of the economy when it hits. In menu cost models, this state is captured by the cross-sectional distribution of price gaps: the misalignment between firms' actual and reset prices. However, this distribution is unobserved. We recast a random menu cost model in state-space form and develop Bayesian methods to recover latent reset prices from observed price paths. Applying the method to almost 35 million UK consumer price observations from 1996 to 2026, we construct a time-varying empirical distribution of price gaps and show that its shape varies substantially over time. Measuring this distribution lets us take the model's state-dependence predictions directly to the data. We find substantial state dependence in shock transmission: at a two-year horizon, the response to the same monetary tightening ranges from almost fully dampened to amplified by 69 percent across pre-shock gap distributions, with similar patterns for oil-price shocks. The distribution also contains information about where inflation is headed. Gap moments predict future inflation beyond standard macroeconomic controls and improve out-of-sample forecasts relative to a random-walk benchmark up to 37 percent at two years.

JEL Codes: E31, C32, C11, E52, E37.

Keywords: Price gaps, Ss models, state-space methods, inflation, monetary policy.

*We are especially grateful to Ricardo Reis for his guidance and encouragement throughout this project. We thank the Centre for Macroeconomics (CFM) for providing access to the LSE Fabian high-performance computing cluster. We also thank Andrea Alati, Isaac Baley, Andres Blanco, Marco Bonomo, Adrien Bussy, Patrick Coen, Thomas Drechsel, EAGLS, Mike Elsby, Friedrich Geiecke, Elena Manresa (discussant), Ben Moll, Kevin Sheedy, and Javier Turen for insightful comments. We thank seminar participants at the Federal Reserve Bank of St. Louis, Insper, LSE, PUC-Rio, the University of Minnesota, the University of Sao Paulo, the Virtual Israel Macro Meeting, and the Wharton School, as well as conference participants at the Barcelona GSE Summer Forum, the "Inflation: Causes, Consequences, and Policy Implications" Workshop in Milan, the NBER Summer Institute, "New Challenges in Monetary Economics" Conference in Mannheim, SED Barcelona, and the 7th Annual Women in Macroeconomics Conference. This paper supersedes earlier work circulated under the title "Frictionless Inflation".

†Insper and Centre for Macroeconomics. Email: mbandeira@insper.edu.br

‡Duke University and CEPR. Email: l.castillo-martinez@duke.edu

§University of Minnesota. Email: wan02274@umn.edu

1 Introduction

The surge and retreat of post-Covid inflation has renewed interest in the non-linear dynamics of price adjustment. Menu cost models are the natural framework for thinking about them: firms tolerate some price misalignment due to adjustment costs, and how many of them adjust in response to a given shock is endogenous. The literature has mostly focused on one dimension of this non-linearity, shock size, showing that large shocks generate disproportionate price responses. But these models feature a second dimension that has received less attention, the state of the economy when the shock arrives. This is captured by the cross-sectional distribution of price gaps: the misalignment between firms' actual and reset prices. We study how this distribution shapes inflation dynamics.

Our starting point is a random menu cost model, in which we characterize the price gap distribution as a source of state dependence shaping shock propagation, optimal policy, and future inflation. Its empirical counterpart is, however, unobserved; constructing it is the main contribution of this paper. To do so, we recast the pricing dynamics in state-space form and develop Bayesian methods to estimate latent reset prices from observed price trajectories. Applying this methodology to UK consumer price microdata covering 1996 to 2026, we document that the gap distribution varies substantially over time, with the economy spending extended periods far from its ergodic benchmark.

We exploit this variation to test the model's state-dependence predictions. When shocks arrive, the distribution determines how hard they hit: monetary policy pass-through varies substantially with the pre-shock gap distribution. At the two-year horizon, the response to a tightening ranges from almost fully dampened to nearly 70 percent amplified relative to a benchmark distribution. We find the same qualitative state dependence in the pass-through of oil-price shocks. Absent new shocks, the price gap distribution contains information about where inflation is headed: gap moments predict future inflation beyond current inflation and other macroeconomic variables, and improve out-of-sample forecasts relative to a random-walk benchmark, with gains reaching 37 percent at two years.

Section 2 develops the framework. We lay out a canonical random menu cost model à la Nakamura and Steinsson (2010) in which firms face a fixed cost to change their price, which collapses to zero with some probability. When a firm adjusts, it resets to a target level that evolves as a random walk with drift; between adjustments, the widening gap between the actual price and this target generates a profit loss. The optimal pricing policy takes the form of an *Ss* rule: the firm leaves its price unchanged as long as the gap remains inside an inaction region, and resets whenever the gap hits a threshold or a free adjustment opportunity arrives. Because the policy parameters are common across firms, the model aggregates cleanly: observed inflation reflects both changes in reset prices and shifts in the cross-sectional distribution of gaps, G_t, which becomes the key state variable.

This distribution captures the full state of price misalignment in the economy at each point in time. Because firms adjust intermittently, G_t encodes information about past shocks that have not yet fully passed through to prices, and its moments have direct macroeconomic content. The mean reflects accumulated inflationary or deflationary pressure not yet released through price changes. The variance captures the welfare cost associated with dispersion in price misalignment across firms. Higher-order moments carry additional information: the distribution may be asymmetric, with more firms on the verge of raising rather than lowering prices, or heavy-tailed, with an unusual concentration of firms near their adjustment thresholds.

We show that the same economy can respond very differently to the same aggregate shock depending on the state of G_t when it hits. If many firms have accumulated large gaps near their adjustment thresholds, even a small shock can generate a wave of price changes. Conversely, if most firms have recently adjusted and their gaps are clustered near the reset point, even a large shock may trigger few adjustments. We illustrate these effects numerically using simulations from the calibrated model. This state dependence can also matter for the design of monetary policy. In a simple exercise, we show that the welfare-maximizing response to the same aggregate shock varies with the initial gap distribution.

Beyond its role in shock propagation, G_t also contains information about the future path of inflation. Under infrequent adjustment, the density across the inaction region acts as a schedule of upcoming price changes: firms near the lower threshold will soon reprice upward, while those deeper inside the inaction region will follow in later periods. The gradual release of this accumulated pricing pressure is not summarized by past inflation, because many distinct gap distributions can generate the same inflation record. We formalize this intuition in a deterministic special case of the model, showing that this predictive information cannot be recovered from a finite history of inflation. The magnitude of this predictive gain, however, is an empirical question that requires measuring the gap distribution from the data.

Price gaps are not directly observable, which brings us to the methodological contribution of the paper. Section 3 recasts the pricing dynamics implied by the random menu cost model in state-space form. Given a panel of observed price trajectories, each price serves as a measurement variable, and the state vector for each trajectory includes its reset price and an indicator for the arrival of a free adjustment opportunity. The *Ss* rule serves as the measurement equation, mapping past observed prices and current latent states into current observed prices. The model's assumptions about the evolution of reset prices and adjustment opportunities determine the transition equations, and an initial distribution over the latent states completes the specification. The goal is to recover the sequence of latent states underlying the observed prices and the model parameters that govern them.

To recover the latent states conditional on a given set of parameters, we solve the forward-filtering backward-sampling (FFBS) recursions implied by the state-space representation. Although the *Ss* rule introduces kinks that rule out a closed-form solution, we exploit the structure of the model to simplify the problem. First, the filtering and smoothing distributions factorize across price trajectories, allowing us to recover latent states separately for each trajectory. Second, the *Ss* rule tightly restricts the admissible latent states within each trajectory. When a firm's price remains unchanged, we know with certainty that no free adjustment opportunity arrived. When prices do change, the new price reveals the reset price exactly. At the same time, the size of the adjustment is informative about what triggered it. If the size of the observed change is smaller than the distance to the corresponding bound, the firm must have received a free adjustment opportunity. If it is large enough that the gap moved outside the inaction region, the firm may have paid the menu cost, though the trigger itself remains unknown. These features mean that for many observations the latent states are either directly observed or restricted to a narrow range. We take advantage of this structure using grid-based methods that concentrate computational effort on the subset of states where inference remains nontrivial.

To jointly estimate the model parameters and latent states, we embed the FFBS algorithm in a Gibbs sampler. The parameters of the reset price process and the free-adjustment probability admit conjugate full conditional distributions and can be updated directly. The latent states and inaction thresholds, by contrast, are tightly linked by the *Ss* rule: changing the thresholds changes which state trajectories are admissible, and vice versa. We therefore update them jointly through a Metropolis-Hastings step, proposing candidate states from the grid-based FFBS approximation and candidate thresholds from truncated normal random walks. A Monte Carlo experiment confirms that the algorithm recovers both the model parameters and the moments of the price-gap distribution with negligible bias and high precision, even in short panels.

Section 4 applies the method to the universe of locally-collected price quotes underlying the UK Consumer Price Index (CPI), covering almost 35 million price observations across 1,337 distinct items. We estimate the model separately by item, allowing parameters to vary across items and over time. The estimated price gaps allow us to construct, for the first time, a time-varying empirical distribution of price gaps for a major economy. The distribution varies substantially over time in ways that are not captured by the variation in inflation itself. Year-on-year inflation ranges from -0.2 to 11.1 percent over the sample, while the median gap ranges from 0 to 8.5 percentage points, the interquartile range from 2 to 30 percentage points, and the 5th-to-95th percentile range from 33 to 86 percentage points.

To illustrate this variation, we compare the gap distribution one year before three extreme inflation outcomes in our sample with a March 2024 benchmark, when the distribution was close to its long-run configuration. September 2007 and April 2014 are particularly revealing: headline inflation was nearly identical, at 1.7 and 1.8 percent, but the gap distributions were mirror images. In September 2007, the distribution had excess mass at negative gaps, with more firms facing pending price increases. Over the following year, the price level rose by 5.1 percent and inflation reached its pre-pandemic maximum. In April 2014, the distribution had excess mass at positive gaps, with more firms facing pending price cuts. Over the following year, prices were essentially flat and inflation reached its sample minimum of -0.2 percent. October 2021 was different: the distribution had excess mass in both tails, signaling substantial pending adjustment in both directions. The subsequent surge in energy prices pushed adjustment to the upside, and the price level rose by 11.1 percent over the following year, when inflation reached its sample maximum. The case studies suggest that the gap distribution contains information about pending price adjustment, but isolating its role requires separating the state of the distribution from the shocks that subsequently arrive.

To separate the two, we use state-dependent local projections and high-frequency identification of aggregate shocks in Section 5. In particular, we interact the shock with the mean, standard deviation, skewness, and kurtosis of the pre-shock gap distribution. Our first shock is monetary policy, identified using the high-frequency surprises of Braun et al. (2025). We find that a higher mean gap amplifies the response to a tightening and dampens the response to an easing, while greater dispersion amplifies the response in either direction. Higher skewness tends to dampen the response to a tightening and amplify the response to an easing, while kurtosis has no systematic effect.

The quantitative differences are large. At our three case-study dates, the response to a tightening is, on average over the first two years, dampened by about 55 percent in September 2007, amplified by about 70 percent in April 2014, and amplified by about 23 percent in October 2021 relative to the March 2024 benchmark. Looking across the full sample makes the extent of the variation even clearer. At the two-year horizon, the pass-through of a tightening ranges from almost fully dampened around the global financial crisis to amplified by 69 percent in early 2018. For an easing, it ranges from 85 percent dampened to 178 percent amplified. The direction of the asymmetry also changes over time: before roughly 2012, easings tend to be amplified and tightenings dampened; over much of the following decade, the pattern reverses.

We find the same qualitative state dependence for oil-price shocks. Pass-through is stronger when the shock moves prices in the same direction as the imbalance already present in the gap distribution, and weaker when it pushes against it; the same broad evolution in pass-through asymmetry also appears over the full sample.

Finally, we show that the relationship between the gap distribution and subsequent inflation visible in the case studies holds systematically over the full sample. Gap moments predict future inflation beyond current inflation and a standard set of macroeconomic controls, with predictive content extending from six months to two years. More importantly, the relationship also survives out of sample. Using only information available at the time of the forecast, adding the gap moments lowers the root mean squared prediction error relative to a random-walk benchmark by 8 percent at twelve months and 37 percent at two years, with statistically significant gains at longer horizons. The gains are present in both low- and high-inflation periods.

Relation to the literature. This paper builds on the tradition of menu cost models (Barro, 1972). We bring state-space methods to this literature. This has two main advantages. First, we can recover the key state variable in these models, the time-varying distribution of price gaps, by estimating latent reset prices from observed price paths alone.[1,2] Second, we can estimate the parameters governing price dynamics directly from the likelihood, rather than by matching a selected set of moments. This avoids having to choose which moments to target, a decision that can materially affect parameter estimates under model misspecification, as shown recently by Andrews and Sanders (2026) in a menu cost application.

Within the menu cost literature, a growing body of work studies non-linear inflation dynamics. Early work stressed the role of shock size, showing that large shocks generate disproportionately large price-level responses (Golosov and Lucas, 2007), with first-order implications for optimal stabilization policy (Karadi et al., 2025). In standard models, however, this channel is quantitatively limited once the prevalence of small price changes is taken into account (Blanco et al., 2024a,b). We study a different source of non-linearity that goes back to the generalized *Ss* framework of Caballero and Engel (1993, 2007): the state of the gap distribution when the shock arrives. By recovering the gap distribution over time, we can study empirically how the initial distribution shapes the transmission of a given shock. This source of non-linearity is complementary to shock size, and the two can reinforce each other.

We also relate to a literature studying state- and sign-dependent effects of monetary policy on inflation. State dependence has been documented along several dimensions, including macroeconomic conditions (Tenreyro and Thwaites, 2016), volatility (Vavra, 2014), financial conditions (Jordà et al., 2020), inflation regimes (Ascari and Haber, 2022), and disagreement about inflation expectations (Falck et al., 2021). A separate literature examines asymmetric effects of monetary tightening and easing (Cover, 1992; Barnichon and Matthes, 2016; Debortoli et al., 2025). We connect these strands by showing that the pass-through of monetary policy to inflation is both state and sign dependent, with both dimensions shaped by the distribution of price gaps.

Finally, we relate to the broad literature on inflation forecasting, surveyed by Faust and

[1] In other *Ss* applications, such as investment and durable goods, gaps between actual and target levels can be constructed from observable fundamentals or proxied by the age of the capital stock (Caballero et al., 1995; Bachmann et al., 2013; Berger and Vavra, 2015). In pricing, by contrast, measuring gaps requires knowledge of firms' optimal prices, which depend on marginal costs that are difficult to observe at the firm level. A recent exception is Gagliardone et al. (2025), who use firm-level cost data and competitors' prices to construct a measure of price gaps and study non-linearities in cost-shock pass-through.

[2] A complementary sufficient-statistics approach circumvents the need to recover the full gap distribution by linking aggregate responses to one-off shocks to steady-state moments of microeconomic adjustment (Alvarez et al., 2016; Baley and Blanco, 2021). In the context of pricing, these predictions have been tested empirically using cross-sectional variation in micro pricing moments (Hong et al., 2023; Alvarez et al., 2025). Our objective is different: we recover the gap distribution as it evolves over time and study how an arbitrary inherited distribution shapes the transmission of subsequent shocks. Accordingly, our empirical exercises exploit time variation in the gap distribution within products, rather than cross-sectional variation in steady-state moments.

Wright (2013). Within this literature, a smaller body of work has begun to exploit micro price data, both for inflation nowcasting (Powell et al., 2018; Beck et al., 2024) and for forecasting future inflation using the cross-sectional distribution of realized price changes (Chen et al., 2026). We contribute to this literature by using moments of the cross-sectional distribution of price gaps as predictors of future inflation.

2 Price gaps: concept and relevance

This section introduces the distribution of price gaps: what it is and why it matters. We begin by laying out the framework we use throughout, a canonical menu cost model with random free adjustment in discrete time (Nakamura and Steinsson, 2010; Auclert et al., 2024).[3] After defining price gaps and their cross-sectional distribution, we turn to why the gap distribution is important, highlighting its role in the non-linear transmission of aggregate shocks and as a source of predictive content for future inflation.

2.1 A simple model of price setting

A firm i produces a differentiated good and sets a log price $p_{i,t}$ in each period $t = 0, 1, 2, \ldots$. Its *reset price*, denoted $p^r_{i,t}$, is the price the firm would choose upon adjustment and evolves according to a random walk with drift:

$$p^r_{i,t} = \mu_t + p^r_{i,t-1} + \varepsilon_{i,t}, \tag{1}$$

where μ_t is the common drift in reset prices and $\varepsilon_{i,t} \sim \mathcal{N}(0, \sigma^2_{\varepsilon,t})$ is an idiosyncratic shock, independent across firms and over time. Changes in μ_t capture the combined effect of aggregate fundamental shocks and the monetary policy response to them. Changes in $\sigma_{\varepsilon,t}$ capture shifts in firm-level volatility. We model shocks to both as unexpected.

Each period, the firm incurs a loss from deviating from its reset price. Adjusting the price requires paying a fixed cost $c_{i,t}$, independently and identically drawn from $\{0, \bar{c}\}$: with probability $\lambda_t \in [0, 1]$, the firm receives a free adjustment ($c_{i,t} = 0$); with probability $1 - \lambda_t$, it must pay $\bar{c} > 0$. The free adjustment probability is potentially subject to unexpected shocks as well, capturing shifts in the technology of price setting.

The firm's problem can be written in terms of the *price gap* $s_{i,t} \equiv p_{i,t} - p^r_{i,t}$. The firm chooses a sequence $\{s_{i,t}\}_{t=0}^{\infty}$ to minimize expected discounted losses:

$$\min_{\{s_{i,t}\}_{t=0}^{\infty}} \mathbf{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\frac{1}{2} s^2_{i,t} + c_{i,t} \cdot \mathbf{1}_{\{s_{i,t} \neq s_{i,t-1} - \mu_t - \varepsilon_{i,t}\}} \right], \tag{2}$$

where the indicator identifies a price adjustment and $c_{i,t}$ is the realized adjustment cost.[4]

[3] Appendix A.1 provides microfoundations for this framework.

[4] Equation (2) is a reduced-form representation of the firm's pricing problem; Appendix A.1 derives it from a standard CES demand structure with random menu costs. This formulation is standard in the analytical

The optimal pricing policy takes the form of a standard Ss rule:[5]

$$s_{i,t} = \begin{cases} 0 & \text{with prob } \lambda_t \text{ or if } s_{i,t-1} - \mu_t - \varepsilon_{i,t} \notin [\underline{s}_t, \bar{s}_t] \\ s_{i,t-1} - \mu_t - \varepsilon_{i,t} & \text{otherwise} \end{cases} \tag{3}$$

The firm adjusts its price in one of two cases: either it receives a free adjustment opportunity, or the price gap falls outside an inaction region $[\underline{s}_t, \bar{s}_t]$, with $\underline{s}_t < 0 < \bar{s}_t$. Upon adjustment, the firm closes its gap entirely, setting its price to the current reset level. Otherwise, the gap evolves passively, drifting with the common reset price drift and the idiosyncratic shock. In general, the reduced-form parameters $[\underline{s}_t, \bar{s}_t]$ vary over time.

2.2 The distribution of price gaps

Let G_t denote the cross-sectional distribution of price gaps across firms at time t, with density g_t over the inaction region $[\underline{s}_t, \bar{s}_t]$ and a mass point at the reset point $s = 0$. This distribution is the key state variable of the model, summarizing the cross-sectional state of price misalignment. In steady state, the drift of non-adjusting gaps and the resetting of adjusting firms balance to produce a stationary distribution with mass concentrated near zero and tails extending toward the inaction thresholds.

While G_t is an infinite-dimensional object, its key features can be read from a small number of moments. The *mean gap*, $\mathbb{E}[s_t] \equiv \int s\, dG_t$, captures the average direction of price misalignment: when negative, firms' prices sit on average below their reset levels, reflecting accumulated inflationary pressure. The *variance*, $\mathbb{V}\text{ar}(s_t)$, measures how dispersed the misalignment is across firms. Under the quadratic loss in (2), the average reduced-form loss from price misalignment is

$$\frac{1}{2} E[s_t^2] = \frac{1}{2} \left[\mathbb{V}\text{ar}(s_t) + \mathbb{E}[s_t]^2 \right]. \tag{4}$$

so that dispersion has direct consequences even when the mean gap is zero. As shown in Appendix A.3, this expression closely approximates the corresponding microfounded welfare loss. The *skewness* captures asymmetry in the distribution. More mass in the lower tail points to latent inflationary pressure, while more mass in the upper tail points to deflationary pressure. The *kurtosis* captures how much weight is placed in the tails, including near the adjustment thresholds.

menu cost literature, e.g., Alvarez et al. (2011), Alvarez et al. (2016) and Auclert et al. (2024).

[5] Auclert et al. (2024) establish the Ss policy analytically at zero drift in the discrete-time model. With positive drift, symmetry is lost and an analogous analytical result is not available. Appendix A.2 shows that the Ss structure continues to hold whenever the value function is strictly quasiconvex; we verify this condition numerically for the calibrations used in our structural exercises. The Ss structure under positive drift is also well established in continuous-time models.

From price gaps to inflation. Since $p_{i,t} = p^r_{i,t} + s_{i,t}$, each firm's inflation rate decomposes as

$$\pi_{i,t} = \pi^r_{i,t} + \Delta s_{i,t}, \tag{5}$$

where $\pi^r_{i,t} \equiv p^r_{i,t} - p^r_{i,t-1} = \mu_t + \varepsilon_{i,t}$ is the firm's *reset inflation* and $\Delta s_{i,t} \equiv s_{i,t} - s_{i,t-1}$ is the wedge introduced by adjustment frictions. Averaging across firms, and using the first-order aggregation of the CES price index derived in Appendix A.1, aggregate inflation satisfies

$$\pi_t = \mu_t + \Delta\mathbb{E}[s_t], \tag{6}$$

where μ_t is aggregate reset inflation and $\Delta\mathbb{E}[s_t]$ is the change in the mean gap.

Equation (6) shows why headline inflation is an incomplete summary of the pricing state. Two economies can have the same π_t but very different gap distributions. In one, inflation may reflect many small price adjustments. In another, it may be driven by a few large adjustments while many prices remain far from their reset levels. The gap distribution thus reveals information about the state of the economy that headline inflation cannot.

2.3 Why the gap distribution matters

We now turn to how we can use the information on the state of the economy that the gap distribution reveals. We show that it determines the non-linear transmission of aggregate shocks to inflation, and that it carries predictive content for future inflation beyond what the inflation history alone can provide.

2.3.1 Non-linearity and the role of initial conditions

In menu cost models, the inflation response to aggregate shocks is inherently non-linear. One well-understood dimension of this non-linearity is shock size: large shocks push more firms past their adjustment thresholds, so that the per-unit inflation response rises with the magnitude of the shock (Golosov and Lucas, 2007). We highlight a complementary dimension that has received less attention: the shape of the gap distribution at the time the shock arrives.

To illustrate both channels, we calibrate the microfounded model in Appendix A.1 following the menu cost calibration of Nakamura and Steinsson (2010): a monthly frequency of price change of 8.7%, a median absolute price change of 8.5%, and a 75% share of free adjustments. Parameter values are reported in Appendix Table A.1. We then subject the economy to an unexpected, one-off level shock to the drift, $\mu_1 = \mu + \Delta\mu$ with $\mu_t = \mu$ for $t \geq 2$, and ask how the impact excess inflation $\pi_1 - \mu$ varies with the size of the shock and the state of G_0 when it hits.

The left panel of Figure 1 isolates the shock-size channel. Starting from the ergodic distribution, we vary the magnitude of the shock and plot the resulting impact excess inflation.

The dashed line shows a linearized benchmark, extrapolated from the slope at small shocks. The relationship is convex: the impact response grows more than proportionally with shock size, reflecting the increasing share of firms pushed past their adjustment thresholds. For small shocks, however, the relationship is approximately linear since the model and the linearized benchmark nearly coincide. This is a well-known result in the literature.

The right panel turns to the less-studied dimension. We fix the shock at 6 pp annualized, still in the approximately linear region of the left panel, and vary only the initial gap distribution G_0. To generate economically meaningful variation in G_0, rather than constructing distributions arbitrarily, we subject the ergodic distribution to a range of large persistent perturbations to the drift and record snapshots along each transition path back to steady state. Each snapshot provides a different initial condition for the shock. These distributions vary in location, dispersion, and tail behavior, while the structural parameters of the model are held fixed.

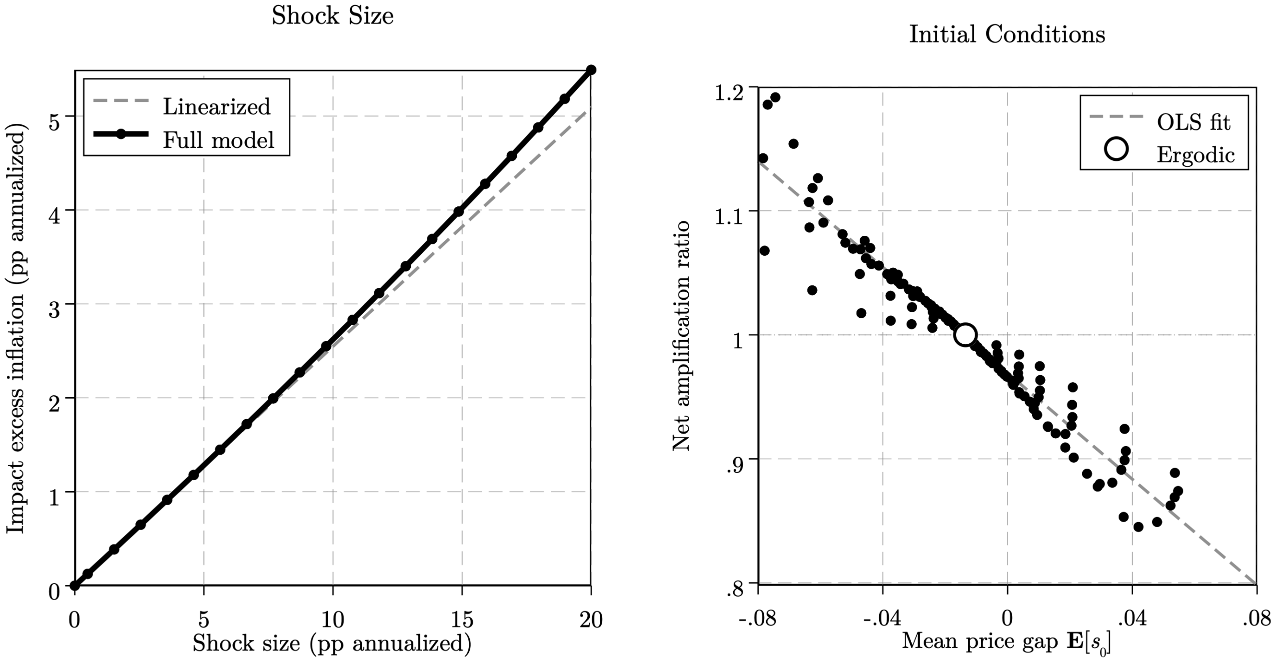
To isolate the genuine interaction between the shock and the initial conditions, we construct the *net amplification ratio*. For each displaced G_0, we compute the impact excess inflation both with and without the shock, subtract the counterfactual to remove the mechanical unwinding of pent-up pressure that G_0 would have generated on its own, and normalize by the ergodic response. The ergodic point equals one by construction.

To summarize each distribution by a single number, we regress the net amplification ratio on each of the first four moments of G_0 separately. The mean gap $E[s_0]$ provides the best univariate fit ($R^2 = 0.93$). The right panel of Figure 1 plots the net amplification ratio against this statistic. The relationship is monotone: a more negative mean gap amplifies the shock (ratio above one), while a more positive mean gap dampens it (ratio below one). Intuitively, a negative mean gap places more firms near the lower adjustment threshold, which is where a positive drift shock triggers repricing. The magnitudes are economically meaningful: with an estimated slope of -2.1, a mean gap 5 percentage points below its ergodic level raises the amplification ratio to about 1.11, an 11% increase in the inflation response. This variation arises even though the shock itself is small enough to generate no non-linearity in the size dimension. With larger shocks, the two channels reinforce each other and the amplification increases.

The displaced distributions in the right panel are generated by past perturbations to the drift μ. These primarily shift the mean of G_0 while leaving its shape relatively unchanged. If instead the displacement comes from past changes in idiosyncratic volatility σ_ε, the mean gap stays near its ergodic level but the variance changes substantially. In that case, the variance of G_0 becomes the relevant summary statistic: higher variance places more mass near the adjustment thresholds, increasing the economy's sensitivity to the shock. Appendix Figure A.1 reports this exercise.

Implications for monetary policy. The state dependence documented above has implications for the design of monetary policy. We illustrate this with a simple policy exercise. We

Figure 1: Non-linearity in the inflation response



Notes: Both panels use the calibration in Appendix Table A.1 and report the impact response to an unexpected one-off level shock to the drift. The left panel varies the size of the shock, starting from the ergodic distribution, and plots the impact excess inflation $\pi_1 - \mu$; the dashed line extrapolates linearly from the slope at small shocks. The right panel holds the shock fixed at 6 pp annualized and varies the initial gap distribution G_0, indexed by the mean price gap $E[s_0]$; it plots the net amplification ratio, which subtracts the no-shock counterfactual and normalizes by the ergodic response. The dashed line is the OLS fit and the star marks the ergodic distribution.

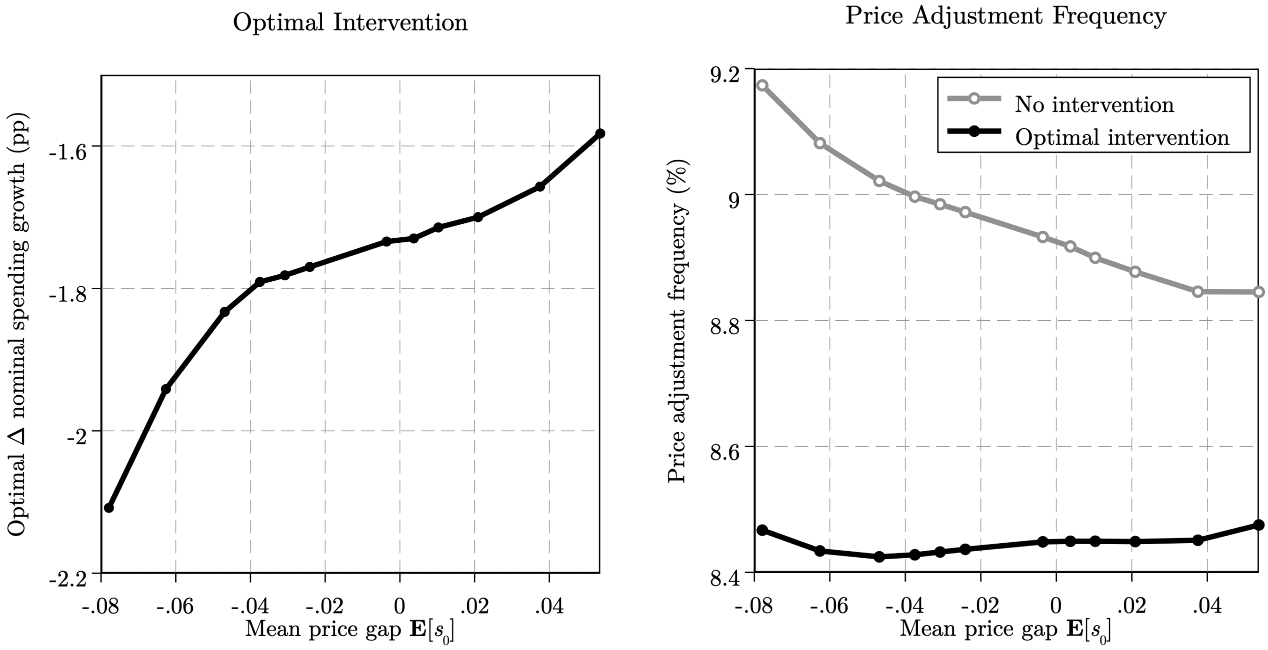
consider a one-off 6 pp annualized increase in nominal spending growth, which in the absence of a subsequent policy intervention maps almost one-for-one into the reset price-drift shock considered above.[6] We assume that monetary policy is predetermined in the period of the shock, $t = 1$, and therefore cannot respond on impact. Once the shock is observed, the central bank announces an intervention for $t = 2$, which firms anticipate when making their $t = 1$ pricing decisions. To keep the exercise simple, we restrict the intervention to $t = 2$, with policy returning to baseline thereafter.

We select twelve initial distributions G_0 from the right panel of Figure 1 and, for each one, report the one-period intervention that maximizes household welfare.[7] The left panel of Figure 2 plots the optimal intervention against the mean initial price gap, $\mathbb{E}[s_0]$. Across the distributions, the optimal change in nominal spending growth ranges from approximately -1.6 to -2.1 percentage points per month, a spread of about 0.5 percentage points despite the identical aggregate shock. This cross-state variation is roughly as large as the monthly

[6]In the microfoundation, nominal spending growth determines desired price drift, ν_t, while reset price drift satisfies $\mu_t = \nu_t + \Delta x_t^*$, where Δx_t^* is the change in the reset point as defined in Appendix A.1. Absent an anticipated intervention, the change in the reset point is negligible, so the two shocks coincide almost exactly. With an anticipated intervention, the reset point responds already on impact and reset price drift therefore adjusts endogenously.

[7]Appendix A.4 describes the welfare criterion for this exercise. The optimal full policy path is left for future research.

Figure 2: Monetary policy and initial conditions



Notes: Both panels use the same calibration, a 6 pp annualized increase in nominal spending growth, and twelve initial distributions from the right panel of Figure 1, indexed by the mean price gap $\mathbb{E}[s_0]$. Policy chooses a one-period intervention in nominal spending growth in $t = 2$. The left panel reports the welfare-maximizing intervention in monthly percentage points. The right panel reports the $t = 2$ price adjustment frequency with and without the optimal intervention.

equivalent of the shock itself. The relationship is monotone: a more negative mean price gap calls for a stronger contraction in nominal spending growth.

The right panel illustrates the mechanism. In the absence of intervention, a more negative mean gap is associated with a higher price adjustment frequency in $t = 2$, reflecting greater pricing pressure inherited from the initial distribution. The optimal intervention is more contractionary precisely in these states and leaves the adjustment frequency nearly constant across initial conditions. Thus, even conditional on the size of the aggregate disturbance, the optimal monetary stance depends on the cross-section of price gaps in which the shock lands.[8]

2.3.2 Predictive content of the gap distribution

The analysis above is conditional on an aggregate shock, but the gap distribution also contains information about future inflation more generally. In menu cost models, price adjustment generates an endogenous component of inflation persistence. Whenever the gap distribution departs from its ergodic shape, whether due to past shocks, policy changes, or other features of the economy's history, it contains information about price adjustments that have yet to occur. In particular, directional imbalances embed inflationary or deflationary

[8] Karadi et al. (2025) show that endogenous price adjustment makes the optimal policy response depend on shock size. Our exercise is complementary: we hold the shock fixed and vary the inherited gap distribution.

pressure that is released gradually as firms adjust. The current shape of G_t can therefore contain a predictable component of future inflation even in the absence of additional aggregate shocks.

A natural question is whether this information is already summarized by observed inflation. From equation (6), aggregate inflation is $\pi_t = \mu_t + \Delta E[s_t]$, a scalar outcome of the much richer cross-sectional state G_t. Observed inflation reflects the aggregate effect of price adjustments realized at a given date, but does not reveal where firms that have not adjusted are located inside the inaction region. More generally, a finite history of inflation contains information about recent price adjustments without necessarily identifying the positions of firms that have remained inactive throughout. Two economies can therefore display the same recent inflation history while embedding different amounts and timing of future pricing pressure.

We formalize this intuition in a deterministic special case of the model. Consider the pure menu cost case, $\lambda = 0$, with negligible idiosyncratic volatility, $\sigma_\varepsilon \to 0$, constant positive drift $\mu > 0$, and fixed adjustment thresholds $\underline{s} < 0 < \bar{s}$. In the absence of adjustment, a firm's gap evolves deterministically according to $s_{i,t+1} = s_{i,t} - \mu$. A firm adjusts when its gap crosses the lower threshold and then resets its gap to zero.

Proposition 1 *Consider the deterministic pure menu cost model.*

(i) Suppose G_t has a smooth density g_t on $[\underline{s}, 0)$. Then, for horizons $1 \leq h \leq |\underline{s}|/\mu$,

$$\pi_{t+h} = |\underline{s}|\, \mu\, g_t(\underline{s} + (h-1)\mu) + O(\mu^2). \tag{7}$$

(ii) The history of aggregate inflation does not recover this information: for any window K shorter than the transit time, $(\bar{s} - \underline{s})/\mu$, across the inaction region, there exist two economies with gap distributions $G_t \neq G_t'$ and identical inflation histories

$$\pi_{t-k} = \pi'_{t-k}, \qquad k = 0, \ldots, K-1,$$

yet different future inflation at some horizon $h \geq 1$.[9]

The proof is in Appendix A.5. The intuition for part (i) is that under constant drift the lower part of the inaction region is partitioned into bands of width μ, each mapping to a future horizon: firms in the h-th band will cross the lower threshold and reset to zero h periods from now. The density thus acts as a schedule of upcoming price adjustments. Part (ii) follows because recent inflation reveals only the cohorts that have adjusted, not the positions of firms still inside the inaction region. In the general model, idiosyncratic shocks and free adjustment opportunities blur the deterministic mapping from current gaps

[9] The transit time is the maximal number of periods a firm can remain in the inaction region without adjusting. Under the calibration in Appendix Table A.1, this is roughly 140 months.

to future adjustment dates, but the underlying logic remains: the location of mass within the inaction region determines the expected timing and size of future price adjustments. The bound on K in Proposition 1 is specific to the deterministic limit.

Proposition 1 therefore shows why the gap distribution can contain predictive information about future inflation that is not captured by observed inflation alone. Whether this additional information is quantitatively useful in the data is an empirical question. To address it, we first need to recover the latent price gaps from observed prices. We turn to this in the next section.

3 State-space methods for a random menu cost model

This section develops a methodology for estimating the random menu cost model using micro price data. We first cast the pricing dynamics of Section 2 in state-space form and then develop a Bayesian procedure to jointly recover the model parameters and latent states. We assess the finite-sample performance of the resulting estimator in a Monte Carlo experiment.

3.1 Random menu cost model in state-space form

To take the pricing dynamics of Section 2.1 to the data, we cast them in state-space form. Consider a panel of observed prices $p_{i,t}$, indexed by i and t. For estimation, we divide the sample period into a sequence of estimation windows, allowing observations associated with a given i to span multiple consecutive windows. Let $\mathcal{W}$ denote the collection of observations associated with a generic estimation window and $\mathcal{I}_{\mathcal{W}}$ the set of indices i represented in that window. For each $i \in \mathcal{I}_{\mathcal{W}}$, let

$$\mathcal{T}_{i,\mathcal{W}} = \{\underline{t}_{i,\mathcal{W}}, \ldots, \bar{t}_{i,\mathcal{W}}\}$$

denote the contiguous sequence of dates for i that fall within $\mathcal{W}$. We can write this collection as

$$\mathcal{W} = \{(i,t) : i \in \mathcal{I}_{\mathcal{W}},\, t \in \mathcal{T}_{i,\mathcal{W}}\},$$

with corresponding, potentially unbalanced, panel of observed prices

$$\mathbf{p}_{\mathcal{W}} = \{p_{i,t} : (i,t) \in \mathcal{W}\}.$$

State-space representation. Given the observed panel $\mathbf{p}_{\mathcal{W}}$, for each $(i,t) \in \mathcal{W}$ let $\mathbf{x}_{i,t} \equiv [p^r_{i,t}, \ell_{i,t}]$ denote the latent state, where $p^r_{i,t}$ is the reset price and $\ell_{i,t}$ indicates the arrival of a free adjustment opportunity. The corresponding collection of latent states is $\mathbf{x}_{\mathcal{W}} = \{\mathbf{x}_{i,t} : (i,t) \in \mathcal{W}\}$. The parameters of the state-space representation are collected in

$$\boldsymbol{\theta}_{\mathcal{W}} = \{\mu_{\mathcal{W}}, \sigma_{\varepsilon,\mathcal{W}}, \lambda_{\mathcal{W}}, \underline{s}_{\mathcal{W}}, \bar{s}_{\mathcal{W}}\},$$

and are assumed to be constant within the estimation window $\mathcal{W}$.

The observed prices evolve according to

$$p_{i,t} = p_{i,t-1} d(\mathbf{x}_{i,t}, p_{i,t-1}, \boldsymbol{\theta}_{\mathcal{W}}) + p^r_{i,t}[1 - d(\mathbf{x}_{i,t}, p_{i,t-1}, \boldsymbol{\theta}_{\mathcal{W}})], \tag{8}$$

where

$$d(\mathbf{x}_{i,t}, p_{i,t-1}, \boldsymbol{\theta}_{\mathcal{W}}) = \mathbb{1}\{p_{i,t-1} - p^r_{i,t} \in [\underline{s}_{\mathcal{W}}, \bar{s}_{\mathcal{W}}]\}(1 - \ell_{i,t}) + \mathbb{1}\{p_{i,t-1} = p^r_{i,t}\}\ell_{i,t}. \tag{9}$$

The latent states evolve according to

$$p^r_{i,t} = \mu_{\mathcal{W}} + p^r_{i,t-1} + \varepsilon_{i,t}, \qquad \varepsilon_{i,t} \sim \mathcal{N}(0, \sigma^2_{\varepsilon,\mathcal{W}}), \tag{10}$$

$$\ell_{i,t} = \mathbb{1}\{\nu_{i,t} \leqslant \lambda_{\mathcal{W}}\}, \qquad \nu_{i,t} \sim \mathcal{U}(0,1), \tag{11}$$

where the innovations $\varepsilon_{i,t}$ and $\nu_{i,t}$ are independent across i and over time. Finally, initial latent states are independent across i conditional on the corresponding initial observed prices and $\boldsymbol{\theta}_{\mathcal{W}}$:

$$f(\{\mathbf{x}_{i,\underline{t}_{i,\mathcal{W}}}\}_{i\in\mathcal{I}_{\mathcal{W}}} \mid \{p_{i,\underline{t}_{i,\mathcal{W}}}\}_{i\in\mathcal{I}_{\mathcal{W}}}, \boldsymbol{\theta}_{\mathcal{W}}) = \prod_{i\in\mathcal{I}_{\mathcal{W}}} f(\mathbf{x}_{i,\underline{t}_{i,\mathcal{W}}} \mid p_{i,\underline{t}_{i,\mathcal{W}}}, \boldsymbol{\theta}_{\mathcal{W}}). \tag{12}$$

This state-space representation is nonlinear and non-Gaussian. The measurement equations (8) and (9) impose the two-sided *Ss* rule: the observed price remains unchanged when no free adjustment opportunity arrives and the implied price gap lies within the inaction region, while upon adjustment it is set equal to the reset price.[10] The transition equations (10) and (11) govern the evolution of the latent reset price and the arrival of free adjustment opportunities, respectively. Equation (12) completes the representation by specifying the distribution of the initial latent states.

The pricing model in Section 2.1 provides one microfoundation for this representation. In that model, the transition equations correspond directly to the laws of motion for the reset price and free adjustment opportunities, while the measurement equation implements the optimal two-sided *Ss* pricing rule. The state-space representation, however, need not be tied to this particular microfoundation. Alternative pricing models that imply the same adjustment rule and latent-state dynamics admit the same empirical representation. Ultimately, what matters for estimation is the state-space structure itself rather than the particular model that rationalizes it.

Relative to the pricing model in Section 2.1, the empirical representation differs in two

[10]Equation (9) also implies inaction when a free adjustment opportunity arrives but the observed price already coincides with the reset price. Given the continuous distribution of reset price innovations in (10), this is a probability-zero event.

ways. First, while the model allows both states and parameters to vary over time, we treat the parameters as locally constant within each estimation window, while allowing the latent states to evolve continuously over time. Second, rather than imposing the model-specific mapping from deeper primitives to the optimal inaction bounds, we estimate the bounds directly as reduced-form parameters. This allows the data to determine the inaction region without imposing cross-parameter restrictions that are specific to a particular microfoundation.[11]

3.2 Bayesian estimation

We estimate the state-space parameters, $\boldsymbol{\theta}_{\mathcal{W}}$, and latent states, $\mathbf{x}_{\mathcal{W}}$, jointly using a Bayesian procedure. We first describe how to sample the latent states conditional on the parameters and observed data, and then embed this step in a Gibbs sampler for the joint posterior.

3.2.1 Sampling the latent states

To sample the latent states conditional on the parameters and observed prices, we use a forward-filtering backward-sampling (FFBS) algorithm (Carter and Kohn, 1994; Frühwirth-Schnatter, 1994). In our setting, this requires addressing two challenges. First, the estimation window may contain a large number of price trajectories, making the joint state space high dimensional. Second, the state-space representation is nonlinear and non-Gaussian, ruling out standard Gaussian filtering and smoothing methods.

The large cross-sectional dimension can be handled by exploiting the conditional independence built into the state-space representation. Conditional on the parameters, the joint filtering and smoothing problem factorizes across price trajectories, allowing the latent-state trajectories to be sampled separately. The following lemma formalizes this result.

Lemma 1 (FFBS factorization). *For each $i \in \mathcal{I}_{\mathcal{W}}$, let $\mathbf{p}_{i,\mathcal{W}} = \{p_{i,t} : t \in \mathcal{T}_{i,\mathcal{W}}\}$ and $\mathbf{x}_{i,\mathcal{W}} = \{\mathbf{x}_{i,t} : t \in \mathcal{T}_{i,\mathcal{W}}\}$ denote the observed-price and latent-state trajectories, respectively. Let $\mathcal{I}_{\mathcal{W},t} = \{i \in \mathcal{I}_{\mathcal{W}} : t \in \mathcal{T}_{i,\mathcal{W}}\}$ denote the price trajectories observed at date t, and $\mathcal{I}^{-}_{\mathcal{W},t} = \{i \in \mathcal{I}_{\mathcal{W},t} : t > \underline{t}_{i,\mathcal{W}}\}$ those also observed at $t-1$. Under the state-space representation in equations* (8)–(12)*, the following factorizations and recursions hold.*

Forward filtering. *The filtered density factorizes as*

$$f(\{\mathbf{x}_{i,t}\}_{i\in\mathcal{I}_{\mathcal{W},t}} \mid \{\mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}\}_{i\in\mathcal{I}_{\mathcal{W},t}}, \boldsymbol{\theta}_{\mathcal{W}}) = \prod_{i\in\mathcal{I}_{\mathcal{W},t}} f(\mathbf{x}_{i,t} \mid \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}, \boldsymbol{\theta}_{\mathcal{W}}), \tag{13}$$

[11] The distinction is similar in spirit to that between DSGE and structural VAR approaches: the latter impose a more limited set of economic restrictions tailored to the empirical objects of interest, without requiring estimation of all the deeper primitives of a fully specified structural model.

while the one-step-ahead predictive density for continuing price trajectories satisfies

$$f(\{\mathbf{x}_{i,t}\}_{i\in\mathcal{I}^-_{\mathcal{W},t}} \mid \{\mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t-1}\}_{i\in\mathcal{I}^-_{\mathcal{W},t}}, \boldsymbol{\theta}_{\mathcal{W}}) = \prod_{i\in\mathcal{I}^-_{\mathcal{W},t}} f(\mathbf{x}_{i,t} \mid \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t-1}, \boldsymbol{\theta}_{\mathcal{W}}). \tag{14}$$

The trajectory-level recursions are

$$f(\mathbf{x}_{i,t} \mid \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t-1}, \boldsymbol{\theta}_{\mathcal{W}}) = \int f(\mathbf{x}_{i,t} \mid \mathbf{x}_{i,t-1}, \boldsymbol{\theta}_{\mathcal{W}}) f(\mathbf{x}_{i,t-1} \mid \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t-1}, \boldsymbol{\theta}_{\mathcal{W}})\, d\mathbf{x}_{i,t-1}, \tag{15}$$

$$f(\mathbf{x}_{i,t} \mid \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}, \boldsymbol{\theta}_{\mathcal{W}}) \propto f(p_{i,t} \mid \mathbf{x}_{i,t}, p_{i,t-1}, \boldsymbol{\theta}_{\mathcal{W}}) f(\mathbf{x}_{i,t} \mid \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t-1}, \boldsymbol{\theta}_{\mathcal{W}}). \tag{16}$$

Backward sampling. *The posterior over latent-state trajectories factorizes as*

$$f(\mathbf{x}_{\mathcal{W}} \mid \mathbf{p}_{\mathcal{W}}, \boldsymbol{\theta}_{\mathcal{W}}) = \prod_{i\in\mathcal{I}_{\mathcal{W}}} f(\mathbf{x}_{i,\mathcal{W}} \mid \mathbf{p}_{i,\mathcal{W}}, \boldsymbol{\theta}_{\mathcal{W}}), \tag{17}$$

and each trajectory is sampled according to

$$f(\mathbf{x}_{i,\mathcal{W}} \mid \mathbf{p}_{i,\mathcal{W}}, \boldsymbol{\theta}_{\mathcal{W}}) = f(\mathbf{x}_{i,\bar{t}_{i,\mathcal{W}}} \mid \mathbf{p}_{i,\mathcal{W}}, \boldsymbol{\theta}_{\mathcal{W}}) \prod_{t=\underline{t}_{i,\mathcal{W}}}^{\bar{t}_{i,\mathcal{W}}-1} f(\mathbf{x}_{i,t} \mid \mathbf{x}_{i,t+1}, \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}, \boldsymbol{\theta}_{\mathcal{W}}), \tag{18}$$

$$f(\mathbf{x}_{i,t} \mid \mathbf{x}_{i,t+1}, \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}, \boldsymbol{\theta}_{\mathcal{W}}) \propto f(\mathbf{x}_{i,t+1} \mid \mathbf{x}_{i,t}, \boldsymbol{\theta}_{\mathcal{W}}) f(\mathbf{x}_{i,t} \mid \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}, \boldsymbol{\theta}_{\mathcal{W}}). \tag{19}$$

The proof is in Appendix B.1. The factorization follows from two features of the state-space representation. Under the measurement equations (8) and (9), each observed price depends only on its own lagged price, latent state, and the common parameters. As a result, each observed price trajectory is linked only to its corresponding latent-state trajectory. Under the transition equations (10) and (11), together with the initial conditions in (12), latent states are initialized and evolve independently across price trajectories. As a result, latent-state sampling can proceed price trajectory by price trajectory rather than jointly over the full panel, reducing a high-dimensional problem to a collection of low-dimensional problems that can be processed in parallel.

Lemma 1 reduces the dimensionality of the full-panel FFBS problem by decomposing it into separate problems for each price trajectory. The corresponding recursions, given by (15)–(16) and (18)–(19), remain nonlinear and non-Gaussian and do not admit closed-form solutions, so they must be approximated numerically. We build on grid-based methods for nonlinear, non-Gaussian state-space models (Kitagawa, 1987; Kramer and Sorenson, 1988) to develop an FFBS procedure tailored to a measurement equation induced by a two-sided *Ss* rule. The approach is particularly well suited to our setting because each price trajectory has only a two-dimensional latent state and the measurement equation sharply restricts the states consistent with the observed prices.

Table 1: Prior distributions for parameters and initial states

Parameter	Prior distribution	Hyperparameters
(a) State-space parameters		
$\mu_{\mathcal{W}}$	$\mathcal{N}(m_\mu^0, v_\mu^0)$	$m_\mu^0 = 0, \quad v_\mu^0 = 1$
$1/\sigma^2_{\varepsilon,\mathcal{W}}$	$\mathcal{G}(s_\sigma^0, r_\sigma^0)$	$s_\sigma^0 = r_\sigma^0 = 10^{-3}$
$\lambda_{\mathcal{W}}$	$\mathcal{B}(s^0_{1,\lambda}, s^0_{2,\lambda})$	$s^0_{1,\lambda} = s^0_{2,\lambda} = 1$
$\underline{s}_{\mathcal{W}}$	$\mathcal{U}(\underline{s}^0_{\text{LB},\mathcal{W}}, \underline{s}^0_{\text{UB},\mathcal{W}})$	$\underline{s}^0_{\text{LB},\mathcal{W}} = \min(-\Delta \mathbf{p}^+_{\mathcal{W}}), \quad \underline{s}^0_{\text{UB},\mathcal{W}} = \max(-\Delta \mathbf{p}^+_{\mathcal{W}})$
$\bar{s}_{\mathcal{W}}$	$\mathcal{U}(\bar{s}^0_{\text{LB},\mathcal{W}}, \bar{s}^0_{\text{UB},\mathcal{W}})$	$\bar{s}^0_{\text{LB},\mathcal{W}} = \min(-\Delta \mathbf{p}^-_{\mathcal{W}}), \quad \bar{s}^0_{\text{UB},\mathcal{W}} = \max(-\Delta \mathbf{p}^-_{\mathcal{W}})$
(b) Initial latent states		
$\mathbf{x}_{i,\underline{t}_{i,\mathcal{W}}} \mid p_{i,\underline{t}_{i,\mathcal{W}}}, \boldsymbol{\theta}_{\mathcal{W}}$	$(p_{i,\underline{t}_{i,\mathcal{W}}}, 1)$ w/ prob. $\kappa_0\omega_0$	$\kappa_0 = \text{freq}(\Delta \mathbf{p}_{\mathcal{W}})$
	$(p_{i,\underline{t}_{i,\mathcal{W}}}, 0)$ w/ prob. $\kappa_0(1-\omega_0)$	$\omega_0 = 0.6$
	$(p^r_{i,\underline{t}_{i,\mathcal{W}}}, 0), \quad p^r_{i,\underline{t}_{i,\mathcal{W}}} \sim \mathcal{U}\big(p_{i,\underline{t}_{i,\mathcal{W}}} - \bar{s}_{\mathcal{W}},\, p_{i,\underline{t}_{i,\mathcal{W}}} - \underline{s}_{\mathcal{W}}\big)$ w/ prob. $1-\kappa_0$	

Notes: For the Gamma prior, s_σ^0 and r_σ^0 denote the shape and rate parameters, respectively. $\Delta \mathbf{p}_{\mathcal{W}}$ denotes the set of observed price changes within $\mathcal{W}$, while $\Delta \mathbf{p}^+_{\mathcal{W}}$ and $\Delta \mathbf{p}^-_{\mathcal{W}}$ denote the corresponding subsets of positive and negative price changes. The initial state is $\mathbf{x}_{i,\underline{t}_{i,\mathcal{W}}} = (p^r_{i,\underline{t}_{i,\mathcal{W}}}, \ell_{i,\underline{t}_{i,\mathcal{W}}})$. The parameter κ_0 governs the probability that the initial price gap is closed, while ω_0 denotes the conditional probability that, given a closed initial gap, the closure reflects a free adjustment opportunity.

During inaction, the absence of a price change implies $\ell_{i,t} = 0$ and restricts the reset price to values consistent with a gap inside the inaction region. Upon adjustment, the observed price reveals the reset price exactly, $p^r_{i,t} = p_{i,t}$. If adjustment occurs from outside the inaction region, either value of $\ell_{i,t}$ is admissible; if it occurs from within the region, the adjustment must reflect a free opportunity and hence $\ell_{i,t} = 1$.[12] We use these restrictions to construct date-specific grids and solve the FFBS recursions. Appendix B.2 provides the details of the grid-based algorithm.

3.2.2 Gibbs sampler for the joint posterior

Building on the latent-state sampler described above, we construct a Gibbs sampler for the joint posterior of the state-space parameters and latent states.

Priors. Table 1 reports the prior distributions. We choose prior functional forms to obtain conjugate full conditionals wherever possible and set the hyperparameters to be weakly informative, so that posterior inference is driven primarily by the likelihood. Two elements of the prior are more directly tied to the data.

First, we construct the prior support for the inaction bounds using the range of observed

[12] At the initial observation, for which no preceding price change is observed, the filtered distribution allows either value of $\ell_{i,t}$ and any reset price consistent with a gap inside the inaction region.

price changes.[13] Once a bound moves beyond the range of observed price changes on the corresponding side of the distribution, observed adjustments provide no further information for classifying whether a price change was free or threshold-triggered. We therefore restrict the prior support to this range to prevent the bounds from drifting into weakly identified tail regions.

Second, the prior for the initial latent states reflects that we do not observe whether the first price in a price trajectory results from an adjustment. We set the probability that the initial price gap is zero equal to the empirical frequency of price changes. Conditional on a zero gap, we assign a 60% probability to a free adjustment, in line with the average degree of Calvoness estimated by Gautier and Le Bihan (2022) using French micro-price data.

Full conditionals. Given the priors in Table 1, the full conditionals for μ, σ_ε^2, and λ belong to standard conditionally conjugate families: Normal-Inverse-Gamma for $(\mu, \sigma_\varepsilon^2)$ and Beta-Bernoulli for λ. Conditional on the latent states, these parameters can be sampled directly.

Because the *Ss* measurement equation does not contain measurement error, the latent states and inaction thresholds are tightly linked. Conditional on the thresholds, the latent states must satisfy the restrictions implied by the observed price changes: inaction without a free adjustment requires the price gap to lie inside the inaction region, while adjustment without a free opportunity requires it to lie outside. Conversely, conditional on the latent states, these restrictions constrain the admissible values of the thresholds. As a result, separate updates of the latent states and thresholds can lead to poor mixing of the Markov chain.

We therefore update the latent states and inaction thresholds jointly. Because their joint full conditional does not admit a closed-form solution, we use a Metropolis–Hastings step within the Gibbs sampler (Geweke and Tanizaki, 2001). We propose the latent states from their smoothed conditional distribution, approximated using the grid-based FFBS procedure described in Section 3.2.1, and the inaction thresholds from truncated normal distributions with the same support as their priors and centered at their previous draws. The proposal standard deviations are tuned during burn-in using a Robbins–Monro stochastic approximation (Robbins and Monro, 1951).[14] We target an acceptance rate of 33 percent, which lies between the commonly used optimal-scaling benchmarks of approximately 23 percent for high-dimensional and 44 percent for univariate Gaussian random-walk Metropolis proposals (Roberts et al., 1997; Roberts and Rosenthal, 2001). Appendix B.3 provides the full Gibbs sampling algorithm, while Appendix B.4 discusses extensions to alternative reset price processes, including autoregressive dynamics, exogenous covariates, and fat-tailed innovations.

[13] In our empirical application, we construct estimation windows to contain at least ten positive and ten negative price changes, ensuring that the data-dependent prior support can be constructed on both sides of the price-change distribution.

[14] In all estimations, we use 8,001 Gibbs draws, discard the first 3,001 as burn-in, and retain the remaining 5,000 for posterior inference.

Table 2: Monte Carlo Results

	True	Bias		Standard Deviation	
		$T = 12$	$T = 24$	$T = 12$	$T = 24$
(a) Parameters					
$\underline{s}$	-20.18	0.00	0.00	0.08	0.04
$\bar{s}$	18.90	0.00	0.00	0.12	0.07
μ	0.28	-0.00	-0.00	0.10	0.05
σ_e	3.85	0.04	0.02	0.14	0.09
λ	6.53	-0.04	-0.03	0.43	0.29
(b) Cross-sectional moments of price gaps*					
Mean	-1.02	0.00	0.00	0.23	0.18
Std Dev	8.25	-0.11	-0.04	0.09	0.06
Skewness	-0.03	-0.01	-0.00	0.02	0.01
Kurtosis	2.66	0.08	0.05	0.04	0.03

Notes: The "True" column in Panel (a) reports DGP parameter values and in Panel (b) reports cross-sectional price gap moments (including zeros) averaged over time periods and across replications. The "Bias" and "Standard Deviation" columns report, respectively, the mean bias and standard deviation of the posterior mean across replications, averaging first over time periods within each replication for Panel (b). True values are in percent, except for the cross-sectional standard deviation of price gaps which is in percentage points and skewness and kurtosis which are scale-free. All Bias and Standard Deviation columns are in percentage points, except for skewness and kurtosis which are scale-free. Results are based on 1,000 Monte Carlo replications, each a balanced panel of 300 price trajectories with $T = 12$ or $T = 24$ periods, with parameter values as specified in Appendix Table A.1. * The average correlation between estimated (posterior mean) and true individual price gaps across Monte Carlo replications is 0.65 for $T = 12$ and 0.67 for $T = 24$.

3.3 Monte Carlo

We use a Monte Carlo experiment to assess how well the algorithm recovers the model parameters and latent states in finite samples. We calibrate the data-generating process using the parameter values in Appendix Table A.1 and consider samples with 300 price trajectories observed over 12 and 24 periods.[15] For each sample size, we generate 1,000 samples from the state-space representation and use the Gibbs sampler in Section 3.2.2 to recover the parameters and latent states. Table 2 reports the results.

Panel (a) reports the bias and standard deviation of the posterior mean estimates of the model parameters across Monte Carlo replications, while panel (b) reports the corresponding statistics for the cross-sectional moments of the price gap distribution. Both the parameters and the moments of the price gap distribution are recovered with negligible bias and

[15] These sample sizes are chosen to be comparable to those in the empirical application. Across estimation windows, 88% contain 12 periods and 5.3% contain 24 periods, while the mean and median numbers of quotelines are 360 and 311, respectively.

high precision. Increasing the sample length from 12 to 24 periods further reduces the dispersion of the estimates, reflecting the additional information contained in the longer panels. Appendix B.5 reports post-estimation diagnostics and additional Monte Carlo experiments under calibrations with different degrees of Calvoness. Overall, the results validate the proposed algorithm and show that both the model parameters and the moments of the latent price gap distribution can be recovered accurately in finite samples.

4 Application to UK micro-price data

This section brings the algorithm to micro-level price data from the UK. We begin by describing the dataset, the cleaning procedure, and the estimation windows. We then present the parameter estimates and examine their empirical properties. Finally, we construct the estimated distribution of price gaps and document its time variation.

4.1 Data description

We use publicly available data on locally collected price quotes underlying the construction of the UK Consumer Price Index (CPI). To produce the CPI, the Office for National Statistics (ONS) collects monthly prices for a broad set of goods and services selected to represent household expenditure across the UK.[16] Prices are collected either centrally or locally. Central collection is used for items with uniform pricing nationwide or for which regional variation can be captured remotely, for example, through the internet, telephone, or email. To protect confidentiality, centrally collected prices that could reveal the identity of the price setter are excluded from the public release.

The remaining items, accounting for roughly 60 percent of the aggregate CPI by weight and spanning 11 of the 12 major COICOP divisions, are collected in person by ONS field agents. Each month, they visit thousands of retail outlets across more than 140 UK locations and collect over 100,000 prices using handheld devices.[17] Our baseline sample consists of these locally collected price quotes from 1996m1 to 2026m1.[18]

Our cleaning procedure has two stages. In the first stage, we apply three filters: we exclude price quotes that did not enter the actual CPI calculation because they failed the ONS internal validation procedures; we remove quotes referring to research items that were not part of the official CPI; and we exclude any quotes that cannot be uniquely identified by the combination of month, item identifier, shop code, and region.[19]

[16] ONS refers to these as representative items. Examples of item descriptions include: "large loaf-white-sliced-800g", "ultra low sulphur petrol 10l", "washing mach no dryer max 1800", "funeral-cremation", "driving lesson 1 hour".

[17] Education is the only COICOP division not represented, as all items in that division are collected centrally.

[18] Bunn and Ellis (2012) were the first to document stylized facts from this dataset, which has since been widely used in the pricing literature. Since this is not the first paper to use the data, we focus on the features most relevant for constructing our final sample.

[19] The validation procedures are described in ONS (2014, Chapter 6). In the ONS internal system, quote-lines are uniquely identified by month, item identifier, shop code, and location. For confidentiality reasons,

In the second stage, we clean the raw data to obtain quote-lines of regular prices by removing product substitutions, imputing regular prices during sale spells, excluding price changes whose absolute log size is greater than 2 or less than 0.01, and splitting quote-lines at gaps to ensure contiguous observations. We also exclude any items that do not have at least ten positive and ten negative price changes over the full sample. Details on these procedures are provided in Appendix C.1. Our final sample includes almost 35 million price quote observations, more than 3.6 million quote-lines, and 1,337 unique items.

As described in Section 3.1, model parameters are window specific. We split each item's sample into a sequence of non-overlapping estimation windows, which we call *item-windows*, and conduct estimation at this level, thereby allowing parameters to vary both across items and over time. Starting from the item's first observation, the first window ends in the following January if it contains at least ten positive and ten negative price changes. Otherwise, it is extended in twelve-month increments until a subsequent January at which this condition is satisfied or the item exits the sample.[20] Subsequent windows are defined in the same way. This guarantees that each item-window used in estimation contains at least ten positive and ten negative price changes, allowing both adjustment thresholds to be identified.

The resulting windows are overwhelmingly annual: across the 1,337 items we obtain 15,664 item-windows, 88 percent of which span exactly twelve months. For the few terminal windows that do not satisfy this requirement (fewer than 1 percent), we estimate only the latent states, holding parameters fixed at the posterior means from the item's last estimable window. Finally, to avoid discontinuities in the estimated reset price trajectories, for quote-lines that continue across item-windows we use the posterior means of the latent states at the end of the preceding window as initial conditions for the latent states in the next window.

4.2 Parameter estimates

Table 3 summarizes the parameter estimates across item-windows. The first two columns report the mean and median of the posterior mean of each parameter, while Appendix Figure A.2 shows their cross-sectional distributions over time. The estimated inaction region is asymmetric, with mean lower and upper bounds of -33 and 50 percent, respectively, and corresponding medians of -22 and 40 percent. Reset prices drift upward by 0.11 percent per month on average, with innovations that have a standard deviation of 8.9 percent. The estimated probability of a free adjustment opportunity averages 7.1 percent. Finally, the mean degree of Calvoness is 66 percent, implying that roughly two-thirds of price changes occur

the published data report region rather than location. As a result, two outlets with the same shop code may be located in different locations within the same region, making their price trajectories indistinguishable in the public data. These account for roughly 3.8 million observations. Following ONS guidance, we exclude these cases and retain only price observations that are uniquely identified by the combination of month, item identifier, shop code, and region.

[20] We end windows in January to align them with the annual rotation of the ONS basket. In our data, most items enter the sample in February and exit in January.

Table 3: Parameter estimates: descriptive statistics and variance decomposition

	Mean	Median	Std. dev.	Within share	Between share
Inaction lower bound ($\underline{s}$)	-32.96	-22.05	33.24	68.61	31.39
Inaction upper bound ($\bar{s}$)	49.51	40.42	34.09	56.13	43.87
Reset price drift (μ)	0.11	0.15	1.08	83.74	16.26
Reset price std. dev. (σ_ε)	8.93	8.06	4.95	37.09	62.91
Free adj. probability (λ)	7.07	5.17	6.41	33.44	66.56
Calvoness	66.47	69.81	26.77	70.86	29.14

Notes: Computed across 15,513 item-window observations for 1,337 items (excluding FFBS-only windows). Total reports the standard deviation in the same units as the parameter, with all parameters multiplied by 100. Within and Between report the shares of total variance, in percent, attributable to within-item variation over time and between-item variation, respectively. Calvoness is defined as the share of price changes occurring inside the inaction region.

inside the estimated inaction region.

Comparisons with existing estimates are not exact because our estimation approach differs from the literature in two important respects: we estimate the semi-structural parameters of the state-space representation in Section 3.1, and we do so using Bayesian methods rather than moment matching. The closest comparison is Gautier and Le Bihan (2022), who estimate a random menu cost model using French CPI micro-price data by simulated method of moments. Despite these differences, the estimates are broadly comparable. Relative to theirs, we find a wider inaction region and somewhat greater reset price volatility, while the reset price drift and probability of free adjustment are nearly identical.[21] Our estimated degree of Calvoness is also close to theirs, and lies within the range of related estimates in the literature, from 50 percent for the UK in Blanco (2021) to 75 percent for the US in Nakamura and Steinsson (2010).

Our estimates display substantial variation across items and over time. The last three columns of Table 3 decompose the variance of each parameter into within-item and between-item components. The inaction bounds and reset price drift exhibit predominantly within-item variation, whereas the standard deviation of reset price innovations and the free adjustment probability vary mainly across items. Thus, the relative importance of time-series and cross-sectional variation differs substantially across parameters.

The estimates also exhibit systematic relationships across parameters. Our semi-structural approach does not impose a mapping from the process parameters to the bounds of the in-

[21] For the standard deviation of reset price innovations, both sets of estimates range from close to zero to around 20 percent. For the inaction bounds, the estimates in Gautier and Le Bihan (2022) span roughly 0 to 40 percent in absolute value, while ours range from about 1 to 199 percent.

Table 4: Cross-parameter relationships

	Dependent variable			
	Inaction lower bound ($\underline{s}$)	Inaction upper bound ($\bar{s}$)	Frequency of price increases	Frequency of price decreases
Reset price drift (μ)	-5.20**	-4.41**	0.77**	-0.51**
	(0.38)	(0.36)	(0.06)	(0.06)
Reset price std. dev. (σ_ε)	-0.73**	2.81**	0.18**	0.17**
	(0.16)	(0.12)	(0.01)	(0.01)
Free adj. probability (λ)	-1.52**	-0.36**	0.44**	0.40**
	(0.12)	(0.11)	(0.02)	(0.02)
Adj. R^2	0.40	0.54	0.84	0.90
Observations	15,513	15,513	15,513	15,513

Notes: Each column reports an item-window level OLS regression of the dependent variable on all the estimated parameters, item fixed effects, and window-start fixed effects. All variables are expressed in percentage points. Standard errors clustered by item in parentheses. Sample: 15,513 item-window observations for 1,337 items (excluding FFBS-only windows). * (**) denotes rejection at the 32% (10%) significance level.

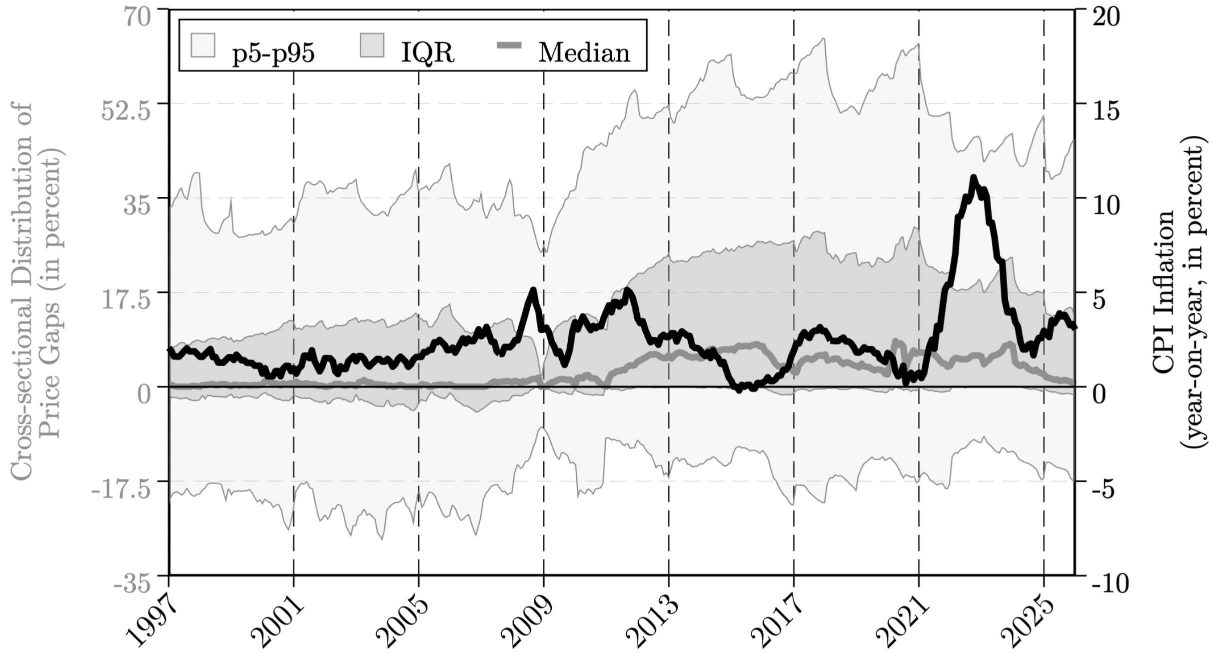
action region. In the structural model, this mapping arises endogenously from firms' optimization. Table 4 examines whether the corresponding comparative statics emerge in the estimated item-window parameters. The first row shows that a higher reset price drift shifts the inaction region downward and is associated with more price increases and fewer price decreases. This pattern is consistent with the asymmetry generated by positive trend inflation in menu cost models (Ball and Mankiw, 1994). The second row shows that greater reset price volatility widens the inaction region and raises the frequency of adjustments in both directions, as in *Ss* price-setting models with higher volatility (Vavra, 2014). Finally, a higher probability of a free adjustment opportunity is associated with a more negative lower bound and with more frequent price increases and decreases, consistent with models featuring random opportunities for costless adjustment (Stokey, 2009). The only exception is the upper-bound response, which is estimated to be slightly negative rather than positive. Overall, 11 of the 12 estimated relationships have the sign predicted by the comparative statics of standard *Ss* models.

4.3 The estimated distribution of price gaps

For each observed price quote $p_{i,t}$, we define the estimated price gap as $\widehat{s}_{i,t} = p_{i,t} - \widehat{p}^r_{i,t}$, where $\widehat{p}^r_{i,t}$ denotes the posterior mean of the reset price.[22] Since the price gap resets to zero

[22]Throughout the remainder of the paper, we use posterior means and abstract from posterior uncertainty in the price-gap moments. The Monte Carlo exercises in Section 3.3 show that, for samples comparable in size to a typical item-window in our application, posterior means of the moments are tightly centered around their true values. For the aggregate moments used in the empirical analysis, Appendix Figure A.3 additionally constructs the standardized moments separately for each posterior draw, after aggregating across items, and

Figure 3: Price gaps and inflation over time



Notes: The thick gray line reports the median of the weighted cross-sectional price-gap distribution; dark and light shaded regions show the interquartile and 5th–95th percentile ranges, respectively (left axis). Each quote-line is weighted by its CPI weight. The black line is published year-on-year CPI inflation (source: ONS; right axis).

upon adjustment for every quote-line, estimated gaps are directly comparable across items and can be pooled to construct an aggregate distribution. Each quote-line enters with its CPI weight, so that the resulting distribution is representative of the price index. We denote this aggregate distribution by $\widehat{G}_t$.[23] Figure 3 plots its evolution from January 1997 to January 2026 alongside published year-on-year CPI inflation, reporting the median together with the interquartile and 5th-to-95th percentile ranges.

The gap distribution exhibits substantial time variation in its shape. Year-on-year inflation itself varies considerably, ranging from -0.2 to 11.1 percent over the sample. Yet the underlying gap distribution exhibits even greater variation, both in its location and dispersion. The median gap ranges from 0 to 8.5 percentage points, the interquartile range from 2 to 30 percentage points, and the 5th-to-95th percentile range from 33 to 86 percentage points.[24] Characterizing the gap distribution through its ergodic counterpart alone would miss this variation entirely.

shows that posterior uncertainty around the resulting mean, standard deviation, skewness, and kurtosis is almost negligible.

[23] While the true structural state of the economy depends on the joint distribution of price gaps and item-specific adjustment rules, pooling the raw gaps yields a tractable, aggregate indicator that effectively summarizes macroeconomic pricing pressure.

[24] The extremes of the median and interquartile range occur during the pandemic months, when ONS price collection was partially suspended, but they are not artifacts of the disruption. Excluding 2020m3–2021m7 leaves the ranges essentially unchanged, with maxima of 8.0 and 29.4 percentage points in December 2023 and October 2017, respectively.

Through the first decade of the sample, the gap distribution fluctuates around a comparatively tight configuration, with the median close to zero and occasional episodes of marked compression, most visibly at the end of 2008, when the VAT cut introduced in response to the global financial crisis triggered broad repricing across items. From 2009 onwards, the distribution is persistently wider and increasingly asymmetric. The upper tail shifts outward over the 2010s, reaching its widest point in late 2017, about a year and a half after the Brexit referendum and the associated sterling depreciation. Over the same period, the median becomes persistently positive, so that most of the mass comes to lie above zero. The Covid-19 pandemic marks the sharpest break in the sample. The distribution widens abruptly in the spring of 2020, remains displaced throughout the inflation surge of 2022–2023, and approaches its pre-pandemic shape only near the end of the sample.

The evolution of the distribution bears no simple relationship to contemporaneous inflation. The tightest and the widest gap distributions in the sample, at the end of 2008 and in late 2017, both occur with inflation at roughly 3 percent. The figure instead suggests that the connection emerges with a delay. Inflation records price changes as they happen, while the distribution accumulates those still to come. The widening of 2020 arrives while inflation is still subdued and precedes the surge of 2022–2023; inflation returns towards target only as the distribution reverts. We explore this relationship more systematically next.

5 Price gaps in action

This section puts the estimated price gap distribution to work. We begin with a snapshot of the distribution before three extreme inflation outcomes in our sample. Conditioning on identified monetary policy and oil shocks, we then examine how the state of the distribution shapes their pass-through to inflation. Unconditionally, we evaluate the predictive content of its moments for inflation and quantify their contribution.

5.1 Price gaps ahead of inflation extremes

We begin with three case studies centered on the inflation extremes in our sample. Two precede the sharpest inflation peaks: year-on-year inflation reached its pre-pandemic peak of 5.1 percent in September 2008 and its overall maximum of 11.1 percent in October 2022. The third precedes the sample minimum of -0.2 percent, reached in April 2015. We examine the gap distribution one year before each of these outcomes and compare it with a benchmark date at which the distribution was close to its long-run configuration.[25] We use March 2024 as the benchmark, by which point the distribution had reverted to its pre-pandemic shape following the displacement documented in Figure 3. Over the following twelve months, cumulative inflation was 2.6 percent, close to its sample average.

[25] We define the benchmark as the in-sample date that minimizes the sum of squares of the standardized aggregate mean, standard deviation, skewness, and kurtosis of the gap distribution.

Figure 4 overlays the gap distribution at each case-study date on the benchmark distribution and reports cumulative inflation over the year that followed. Appendix Table A.5 reports the corresponding summary statistics and standardized moments.

September 2007. Twelve months before the biggest pre-pandemic inflation peak, headline inflation stood at 1.7 percent. The gap distribution deviated from the benchmark on a single side, with a shoulder of prices that had fallen behind their reset values, while the upper tail was in line with its long-run counterpart. The pending adjustments were price increases. Over the following year, the price level rose by 5.1 percent, twice the benchmark path, and year-on-year inflation reached its pre-pandemic peak.

April 2014. Twelve months before the deepest inflation trough, the picture was the mirror image. Headline inflation was 1.8 percent, indistinguishable from September 2007. The deviation now sat entirely on the positive side, with substantially more mass at large positive gaps and a negative range that matched the benchmark. The pending adjustments were price cuts. Over the following year, prices barely moved, and if anything declined. The shortfall relative to trend was substantial, as the benchmark path accumulated 2.6 percent of inflation over the same horizon, and year-on-year inflation reached its sample minimum.

October 2021. In contrast with the two previous episodes, in October 2021 the price gap distribution offered no such one-sided reading. Inflation stood at 4.1 percent, already high by recent historical standards, though well below the rates that would follow. The distribution carried extra mass in both tails relative to the benchmark, a legacy of the large relative price movements of the pandemic, and was primed for large price adjustments in either direction. The surge in energy prices in 2022 pushed the adjustment to the upside. Over the following year, the price level rose by 11.1 percent, more than four times the benchmark path, and year-on-year inflation reached its sample maximum.

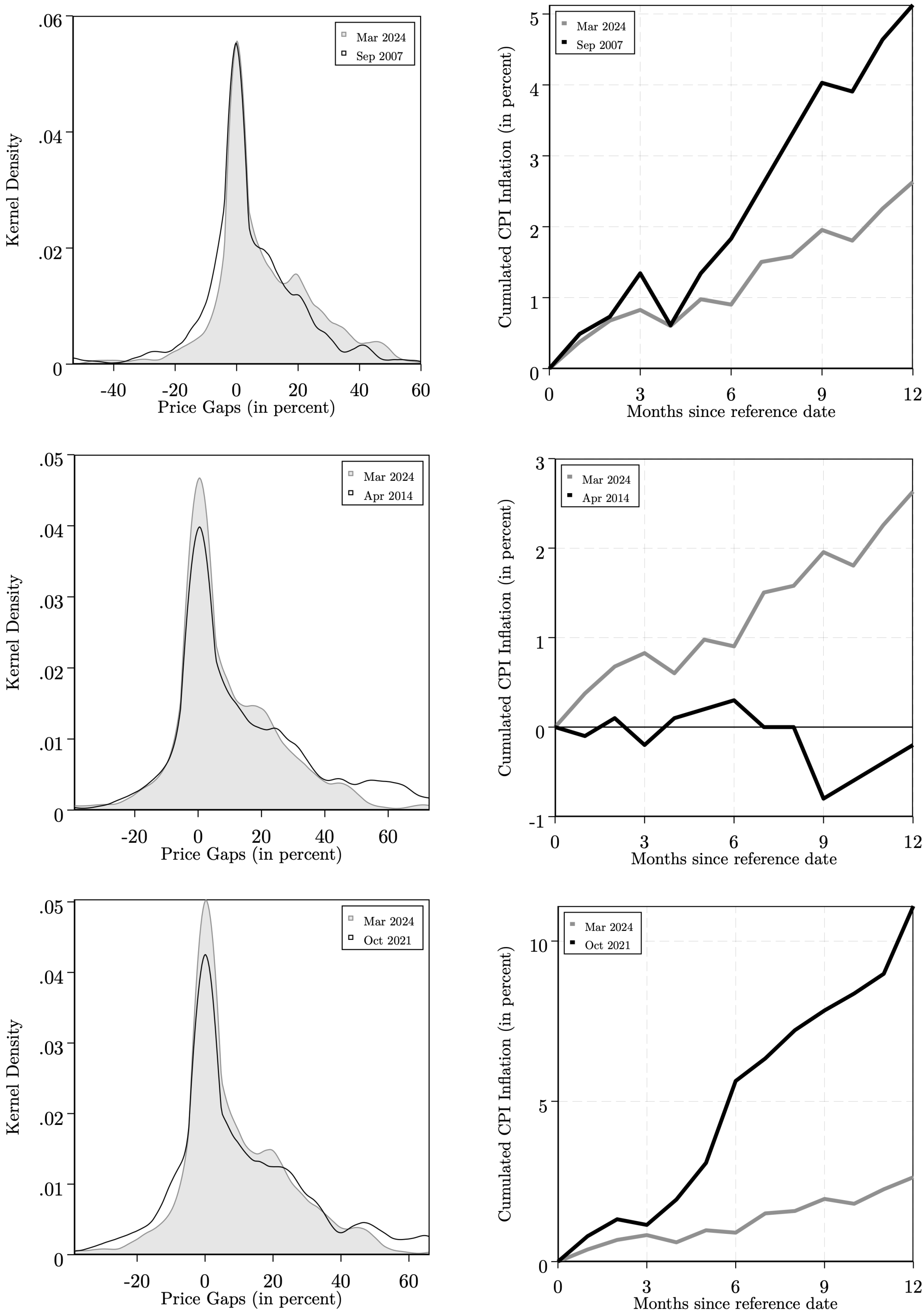
These comparisons are purely descriptive and cannot decompose realized inflation into the unwinding of the pressure in place at the outset and shocks that arrived along the way.[26] We now make progress on the latter by turning to externally identified aggregate shocks and testing whether their propagation depends on the state of the distribution.

5.2 State-dependent shock propagation

The theory predicts that the transmission of an aggregate shock depends on the state of the price gap distribution when the shock arrives. As discussed in Section 2.3.1, transmission should be stronger when the distribution contains more latent adjustment pressure in the direction of the shock. We examine this prediction using identified monetary policy and oil shocks.

[26] While our framework does not accommodate aggregate uncertainty per se, shocks are captured ex post through changes in the trend of reset prices, idiosyncratic volatility, the free-adjustment probability, and the associated adjustment bands.

Figure 4: Price gaps ahead of inflation extremes



Notes: The left column reports kernel density estimates of the weighted cross-sectional price-gap distribution at each reference date (black line) and at the March 2024 benchmark (shaded gray). The right column reports cumulative CPI inflation over the twelve months following each reference date (black) and following March 2024 (gray). The rows correspond to September 2007, April 2014, and October 2021, respectively. Each quote-line is weighted by its CPI weight.

5.2.1 Empirical strategy

Our empirical strategy combines within-class time variation with instrumental variables in a state-dependent panel local projection framework.[27] More concretely, for each horizon h, we estimate

$$\begin{aligned}\log P_{c,t+h} = {} & \alpha_h \Delta x_t + \beta_h^+ \Delta x_t^+ \mathrm{m}_{c,t-1} + \beta_h^- \Delta x_t^- \mathrm{m}_{c,t-1} + \gamma_h \Delta x_t \mathrm{sd}_{c,t-1} \\ & + \zeta_h^+ \Delta x_t^+ \mathrm{sk}_{c,t-1} + \zeta_h^- \Delta x_t^- \mathrm{sk}_{c,t-1} + \eta_h \Delta x_t \mathrm{k}_{c,t-1} + \mathbf{w}_{c,t}' \boldsymbol{\Gamma}_h + e_{c,t+h}, \qquad (20)\end{aligned}$$

where $\log P_{c,t+h}$ is the log of the published class price index, Δx_t is the endogenous impulse variable, $\Delta x_t^+ \equiv \max\{\Delta x_t, 0\}$ and $\Delta x_t^- \equiv \min\{\Delta x_t, 0\}$ are its positive and negative parts, and $\mathrm{m}_{c,t-1}$, $\mathrm{sd}_{c,t-1}$, $\mathrm{sk}_{c,t-1}$, and $\mathrm{k}_{c,t-1}$ denote the standardized mean, standard deviation, skewness, and kurtosis of the estimated price gap distribution for class c in the previous month.[28] The control vector $\mathbf{w}_{c,t}$ includes class fixed effects and twelve lags of the dependent variable, the endogenous impulse variable, the standardized gap moments, and a set of macroeconomic variables typically included in VARs for the UK economy.[29] We instrument all terms involving the endogenous impulse variable. For each instrument, we apply the same sign decomposition and interactions with the lagged price gap moments.[30] Standard errors are clustered by product class.

The interactions between the impulse variable and the moments of the estimated price gap distribution allow the response to vary with the initial state of the distribution. We allow the interactions with the mean and skewness to depend on the sign of the impulse because these are odd moments. They capture the direction in which pricing pressure has accumulated, so their effect is expected to depend on whether the impulse pushes prices with or against that pressure. By contrast, the standard deviation and kurtosis are even moments. They capture features of the distribution that are invariant to a reversal in sign, such as the extent to which mass is concentrated away from the center and toward the adjustment thresholds, and therefore enter symmetrically.

5.2.2 Monetary policy

We first consider monetary policy shocks, the primary aggregate disturbance studied in the menu cost literature. An important feature of our sample is that the Bank Rate was con-

[27] A product class is a COICOP4 category. Although gaps are constructed at the individual quote-line level, we estimate the specification at the class level because product substitutions make the panel increasingly unbalanced at finer levels of disaggregation and leave fewer observations from which to construct the moments of the price gap distribution.

[28] Each moment is standardized using the class's own time-series mean and standard deviation.

[29] Following Cesa-Bianchi et al. (2019) and Braun et al. (2025), we include log GDP, the log nominal sterling exchange rate index, corporate bond spreads, the unemployment rate, and the log FTSE. More details are provided in Appendix C.3.

[30] More precisely, for each instrument $z_{i,t}$, the instrument set includes $z_{i,t}$, $z_{i,t}^+ \mathrm{m}_{c,t-1}$, $z_{i,t}^- \mathrm{m}_{c,t-1}$, $z_{i,t} \mathrm{sd}_{c,t-1}$, $z_{i,t}^+ \mathrm{sk}_{c,t-1}$, $z_{i,t}^- \mathrm{sk}_{c,t-1}$, and $z_{i,t} \mathrm{k}_{c,t-1}$.

strained by the effective lower bound for an extended period. This motivates our choice of both the impulse variable and the external instruments. To capture the stance of monetary policy, we use changes in the one-year gilt yield as the impulse variable rather than changes in the Bank Rate.[31] As instruments, we use the target, path, and QE factors estimated by Braun et al. (2025), which capture different dimensions of monetary policy over this period.[32]

Coefficient estimates. Table 5 reports the estimated coefficients and hypothesis tests at selected horizons.[33] The baseline coefficient captures the response at a distribution in which each of the four moments is equal to its time-series mean. It is negative and statistically significant at the 10 percent level at all horizons and increases in magnitude through 18 months, implying that a monetary tightening lowers the price level relative to a no-shock counterfactual, with the effect building gradually over time. Despite differences in methodology and sample period, the magnitude is broadly consistent with the response of consumer prices reported by Braun et al. (2025).[34]

State dependence operates through the interaction terms. We first consider the mean. Holding the other moments fixed, a higher mean price gap implies that prices are, on average, further above their reset prices, generating greater downward pressure on prices. Consistent with this interpretation, the interaction with monetary tightening is negative, so a higher mean amplifies the decline in prices following a tightening; whereas the interaction with monetary easing is positive, dampening the increase in prices following an easing. The difference between the two mean interaction coefficients is statistically significant through 18 months, as shown by the rejection of the no-mean-asymmetry hypothesis in Panel (b).

Turning to the remaining moments, the interaction with the standard deviation is negative at all horizons and statistically significant at several of them. Holding the other moments fixed, a higher standard deviation corresponds to a wider gap distribution, placing more price setters closer to their adjustment thresholds. The negative interaction therefore implies that greater dispersion amplifies the price response to both monetary tightenings and easings. Given the bounded nature of the inaction region, holding the other moments fixed, a higher skewness places relatively more mass on the left-hand side of the gap distribution. Consistent with this, higher skewness tends to dampen the response to a monetary

[31] The use of 1y yields to capture monetary policy stance has become common in monetary VARs that span the effective-lower-bound period (e.g., Gertler and Karadi, 2015; Cesa-Bianchi et al., 2019).

[32] Panel (a) of Appendix Figure A.4 reports the weak-instrument test of Lewis and Mertens (2026). The null of weak instruments is rejected at the 10% significance level for a relative-bias threshold of approximately 30%. Appendix Figures A.5, A.6, and A.7 report identification-robust inference following Andrews (2018) alongside conventional inference for all horizons and hypothesis tests. Conclusions are mostly unchanged under identification-robust inference.

[33] Coefficient estimates and hypothesis tests for all horizons are reported in Appendix Figures A.5, A.6, and A.7.

[34] For comparison, rescaling the impulse responses from the SVAR estimated by Braun et al. (2025) over 1997–2019 and identified using the Path factor, which has its peak loading around the one-year maturity, implies that a 25 basis point increase in the one-year gilt yield lowers consumer prices by approximately 0.75–1 percent after two years.

Table 5: State-dependent monetary policy transmission

	Dependent variable: $\log P_{c,t+h}$					
Horizon	0	1	6	12	18	24
(a) Local Projection coefficients						
Δx_t	-0.29**	-0.34**	-0.55**	-0.85**	-1.15**	-1.07**
	(0.09)	(0.14)	(0.11)	(0.16)	(0.20)	(0.24)
$\Delta x_t^+ \times \mathrm{m}_{c,t-1}$	-0.29**	-0.29**	-0.47*	-0.67**	-0.44*	-0.33
	(0.14)	(0.13)	(0.31)	(0.33)	(0.37)	(0.45)
$\Delta x_t^- \times \mathrm{m}_{c,t-1}$	-0.03	0.00	0.40**	0.42**	0.71**	0.68**
	(0.07)	(0.09)	(0.15)	(0.22)	(0.30)	(0.38)
$\Delta x_t \times \mathrm{sd}_{c,t-1}$	-0.12**	-0.13**	-0.26**	-0.18*	-0.23*	-0.14
	(0.05)	(0.06)	(0.12)	(0.14)	(0.17)	(0.20)
$\Delta x_t^+ \times \mathrm{sk}_{c,t-1}$	0.16*	0.10	0.07	0.40*	0.01	0.35
	(0.11)	(0.17)	(0.24)	(0.32)	(0.37)	(0.36)
$\Delta x_t^- \times \mathrm{sk}_{c,t-1}$	-0.02	-0.05*	-0.10*	-0.07	-0.05	-0.05
	(0.03)	(0.04)	(0.07)	(0.09)	(0.13)	(0.15)
$\Delta x_t \times \mathrm{k}_{c,t-1}$	-0.08**	0.04	0.05	-0.07	-0.06	-0.02
	(0.03)	(0.04)	(0.08)	(0.09)	(0.10)	(0.12)
(b) Hypothesis tests (p-values)						
No mean asymmetry	0.06	0.05	0.02	0.01	0.05	0.17
No skew. asymmetry	0.11	0.38	0.52	0.17	0.88	0.35
No state dependency	0.10	0.02	0.17	0.14	0.18	0.33
Observations	25,023	24,958	24,633	24,243	23,853	23,463

Notes: This table presents estimates of (20) with Δx_t equal to the change in the 1y Gilt and the instrument set given by the target, path and QE factors from Braun et al. (2025). Panel (a) reports the local projection coefficients scaled to a 25 bps change in the 1y Gilt with standard errors clustered at the class level in parentheses. Panel (b) reports p-values for: no mean asymmetry ($H_0\colon \beta_h^+ = \beta_h^-$), no skewness asymmetry ($H_0\colon \zeta_h^+ = \zeta_h^-$), and no state dependency ($H_0\colon$ all six interaction coefficients equal to zero). * (**) denotes rejection at the 32% (10%) significance level.

tightening and amplify the response to a monetary easing. However, the individual interaction coefficients are statistically significant at only a handful of horizons, and the hypothesis of no skewness asymmetry cannot be rejected at most horizons. Finally, holding the other moments constant, higher kurtosis places more mass near the center and in the tails of the gap distribution. Its interaction displays no consistent sign across horizons and is generally not statistically significant; the only clear exception is the impact response, where higher kurtosis amplifies the response to monetary policy shocks of either sign.

Taken together, the signs of the interaction coefficients are consistent with the underlying economics of the model. The joint tests provide additional evidence of state dependence, rejecting the null that all interaction terms are jointly zero at the 10 percent level on impact and after one month, with p-values remaining below 0.20 through 18 months. Thus, the evidence for state dependence is not driven solely by the significance of individual coefficients.

Impulse responses at the case-study dates. To gauge the quantitative implications of the local projection estimates, Figure 5 reports the impulse responses to a 25 basis point increase in the one-year gilt yield evaluated at the three case-study distributions from Section 5.1, together with the corresponding response at the March 2024 benchmark distribution.[35]

The responses line up closely with the shape of the gap distributions documented in Section 5.1. In September 2007, the distribution had a shoulder of prices below their reset values and less mass on the positive side than the benchmark. Consistent with there being less downward adjustment pressure to reinforce a tightening, the response is dampened relative to the benchmark, by an average of 55 percent across horizons from impact to 24 months. The difference is statistically significant at the 32 percent level through approximately 18 months. The other two case-study dates display the opposite pattern. In April 2014, substantially more mass lies at large positive gaps, and the response to a tightening is amplified by an average of around 70 percent relative to the benchmark. The October 2021 distribution also has excess mass at positive gaps, as part of its two-sided displacement, and the response is amplified by an average of around 23 percent. The larger amplification in April 2014 is consistent with the more pronounced displacement on the positive side of the distribution. In both cases, the differences from the benchmark are statistically significant at the 32 or 10 percent level through approximately 21 months.

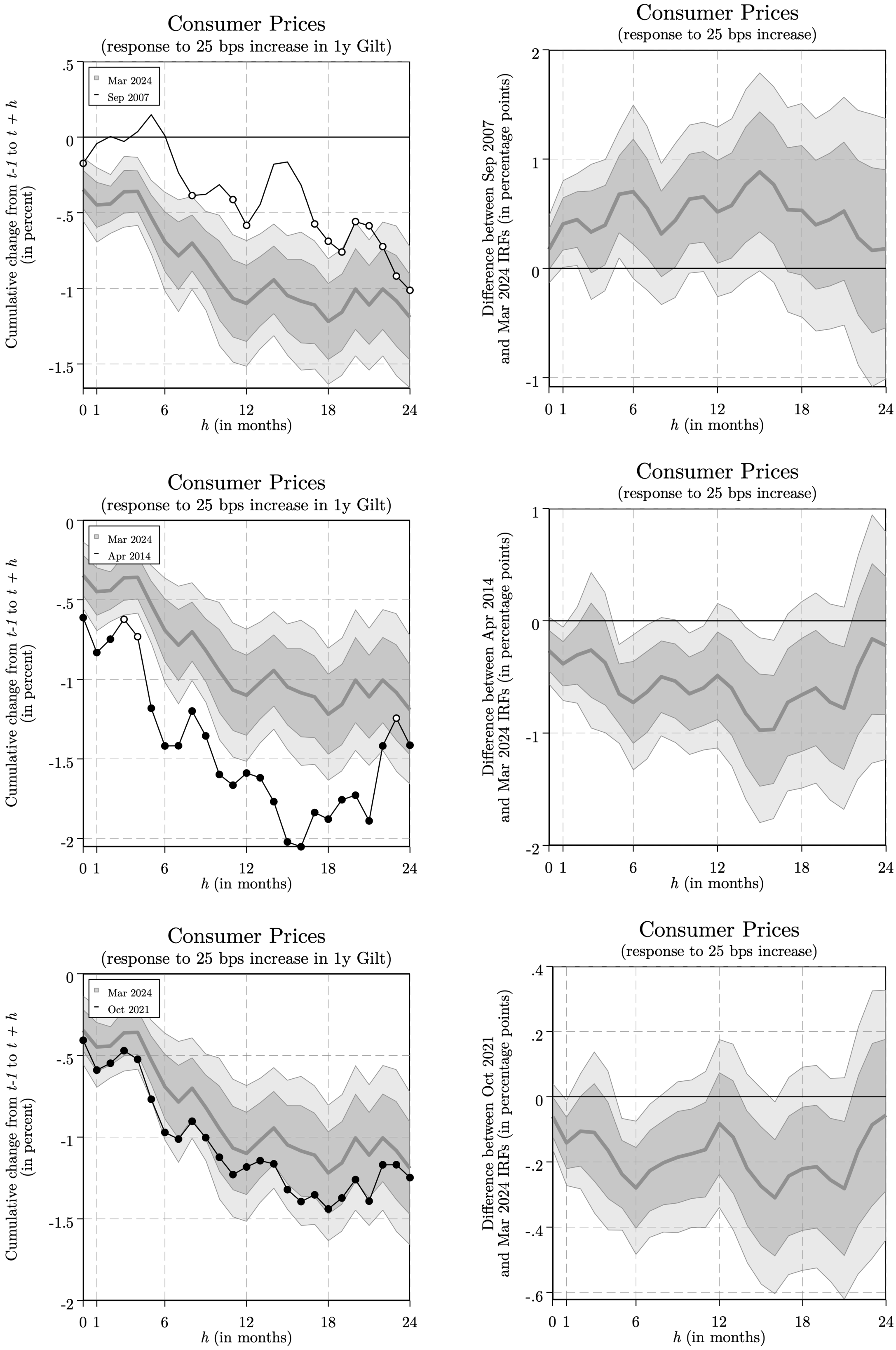
Appendix Figure A.8 repeats the exercise for a 25 basis point monetary easing. The pattern shifts in the expected direction: the response is amplified relative to the benchmark in September 2007 and October 2021, by 132 and 48 percent on average, respectively, but is essentially unchanged in April 2014. Overall, the estimates are consistent with stronger monetary policy transmission when the shock pushes prices in the same direction as the imbalance already present in the gap distribution.

Monetary policy pass-through over time. Beyond the three case-study dates, the price gap distribution varies considerably over time, as documented in Figure 3, implying corresponding variation in monetary policy transmission. Figure 6 traces this variation over the full sample, reporting the estimated pass-through of a monetary tightening and easing at the two-year horizon relative to the March 2024 benchmark.

The variation in pass-through induced by changes in the price gap distribution is substantial. Relative to the March 2024 benchmark, the response to a tightening ranges from being almost fully dampened during a handful of months around the global financial crisis

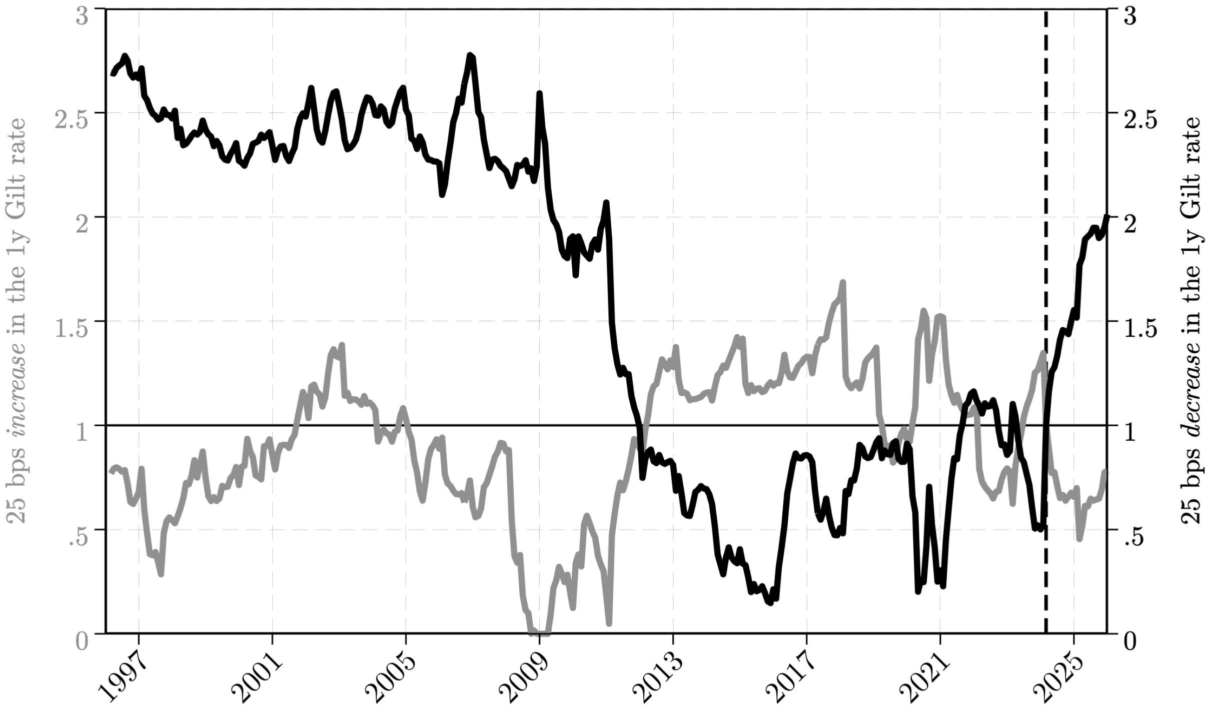
[35] The specification is linear in the shock, so the magnitudes reported here are local to a 25 basis point move. In the model, the shock-size and initial-condition channels reinforce each other (Section 2.3.1). The proportional comparison across states does not depend on the scale of the baseline response.

Figure 5: Responses to a monetary tightening at selected dates



Notes: The left column plots the estimated impulse response of cumulative inflation to a 25 bps increase in the 1y Gilt rate from (20), with Δx_t equal to the change in the 1y Gilt instrumented by the target, path, and QE factors from Braun et al. (2025). The grey line with shaded bands is the response evaluated at the benchmark distribution (March 2024); the black line is the response evaluated at the alternative date (September 2007, April 2014, and October 2021 in rows 1, 2, and 3, respectively). The right column plots the difference between the impulse response at the alternative date and at the benchmark date, with shaded areas denoting the 68% and 90% confidence bands. A white (black) filled circle indicates that the difference is statistically significant at the 32% (10%) level.

Figure 6: Monetary policy passthrough at the 2-year horizon



Notes: The figure plots the estimated cumulative inflation passthrough to monetary policy shocks 24 months ahead, relative to the benchmark distribution (2024m3). The grey line (left axis) and the black line (right axis) show the passthrough for a 25 bps increase and decrease in the 1y Gilt rate, respectively. Both series are normalised to 1 at 2024m3 (vertical dashed line): values above 1 indicate a larger-than-reference passthrough, values below 1 a smaller one. The passthrough is set to zero before normalisation if the estimated response has the wrong sign (6 instances: 2008m10 and 2008m12–2009m4). The passthrough is estimated based on coefficients from (20) with Δx_t equal to the change in the 1y Gilt and the instrument set given by the target, path and QE factors from Braun et al. (2025) and the standardized moments of the aggregate price gaps distribution over time.

to being amplified by as much as 69 percent in early 2018. The response to an easing ranges from being dampened by 85 percent to being amplified by 178 percent, with the largest amplification occurring near the beginning of the sample. The figure also reveals a broad shift in the direction of the asymmetry over time. Until around 2012, easings tend to be amplified while tightenings are dampened. The pattern reverses over much of the following decade, with stronger transmission of tightenings and weaker transmission of easings, before moving back toward the earlier configuration near the end of the sample. This middle period overlaps with the effective-lower-bound era and the subsequent period of subdued inflation and growth, suggesting that the gap distribution reflects these broader macroeconomic conditions in ways that matter for monetary policy transmission.

5.2.3 Oil prices

Monetary policy is one force behind movements in reset prices, but direct cost shocks such as changes in oil prices move them as well. To assess whether the same state dependence extends to a distinct source of variation in reset prices, we reestimate (20) using changes in oil prices as the impulse variable and the oil supply news shocks from Känzig (2021) as instruments. The coefficient estimates and hypothesis tests, reported in Appendix Figures A.9, A.10, and A.7, reproduce the main qualitative patterns found for monetary policy.[36]

Figure 7 reports the responses to a 10 percent increase in oil prices evaluated at the three case-study distributions and at the March 2024 benchmark. The results closely parallel those for monetary policy. In September 2007, when the distribution had excess mass below reset prices, the response to an oil-price increase is amplified relative to the benchmark. In April 2014, when the imbalance lies on the opposite side of the distribution, the response is dampened. In October 2021, when the distribution is displaced in both directions, the response is again amplified. The differences from the benchmark are less precisely estimated than for monetary policy, but the pattern across the three distributions is the same. Thus, a distinct aggregate shock generates the same qualitative pattern of state dependence across the three gap distributions.

Appendix Figure A.11 reports the corresponding responses to a 10 percent decrease in oil prices. As expected, the pattern across the first two case-study dates reverses: the response is dampened in September 2007 and amplified in April 2014, while it remains amplified in October 2021. Appendix Figure A.12 shows that the similarity with monetary policy extends beyond these three dates, with the direction of pass-through asymmetry displaying a similar broad evolution over the sample.

These results suggest that the gap distribution can serve as an indicator of the economy's sensitivity to new shocks. A displaced distribution signals that latent pricing pressure has built up, so that even a moderate shock can generate a large price response. The same features of the distribution that generate this state dependence should also contain unconditional information about future inflation, well beyond the three dates considered in Section 5.1. We explore this next.

5.3 Predictive content of price gaps

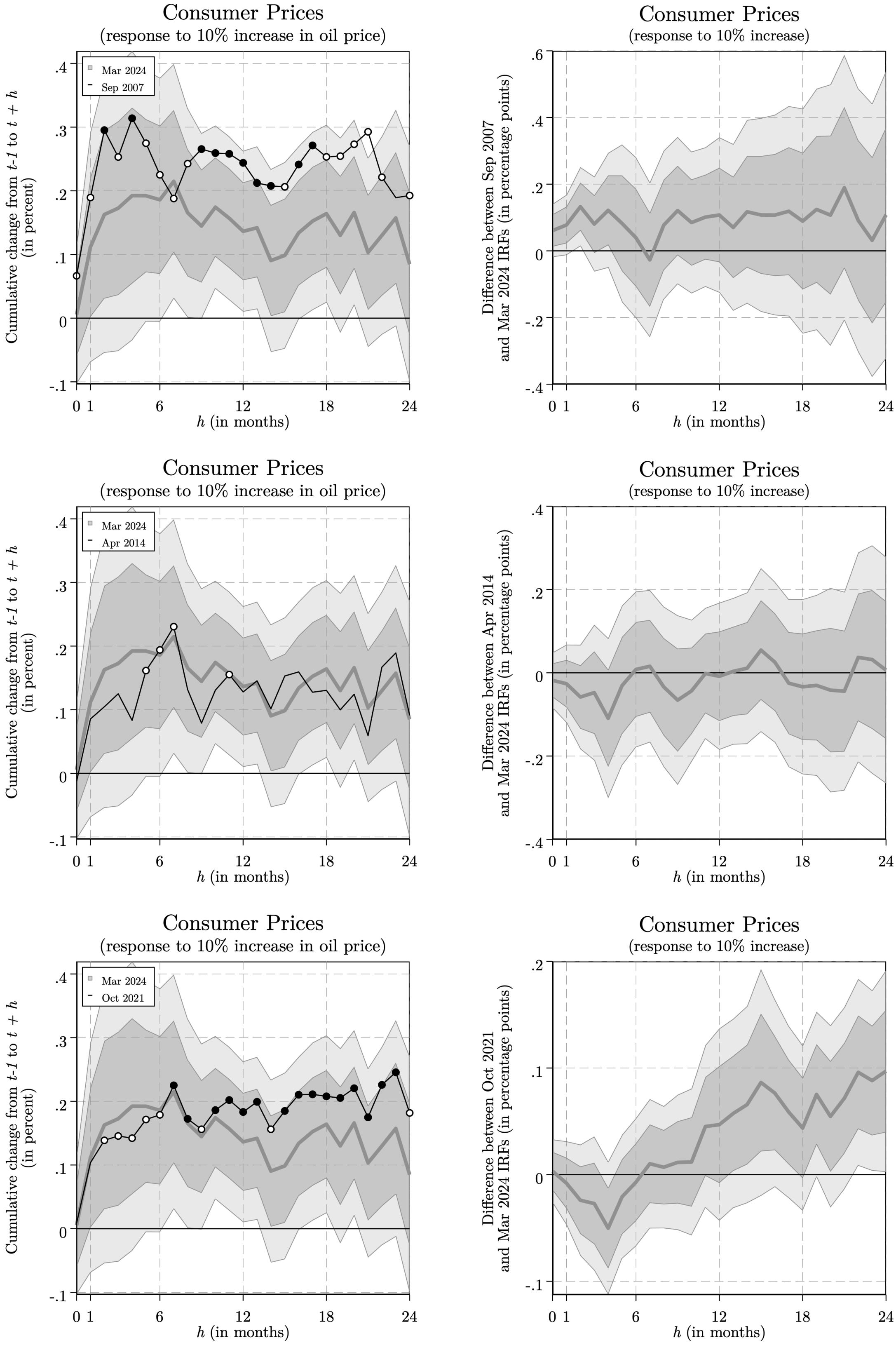
We study the predictive content of price gaps in two steps. We first establish the in-sample relationship between the price gap distribution and future inflation, and then evaluate its out-of-sample forecasting performance.

5.3.1 In-sample predictive content

Figure 4 suggests that the shape of the price gap distribution contains predictive information about future inflation. We now examine this relationship more systematically and assess whether it survives after controlling for standard predictors of inflation. Let π^y_{t+h} denote year-on-year inflation h months ahead. We then estimate

[36]Panel (b) of Appendix Figure A.4 reports the weak-instrument test of Lewis and Mertens (2026), which rejects the null of weak instruments at the 10% significance level for a relative-bias threshold of approximately 10%. Appendix Figures A.9, A.10, and A.7 also report identification-robust inference following Andrews (2018) alongside conventional inference for all horizons and hypothesis tests. Conclusions are mostly unchanged under identification-robust inference.

Figure 7: Responses to an oil price increase at selected dates



Notes: The left column plots the estimated impulse response of cumulative inflation to a 10% increase in oil prices from (20), with Δx_t equal to the log change in oil prices instrumented by oil supply news shocks from Känzig (2021). The grey line with shaded bands is the response evaluated at the benchmark distribution (March 2024); the black line is the response evaluated at the alternative date (September 2007, April 2014, and October 2021 in rows 1, 2, and 3, respectively). The right column plots the difference between the impulse response at the alternative date and at the benchmark date, with shaded areas denoting the 68% and 90% confidence bands. A white (black) filled circle indicates that the difference is statistically significant at the 32% (10%) level.

Table 6: Price gap moments and inflation

	Dependent variable: year-on-year inflation 12 months ahead						
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
(a) Price gap coefficients							
m_t	-0.12*	-0.20**	-0.19*	-0.25**	-0.10	-0.29**	-0.36**
	(0.10)	(0.10)	(0.12)	(0.13)	(0.13)	(0.12)	(0.12)
sd_t	-0.01	0.12*	-0.05	0.16**	0.06	0.00	0.12*
	(0.10)	(0.10)	(0.12)	(0.06)	(0.10)	(0.11)	(0.08)
sk_t	1.99**	1.45**	1.52**	1.27**	2.28**	0.93**	0.20
	(0.67)	(0.63)	(0.65)	(0.63)	(0.74)	(0.55)	(0.80)
k_t	-0.29*	-0.29*	-0.27*	-0.31*	-0.09	-0.32*	-0.20*
	(0.24)	(0.25)	(0.23)	(0.26)	(0.19)	(0.26)	(0.18)
(b) Joint significance (p-values)							
Price gaps	0.03	0.14	0.07	0.04	0.02	0.05	0.03
Controls		0.01	0.20	0.00	0.11	0.02	0.00
All vars	0.03	0.00	0.05	0.00	0.03	0.03	0.00
Adj. R^2	0.08	0.23	0.11	0.38	0.20	0.15	0.47
Observations	348	337	348	348	325	337	325

Notes: Panel (a) reports regression coefficients for price gap moments with Newey-West HAC standard errors (12 lags) in parentheses. The dependent variable is year-on-year inflation 12 months ahead. Column (1) includes no additional controls; (2) adds published year-on-year inflation π_t^y; (3) adds the 1-year Gilt yield; (4) adds log WTI spot price and log EU gas price; (5) adds year-on-year real GDP growth; (6) adds the log sterling exchange rate index; (7) includes all of the above. * (**) denotes rejection at the 32% (10%) significance level. Panel (b) reports p-values for joint significance of price gaps, additional controls, and all variables.

$$\pi_{t+h}^y = \alpha + \beta' \mathcal{M}_t + \gamma' \mathbf{X}_t + e_t, \tag{21}$$

where $\mathcal{M}_t$ contains the four standardized moments of the price gap distribution and $\mathbf{X}_t$ contains additional controls. We begin with the price gap moments alone, then add standard inflation predictors one at a time, before including them jointly. Table 6 reports the results for year-on-year inflation 12 months ahead.

Column (1) formalizes the patterns illustrated by the case studies in Section 5.1. Consistent with the theory and our previous results, a higher mean gap and higher kurtosis predict lower future inflation, while higher skewness predicts higher future inflation. The standard deviation does not enter significantly. The four price gap moments are jointly significant and account for around 8 percent of the variation in year-on-year inflation twelve months ahead.

A natural concern is that the gap distribution may capture the effects of other variables that jointly shape current price gaps and future inflation. The remaining columns examine this possibility by adding a different control in each specification: current year-on-year inflation in column (2), the one-year gilt yield in column (3), oil and gas prices in column (4), year-on-year real GDP growth in column (5), the sterling exchange rate index in column (6), and all controls jointly in column (7). Three results stand out. First, the explanatory power of the regressions rises substantially as controls are added, reaching 47 percent in column (7). This confirms that the controls capture important variation in future inflation. Second, the signs of the price gap coefficients remain broadly stable across specifications. This suggests that no single control is driving the relationship. Third, and most importantly, the gap moments are jointly significant at the 10 percent level in six of the seven specifications, with a p-value of 0.14 in the remaining one.[37]

Taken together, these results show that the predictive content of the price gap distribution extends beyond the three case studies and is robust to controlling for standard macroeconomic observables. Having established this relationship in sample, we now turn to evaluating whether it translates into improved out-of-sample inflation forecasts.

5.3.2 Out-of-sample forecasting

We estimate equation (21) without controls at horizons $h \in \{12, 15, 18, 21, 24\}$ months and evaluate forecast performance for target dates from February 2016 to January 2026, relative to the widely used random-walk benchmark of Atkeson and Ohanian (2001). To assess whether forecast performance differs across inflation environments, we also split the evaluation sample into two equal periods. In the first half, year-on-year inflation averaged 1.7 percent with a standard deviation of 0.9 percentage points; we label this the low-inflation period. In the second half, inflation averaged 5 percent with a standard deviation of 3.1 percentage points; we label this the high inflation period.

A genuine out-of-sample exercise requires estimates of the price gap distribution that use only information available at the time of the forecast. To construct these *real-time* estimates, we treat completed and ongoing item-windows differently. For item-windows completed by a given forecast date, we use their estimated gaps. When the forecast date falls within an ongoing item-window, we instead filter the latent states using data only through that date, while holding the item-specific model parameters fixed at their estimates from the most recently completed window.[38] This avoids repeatedly re-estimating the model as the current window expands and, importantly, estimating window-specific parameters when only a small portion of the window is available. Finally, we aggregate these item-level estimates to construct the real-time moments of the price gap distribution used for forecasting.[39]

[37]Similar results hold at horizons of 6, 18, and 24 months. Results available upon request.

[38]Newly introduced items are excluded from the real-time aggregate until their first window is completed.

[39]For each horizon h, the forecast for the target date is formed at $\tau - h$. To ensure a common set of target dates across horizons, we estimate the regression coefficients using only observations whose inflation outcome

Table 7: RMSPE ratio: Price Gaps *vs.* Random-Walk

Horizon	OOS Forecast Evaluation Sample		
	Full sample	Low inflation	High inflation
12	0.92	1.13	0.91
15	0.78*	1.03	0.76*
18	0.70**	0.92	0.68**
21	0.65**	0.76*	0.64**
24	0.63**	0.66*	0.63**

Notes: RMSPE ratio of Price Gaps relative to the Random-Walk benchmark. Forecasts are of year-on-year inflation h months ahead. Values below 1 indicate the competing model outperforms the benchmark. Full sample: 2016m2–2026m1; Low inflation: 2016m2–2021m1; High inflation: 2021m2–2026m1. Diebold-Mariano test of equal predictive accuracy with Newey-West HAC standard errors (lag $= h$). * (**) denotes rejection at the 32% (10%) significance level.

Table 7 reports the ratio of the root mean squared prediction error (RMSPE) of the price-gap forecast to that of the random-walk benchmark. Over the full evaluation sample, the forecasting gains increase with the horizon, from an 8 percent reduction in RMSPE at 12 months to 37 percent at 24 months, with the gains statistically significant at the longer horizons. The price-gap forecast also outperforms the benchmark in both inflation environments. During the high-inflation period, the reduction in RMSPE reaches 37 percent at 24 months. During the low-inflation period, gains emerge from 18 months onward, reaching 34 percent at 24 months, while the random-walk benchmark is harder to beat at shorter horizons when inflation is stable.

Results are robust to alternative forecasting choices. Appendix Table A.6 shows that forecasting gains are similar when the target is quarterly inflation rather than year-on-year inflation, with average monthly inflation showing a similar pattern over the full sample and high-inflation period. Appendix Table A.7 considers autoregressive benchmarks instead. Relative to an autoregressive model with coefficients fixed at the beginning of the evaluation sample, the price-gap forecast performs better at longer horizons in the full sample, though the fixed benchmark is more competitive during the low-inflation period. The gains are smaller relative to a recursively estimated autoregressive model, but the price-gap forecasts remain competitive and begin to outperform it at longer horizons. Finally, Appendix Table A.8 shows that alternative measures constructed using only the estimated reset prices, including reset price inflation as in Bils et al. (2012) and pricing-pressure indices in the spirit of Caballero and Engel (2007), also outperform the random-walk benchmark.

is realized by the first forecast origin, February 2016 minus h months. The coefficients are then held fixed throughout the evaluation period, while the real-time price gap moments are updated as new micro price data become available.

6 Conclusion

The cross-sectional distribution of price gaps is the key state variable of menu cost models, but it is unobserved. This paper develops state-space methods that exploit the structure of the *Ss* adjustment rule to recover the entire distribution from observed prices alone. Taking these methods to UK CPI microdata from 1996 to 2026, we find the economy spends extended periods far from its ergodic distribution, and we show two ways in which this matters for inflation dynamics. First, the distribution shapes the *propagation* of shocks: the same monetary policy shock can generate very different price responses depending on the distribution in place when it hits. Second, the distribution delivers *prediction*: gap moments improve upon past inflation in-sample and outperform the standard past-inflation benchmark out-of-sample. Together, these results show that the price gap distribution contains information about inflation dynamics that headline inflation alone cannot reveal.

Our findings point to several directions for future work. One is monetary policy. If the same shock propagates differently depending on the distribution, then the state of price gaps should matter for policy design. Our simple policy exercise points in this direction, and characterizing the optimal full policy path and how the distribution should enter policy rules is a natural next step. The methodology itself also has broader reach. Nothing in the state-space approach is specific to this simple pricing model. Within pricing, our algorithm could be extended to richer menu cost models with aggregate uncertainty or strategic complementarities. Beyond pricing, the same approach could be adapted to other settings with lumpy adjustment and latent targets, such as investment, durable goods purchases, and inventory management.

References

Alvarez, Fernando, Hervé Le Bihan, and Francesco Lippi. 2016. "The Real Effects of Monetary Shocks in Sticky Price Models: A Sufficient Statistic Approach." *American Economic Review* 106 (10): 2817 – 2851.

Alvarez, Fernando E., Francesco Lippi, and Luigi Paciello. 2011. "Optimal Price Setting with Observation and Menu Costs." *Quarterly Journal of Economics* 126 (4): 1909–1960.

Alvarez, Fernando, Andrea Ferrara, Erwan Gautier, Hervé Le Bihan, and Francesco Lippi. 2025. "Empirical Investigation of a Sufficient Statistic for Monetary Shocks." *The Review of Economic Studies* 92 (4): 2165–2196.

Alvarez, Fernando, and Francesco Lippi. 2014. "Price Setting With Menu Cost for Multiproduct Firms." *Econometrica* 82 (1): 89–135.

Andrews, Isaiah. 2018. "Valid Two-Step Identification-Robust Confidence Sets for GMM." *Review of Economics and Statistics* 100 (2): 337–348.

Andrews, Isaiah, and Bas Sanders. 2026. "The Incredible Flexibility of Moment Matching." Working Paper35378, National Bureau of Economic Research.

Ascari, Guido, and Timo Haber. 2022. "Non-Linearities, State-Dependent Prices and the Transmission Mechanism of Monetary Policy." *The Economic Journal* 132 (641): 37–57.

Atkeson, Andrew, and Lee E. Ohanian. 2001. "Are Phillips Curves Useful for Forecasting Inflation?" *Federal Reserve Bank of Minneapolis Quarterly Review* 25 (1): 2–11.

Auclert, Adrien, Rodolfo Rigato, Matthew Rognlie, and Ludwig Straub. 2024. "New Pricing Models, Same Old Phillips Curves?" *The Quarterly Journal of Economics* 139 (1): 121–186.

Bachmann, Rüdiger, Ricardo J. Caballero, and Eduardo M. R. A. Engel. 2013. "Aggregate Implications of Lumpy Investment: New Evidence and a DSGE Model." *American Economic Journal: Macroeconomics* 5 (4): 29–67.

Baley, Isaac, and Andrés Blanco. 2021. "Aggregate Dynamics in Lumpy Economies." *Econometrica* 89 (3): 1235–1264.

Ball, Laurence, and N. Gregory Mankiw. 1994. "Asymmetric Price Adjustment and Economic Fluctuations." *The Economic Journal* 104 (423): 247–261.

Barnichon, Régis, and Christian Matthes. 2016. "Gaussian Mixture Approximations of Impulse Responses and the Nonlinear Effects of Monetary Shocks." Working Paper16-08, Federal Reserve Bank of Richmond.

Barro, Robert J. 1972. "A Theory of Monopolistic Price Adjustment." *Review of Economic Studies* 39 (1): 17–26.

Beck, Günter W., Kai Carstensen, Jan-Oliver Menz, Richard Schnorrenberger, and Elisabeth Wieland. 2024. "Nowcasting Consumer Price Inflation Using High-Frequency Scanner Data: Evidence from Germany." Working Paper Series2930, European Central Bank.

Berger, David, and Joseph Vavra. 2015. "Consumption Dynamics During Recessions." *Econometrica* 83 (1): 101–154.

Bils, Mark, Peter J. Klenow, and Benjamin A. Malin. 2012. "Reset Price Inflation and the Impact of Monetary Policy Shocks." *American Economic Review* 102 (6): 2798–2825.

Blanco, Andrés, Corina Boar, Callum Jones, and Virgiliu Midrigan. 2024a. "Nonlinear Dynamics in Menu Cost Economies? Evidence from U.S. Data." NBER Working Paper 32748.

Blanco, Andrés, Corina Boar, Callum Jones, and Virgiliu Midrigan. 2024b. "Nonlinear Inflation Dynamics in Menu Cost Economies." NBER Working Paper 32094.

Blanco, Andrés. 2021. "Optimal Inflation Target in an Economy with Menu Costs and a Zero Lower Bound." *American Economic Journal: Macroeconomics* 13 (3): 108–141.

Braun, Robin, Silvia Miranda-Agrippino, and Tuli Saha. 2025. "Measuring Monetary Policy in the UK: The UK Monetary Policy Event-Study Database." *Journal of Monetary Economics* 149 103645.

Bunn, Philip, and Colin Ellis. 2012. "Examining the Behaviour of Individual UK Consumer Prices." *The Economic Journal* 122 (558): F35 – F55.

Caballero, Ricardo, and Eduardo Engel. 2007. "Price Stickiness in Ss Models: New Interpretations of Old Results." *Journal of Monetary Economics* 54 (Supplement): 100–121.

Caballero, Ricardo, Eduardo Engel, and John Haltiwanger. 1995. "Plant-Level Adjustment and Aggregate Investment Dynamics." *Brookings Papers on Economic Activity* 2.

Caballero, Ricardo J., and Eduardo M. R. A. Engel. 1993. "Microeconomic Adjustment Hazards and Aggregate Dynamics." *Quarterly Journal of Economics* 108 (2): 359–383.

Carter, C.K., and R. Kohn. 1994. "On Gibbs sampling for state space models." *Biometrika* 81 (3): 541–553.

Cesa-Bianchi, Ambrogio, Gregory Thwaites, and Alejandro Vicondoa. 2019. "Monetary Policy Transmission in the United Kingdom: A High Frequency Identification Approach." October, Working Paper.

Chen, Catherine, Chen Gao, Jonathon Hazell, Lihua Lei, and Chen Lian. 2026. "Forecasting Inflation with Microdata: An Adaptive Machine Learning Approach." *arXiv preprint arXiv:2607.12345*.

Cover, James Peery. 1992. "Asymmetric Effects of Positive and Negative Money-Supply Shocks." *The Quarterly Journal of Economics* 107 (4): 1261–1282.

Debortoli, Davide, Mario Forni, Luca Gambetti, and Luca Sala. 2025. "Nonlinear Monetary Policy Tradeoffs." *The Economic Journal*, Advance online publication.

Falck, Elisabeth, Mathias Hoffmann, and Patrick Hürtgen. 2021. "Disagreement about Inflation Expectations and Monetary Policy Transmission." *Journal of Monetary Economics* 118 15–31.

Faust, Jon, and Jonathan H. Wright. 2013. "Forecasting Inflation." In *Handbook of Economic Forecasting*, edited by Graham Elliott, and Allan Timmermann, Volume 2A. 2–56, Elsevier.

Frühwirth-Schnatter, Sylvia. 1994. "Data Augmentation and Dynamic Linear Models." *Journal of Time Series Analysis* 15 (2): 183–202.

Gagliardone, Luca, Mark Gertler, Simone Lenzu, and Joris Tielens. 2025. "Micro and Macro Cost-Price Dynamics in Normal Times and During Inflation Surges." NBER Working Paper 33478.

Gautier, Erwan, and Hervé Le Bihan. 2022. "Shocks versus Menu Costs: Patterns of Price Rigidity in an Estimated Multisector Menu-Cost Model." *Review of Economics and Statistics* 104 (4): 668–685.

Gertler, Mark, and Peter Karadi. 2015. "Monetary Policy Surprises, Credit Costs, and Economic Activity." *American Economic Journal: Macroeconomics* 7 (1): 44–76.

Geweke, John, and Hisashi Tanizaki. 2001. "Bayesian Estimation of State-Space Models Using the Metropolis-Hastings Algorithm within Gibbs Sampling." *Computational Statistics & Data Analysis* 37 (2): 151–170.

Golosov, Mikhail, and Robert E. Lucas. 2007. "Menu Costs and Phillips Curves." *Journal of Political Economy* 115 (2): 171–199.

Hong, Gee Hee, Matthew Klepacz, Ernesto Pasten, and Raphael Schoenle. 2023. "The Real Effects of Monetary Shocks: Evidence from Micro Pricing Moments." *Journal of Monetary Economics* 139 1–20.

Jordà, Òscar, Moritz Schularick, and Alan M. Taylor. 2020. "The Effects of Quasi-Random Monetary Experiments." *Journal of Monetary Economics* 112 22–40.

Känzig, Diego R. 2021. "The Macroeconomic Effects of Oil Supply News: Evidence from OPEC Announcements." *American Economic Review* 111 (4): 1092–1125.

Karadi, Peter, Anton Nakov, Galo Nuño, Ernesto Pastén, and Dominik Thaler. 2025. "Strike While the Iron is Hot: Optimal Monetary Policy with a Nonlinear Phillips Curve." ECB Working Paper 3068.

Karadi, Peter, and Adam Reiff. 2019. "Menu Costs, Aggregate Fluctuations, and Large Shocks." *American Economic Journal: Macroeconomics* 11 (3): 111–146.

Kitagawa, Genshiro. 1987. "Non-Gaussian State-Space Modeling of Nonstationary Time Series." *Journal of the American Statistical Association* 82 (400): 1032–1041.

Kramer, Stuart C., and Harold W. Sorenson. 1988. "Recursive Bayesian Estimation Uusing Piecewise Constant Approximations." *Automatica* 24 (6): 789–801.

Lewis, Daniel J., and Karel Mertens. 2026. "A Robust Test for Weak Instruments for 2SLS with Multiple Endogenous Regressors." *The Review of Economic Studies*.

Nakamura, Emi, and Jón Steinsson. 2010. "Monetary Non-Neutrality in a Mutisector Menu Cost Model." *The Quarterly Journal of Economics* 125 (3): 961–1013.

Office for National Statistics. 2019. *Consumer Prices Indices Technical Manual, 2019*. Office for National Statistics, Newport, UK.

ONS. 2014. *Consumer Price Indices: Technical Manual*. Ofiice for National Statistics, 2014th edition.

Powell, Ben, Guy Nason, Duncan Elliott, Matthew Mayhew, Jennifer Davies, and Joe Winton. 2018. "Tracking and Modelling Prices Using Web-Scraped Price Microdata: Towards Automated Daily Consumer Price Index Forecasting." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 181 (3): 737–756.

Robbins, Herbert, and Sutton Monro. 1951. "A Stochastic Approximation Method." *Annals of Mathematical Statistics* 22 (3): 400–407.

Roberts, Gareth O., Andrew Gelman, and Walter R. Gilks. 1997. "Weak Convergence and Optimal Scaling of Random Walk Metropolis Algorithms." *The Annals of Applied Probability* 7 (1): 110–120.

Roberts, Gareth O., and Jeffrey S. Rosenthal. 2001. "Optimal Scaling for Various Metropolis-Hastings Algorithms." *Statistical Science* 16 (4): 351–367.

Stokey, Nancy L. 2009. *The Economics of Inaction: Stochastic Control Models with Fixed Costs*. 1st edition, Princeton University Press.

Tenreyro, Silvana, and Gregory Thwaites. 2016. "Pushing on a String: US Monetary Policy Is Less Powerful in Recessions." *American Economic Journal: Macroeconomics* 8 (4): 43–74.

Vavra, Joseph. 2014. "Inflation Dynamics and Time-Varying Volatility: New Evidence and an Ss Interpretation." *The Quarterly Journal of Economics* 129 (1): 215–258.

Appendix to

Price Gaps and Inflation Dynamics

by Miguel Bandeira, Laura Castillo-Martínez and Shiyuan Wang

A Model Appendix

This appendix provides the theoretical foundations for the reduced-form model used in the main text. Appendix A.1 presents the microfoundations: the household and firm problems, the law of motion for the desired price, the firm's pricing problem, and aggregation. Appendix A.2 establishes conditions under which the optimal pricing policy takes the *Ss* form. Appendix A.3 maps the desired price notation used in Appendices A.1–A.2 to the reset price notation used in the main text. Appendix A.4 describes the monetary policy exercise and welfare criterion. Appendix A.5 provides the proof of Proposition 1.

Notation. Appendices A.1, A.2, and A.4 work with the *desired price gap* $x_{i,t} \equiv p_{i,t} - p^*_{i,t}$, where $p^*_{i,t}$ is the firm's static profit-maximizing log price. The cross-sectional distribution of $x_{i,t}$ is denoted H_t with density h_t. Appendix A.3 establishes the connection to the *reset price gap* $s_{i,t} \equiv p_{i,t} - p^r_{i,t}$ and the distribution G_t used in the main text.

A.1 Microfoundations

Households. A representative household maximizes expected discounted lifetime utility:

$$\mathbf{E}_0 \sum_{t=0}^{\infty} \beta^t \left[\log C_t - N_t\right],$$

where C_t is consumption, N_t is labor supply, and $\beta \in (0, 1)$ is the discount factor. Consumption is a CES aggregate of differentiated varieties:

$$C_t = \left(\int_0^1 c_t(i)^{\frac{\epsilon-1}{\epsilon}} \, di \right)^{\frac{\epsilon}{\epsilon-1}},$$

where $\epsilon > 1$ is the elasticity of substitution. The household faces a cash-in-advance constraint:

$$P_t C_t = M_t,$$

where P_t is the aggregate price index and M_t is aggregate nominal spending, which is controlled by monetary policy.

The household's optimality conditions yield a labor supply condition and a demand function for each variety:

$$W_t = M_t, \tag{22}$$

$$c_t(i) = \left(\frac{P_t(i)}{P_t}\right)^{-\epsilon} C_t, \tag{23}$$

where (22) combines the intratemporal optimality condition $W_t = P_t C_t$ with the cash-in-advance constraint.

Firms. Each firm $i \in [0,1]$ produces a differentiated variety using labor with a linear technology:

$$y_t(i) = A_t(i) \cdot N_t(i),$$

where $A_t(i)$ is firm productivity. It has an aggregate and an idiosyncratic component, $\log A_t(i) = \log A_t + \log \tilde{A}_t(i)$, both of which follow random walks:

$$\log A_t = \log A_{t-1} + a_t, \qquad \log \tilde{A}_t(i) = \log \tilde{A}_{t-1}(i) + a_{i,t}, \quad a_{i,t} \overset{iid}{\sim} \mathcal{N}(0, \sigma_{a,t}^2),$$

with a_t common across firms and $a_{i,t}$ independently distributed across firms and over time. The volatility of idiosyncratic shocks $\sigma_{a,t}$ is itself allowed to vary over time. The nominal marginal cost of firm i is therefore:

$$MC_t(i) = \frac{W_t}{A_t(i)} = \frac{M_t}{A_t(i)},$$

where the second equality uses (22). The static profit-maximizing (desired) price is:

$$P_t^*(i) = \frac{\epsilon}{\epsilon - 1} \cdot \frac{M_t}{A_t(i)}, \tag{24}$$

where $\epsilon/(\epsilon - 1)$ is the constant desired markup over marginal cost.

Desired price dynamics. Taking logs of (24) and first-differencing yields:

$$p_t^*(i) - p_{t-1}^*(i) = \Delta \log M_t - a_t - a_{i,t}.$$

Define the common drift of the desired price as the growth in nominal spending net of aggregate productivity growth, $\nu_t \equiv \Delta \log M_t - a_t$, and the reduced-form idiosyncratic shock as $\varepsilon_{i,t} \equiv -a_{i,t}$. Then the above equation becomes:

$$p_t^*(i) = \nu_t + p_{t-1}^*(i) + \varepsilon_{i,t}, \quad \varepsilon_{i,t} \sim \mathcal{N}(0, \sigma_{\varepsilon,t}^2), \tag{25}$$

where $\sigma_{\varepsilon,t} = \sigma_{a,t}$.

The firm's pricing problem. In each period, firm i chooses whether to adjust its price $P_t(i)$, trading off the gain from realigning with the desired price against the random menu cost.

Using the demand function (23), the production technology, and the desired price (25), the firm's normalized real profit in period t can be decomposed as:

$$\frac{\Pi_t(i)}{P_t C_t} = B_t(i) \cdot \tilde{\pi}(x_{i,t}), \tag{26}$$

where $x_{i,t} \equiv p_{i,t} - p^*_{i,t}$ is the desired price gap,

$$\tilde{\pi}(x) \equiv e^{(1-\epsilon)x} - \frac{1}{\mathcal{M}} e^{-\epsilon x} \tag{27}$$

is the profit expressed purely as a function of the gap with $\mathcal{M} \equiv \epsilon/(\epsilon-1)$ denoting the desired markup, and $B_t(i) \equiv (P^*_t(i)/P_t)^{1-\epsilon} > 0$ is a time-varying weight that depends on aggregate conditions and the firm's idiosyncratic productivity but not on its pricing decision.[40]

Following Alvarez and Lippi (2014), we assume that the resource cost of price adjustment is proportional to the firm's desired profit $\tilde{\pi}(0) \cdot B_t(i) \cdot P_t C_t = \frac{1}{\epsilon} B_t(i) \cdot P_t C_t$. The normalized adjustment cost is therefore $B_t(i) \cdot c_{i,t}$, where $c_{i,t} \in \{0, \hat{c}\}$ with $\Pr(c_{i,t} = 0) = \lambda_t$.[41]

The firm minimizes expected discounted profit losses net of adjustment costs:

$$\min_{\{x_{i,t}\}_{t=0}^{\infty}} \mathbf{E}_0 \sum_{t=0}^{\infty} \beta^t B_t(i) \left[L(x_{i,t}) + c_{i,t} \cdot \mathbf{1}_{\{x_{i,t} \neq x_{i,t-1} - \nu_t - \varepsilon_{i,t}\}} \right], \tag{28}$$

where $L(x) \equiv \tilde{\pi}(0) - \tilde{\pi}(x)$ is the per-period profit loss from maintaining a desired price gap x instead of the optimal gap 0.

Reduction to the quadratic formulation. By construction, $x = 0$ maximizes $\tilde{\pi}$, so $\tilde{\pi}'(0) = 0$. A second-order Taylor expansion of the profit loss around the optimal gap gives:

$$L(x) = \tilde{\pi}(0) - \tilde{\pi}(x) \approx \frac{\epsilon - 1}{2} x^2. \tag{29}$$

Substituting into (28), the firm's problem becomes:

$$\min_{\{x_{i,t}\}_{t=0}^{\infty}} \mathbf{E}_0 \sum_{t=0}^{\infty} \beta^t B_t(i) \left[\frac{\epsilon - 1}{2} x_{i,t}^2 + c_{i,t} \cdot \mathbf{1}_{\{x_{i,t} \neq x_{i,t-1} - \nu_t - \varepsilon_{i,t}\}} \right]. \tag{30}$$

To obtain the canonical quadratic pricing problem, we abstract from variation in the multiplicative scale factor $B_t(i)$ across adjacent periods and approximate

$$\frac{B_{t+1}(i)}{B_t(i)} \simeq 1.$$

[40] The derivation of (26) proceeds by writing $P_t(i)/P_t = (P^*_t(i)/P_t)\, e^{x_{i,t}}$, substituting into the profit expression, and using $MC_t(i) = P^*_t(i)/\mathcal{M}$ to simplify.

[41] A similar scaling of fixed adjustment costs by firm size is standard in the lumpy investment literature (Baley and Blanco, 2021).

Dividing by $B_t(i)(\epsilon-1)$ then yields a problem with flow loss $\frac{1}{2}x_{i,t}^2$, menu cost $\bar{c} \equiv \hat{c}/(\epsilon-1)$, and discount factor β.

Bellman equation and ***Ss*** **rule.** Under this approximation, the rescaled value function $V_t(x) \equiv J_t(x)/[B_t(i)(\epsilon-1)]$ satisfies

$$V_t(x) = \frac{1}{2}\,x^2 + \beta\,\mathbf{E}_t\left[(1-\lambda_{t+1})\min\Big(V_{t+1}(x-\nu_{t+1}-\varepsilon),\ \bar{c}+\min_{x'} V_{t+1}(x')\Big) + \lambda_{t+1}\min_{x'} V_{t+1}(x')\right], \tag{31}$$

where $\varepsilon \sim \mathcal{N}(0, \sigma^2_{\varepsilon,t+1})$. In a stationary environment with constant $(\nu_t, \sigma_{\varepsilon,t}, \lambda_t)$, the value function is time invariant.[42]

As discussed in Appendix A.2, the solution to (31) takes the form of an *Ss* rule: for each drift ν_t, there exist thresholds $\underline{x}_t < x_t^* < \bar{x}_t$ such that the firm resets its desired price gap to x_t^* whenever it falls outside $[\underline{x}_t, \bar{x}_t]$ or when adjustment is free.

Aggregation. Let H_t denote the cross-sectional distribution of desired price gaps $x_{i,t} \equiv p_{i,t} - p_{i,t}^*$ across firms, with density h_t. Given the *Ss* pricing rule, H_t evolves according to:

$$H_t = \mathcal{T}(H_{t-1};\, \nu_t, \sigma_{\varepsilon,t}, \underline{x}_t, \bar{x}_t, x_t^*),$$

where $\mathcal{T}$ maps the distribution from period $t-1$ to period t, given the drift, shock volatility, and the parameters of the pricing rule. The CES price index implies the consistency condition:

$$1 = \int_0^1 B_t(i)\, e^{(1-\epsilon)x_{i,t}}\, di,$$

where $B_t(i) = (P_t^*(i)/P_t)^{1-\epsilon}$ is the firm-specific weight from (26). A first-order log-linearization of the CES price index around the symmetric allocation gives

$$\log P_t \simeq \int_0^1 p_{i,t}\, di.$$

Using

$$p_{i,t} = p_{i,t}^* + x_{i,t},$$

and the desired-price process in (25), aggregate inflation therefore satisfies, to first order,

$$\pi_t \equiv \Delta \log P_t \simeq \nu_t + \int_0^1 (x_{i,t} - x_{i,t-1})\ di,$$

where ν_t is the common drift of the desired price. Finally, the aggregate resource constraint is:

$$N_t = \frac{1}{\mathcal{M}} \int_0^1 B_t(i)\, e^{-\epsilon x_{i,t}}\, di + \hat{c}\int_0^1 B_t(i)\, \mathbf{1}_{\{c_{i,t}=\hat{c},\ \text{adjust}\}}(i)\, di,$$

[42] If changes in $(\nu_t, \sigma_{\varepsilon,t}, \lambda_t)$ are unanticipated, the firm treats the current environment as permanent and the time-invariant Bellman equation applies conditional on it. When the firm instead anticipates the future path of these parameters, V_t is genuinely non-stationary; see the non-stationary case in Appendix A.2.

where the first term is aggregate production labor and the second is the real resource cost of price adjustment, with $\hat{c}$ denoting the structural menu cost. The resource cost is incurred only by firms that draw a positive menu cost ($c_{i,t} = \hat{c}$, probability $1 - \lambda_t$) and whose gap falls outside the inaction region; firms receiving free adjustment opportunities pay no resource cost.

A.2 The *Ss* pricing rule

This appendix establishes conditions under which the optimal pricing policy for (31) takes the *Ss* form. The argument follows the approach in Auclert et al. (2024) (Online Appendix B.2), who prove the result analytically for $\nu = 0$, and extends it to $\nu > 0$ via a quasiconvexity condition.

A.2.1 Stationary case

Consider the stationary version of (31) with constant drift ν and discount factor β:

$$V(x) = \frac{1}{2}x^2 + \beta\Big\{(1-\lambda)\,\mathbf{E}\Big[\min\big(V(x-\nu-\varepsilon),\ \bar{c} + V(x^*)\big)\Big] + \lambda\, V(x^*)\Big\}, \qquad (32)$$

where $\varepsilon \sim N(0, \sigma_\varepsilon^2)$ and $x^* = \arg\min V$. The Bellman operator is a contraction (modulus β), so V exists and is unique. Standard arguments yield the following properties:

(i) $V(x) \geq \frac{1}{2}x^2$ for all x, and $V(x^*) \leq \beta(1-\lambda)\bar{c}/(1-\beta)$.

(ii) Setting $M = \sqrt{2\bar{c}/(1-\beta)}$, the firm strictly prefers to adjust for $|x| > M$, so it suffices to track V on $[-M, M]$.

(iii) Since the expectation in (32) is a convolution with a Gaussian kernel, V is C^∞ whenever $\sigma_\varepsilon > 0$.

(iv) Differentiating under the integral gives

$$V'(x) = x + \beta(1-\lambda)\int_{\underline{x}}^{\bar{x}} V'(y)\, f(x-\nu-y)\, dy, \qquad (33)$$

where f is the density of ε and $[\underline{x}, \bar{x}]$ is the inaction set. Since the integral is bounded, $V'(x)/x \to 1$ as $|x| \to \infty$, so V has at least one interior minimizer.

Quasiconvexity implies *Ss*. Say V is *strictly quasiconvex* if it has a unique minimizer x^* and is strictly decreasing on $(-\infty, x^*]$, strictly increasing on $[x^*, \infty)$. If V is strictly quasiconvex, then every sublevel set $\{x : V(x) \leq c\}$ is a closed interval. In particular, the inaction set $\mathcal{I} = \{x : V(x) \leq \bar{c} + V(x^*)\}$ is a closed interval $[\underline{x}, \bar{x}]$ with $\underline{x} < x^* < \bar{x}$ (since $\bar{c} > 0$). Outside this interval the firm strictly prefers to pay $\bar{c}$ and reset to the unique minimizer x^*; inside, it weakly prefers inaction. This is the *Ss* rule.

Note that quasiconvexity is weaker than convexity: V can have $V'' < 0$ near the thresholds, as long as it remains monotone on each side of x^*.

Verification. When $\nu = 0$, Auclert et al. (2024) show that V is symmetric ($V(x) = V(-x)$) and $V'(x) > 0$ for $x > 0$, which gives strict quasiconvexity with $x^* = 0$. When $\nu > 0$, symmetry breaks and the direct analytical proof is not available. Quasiconvexity can be checked ex-post from the computed value function.

A.2.2 Non-stationary case

Under MIT shocks to $(\nu_t, \sigma_{\varepsilon,t}, \lambda_t)$, the properties (i)–(iv) and the quasiconvexity-implies-Ss argument apply to V_t at each date t, treating the continuation value V_{t+1} as given and replacing $(\nu, \sigma_\varepsilon, \lambda)$ with $(\nu_{t+1}, \sigma_{\varepsilon,t+1}, \lambda_{t+1})$. Hence the Ss structure holds at date t whenever V_t is strictly quasiconvex.

If the shock is permanent, V_t is simply the stationary fixed point at the current environment, and verification reduces to the stationary case. If the environment reverts along a deterministic path, the firm anticipates future changes and V_t is genuinely non-stationary. In this case V_t is obtained by backward induction from the terminal stationary solution V_∞, and quasiconvexity should be checked at each backward step.

A.3 Connection to the main text

The main text works with the *reset price* and the *reset price gap*, rather than the desired price and its gap. This appendix establishes the mapping and derives equations (1)–(3) and (4) of the main text.

Reset price. When a firm adjusts, it sets its desired price gap to the optimal reset point $x_t^* \equiv \arg\min V_t$, which is generically nonzero under positive drift, though quantitatively small ($x^* \approx 0.013$ under the calibration in Table A.1). The price the firm chooses upon adjustment is therefore $p^r_{i,t} \equiv p^*_{i,t} + x^*_t$, which we call the *reset price*. First-differencing and using (25):

$$p^r_{i,t} - p^r_{i,t-1} = \underbrace{(\nu_t + \varepsilon_{i,t})}_{\Delta p^*_{i,t}} + \Delta x^*_t = \mu_t + \varepsilon_{i,t},$$

where

$$\mu_t \;\equiv\; \nu_t + \Delta x^*_t \tag{34}$$

is the *reset price drift*—the drift used throughout the main text. This is equation (1). The reset price drift μ_t differs from the desired price drift ν_t whenever the optimal reset point x^*_t changes. Under the unexpected one-off shocks considered in Figure 1, this change is negligible, so that $\mu_t \approx \nu_t$.

Reset-price gap and the firm's problem. Define the *reset price gap* $s_{i,t} \equiv p_{i,t} - p^r_{i,t} = x_{i,t} - x^*_t$. In the absence of adjustment, the gap evolves as:

$$s^{\text{na}}_{i,t} = \underbrace{x_{i,t-1} - \nu_t - \varepsilon_{i,t}}_{\text{no-adjustment } x_{i,t}} - x^*_t = s_{i,t-1} + x^*_{t-1} - \nu_t - \varepsilon_{i,t} - x^*_t = s_{i,t-1} - \mu_t - \varepsilon_{i,t},$$

using $x_{i,t-1} = s_{i,t-1} + x^*_{t-1}$ and $\mu_t = \nu_t + \Delta x^*_t$. The flow loss in the rescaled problem transforms as

$$\frac{1}{2}x_{i,t}^2 = \frac{1}{2}(s_{i,t} + x^*_t)^2 = \frac{1}{2}s_{i,t}^2 + x^*_t s_{i,t} + \frac{1}{2}(x^*_t)^2.$$

The final term is constant with respect to the firm's pricing decision. For the reduced-form representation used in the main text, we also neglect the linear term $x^*_t s_{i,t}$, which is small under our calibration ($x^*_t \approx 0.013$). This yields the centered quadratic loss $\frac{1}{2}s_{i,t}^2$ in equation (2). Importantly, this approximation is not used to determine the reset point. By definition, $s_{i,t} \equiv x_{i,t} - x^*_t$, so an adjustment to the optimal desired-price gap x^*_t corresponds exactly to resetting the reset-price gap to zero.

Ss **rule.** The *Ss* rule resets the desired-price gap to x^*_t. Since $s_{i,t} \equiv x_{i,t} - x^*_t$, this corresponds exactly to resetting the reset-price gap to zero: $s_{i,t} = 0$. The inaction region becomes $[\underline{s}_t, \bar{s}_t]$ with

$$\underline{s}_t \equiv \underline{x}_t - x^*_t < 0 < \bar{x}_t - x^*_t \equiv \bar{s}_t.$$

This is equation (3) in the main text.

Distribution. The cross-sectional distribution of reset price gaps, denoted G_t with density g_t in the main text, is H_t shifted by x^*_t:

$$g_t(s) = h_t(s + x^*_t).$$

In particular, $E[s_t] = E[x_t] - x^*_t$ and $\mathrm{Var}(s_t) = \mathrm{Var}(x_t)$.

Price misalignment. A second-order expansion of household welfare around the flexible price allocation implies that the welfare loss associated with price misalignment is proportional to $\frac{1}{2}E[x_{i,t}^2]$. Using $x_{i,t} = s_{i,t} + x^*_t$ and the fact that x^*_t is small under our calibration, this is well approximated by

$$\frac{1}{2}E[s_{i,t}^2] = \frac{1}{2}\left[\mathrm{Var}(s_{i,t}) + E[s_{i,t}]^2\right],$$

which corresponds to equation (4) in the main text.

A.4 Monetary policy exercise

For the policy exercise, we abstract from firm-level variation in $B_t(i)$ and set $B_t(i) = B_t$ for all firms. Let

$$Q_t \equiv \int e^{(1-\epsilon)x} dH_t(x), \qquad D_t \equiv \int e^{-\epsilon x} dH_t(x),$$

where H_t is the cross-sectional distribution of desired price gaps. The price-index condition above then implies $B_t Q_t = 1$, so $B_t = Q_t^{-1}$.

We allow for a constant employment subsidy and define

$$\kappa \equiv \frac{\epsilon}{\epsilon - 1}(1 - \tau).$$

Using the cash-in-advance constraint and normalizing aggregate productivity to one, aggregate consumption and labor are

$$C_t = \frac{Q_t^{1/(\epsilon-1)}}{\kappa}, \qquad N_t = \frac{D_t}{\kappa Q_t} + \frac{\hat{c}}{Q_t}(f_t - \lambda),$$

where f_t is the fraction of firms that adjust and $\hat{c} = (\epsilon - 1)\bar{c}$ is the structural menu cost. Period utility is therefore

$$u_t = \log C_t - \frac{D_t}{\kappa Q_t} - \frac{\hat{c}}{Q_t}(f_t - \lambda).$$

We choose the employment subsidy so that the consumption-production margin has no first-order distortion at the calibrated ergodic steady state. This requires

$$\kappa = \frac{\bar{D}}{\bar{Q}}$$

where bars denote values under the ergodic distribution.

The economy enters period $t = 1$ with initial distribution G_0 and is hit by the one-off nominal spending growth shock described in the main text. Nominal spending growth is predetermined at $t = 1$. After observing the shock, the central bank chooses its deviation from baseline in period $t = 2$, which firms anticipate when making their $t = 1$ pricing decisions. Nominal spending growth returns to baseline from $t = 3$ onward. For each G_0, we choose the period-2 intervention to maximize

$$\sum_{t=1}^{\infty} \beta^{t-1} u_t,$$

subject to the firm's pricing problem and the law of motion for the gap distribution.

A.5 Proof of Proposition 1

Proof. Part (i). Define

$$B_h \equiv [\underline{s} + (h-1)\mu, \underline{s} + h\mu).$$

For $1 \leq h \leq |\underline{s}|/\mu$, no firm adjusting after date t can reach the lower threshold a second time within the relevant horizon: upon adjustment a firm resets its gap to zero, and from there reaching $\underline{s}$ takes more than $|\underline{s}|/\mu$ periods. A firm with gap s at date t that has not yet adjusted has gap $s - h\mu$ at date $t + h$. It therefore crosses the lower threshold at $t + h$ if and only if

$$s - h\mu < \underline{s} \leq s - (h-1)\mu,$$

or equivalently $s \in B_h$.

Upon adjustment, a firm with initial gap $s \in B_h$ raises its price by $h\mu - s$. Non-adjusting firms leave their prices unchanged. Hence aggregate inflation at $t+h$ is

$$\pi_{t+h} = \int_{B_h} (h\mu - s)\, dG_t(s).$$

If G_t admits a smooth density g_t over B_h, the mass of firms in this band is

$$\int_{B_h} g_t(s)\, ds = g_t\big(\underline{s} + (h-1)\mu\big)\,\mu + O(\mu^2).$$

Moreover, for $s \in B_h$,

$$h\mu - s = |\underline{s}| + O(\mu).$$

Combining these expressions yields

$$\pi_{t+h} = |\underline{s}|\, \mu\, g_t\big(\underline{s} + (h-1)\mu\big) + O(\mu^2),$$

which proves (7).

Part (ii). Let

$$T \equiv \frac{\bar{s} - \underline{s}}{\mu}$$

denote the transit time across the inaction region, and fix $K < T$. Since

$$\bar{s} - K\mu > \underline{s},$$

we can choose two points

$$\underline{s} < s_1 < s_2 < \min\{\underline{s} + \mu,\, \bar{s} - K\mu\}.$$

Consider two economies with identical parameters whose gap distributions at date $t-K$ differ only in that the second relocates a small mass $\delta > 0$ from $s_2 + K\mu$ to $s_1 + K\mu$. The same construction can be implemented over arbitrarily small neighborhoods of these points.

Over the following K periods, the gaps of the relocated firms drift deterministically from $s_j + K\mu$ to s_j, for $j = 1, 2$. By construction, they remain inside the inaction region throughout and therefore never adjust. All other firms behave identically across the two economies. The sequence of price adjustments, and hence aggregate inflation, is therefore identical over the K-period window:

$$\pi_{t-k} = \pi'_{t-k}, \qquad k = 0, \ldots, K-1.$$

Nevertheless, the gap distributions at date t differ, since the second economy has additional mass at s_1 and correspondingly less mass at s_2.

Because s_1 and s_2 both lie in

$$B_1 = [\underline{s},\, \underline{s} + \mu),$$

both groups cross the lower threshold at $t+1$. Their price increases, however, differ. The mass relocated from s_2 to s_1 increases inflation at $t+1$ by

$$\pi'_{t+1} - \pi_{t+1} = \delta\big[(\mu - s_1) - (\mu - s_2)\big] = \delta(s_2 - s_1) > 0.$$

Thus the two economies have identical finite inflation histories but different current gap distributions and different future inflation, proving part (ii). $\square$

B State Space Methods Appendix

B.1 Forward Filtering Backward Sampling: Proof

Proof. We begin by establishing the conditional-independence properties implied by the state-space representation, and then derive the forward-filtering and backward-sampling recursions. To simplify notation, define

$$\mathbf{x}_t^- \equiv \{\mathbf{x}_{i,t}\}_{i\in\mathcal{I}_{\mathcal{W},t}^-}, \qquad \mathbf{x}_{t-1}^- \equiv \{\mathbf{x}_{i,t-1}\}_{i\in\mathcal{I}_{\mathcal{W},t}^-}, \qquad \mathbf{p}_{t-1}^- \equiv \{\mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t-1}\}_{i\in\mathcal{I}_{\mathcal{W},t}^-}.$$

Conditional-independence properties. First, for price trajectories that continue from $t-1$ to t, the state transition density factorizes across i:

$$f(\mathbf{x}_t^- \mid \mathbf{x}_{t-1}^-, \mathbf{p}_{t-1}^-, \boldsymbol{\theta}_{\mathcal{W}}) = \prod_{i\in\mathcal{I}_{\mathcal{W},t}^-} f(\mathbf{x}_{i,t} \mid \mathbf{x}_{i,t-1}, \boldsymbol{\theta}_{\mathcal{W}}).$$

This follows from equations (10)–(11): conditional on its own lagged state and $\boldsymbol{\theta}_{\mathcal{W}}$, the current state of each price trajectory is independent of all other states and observed prices. Second, the measurement density also factorizes across continuing price trajectories:

$$f(\{p_{i,t}\}_{i\in\mathcal{I}_{\mathcal{W},t}^-} \mid \mathbf{x}_t^-, \mathbf{p}_{t-1}^-, \boldsymbol{\theta}_{\mathcal{W}}) = \prod_{i\in\mathcal{I}_{\mathcal{W},t}^-} f(p_{i,t} \mid \mathbf{x}_{i,t}, p_{i,t-1}, \boldsymbol{\theta}_{\mathcal{W}}).$$

This follows from equations (8)–(9), since conditional on $\mathbf{x}_{i,t}$, $p_{i,t-1}$, and $\boldsymbol{\theta}_{\mathcal{W}}$, the observed price $p_{i,t}$ is independent of all other states and observed prices. Finally, the initial filtered densities factorize across price trajectories:

$$f(\{\mathbf{x}_{i,\underline{t}_{i,\mathcal{W}}}\}_{i\in\mathcal{I}_{\mathcal{W}}} \mid \{p_{i,\underline{t}_{i,\mathcal{W}}}\}_{i\in\mathcal{I}_{\mathcal{W}}}, \boldsymbol{\theta}_{\mathcal{W}}) = \prod_{i\in\mathcal{I}_{\mathcal{W}}} f(\mathbf{x}_{i,\underline{t}_{i,\mathcal{W}}} \mid p_{i,\underline{t}_{i,\mathcal{W}}}, \boldsymbol{\theta}_{\mathcal{W}}),$$

by equation (12).

Forward filtering. Suppose that the filtered density factorizes at date $t-1$. Then, for the subset of price trajectories that continue to date t,

$$f(\mathbf{x}_{t-1}^- \mid \mathbf{p}_{t-1}^-, \boldsymbol{\theta}_{\mathcal{W}}) = \prod_{i\in\mathcal{I}_{\mathcal{W},t}^-} f(\mathbf{x}_{i,t-1} \mid \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t-1}, \boldsymbol{\theta}_{\mathcal{W}}).$$

By the Chapman–Kolmogorov equation,

$$
\begin{aligned}
f(\mathbf{x}_t^- \mid \mathbf{p}_{t-1}^-, \boldsymbol{\theta}_{\mathcal{W}}) &= \int f(\mathbf{x}_t^- \mid \mathbf{x}_{t-1}^-, \mathbf{p}_{t-1}^-, \boldsymbol{\theta}_{\mathcal{W}}) f(\mathbf{x}_{t-1}^- \mid \mathbf{p}_{t-1}^-, \boldsymbol{\theta}_{\mathcal{W}})\, d\mathbf{x}_{t-1}^- \\
&= \prod_{i \in \mathcal{I}_{\mathcal{W},t}^-} \int f(\mathbf{x}_{i,t} \mid \mathbf{x}_{i,t-1}, \boldsymbol{\theta}_{\mathcal{W}}) f(\mathbf{x}_{i,t-1} \mid \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t-1}, \boldsymbol{\theta}_{\mathcal{W}})\, d\mathbf{x}_{i,t-1} \\
&= \prod_{i \in \mathcal{I}_{\mathcal{W},t}^-} f(\mathbf{x}_{i,t} \mid \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t-1}, \boldsymbol{\theta}_{\mathcal{W}}). \qquad (35)
\end{aligned}
$$

This establishes equation (14), with each factor given by the trajectory-level recursion in equation (15). For the filtering update, Bayes' rule gives

$$
\begin{aligned}
f(\mathbf{x}_t^- \mid \{\mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}\}_{i \in \mathcal{I}_{\mathcal{W},t}^-}, \boldsymbol{\theta}_{\mathcal{W}}) &\propto f(\{p_{i,t}\}_{i \in \mathcal{I}_{\mathcal{W},t}^-} \mid \mathbf{x}_t^-, \mathbf{p}_{t-1}^-, \boldsymbol{\theta}_{\mathcal{W}}) f(\mathbf{x}_t^- \mid \mathbf{p}_{t-1}^-, \boldsymbol{\theta}_{\mathcal{W}}) \\
&= \prod_{i \in \mathcal{I}_{\mathcal{W},t}^-} f(p_{i,t} \mid \mathbf{x}_{i,t}, p_{i,t-1}, \boldsymbol{\theta}_{\mathcal{W}}) f(\mathbf{x}_{i,t} \mid \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t-1}, \boldsymbol{\theta}_{\mathcal{W}}). \qquad (36)
\end{aligned}
$$

Since the normalizing constant also factorizes across i, it follows that

$$
f(\mathbf{x}_t^- \mid \{\mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}\}_{i \in \mathcal{I}_{\mathcal{W},t}^-}, \boldsymbol{\theta}_{\mathcal{W}}) = \prod_{i \in \mathcal{I}_{\mathcal{W},t}^-} f(\mathbf{x}_{i,t} \mid \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}, \boldsymbol{\theta}_{\mathcal{W}}),
$$

where each factor satisfies equation (16). To incorporate price trajectories whose first observation occurs at t, define

$$
\mathcal{I}_{\mathcal{W},t}^0 \equiv \mathcal{I}_{\mathcal{W},t} \setminus \mathcal{I}_{\mathcal{W},t}^-.
$$

For each $i \in \mathcal{I}_{\mathcal{W},t}^0$, the filtered density is given by its initial filtered distribution,

$$
f(\mathbf{x}_{i,t} \mid p_{i,t}, \boldsymbol{\theta}_{\mathcal{W}}).
$$

By equation (12) and the conditional-independence properties established above, these densities factorize across entering price trajectories and are independent of those for continuing trajectories. Combining the two sets therefore gives

$$
f(\{\mathbf{x}_{i,t}\}_{i \in \mathcal{I}_{\mathcal{W},t}} \mid \{\mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}\}_{i \in \mathcal{I}_{\mathcal{W},t}}, \boldsymbol{\theta}_{\mathcal{W}}) = \prod_{i \in \mathcal{I}_{\mathcal{W},t}} f(\mathbf{x}_{i,t} \mid \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}, \boldsymbol{\theta}_{\mathcal{W}}).
$$

Since the factorization holds at the initial observation of each price trajectory by equation (12), the result follows recursively for all dates, establishing equation (13).

Backward sampling. Conditional on the observed prices, the posterior kernel for the complete latent-state trajectories can be written as

$$f(\mathbf{x}_{\mathcal{W}} \mid \mathbf{p}_{\mathcal{W}}, \boldsymbol{\theta}_{\mathcal{W}}) \propto \prod_{i \in \mathcal{I}_{\mathcal{W}}} \Bigg[f(\mathbf{x}_{i,\underline{t}_{i,\mathcal{W}}} \mid p_{i,\underline{t}_{i,\mathcal{W}}}, \boldsymbol{\theta}_{\mathcal{W}}) \\ \times \prod_{t=\underline{t}_{i,\mathcal{W}}+1}^{\bar{t}_{i,\mathcal{W}}} f(\mathbf{x}_{i,t} \mid \mathbf{x}_{i,t-1}, \boldsymbol{\theta}_{\mathcal{W}}) f(p_{i,t} \mid \mathbf{x}_{i,t}, p_{i,t-1}, \boldsymbol{\theta}_{\mathcal{W}}) \Bigg]. \tag{37}$$

The first factor is the initial filtered density, while the remaining factors are the state transition and measurement densities. Since the right-hand side is a product across price trajectories, its normalizing constant also factorizes across i. Hence,

$$f(\mathbf{x}_{\mathcal{W}} \mid \mathbf{p}_{\mathcal{W}}, \boldsymbol{\theta}_{\mathcal{W}}) = \prod_{i \in \mathcal{I}_{\mathcal{W}}} f(\mathbf{x}_{i,\mathcal{W}} \mid \mathbf{p}_{i,\mathcal{W}}, \boldsymbol{\theta}_{\mathcal{W}}),$$

which establishes equation (17). For a given price trajectory i, apply the chain rule backward:

$$f(\mathbf{x}_{i,\mathcal{W}} \mid \mathbf{p}_{i,\mathcal{W}}, \boldsymbol{\theta}_{\mathcal{W}}) = f(\mathbf{x}_{i,\bar{t}_{i,\mathcal{W}}} \mid \mathbf{p}_{i,\mathcal{W}}, \boldsymbol{\theta}_{\mathcal{W}}) \times \prod_{t=\underline{t}_{i,\mathcal{W}}}^{\bar{t}_{i,\mathcal{W}}-1} f(\mathbf{x}_{i,t} \mid \mathbf{x}_{i,t+1:\bar{t}_{i,\mathcal{W}}}, \mathbf{p}_{i,\mathcal{W}}, \boldsymbol{\theta}_{\mathcal{W}}). \tag{38}$$

By the Markov property of the latent states and the conditional independence of future observed prices,

$$f(\mathbf{x}_{i,t} \mid \mathbf{x}_{i,t+1:\bar{t}_{i,\mathcal{W}}}, \mathbf{p}_{i,\mathcal{W}}, \boldsymbol{\theta}_{\mathcal{W}}) = f(\mathbf{x}_{i,t} \mid \mathbf{x}_{i,t+1}, \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}, \boldsymbol{\theta}_{\mathcal{W}}),$$

which establishes equation (18). Finally, by Bayes' rule,

$$f(\mathbf{x}_{i,t} \mid \mathbf{x}_{i,t+1}, \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}, \boldsymbol{\theta}_{\mathcal{W}}) = \frac{f(\mathbf{x}_{i,t+1} \mid \mathbf{x}_{i,t}, \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}, \boldsymbol{\theta}_{\mathcal{W}}) f(\mathbf{x}_{i,t} \mid \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}, \boldsymbol{\theta}_{\mathcal{W}})}{f(\mathbf{x}_{i,t+1} \mid \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}, \boldsymbol{\theta}_{\mathcal{W}})}. \tag{39}$$

The state transition equations imply

$$f(\mathbf{x}_{i,t+1} \mid \mathbf{x}_{i,t}, \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}, \boldsymbol{\theta}_{\mathcal{W}}) = f(\mathbf{x}_{i,t+1} \mid \mathbf{x}_{i,t}, \boldsymbol{\theta}_{\mathcal{W}}).$$

Since the denominator does not depend on $\mathbf{x}_{i,t}$,

$$f(\mathbf{x}_{i,t} \mid \mathbf{x}_{i,t+1}, \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}, \boldsymbol{\theta}_{\mathcal{W}}) \propto f(\mathbf{x}_{i,t+1} \mid \mathbf{x}_{i,t}, \boldsymbol{\theta}_{\mathcal{W}}) f(\mathbf{x}_{i,t} \mid \mathbf{p}_{i,\underline{t}_{i,\mathcal{W}}:t}, \boldsymbol{\theta}_{\mathcal{W}}),$$

which establishes equation (19). Equations (15)–(16) and (18)–(19) are the trajectory-level forward-filtering and backward-sampling recursions used by the FFBS algorithm. □

B.2 Grid-based FFBS

This appendix provides the full grid-based FFBS algorithm used to sample the latent states conditional on the state-space parameters and observed prices. By Lemma 1, the filtering and smoothing distributions factorize across price trajectories. We therefore describe the algorithm for a single price trajectory i, with the same procedure applied independently to each $i \in \mathcal{I}_{\mathcal{W}}$.

Grids. For each price trajectory i, denote by

$$\mathbf{p}_{i,\mathcal{W}}^{t} \equiv \{p_{i,\tau} : \tau \in \mathcal{T}_{i,\mathcal{W}},\ \tau \leqslant t\}$$

the observed price history through date $t \in \mathcal{T}_{i,\mathcal{W}}$.

For a given grid rule r, define the grid function $\mathbf{G}^r(\mathbf{p}_{i,\mathcal{W}}^{t}, \boldsymbol{\theta}_{\mathcal{W}}; t)$ as

$$\mathbf{G}^r(\mathbf{p}_{i,\mathcal{W}}^{t}, \boldsymbol{\theta}_{\mathcal{W}}; t) = \begin{cases} \begin{bmatrix} \mathbf{g}^r(p_{i,t} - \bar{s}_{\mathcal{W}}, p_{i,t} - \underline{s}_{\mathcal{W}}, n_r) & \mathbf{0}_{n_r} \\ p_{i,t} & 1 \\ p_{i,t} & 0 \end{bmatrix} & \text{if } t = \underline{t}_{i,\mathcal{W}} \\ \begin{bmatrix} \mathbf{g}^r(p_{i,t} - \bar{s}_{\mathcal{W}}, p_{i,t} - \underline{s}_{\mathcal{W}}, n_r) & \mathbf{0}_{n_r} \end{bmatrix} & \text{if } t > \underline{t}_{i,\mathcal{W}} \text{ and } \Delta p_{i,t} = 0 \\ \begin{bmatrix} p_{i,t} & 0 \\ p_{i,t} & 1 \end{bmatrix} & \text{if } t > \underline{t}_{i,\mathcal{W}} \text{ and } -\Delta p_{i,t} \notin [\underline{s}_{\mathcal{W}}, \bar{s}_{\mathcal{W}}] \\ \begin{bmatrix} p_{i,t} & 1 \end{bmatrix} & \text{if } t > \underline{t}_{i,\mathcal{W}},\ -\Delta p_{i,t} \in [\underline{s}_{\mathcal{W}}, \bar{s}_{\mathcal{W}}] \text{ and } \Delta p_{i,t} \neq 0. \end{cases} \tag{40}$$

where $\mathbf{g}^r(a, b, n_r)$ denotes an $n_r \times 1$ vector of grid points on $[a, b]$ constructed according to grid rule r.

Define the associated weight function $\mathbf{W}^r(\mathbf{p}_{i,\mathcal{W}}^{t}, \boldsymbol{\theta}_{\mathcal{W}}; t)$ as

$$
\mathbf{W}^r(\mathbf{p}_{i,\mathcal{W}}^t,\boldsymbol{\theta}_{\mathcal{W}};t)=\begin{cases}\begin{bmatrix}\mathbf{w}^r(p_{i,t}-\bar{s}_{\mathcal{W}},p_{i,t}-\underline{s}_{\mathcal{W}},n_r)\\ 1\\ 1\end{bmatrix} & \text{if } t=\underline{t}_{i,\mathcal{W}}\\[2ex] \mathbf{w}^r(p_{i,t}-\bar{s}_{\mathcal{W}},p_{i,t}-\underline{s}_{\mathcal{W}},n_r) \;\; \text{if } t>\underline{t}_{i,\mathcal{W}} \text{ and } \Delta p_{i,t}=0\\[2ex] \begin{bmatrix}1\\ 1\end{bmatrix} \;\; \text{if } t>\underline{t}_{i,\mathcal{W}} \text{ and } -\Delta p_{i,t}\notin[\underline{s}_{\mathcal{W}},\bar{s}_{\mathcal{W}}]\\[2ex] 1 \;\; \text{if } t>\underline{t}_{i,\mathcal{W}},\ -\Delta p_{i,t}\in[\underline{s}_{\mathcal{W}},\bar{s}_{\mathcal{W}}] \text{ and } \Delta p_{i,t}\neq 0.\end{cases} \tag{41}
$$

where $\mathbf{w}^r(a,b,n_r)$ denotes an $n_r\times 1$ vector of weights associated with the grid points $\mathbf{g}^r(a,b,n_r)$ under grid rule r. For FFBS purposes, we consider two grid rules, $r\in\{\text{GL},\text{ES}\}$.[43] Under the Gauss–Legendre rule, $\mathbf{g}^{\text{GL}}$ contains the Gauss–Legendre nodes on $[a,b]$ and $\mathbf{w}^{\text{GL}}$ the associated quadrature weights. Under the equally spaced rule, $\mathbf{g}^{\text{ES}}$ contains n_{ES} equally spaced points on $[a,b]$ and $\mathbf{w}^{\text{ES}}=\mathbf{1}_{n_{\text{ES}}}$.[44] For notational convenience, we henceforth write $\mathbf{G}_{i,t}^r\equiv\mathbf{G}^r(\mathbf{p}_{i,\mathcal{W}}^t,\boldsymbol{\theta}_{\mathcal{W}};t)$ and $\mathbf{W}_{i,t}^r\equiv\mathbf{W}^r(\mathbf{p}_{i,\mathcal{W}}^t,\boldsymbol{\theta}_{\mathcal{W}};t)$.

Initialization and transition matrices. Before describing the FFBS recursions, we define the initial filtered distribution and the transition matrices used to propagate probability mass across grid points. For each grid rule r, define $\mathbf{F}_{i,t}^r\equiv\mathbf{F}^r(\mathbf{p}_{i,\mathcal{W}}^t,\boldsymbol{\theta}_{\mathcal{W}};t)$ as the filtered distribution at time t evaluated on the grid $\mathbf{G}^r(\mathbf{p}_{i,\mathcal{W}}^t,\boldsymbol{\theta}_{\mathcal{W}};t)$, with its j-th element corresponding to the j-th grid point. At the initial date, the prior in Table 1 implies

$$
\mathbf{F}^r(\mathbf{p}_{i,\mathcal{W}}^{\underline{t}_{i,\mathcal{W}}},\boldsymbol{\theta}_{\mathcal{W}};\underline{t}_{i,\mathcal{W}})=\begin{bmatrix}\dfrac{(1-\kappa_0)}{\mathcal{Z}^r}\,\mathbf{1}_{n_r}\\ \kappa_0\,\omega_0\\ \kappa_0\,(1-\omega_0)\end{bmatrix}, \tag{42}
$$

where $\mathcal{Z}^{\text{GL}}=\bar{s}_{\mathcal{W}}-\underline{s}_{\mathcal{W}}$ and $\mathcal{Z}^{\text{ES}}=n_{\text{ES}}$.

[43] The two grids serve different purposes. The Gauss–Legendre grid provides an accurate numerical approximation to the integrals in the forward-filtering recursion, while the equally spaced grid provides a convenient discrete representation of the filtered distribution for backward sampling.

[44] In the baseline implementation, we use 10 Gauss–Legendre nodes. For the equally spaced grid, we target a spacing of 0.002 in logs and choose the number of nodes so that the grid spans the relevant interval exactly, including both endpoints. To limit computational burden, we cap the equally spaced grid at 4,001 points. If this cap would be exceeded, the target spacing is increased in multiples of 0.002 until the cap is satisfied, up to a maximum target spacing of 0.05.

We next define the state transition matrix used in the filtering and backward-sampling recursions. For any two collections of state points $\mathbf{G}^{r_1}$ and $\mathbf{G}^{r_2}$, let $\mathbf{B}(\mathbf{G}^{r_1}, \mathbf{G}^{r_2}, \boldsymbol{\theta}_{\mathcal{W}})$ denote the state transition matrix, with (j,k)th element

$$\left[\mathbf{B}(\mathbf{G}^{r_1}, \mathbf{G}^{r_2}, \boldsymbol{\theta}_{\mathcal{W}})\right]_{jk} = f\left(\mathbf{x}_{i,t} = \mathbf{G}^{r_1}(j,\cdot) \mid \mathbf{x}_{i,t-1} = \mathbf{G}^{r_2}(k,\cdot); \boldsymbol{\theta}_{\mathcal{W}}\right). \tag{43}$$

Given the baseline state transition equations (10) and (11),

$$p^r_{i,t} \mid p^r_{i,t-1}, \boldsymbol{\theta}_{\mathcal{W}} \sim \mathcal{N}\left(p^r_{i,t-1} + \mu_{\mathcal{W}}, \sigma^2_{\varepsilon,\mathcal{W}}\right), \qquad \ell_{i,t} \mid \boldsymbol{\theta}_{\mathcal{W}} \sim \mathcal{B}(\lambda_{\mathcal{W}}).$$

Hence, the state transition matrix is given by

$$\left[\mathbf{B}(\mathbf{G}^{r_1}, \mathbf{G}^{r_2}, \boldsymbol{\theta}_{\mathcal{W}})\right]_{jk} = \phi\left([\mathbf{G}^{r_1}]_{j,1}; [\mathbf{G}^{r_2}]_{k,1} + \mu_{\mathcal{W}}, \sigma^2_{\varepsilon,\mathcal{W}}\right) \lambda_{\mathcal{W}}^{[\mathbf{G}^{r_1}]_{j,2}} (1-\lambda_{\mathcal{W}})^{1-[\mathbf{G}^{r_1}]_{j,2}}, \tag{44}$$

where $\phi(\cdot; m, v)$ denotes the normal density with mean m and variance v.

Grid-based FFBS algorithm. To draw $\mathbf{x}^*_{i,\mathcal{W}}$ from $f(\mathbf{x}_{i,\mathcal{W}} \mid \mathbf{p}_{i,\mathcal{W}}, \boldsymbol{\theta}_{\mathcal{W}})$, proceed as follows.

1. **Forward filtering.** Starting from the initial condition in (42), solve recursively for the sequence of filtered distributions on the Gauss–Legendre grids using

$$\mathbf{F}^{\mathrm{GL}}_{i,t} = \frac{\mathbf{B}(\mathbf{G}^{\mathrm{GL}}_{i,t}, \mathbf{G}^{\mathrm{GL}}_{i,t-1}, \boldsymbol{\theta}_{\mathcal{W}})\left(\mathbf{W}^{\mathrm{GL}}_{i,t-1} \odot \mathbf{F}^{\mathrm{GL}}_{i,t-1}\right)}{\left(\mathbf{W}^{\mathrm{GL}}_{i,t}\right)' \mathbf{B}(\mathbf{G}^{\mathrm{GL}}_{i,t}, \mathbf{G}^{\mathrm{GL}}_{i,t-1}, \boldsymbol{\theta}_{\mathcal{W}})\left(\mathbf{W}^{\mathrm{GL}}_{i,t-1} \odot \mathbf{F}^{\mathrm{GL}}_{i,t-1}\right)} \quad \text{for } t > \underline{t}_{i,\mathcal{W}}. \tag{45}$$

   where $\odot$ denotes the Hadamard product. Next, solve for the sequence of filtered distributions on the equally spaced grids using

$$\mathbf{F}^{\mathrm{ES}}_{i,t} = \frac{\mathbf{B}(\mathbf{G}^{\mathrm{ES}}_{i,t}, \mathbf{G}^{\mathrm{GL}}_{i,t-1}, \boldsymbol{\theta}_{\mathcal{W}})\left(\mathbf{W}^{\mathrm{GL}}_{i,t-1} \odot \mathbf{F}^{\mathrm{GL}}_{i,t-1}\right)}{\left(\mathbf{W}^{\mathrm{ES}}_{i,t}\right)' \mathbf{B}(\mathbf{G}^{\mathrm{ES}}_{i,t}, \mathbf{G}^{\mathrm{GL}}_{i,t-1}, \boldsymbol{\theta}_{\mathcal{W}})\left(\mathbf{W}^{\mathrm{GL}}_{i,t-1} \odot \mathbf{F}^{\mathrm{GL}}_{i,t-1}\right)} \quad \text{for } t > \underline{t}_{i,\mathcal{W}}. \tag{46}$$

2. **Backward sampling.** Draw the terminal state as

$$\mathbf{x}^*_{i,\bar{t}_{i,\mathcal{W}}} = \mathbf{G}^{\mathrm{ES}}_{i,\bar{t}_{i,\mathcal{W}}}\left(j^*_{\bar{t}_{i,\mathcal{W}}}, \cdot\right) \quad \text{where } j^*_{\bar{t}_{i,\mathcal{W}}} \sim \text{Categorical}\left(\mathbf{F}^{\mathrm{ES}}_{i,\bar{t}_{i,\mathcal{W}}}\right). \tag{47}$$

   Then, recursively for $t = \bar{t}_{i,\mathcal{W}} - 1, \ldots, \underline{t}_{i,\mathcal{W}}$, compute

$$\mathbf{S}^{\mathrm{ES}}_{i,t}\left(\mathbf{x}^*_{i,t+1}\right) = \frac{\mathbf{B}(\mathbf{x}^*_{i,t+1}, \mathbf{G}^{\mathrm{ES}}_{i,t}, \boldsymbol{\theta}_{\mathcal{W}})' \odot \mathbf{F}^{\mathrm{ES}}_{i,t}}{\mathbf{1}'\left[\mathbf{B}(\mathbf{x}^*_{i,t+1}, \mathbf{G}^{\mathrm{ES}}_{i,t}, \boldsymbol{\theta}_{\mathcal{W}})' \odot \mathbf{F}^{\mathrm{ES}}_{i,t}\right]}, \tag{48}$$

   and draw

$$\mathbf{x}^*_{i,t} = \mathbf{G}^{\mathrm{ES}}_{i,t}\left(j^*_t, \cdot\right) \quad \text{where } j^*_t \sim \text{Categorical}\left(\mathbf{S}^{\mathrm{ES}}_{i,t}(\mathbf{x}^*_{i,t+1})\right). \tag{49}$$

Price trajectories spanning multiple windows. The algorithm above describes the case in which a price trajectory begins within $\mathcal{W}$, so that the forward filter is initialized using the prior distribution in Table 1. For price trajectories that continue from a preceding window, we instead initialize the latent state at $t = \underline{t}_{i,\mathcal{W}} - 1$ using a distribution degenerate at the posterior mean of the terminal latent state estimated in the preceding window.[45] In this case, the initial grid contains only this state, with associated weight and filtered probability both equal to one. Starting from this alternative initial condition, the forward-filtering recursions in (45) and (46) proceed for $t \geqslant \underline{t}_{i,\mathcal{W}}$ and the backward-sampling recursions in (47)–(49) proceed unchanged.

B.3 Full Gibbs Sampler

Let $\boldsymbol{\Theta}_{\mathcal{W}} \equiv \{\boldsymbol{\theta}_{\mathcal{W}}, \mathbf{x}_{\mathcal{W}}\}$ collect the state-space parameters and latent states. This appendix provides the Gibbs sampling algorithm used to draw from the joint posterior $f(\boldsymbol{\Theta}_{\mathcal{W}} \mid \mathbf{p}_{\mathcal{W}})$. Let $m = 1, \ldots, M$ index Gibbs draws, and denote the mth draw by $\boldsymbol{\Theta}_{\mathcal{W}}^{(m)}$.

1. **Initial draws.** Initialize the parameters and latent states as follows:[46]

   (a) **Inaction thresholds:** For $b \in \{L, U\}$, define

   $$\mathcal{D}_L = \{|\Delta p_{i,t}| : \Delta p_{i,t} > 0\}, \qquad \mathcal{D}_U = \{|\Delta p_{i,t}| : \Delta p_{i,t} < 0\}.$$

   Partition each $\mathcal{D}_b$ into histogram intervals $\mathcal{H}_{b,j} = [h_{b,j}, h_{b,j+1})$, and let $H_{b,j}$ denote the number of observations in $\mathcal{H}_{b,j}$.[47] Define

   $$j_b^* = \arg\max_j \{H_{b,j+1} - H_{b,j}\}.$$

   Initialize

   $$\underline{s}_{\mathcal{W}}^{(1)} = -\frac{\sum_{x \in \mathcal{D}_L} x\, \mathbb{1}\{x \in \mathcal{H}_{L,j_L^*+1}\}}{\sum_{x \in \mathcal{D}_L} \mathbb{1}\{x \in \mathcal{H}_{L,j_L^*+1}\}}, \qquad \bar{s}_{\mathcal{W}}^{(1)} = \frac{\sum_{x \in \mathcal{D}_U} x\, \mathbb{1}\{x \in \mathcal{H}_{U,j_U^*+1}\}}{\sum_{x \in \mathcal{D}_U} \mathbb{1}\{x \in \mathcal{H}_{U,j_U^*+1}\}}.$$

   (b) **Reset-price drift:** For each price trajectory i, let

   $$\tau_{i,1} < \tau_{i,2} < \cdots < \tau_{i,J_i}$$

[45] The posterior mean of $\ell_{i,t}$ need not be binary. This does not affect the recursion because $\ell_{i,t-1}$ does not enter the state transition density: $p_{i,t}^r$ depends only on $p_{i,t-1}^r$, while $\ell_{i,t}$ is drawn independently from $\mathcal{B}(\lambda_{\mathcal{W}})$.

[46] When a data-driven initial value cannot be computed because the relevant sample information is insufficient or degenerate, we use a deterministic fallback value. For the inaction thresholds and reset-price drift, this is the corresponding prior mean. For the reset-price standard deviation, we use the value implied by the prior mean of the precision. If either inaction threshold cannot be initialized from the data, we instead initialize the free-adjustment probability at 0.6 times the observed frequency of price changes.

[47] For each $b \in \{L, U\}$, let $N_b = |\mathcal{D}_b|$ and $M_b = \min\{N_b, \max\{30, \lceil \sqrt{N_b} \rceil\}\}$. We set the bin width to $h_b = \max\{0.01, (\max \mathcal{D}_b - \min \mathcal{D}_b)/M_b\}$ and construct the bin edges in increments of h_b, starting at $\lfloor \min \mathcal{D}_b / h_b \rfloor h_b$ and ending at $\lceil \max \mathcal{D}_b / h_b \rceil h_b$.

denote the dates at which the observed price changes. For $j = 2, \ldots, J_i$, define

$$\Delta p^{\mathrm{sp}}_{i,j} = p_{i,\tau_{i,j}} - p_{i,\tau_{i,j-1}}, \qquad d_{i,j} = \tau_{i,j} - \tau_{i,j-1}.$$

Then set

$$\mu^{(1)}_{\mathcal{W}} = \frac{1}{N_{\mathrm{sp}}} \sum_i \sum_{j=2}^{J_i} \frac{\Delta p^{\mathrm{sp}}_{i,j}}{d_{i,j}}, \qquad N_{\mathrm{sp}} = \sum_i (J_i - 1).$$

(c) **Reset-price standard deviation:** Using the same complete price spells, set

$$\sigma^{(1)}_{\varepsilon,\mathcal{W}} = \left[\frac{1}{N_{\mathrm{sp}} - 1} \sum_i \sum_{j=2}^{J_i} \left(\frac{\Delta p^{\mathrm{sp}}_{i,j} - \mu^{(1)}_{\mathcal{W}} d_{i,j}}{\sqrt{d_{i,j}}} - \bar{\varepsilon} \right)^2 \right]^{1/2},$$

where

$$\bar{\varepsilon} = \frac{1}{N_{\mathrm{sp}}} \sum_i \sum_{j=2}^{J_i} \frac{\Delta p^{\mathrm{sp}}_{i,j} - \mu^{(1)}_{\mathcal{W}} d_{i,j}}{\sqrt{d_{i,j}}}.$$

(d) **Free-adjustment probability:** Using $\underline{s}^{(1)}_{\mathcal{W}}$ and $\bar{s}^{(1)}_{\mathcal{W}}$, define

$$n_{\mathrm{in}} = \sum_{i,t} \mathbb{1} \left\{ \Delta p_{i,t} \neq 0,\ -\bar{s}^{(1)}_{\mathcal{W}} \leqslant \Delta p_{i,t} \leqslant -\underline{s}^{(1)}_{\mathcal{W}} \right\},$$

$$n_{\mathrm{out}} = \sum_{i,t} \mathbb{1} \left\{ \Delta p_{i,t} \neq 0,\ \Delta p_{i,t} < -\bar{s}^{(1)}_{\mathcal{W}} \text{ or } \Delta p_{i,t} > -\underline{s}^{(1)}_{\mathcal{W}} \right\}.$$

Initialize $\lambda^{(1)}_{\mathcal{W}}$ with the sample analogue of the Bernoulli probability of a free adjustment opportunity,

$$\lambda^{(1)}_{\mathcal{W}} = \frac{n_{\mathrm{in}}}{n_{\mathrm{valid}} - n_{\mathrm{out}}},$$

where n_{valid} is the number of observations for which $\Delta p_{i,t}$ is defined.

(e) **Latent states:** Initialize the latent states as

$$p^{r(1)}_{i,t} = p_{i,t},$$

and, whenever $\Delta p_{i,t}$ is defined,

$$\ell^{(1)}_{i,t} = \mathbb{1} \left\{ \Delta p_{i,t} \neq 0,\ -\bar{s}^{(1)}_{\mathcal{W}} \leqslant \Delta p_{i,t} \leqslant -\underline{s}^{(1)}_{\mathcal{W}} \right\}.$$

For observations for which $\Delta p_{i,t}$ is not defined, set $\ell^{(1)}_{i,t} = 0$.

2. **Subsequent draws.** For $m > 1$, draw parameters and latent states from their full

conditional distributions as follows:

(a) **Reset-price drift and standard deviation:** Define

$$\Delta p_{i,t}^{r(m-1)} \equiv p_{i,t}^{r(m-1)} - p_{i,t-1}^{r(m-1)}, \qquad t > \underline{t}_{i,\mathcal{W}},$$

and let $\mathbf{y}^{(m-1)}$ collect all $\Delta p_{i,t}^{r(m-1)}$ across $i \in \mathcal{I}_{\mathcal{W}}$ and $t > \underline{t}_{i,\mathcal{W}}$, stacked vertically. Then

$$\mathbf{y}^{(m-1)} = \mathbf{X}^{(m-1)}\mu_{\mathcal{W}} + \boldsymbol{\varepsilon}, \qquad \boldsymbol{\varepsilon} \sim \mathcal{N}\left(\mathbf{0}, \sigma_{\varepsilon,\mathcal{W}}^2 \mathbf{I}\right).$$

Here, $\mathbf{X}^{(m-1)} = \mathbf{1}$. Given the prior $\mu_{\mathcal{W}} \sim \mathcal{N}(m_\mu^0, v_\mu^0)$, draw

$$\mu_{\mathcal{W}}^{(m)} \mid \sigma_{\varepsilon,\mathcal{W}}^{(m-1)}, \lambda_{\mathcal{W}}^{(m-1)}, \underline{s}_{\mathcal{W}}^{(m-1)}, \bar{s}_{\mathcal{W}}^{(m-1)}, \mathbf{x}_{\mathcal{W}}^{(m-1)}, \mathbf{p}_{\mathcal{W}} \sim \mathcal{N}\left(m_{\mu,\text{post}}^{(m)}, v_{\mu,\text{post}}^{(m)}\right),$$

where

$$v_{\mu,\text{post}}^{(m)} = \left[\frac{\mathbf{X}^{(m-1)\prime}\mathbf{X}^{(m-1)}}{\left(\sigma_{\varepsilon,\mathcal{W}}^{(m-1)}\right)^2} + \frac{1}{v_\mu^0}\right]^{-1},$$

and

$$m_{\mu,\text{post}}^{(m)} = v_{\mu,\text{post}}^{(m)}\left[\frac{m_\mu^0}{v_\mu^0} + \frac{\mathbf{X}^{(m-1)\prime}\mathbf{y}^{(m-1)}}{\left(\sigma_{\varepsilon,\mathcal{W}}^{(m-1)}\right)^2}\right].$$

Next, define $\gamma_{\mathcal{W}} \equiv \sigma_{\varepsilon,\mathcal{W}}^{-2}$. Given the prior $\gamma_{\mathcal{W}} \sim \mathcal{G}(s_\sigma^0, r_\sigma^0)$, draw

$$\gamma_{\mathcal{W}}^{(m)} \mid \mu_{\mathcal{W}}^{(m)}, \lambda_{\mathcal{W}}^{(m-1)}, \underline{s}_{\mathcal{W}}^{(m-1)}, \bar{s}_{\mathcal{W}}^{(m-1)}, \mathbf{x}_{\mathcal{W}}^{(m-1)}, \mathbf{p}_{\mathcal{W}} \sim \mathcal{G}\left(s_{\sigma,\text{post}}^{(m)}, r_{\sigma,\text{post}}^{(m)}\right),$$

where

$$s_{\sigma,\text{post}}^{(m)} = s_\sigma^0 + \frac{n_{\text{reg}}}{2},$$

and

$$r_{\sigma,\text{post}}^{(m)} = r_\sigma^0 + \frac{1}{2}\left(\mathbf{y}^{(m-1)} - \mathbf{X}^{(m-1)}\mu_{\mathcal{W}}^{(m)}\right)'\left(\mathbf{y}^{(m-1)} - \mathbf{X}^{(m-1)}\mu_{\mathcal{W}}^{(m)}\right),$$

with n_{reg} denoting the number of observations in $\mathbf{y}^{(m-1)}$. Finally, set

$$\sigma_{\varepsilon,\mathcal{W}}^{(m)} = \left(\gamma_{\mathcal{W}}^{(m)}\right)^{-1/2}.$$

(b) **Free-adjustment probability:** Given the prior $\lambda_{\mathcal{W}} \sim \mathcal{B}(s_{1,\lambda}^0, s_{2,\lambda}^0)$, draw

$$\lambda_{\mathcal{W}}^{(m)} \mid \mu_{\mathcal{W}}^{(m)}, \sigma_{\varepsilon,\mathcal{W}}^{(m)}, \underline{s}_{\mathcal{W}}^{(m-1)}, \bar{s}_{\mathcal{W}}^{(m-1)}, \mathbf{x}_{\mathcal{W}}^{(m-1)}, \mathbf{p}_{\mathcal{W}} \sim \mathcal{B}\left(s_{1,\lambda,\text{post}}^{(m)}, s_{2,\lambda,\text{post}}^{(m)}\right),$$

where

$$s_{1,\lambda,\text{post}}^{(m)} = s_{1,\lambda}^0 + \sum_{i \in \mathcal{I}_{\mathcal{W}}} \sum_{t \in \mathcal{T}_{i,\mathcal{W}}} \ell_{i,t}^{(m-1)},$$

and

$$s_{2,\lambda,\text{post}}^{(m)} = s_{2,\lambda}^0 + \sum_{i \in \mathcal{I}_\mathcal{W}} \sum_{t \in \mathcal{T}_{i,\mathcal{W}}} \left(1 - \ell_{i,t}^{(m-1)}\right).$$

(c) **Inaction thresholds and latent states:** Draw from

$$f\left(\underline{s}_\mathcal{W}, \bar{s}_\mathcal{W}, \mathbf{x}_\mathcal{W} \mid \mu_\mathcal{W}^{(m)}, \sigma_{\varepsilon,\mathcal{W}}^{(m)}, \lambda_\mathcal{W}^{(m)}, \mathbf{p}_\mathcal{W}\right)$$

using a Metropolis–Hastings step as follows:

- Draw the candidate lower threshold

$$\underline{s}_\mathcal{W}^* \sim \mathcal{TN}\left(\underline{s}_\mathcal{W}^{(m-1)}, \left(\sigma_{\underline{s},\text{prop}}^{(m)}\right)^2; \underline{s}_{\text{LB},\mathcal{W}}^0, \underline{s}_{\text{UB},\mathcal{W}}^0\right),$$

where $\mathcal{TN}(\mu, \sigma^2; a, b)$ denotes a normal distribution with mean μ and variance σ^2, truncated to $[a, b]$.

- Draw the candidate upper threshold

$$\bar{s}_\mathcal{W}^* \sim \mathcal{TN}\left(\bar{s}_\mathcal{W}^{(m-1)}, \left(\sigma_{\bar{s},\text{prop}}^{(m)}\right)^2; \bar{s}_{\text{LB},\mathcal{W}}^0, \bar{s}_{\text{UB},\mathcal{W}}^0\right).$$

Denote the corresponding proposal densities by $q_{\underline{s}}^{(m)}$ and $q_{\bar{s}}^{(m)}$, respectively.

- Draw the candidate latent states $\mathbf{x}_\mathcal{W}^*$ by drawing, independently for each $i \in \mathcal{I}_\mathcal{W}$,

$$\mathbf{x}_{i,\mathcal{W}}^* \sim f\left(\mathbf{x}_{i,\mathcal{W}} \mid \mathbf{p}_{i,\mathcal{W}}, \mu_\mathcal{W}^{(m)}, \sigma_{\varepsilon,\mathcal{W}}^{(m)}, \lambda_\mathcal{W}^{(m)}, \underline{s}_\mathcal{W}^*, \bar{s}_\mathcal{W}^*\right)$$

using the grid-based FFBS algorithm in Appendix B.2.

- Compute the Metropolis–Hastings acceptance probability

$$\alpha^{(m)} = \min\left\{1, \frac{\mathcal{L}_\mathcal{W}^*}{\mathcal{L}_\mathcal{W}^{(m-1)}} \frac{f\left(\underline{s}_\mathcal{W}^*\right)}{f\left(\underline{s}_\mathcal{W}^{(m-1)}\right)} \frac{f\left(\bar{s}_\mathcal{W}^*\right)}{f\left(\bar{s}_\mathcal{W}^{(m-1)}\right)} \frac{q_{\underline{s}}^{(m)}\left(\underline{s}_\mathcal{W}^{(m-1)} \mid \underline{s}_\mathcal{W}^*\right)}{q_{\underline{s}}^{(m)}\left(\underline{s}_\mathcal{W}^* \mid \underline{s}_\mathcal{W}^{(m-1)}\right)} \frac{q_{\bar{s}}^{(m)}\left(\bar{s}_\mathcal{W}^{(m-1)} \mid \bar{s}_\mathcal{W}^*\right)}{q_{\bar{s}}^{(m)}\left(\bar{s}_\mathcal{W}^* \mid \bar{s}_\mathcal{W}^{(m-1)}\right)}\right\},$$

where

$$\mathcal{L}_\mathcal{W}^* \equiv f\left(\mathbf{p}_\mathcal{W} \mid \mu_\mathcal{W}^{(m)}, \sigma_{\varepsilon,\mathcal{W}}^{(m)}, \lambda_\mathcal{W}^{(m)}, \underline{s}_\mathcal{W}^*, \bar{s}_\mathcal{W}^*\right),$$

and

$$\mathcal{L}_\mathcal{W}^{(m-1)} \equiv f\left(\mathbf{p}_\mathcal{W} \mid \mu_\mathcal{W}^{(m)}, \sigma_{\varepsilon,\mathcal{W}}^{(m)}, \lambda_\mathcal{W}^{(m)}, \underline{s}_\mathcal{W}^{(m-1)}, \bar{s}_\mathcal{W}^{(m-1)}\right)$$

denote the marginal likelihoods evaluated at the candidate and current thresholds, respectively.[48]

[48]Because the latent-state proposal is the conditional smoothing distribution generated by FFBS, its density appears in both the target ratio and the reverse-to-forward proposal ratio and cancels, leaving only the marginal likelihood, threshold priors, and threshold proposal densities in the acceptance probability.

For any

$$\boldsymbol{\theta}_{\mathcal{W}} = \{\mu_{\mathcal{W}}^{(m)}, \sigma_{\varepsilon,\mathcal{W}}^{(m)}, \lambda_{\mathcal{W}}^{(m)}, \underline{s}_{\mathcal{W}}, \bar{s}_{\mathcal{W}}\},$$

define the forward-filter normalizing constant

$$Z_{i,t}(\boldsymbol{\theta}_{\mathcal{W}}) \equiv \big(\mathbf{W}_{i,t}^{\mathrm{GL}}\big)'\mathbf{B}\big(\mathbf{G}_{i,t}^{\mathrm{GL}}, \mathbf{G}_{i,t-1}^{\mathrm{GL}}, \boldsymbol{\theta}_{\mathcal{W}}\big)\big(\mathbf{W}_{i,t-1}^{\mathrm{GL}} \odot \mathbf{F}_{i,t-1}^{\mathrm{GL}}\big),$$

which is the normalizing constant in (45). Under the grid approximation, the corresponding marginal likelihood is

$$\log \mathcal{L}_{\mathcal{W}}\big(\underline{s}_{\mathcal{W}}, \bar{s}_{\mathcal{W}}\big) = \sum_{i \in \mathcal{I}_{\mathcal{W}}} \sum_{t > \underline{t}_{i,\mathcal{W}}} \log Z_{i,t}(\boldsymbol{\theta}_{\mathcal{W}}).$$

Evaluating this expression at $\big(\underline{s}_{\mathcal{W}}^{*}, \bar{s}_{\mathcal{W}}^{*}\big)$ and at $\big(\underline{s}_{\mathcal{W}}^{(m-1)}, \bar{s}_{\mathcal{W}}^{(m-1)}\big)$ gives $\mathcal{L}_{\mathcal{W}}^{*}$ and $\mathcal{L}_{\mathcal{W}}^{(m-1)}$, respectively.[49]

- Draw $u^{(m)} \sim \mathcal{U}(0,1)$.
- If $u^{(m)} \leqslant \alpha^{(m)}$, accept and set $\big(\underline{s}_{\mathcal{W}}^{(m)}, \bar{s}_{\mathcal{W}}^{(m)}, \mathbf{x}_{\mathcal{W}}^{(m)}\big) = \big(\underline{s}_{\mathcal{W}}^{*}, \bar{s}_{\mathcal{W}}^{*}, \mathbf{x}_{\mathcal{W}}^{*}\big)$.
- Otherwise, reject and set $\big(\underline{s}_{\mathcal{W}}^{(m)}, \bar{s}_{\mathcal{W}}^{(m)}, \mathbf{x}_{\mathcal{W}}^{(m)}\big) = \big(\underline{s}_{\mathcal{W}}^{(m-1)}, \bar{s}_{\mathcal{W}}^{(m-1)}, \mathbf{x}_{\mathcal{W}}^{(m-1)}\big)$.

3. **Iteration.** Repeat Steps 2(a)–2(c) until $m = M$.[50]

Proposal fine-tuning. We adapt the standard deviations of the truncated-normal proposals during burn-in using a Robbins–Monro stochastic approximation (Robbins and Monro, 1951). We initialize the proposal standard deviations as

$$\sigma_{\underline{s},\mathrm{prop}}^{(m)} = 0.025\left(\underline{s}_{\mathrm{UB},\mathcal{W}}^{0} - \underline{s}_{\mathrm{LB},\mathcal{W}}^{0}\right), \qquad \sigma_{\bar{s},\mathrm{prop}}^{(m)} = 0.025\left(\bar{s}_{\mathrm{UB},\mathcal{W}}^{0} - \bar{s}_{\mathrm{LB},\mathcal{W}}^{0}\right)$$

during the initial exploration period of 500 draws. Thereafter, adaptation proceeds in windows of 50 draws. Let $\widehat{\alpha}_j$ denote the Metropolis–Hastings acceptance rate in adaptation window j, and let k_j denote the common scaling factor applied to both proposal standard deviations, with $k_1 = 1$. At the end of each adaptation window, update

$$\log k_{j+1} = \log k_j + 0.5\big(\widehat{\alpha}_j - 0.33\big),$$

and, for each Gibbs draw m in the following adaptation window, set

$$\sigma_{\underline{s},\mathrm{prop}}^{(m)} = k_{j+1}0.025\left(\underline{s}_{\mathrm{UB},\mathcal{W}}^{0} - \underline{s}_{\mathrm{LB},\mathcal{W}}^{0}\right), \qquad \sigma_{\bar{s},\mathrm{prop}}^{(m)} = k_{j+1}0.025\left(\bar{s}_{\mathrm{UB},\mathcal{W}}^{0} - \bar{s}_{\mathrm{LB},\mathcal{W}}^{0}\right).$$

[49] For a price trajectory that continues from a preceding window, so that the state at $\underline{t}_{i,\mathcal{W}} - 1$ is supplied as an initial condition, the corresponding inner sum instead runs over $t \geqslant \underline{t}_{i,\mathcal{W}}$.

[50] We use 8,001 Gibbs draws in total, discard the first 3,001 as burn-in, and retain the remaining 5,000 draws for posterior inference.

We target an acceptance rate of 0.33, which lies between the commonly used optimal-scaling benchmarks of approximately 0.234 for high-dimensional and 0.44 for univariate Gaussian random-walk Metropolis proposals (Roberts et al., 1997; Roberts and Rosenthal, 2001). We use 50 adaptation windows, so that proposal fine-tuning ends with the burn-in period. The proposal distributions are then held fixed at their final tuned values for all retained posterior draws.

B.4 Extensions to alternative reset-price processes

The baseline algorithm can be extended to accommodate more general reset-price processes. In particular, the reset-price transition in equation (10) can be replaced by

$$p^r_{i,t} = \rho_W p^r_{i,t-1} + \mathbf{z}'_{i,t}\boldsymbol{\beta}_W + \varepsilon_{i,t},$$

where $\mathbf{z}_{i,t}$ may include a constant, observed exogenous covariates that may be individual-specific or aggregate, and known transformations of these variables. The process for $\varepsilon_{i,t}$ can also be replaced by a mixture of normal distributions, so that

$$\varepsilon_{i,t} \mid q_{i,t} = k \sim \mathcal{N}(0, \sigma^2_{k,W}), \qquad k \in \{1, 2\},$$

with

$$\Pr(q_{i,t} = 1) = \omega_W, \qquad \Pr(q_{i,t} = 2) = 1 - \omega_W.$$

This specification allows for fat-tailed idiosyncratic shocks, as in Karadi and Reiff (2019). As in the baseline specification, innovations and mixture indicators are independent across price trajectories and time. Relative to the baseline specification, the parameters governing the reset-price process, $\{\mu_W, \sigma_{\varepsilon,W}\}$, are now replaced by

$$\boldsymbol{\vartheta}_W = \{\rho_W, \boldsymbol{\beta}_W, \sigma_{1,W}, \sigma_{2,W}, \omega_W\}.$$

Note that the baseline reset-price process in equation (10) is nested within this specification by setting $\mathbf{z}_{i,t} = 1$, $\boldsymbol{\beta}_W = \mu_W$, $\rho_W = 1$, $\omega_W = 1$, and $\sigma_{1,W} = \sigma_{\varepsilon,W}$.

FFBS algorithm. Accommodating this more general reset-price process requires only modifying the transition probabilities entering the matrix $\mathbf{B}$. In particular, equation (44) is replaced by

$$\begin{aligned}\left[\mathbf{B}(\mathbf{G}^{r_1}, \mathbf{G}^{r_2}, \boldsymbol{\theta}_W, \mathbf{z}_{i,t})\right]_{jk} = &\Big[\omega_W \phi\big([\mathbf{G}^{r_1}]_{j,1}; \rho_W[\mathbf{G}^{r_2}]_{k,1} + \mathbf{z}'_{i,t}\boldsymbol{\beta}_W, \sigma^2_{1,W}\big) \\ &+ (1-\omega_W)\phi\big([\mathbf{G}^{r_1}]_{j,1}; \rho_W[\mathbf{G}^{r_2}]_{k,1} + \mathbf{z}'_{i,t}\boldsymbol{\beta}_W, \sigma^2_{2,W}\big)\Big] \\ &\times \lambda_W^{[\mathbf{G}^{r_1}]_{j,2}}(1-\lambda_W)^{1-[\mathbf{G}^{r_1}]_{j,2}}.\end{aligned} \tag{50}$$

Thus, the forward-filtering and backward-sampling recursions remain unchanged, with the transition probabilities in (50) replacing those in (44). Because the reset-price process remains first-order Markov and the covariates in $\mathbf{z}_{i,t}$ are observed, this extension does not increase the dimension of the latent state or the grids used by the FFBS algorithm.

Gibbs sampler. The parameters of the generalized reset-price process can be sampled using conditionally conjugate updates. Define

$$\widetilde{\boldsymbol{\beta}}_{\mathcal{W}} \equiv \begin{bmatrix} \rho_{\mathcal{W}} \\ \boldsymbol{\beta}_{\mathcal{W}} \end{bmatrix}, \qquad \widetilde{\mathbf{z}}_{i,t} \equiv \begin{bmatrix} p^r_{i,t-1} \\ \mathbf{z}_{i,t} \end{bmatrix},$$

so that

$$p^r_{i,t} = \widetilde{\mathbf{z}}'_{i,t} \widetilde{\boldsymbol{\beta}}_{\mathcal{W}} + \varepsilon_{i,t}.$$

Let $\gamma_{k,\mathcal{W}} \equiv \sigma^{-2}_{k,\mathcal{W}}$ for $k \in \{1,2\}$, and impose the conjugate priors[51,52]

$$\widetilde{\boldsymbol{\beta}}_{\mathcal{W}} \sim \mathcal{N}(\mathbf{m}^0_\beta, \mathbf{V}^0_\beta), \qquad \gamma_{k,\mathcal{W}} \sim \mathcal{G}(s^0_{\sigma,k}, r^0_{\sigma,k}), \quad k \in \{1,2\}, \qquad \omega_{\mathcal{W}} \sim \mathcal{B}(s^0_{1,\omega}, s^0_{2,\omega}).$$

Step 2(a) of the baseline Gibbs sampler is then replaced by the following draws:

1. **Mixture indicators.** For each $i \in \mathcal{I}_{\mathcal{W}}$ and $t > \underline{t}_{i,\mathcal{W}}$, draw

$$q^{(m)}_{i,t} \mid \widetilde{\boldsymbol{\beta}}^{(m-1)}_{\mathcal{W}}, \sigma^{(m-1)}_{1,\mathcal{W}}, \sigma^{(m-1)}_{2,\mathcal{W}}, \omega^{(m-1)}_{\mathcal{W}}, \mathbf{x}^{(m-1)}_{\mathcal{W}} \sim \text{Categorical}\left(\pi^{(m)}_{i,t,1}, \pi^{(m)}_{i,t,2}\right),$$

where, for $k \in \{1,2\}$,

$$\pi^{(m)}_{i,t,k} = \frac{\omega^{(m-1)}_{k,\mathcal{W}} \phi\left(p^{r(m-1)}_{i,t}; \widetilde{\mathbf{z}}^{(m-1)\prime}_{i,t} \widetilde{\boldsymbol{\beta}}^{(m-1)}_{\mathcal{W}}, (\sigma^{(m-1)}_{k,\mathcal{W}})^2\right)}{\sum_{\ell=1}^{2} \omega^{(m-1)}_{\ell,\mathcal{W}} \phi\left(p^{r(m-1)}_{i,t}; \widetilde{\mathbf{z}}^{(m-1)\prime}_{i,t} \widetilde{\boldsymbol{\beta}}^{(m-1)}_{\mathcal{W}}, (\sigma^{(m-1)}_{\ell,\mathcal{W}})^2\right)},$$

with $\omega_{1,\mathcal{W}} \equiv \omega_{\mathcal{W}}$ and $\omega_{2,\mathcal{W}} \equiv 1 - \omega_{\mathcal{W}}$. Define

$$n^{(m)}_k = \sum_{i \in \mathcal{I}_{\mathcal{W}}} \sum_{t > \underline{t}_{i,\mathcal{W}}} \mathbb{1}\{q^{(m)}_{i,t} = k\}, \qquad k \in \{1,2\}.$$

2. **Reset-price coefficients.** Conditional on the mixture indicators, draw

$$\widetilde{\boldsymbol{\beta}}^{(m)}_{\mathcal{W}} \mid \mathbf{q}^{(m)}_{\mathcal{W}}, \gamma^{(m-1)}_{1,\mathcal{W}}, \gamma^{(m-1)}_{2,\mathcal{W}}, \mathbf{x}^{(m-1)}_{\mathcal{W}} \sim \mathcal{N}\left(\mathbf{m}^{(m)}_\beta, \mathbf{V}^{(m)}_\beta\right),$$

[51] In line with the weakly informative priors in Table 1, we set $\mathbf{m}^0_\beta = \mathbf{0}$, $\mathbf{V}^0_\beta = 10^2\mathbf{I}$, $s^0_{\sigma,k} = r^0_{\sigma,k} = 10^{-3}$ for $k \in \{1,2\}$, and $s^0_{1,\omega} = s^0_{2,\omega} = 1$.

[52] The two mixture components are observationally equivalent up to relabeling. We therefore use exchangeable priors across components and relabel each draw, when necessary, so that $\sigma_{1,\mathcal{W}} < \sigma_{2,\mathcal{W}}$, with the corresponding relabeling of the mixture indicators and replacement of $\omega_{\mathcal{W}}$ by $1 - \omega_{\mathcal{W}}$.

where

$$\mathbf{V}_{\beta}^{(m)} = \left[(\mathbf{V}_{\beta}^{0})^{-1} + \sum_{i \in \mathcal{I}_{\mathcal{W}}} \sum_{t > \underline{t}_{i,\mathcal{W}}} \gamma_{q_{i,t}^{(m)},\mathcal{W}}^{(m-1)} \widetilde{\mathbf{z}}_{i,t}^{(m-1)} \widetilde{\mathbf{z}}_{i,t}^{(m-1)\prime} \right]^{-1},$$

and

$$\mathbf{m}_{\beta}^{(m)} = \mathbf{V}_{\beta}^{(m)} \left[(\mathbf{V}_{\beta}^{0})^{-1} \mathbf{m}_{\beta}^{0} + \sum_{i \in \mathcal{I}_{\mathcal{W}}} \sum_{t > \underline{t}_{i,\mathcal{W}}} \gamma_{q_{i,t}^{(m)},\mathcal{W}}^{(m-1)} \widetilde{\mathbf{z}}_{i,t}^{(m-1)} p_{i,t}^{r(m-1)} \right].$$

3. **Mixture standard deviations.** For each $k \in \{1, 2\}$, define

$$\mathrm{SSR}_{k,\mathcal{W}}^{(m)} = \sum_{i \in \mathcal{I}_{\mathcal{W}}} \sum_{t > \underline{t}_{i,\mathcal{W}}} \mathbb{1}\{q_{i,t}^{(m)} = k\} \left[p_{i,t}^{r(m-1)} - \widetilde{\mathbf{z}}_{i,t}^{(m-1)\prime} \widetilde{\boldsymbol{\beta}}_{\mathcal{W}}^{(m)} \right]^2.$$

Then draw

$$\gamma_{k,\mathcal{W}}^{(m)} \mid \widetilde{\boldsymbol{\beta}}_{\mathcal{W}}^{(m)}, \mathbf{q}_{\mathcal{W}}^{(m)}, \mathbf{x}_{\mathcal{W}}^{(m-1)} \sim \mathcal{G}\left(s_{\sigma,k}^{0} + \frac{n_k^{(m)}}{2},\ r_{\sigma,k}^{0} + \frac{\mathrm{SSR}_{k,\mathcal{W}}^{(m)}}{2} \right),$$

and set

$$\sigma_{k,\mathcal{W}}^{(m)} = \left(\gamma_{k,\mathcal{W}}^{(m)} \right)^{-1/2}.$$

4. **Mixture probability.** Draw

$$\omega_{\mathcal{W}}^{(m)} \mid \mathbf{q}_{\mathcal{W}}^{(m)} \sim \mathcal{B}\left(s_{1,\omega}^{0} + n_1^{(m)},\ s_{2,\omega}^{0} + n_2^{(m)} \right).$$

Conditional on the mixture indicators, the generalized reset-price process reduces to a heteroskedastic Gaussian regression, so all parameter updates remain available in closed form.

B.5 Additional Monte Carlo Results

This appendix reports additional results from the Monte Carlo exercise in Section 3.3. We first examine post-estimation diagnostics for the Gibbs sampler and then assess estimator performance under alternative data-generating processes.

Post-estimation diagnostics. Table A.2 reports post-burn-in diagnostics for the Gibbs sampler. Panel (a) shows that the Metropolis–Hastings acceptance rate averages about 34 percent for both $T = 12$ and $T = 24$, close to the 33 percent target, with relatively little variation across replications. This confirms that the proposal fine-tuning procedure reliably brings acceptance rates close to their target. Panel (b) reports the Monte Carlo standard error of the posterior mean for each parameter. These errors are small for all parameters and generally decline with the length of the panel, indicating that simulation error in the posterior means is negligible and that the retained draws provide sufficient effective sample sizes. These diagnostics complement the negligible bias and high precision documented in Section 3.3 by showing that the sampler also performs well numerically.

Alternative data-generating processes. We next assess estimator performance under calibrations closer to the menu-cost and Calvo extremes, where the relative scarcity of different types of price changes may make the inaction thresholds harder to identify. Starting from the baseline calibration, we vary only the probability of a free adjustment opportunity, λ, while holding the remaining parameters fixed. The Low Calvoness calibration sets $\lambda = 0.25\%$, generating a degree of Calvoness of approximately 5 percent, so that very few observed price changes occur from within the inaction region. The High Calvoness calibration instead sets $\lambda = 32.63\%$, five times its baseline value, generating a degree of Calvoness of approximately 95 percent and leaving relatively few price changes triggered by crossing the inaction thresholds.

Table A.3 reports the bias of posterior mean estimates under these two data-generating processes. For each calibration, we report results both when all parameters are estimated jointly (Full) and when the inaction bands are fixed at their true values (Fixed). The latter provides a direct check on whether difficulty estimating the bands is responsible for any deterioration in the recovery of the remaining parameters and latent price gaps. Bias remains small in both environments. Under Low Calvoness, the aggregate price-gap moments and most parameters are recovered with negligible bias. The free adjustment probability (λ) exhibits a larger relative bias because its true value is close to zero, but this error remains small in absolute terms (+0.20 percentage points) and does not distort the aggregate moments. Under High Calvoness, the bias in the estimated inaction thresholds increases somewhat, as expected given the scarcity of threshold-triggered price changes, but remains small in absolute terms. Fixing the bands has little effect on the biases of the remaining parameters or the moments of the price-gap distribution. In particular, the mean, standard deviation, skewness, and kurtosis of price gaps remain accurately recovered in both calibrations. These results indicate that the estimator continues to perform well even when the data-generating process lies close to either pricing extreme.

C Data Appendix

C.1 Data cleaning: product substitutions, sales, outliers and quote-line gaps

The model in Section 2 does not feature sales or product substitutions and, hence, the state-space representation above should be interpreted as a data generating process for *regular prices*. To ensure consistency between the model and the data, we remove product substitutions and temporary sales from the data in order to obtain quote-lines of regular prices.

Product substitutions. For each outlet where prices are collected, price collectors begin by selecting a single product - among those that match the item's specification - that best reflects what consumers typically purchase in that area. Once selected, the same product is priced each month at that outlet. However, substitutions are necessary if the original product is no longer available or if its physical characteristics, such as weight or size, change. In such cases, the price collector flags the quote accordingly. We use these flags to identify product substitutions.

Product sales. The prices collected by the price collectors are shelf prices, meaning they may reflect temporary sales at the time of collection. If a price is affected by a sale, it is flagged using a sales indicator. We use these sales flags to construct quote-lines of regular prices. First, we treat any price not flagged as a sale as the regular price for that period. Second, if a price is flagged as a sale, we use the most recent regular price observed before the sale. If no such prior regular price exists, we use the first regular price observed after the sale. Third, if no regular prices are observed either before or after the sale, we treat the quote-line as composed exclusively of sale prices and exclude it from the sample.

Outliers. After cleaning the data for product substitutions and sales, we remove price changes for which the absolute log price change is smaller than 0.01 or larger than 2.

Quote-line gaps. Following these cleaning steps, quote-lines may contain gaps because a price quote failed ONS internal validation, because of a product substitution, or because an outlier price change was removed. Since our estimation is applied to price trajectories of contiguous observations, we split quote-lines at each such gap into separate segments of contiguous regular-price observations. We then exclude any resulting singleton quote-lines from the sample.

C.2 Estimation windows

This appendix details the construction of the item-windows used in the estimation of Section 4. For item j, let

$$\mathbf{p}^j = \{p_{i,t} : i \in \mathcal{I}_j,\ t \in \mathcal{T}_i\}$$

denote the collection of observed prices belonging to the item, and let $\underline{t}_j$ and $\bar{t}_j$ denote the first and last dates at which the item appears in the sample. We partition the interval

$\{\underline{t}_j, \ldots, \bar{t}_j\}$ sequentially into non-overlapping estimation windows

$$\mathcal{T}_{j,k} = \{\underline{t}_{j,k}, \ldots, \bar{t}_{j,k}\}, \qquad k = 1, \ldots, K_j.$$

Window construction. Set $\underline{t}_{j,1} = \underline{t}_j$. For each window k, let $\tau_{j,k}^{(0)}$ denote the first January strictly following $\underline{t}_{j,k}$.[53] Define the sequence of candidate endpoints by

$$\tau_{j,k}^{(r)} = \tau_{j,k}^{(0)} + 12r, \qquad r = 0, 1, 2, \ldots,$$

where additions are measured in months. For each candidate endpoint $\tau_{j,k}^{(r)}$, let $N_{j,k}^{+}(r)$ and $N_{j,k}^{-}(r)$ denote the number of positive and negative price changes, respectively, between $\underline{t}_{j,k}$ and $\tau_{j,k}^{(r)}$, excluding price changes recorded in the first month of the window. We set

$$r_{j,k}^{*} = \min\left\{r \geqslant 0 : N_{j,k}^{+}(r) \geqslant 10, \quad N_{j,k}^{-}(r) \geqslant 10\right\},$$

and, whenever $\tau_{j,k}^{(r_{j,k}^{*})} \leqslant \bar{t}_j$, set

$$\bar{t}_{j,k} = \tau_{j,k}^{(r_{j,k}^{*})}.$$

The next window begins in the following month,

$$\underline{t}_{j,k+1} = \bar{t}_{j,k} + 1.$$

Since non-terminal windows end in January, subsequent windows begin in February. If the item exits the sample before another candidate window satisfies both thresholds, the final window ends at $\bar{t}_j$. Price changes recorded in the first month of each window are excluded from the counts because they reflect adjustments relative to prices observed in the preceding window and could otherwise generate windows containing very little information about the adjustment thresholds.

Terminal windows. If the final window contains at least ten positive and ten negative price changes, it is estimated in the same way as all preceding windows. Otherwise, we do not re-estimate the state-space parameters. Instead, we fix them at the posterior means from the item's most recent estimable window and run only the forward-filtering backward-sampling step to recover the latent states.

The resulting windows are overwhelmingly annual: across the 1,337 items we obtain 15,664 item-windows, 88 percent of which span exactly twelve months. For the 151 terminal windows that do not satisfy the minimum-price-change requirement, fewer than 1 percent of the total, parameters are held fixed as described above and only the latent states are recovered.

[53] We align windows with January and February because the ONS typically updates the set of items collected at the start of the year. In our data, most items that enter the sample do so in February, while most items that exit do so in January.

C.3 Macroeconomic Data

All macroeconomic data used in Section 5 are described in Table A.4.

C.4 Reset-price inflation and pricing-pressure indices

This appendix describes the construction of two alternative measures based on the estimated reset prices that we use in the forecasting exercise as alternatives to the moments of the price-gap distribution: reset-price inflation and pricing-pressure indices.

Reset-price inflation. To construct reset-price inflation in the spirit of Bils et al. (2012), we adapt the methodology used to construct the official CPI, as described in Office for National Statistics (2019), replacing observed prices with estimated reset prices.

Elementary aggregates. The lowest level at which individual price quotes are aggregated is the stratum. We define a stratum as a unique combination of an item and its applicable stratification dimensions. Depending on the item, these dimensions are region, shop type, both region and shop type, or neither.[54] Within each stratum, the CPI primarily aggregates individual price relatives using the geometric mean, or Jevons formula. We apply the same aggregation to changes in estimated reset prices. In particular, let $\mathcal{I}_{s,t|t-1}$ denote the set of quote-lines in stratum s for which reset prices are available in both $t-1$ and t. The month-on-month reset-price elementary aggregate is

$$\mathcal{I}^r_{s,t|t-1} = \exp\left\{ \frac{1}{N_{s,t|t-1}} \sum_{i \in \mathcal{I}_{s,t|t-1}} \Delta \widehat{p}^r_{i,t} \right\}, \tag{51}$$

where $N_{s,t|t-1}$ is the corresponding number of matched quotes, incorporating the ONS central shop weights as replication factors, and $\Delta \widehat{p}^r_{i,t}$ is the estimated change in the reset price of quote-line i.[55] We use the previous month as the base period and chain-link the resulting elementary aggregates over time.[56]

Aggregation. We aggregate the reset-price elementary indices following the same hierarchy and weights used for observed prices. Elementary aggregates are first combined using stratum weights to obtain item-level indices. Item-level indices are then aggregated using the corresponding CPI expenditure weights to obtain COICOP indices and, ultimately, an aggregate reset-price index. Reset-price inflation is computed from this aggregate index in the same way as observed CPI inflation.

[54] There are 12 regions and two shop types: multiple retailers, defined as those with 10 or more outlets, and independent retailers, defined as those with fewer than 10 outlets.

[55] For the forecasting exercises, $\Delta \widehat{p}^r_{i,t}$ is computed using only information available through date t, using the corresponding real-time reset-price vintages.

[56] This differs from the construction of the published CPI, which generally uses the previous January as the base period. Because quote-lines are split at product substitutions and outlier price changes, some observations do not have a corresponding reset price in the previous January. Using January as the base would therefore require either assigning a base-period reset price or dropping these observations. We instead use the previous month as the base period and chain-link the indices over time.

Pricing-pressure indices. Our second alternative measure is inspired by the generalized-*Ss* framework of Caballero and Engel (2007), in which aggregate price adjustment is obtained by integrating price gaps against their adjustment hazard and cross-sectional distribution. We adapt this idea to our estimated model by evaluating the implied adjustment hazard over the one-period-ahead predictive distribution of the counterfactual price gap. For quote-line i, let $\widetilde{s}_{i,t+1}$ denote the counterfactual gap at $t+1$ if the price is not adjusted. Conditional on information at t,

$$\widetilde{s}_{i,t+1} = s_{i,t} - \mu - \varepsilon_{i,t+1}, \qquad \widetilde{s}_{i,t+1} \mid \mathcal{F}_t \sim \mathcal{N}(m_{i,t}, \sigma_\varepsilon^2), \qquad m_{i,t} \equiv s_{i,t} - \mu. \tag{52}$$

To simplify notation, define $a_{i,t} \equiv (\underline{s} - m_{i,t})/\sigma_\varepsilon$ and $b_{i,t} \equiv (\bar{s} - m_{i,t})/\sigma_\varepsilon$, and let $\phi(\cdot)$ and $\Phi(\cdot)$ denote the standard normal density and distribution functions, respectively. The expected price adjustment can then be decomposed into three components:

$$\mathrm{PP}^L_{i,t} = -m_{i,t}\Phi(a_{i,t}) + \sigma_\varepsilon\phi(a_{i,t}), \tag{53}$$

$$\mathrm{PP}^C_{i,t} = \lambda\left[-m_{i,t}\{\Phi(b_{i,t}) - \Phi(a_{i,t})\} + \sigma_\varepsilon\{\phi(b_{i,t}) - \phi(a_{i,t})\}\right], \tag{54}$$

$$\mathrm{PP}^U_{i,t} = -m_{i,t}\{1 - \Phi(b_{i,t})\} - \sigma_\varepsilon\phi(b_{i,t}), \tag{55}$$

corresponding, respectively, to upward adjustment below the lower threshold, free adjustment inside the inaction region, and downward adjustment above the upper threshold. The lower-threshold component is weakly positive, the upper-threshold component is weakly negative, and the interior component can have either sign.[57]

Aggregation. We aggregate each component to the economy-wide level using the combined CPI weights for each quote-line. Let $w_{i,t}$ denote the product of the corresponding shop, stratum, and item weights. For $k \in \{L, C, U\}$, the aggregate pricing-pressure index is

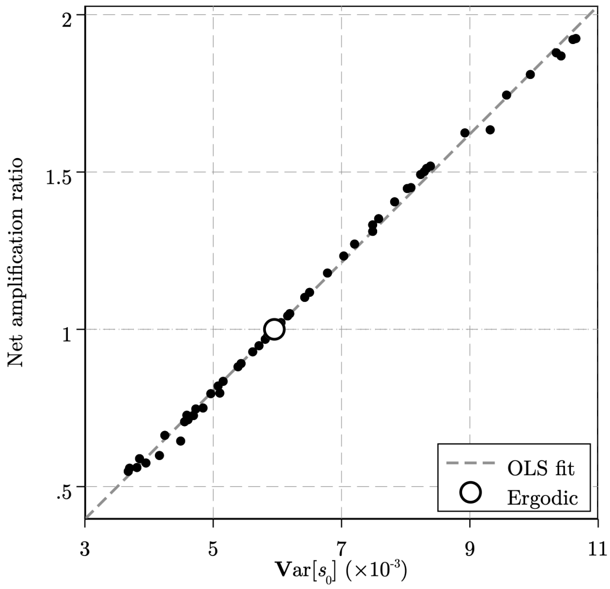
$$\mathrm{PP}^k_t = \frac{\sum_i w_{i,t}\mathrm{PP}^k_{i,t}}{\sum_i w_{i,t}}. \tag{56}$$

These three aggregate components are the pricing-pressure measures used in the forecasting exercise.

[57] For the forecasting exercises, all inputs are constructed using only information available through date t. The price gap $s_{i,t}$ is taken from the corresponding real-time vintage, while μ, σ_ε, λ, $\underline{s}$, and $\bar{s}$ are set to their estimates from the most recently completed estimation window available at t.

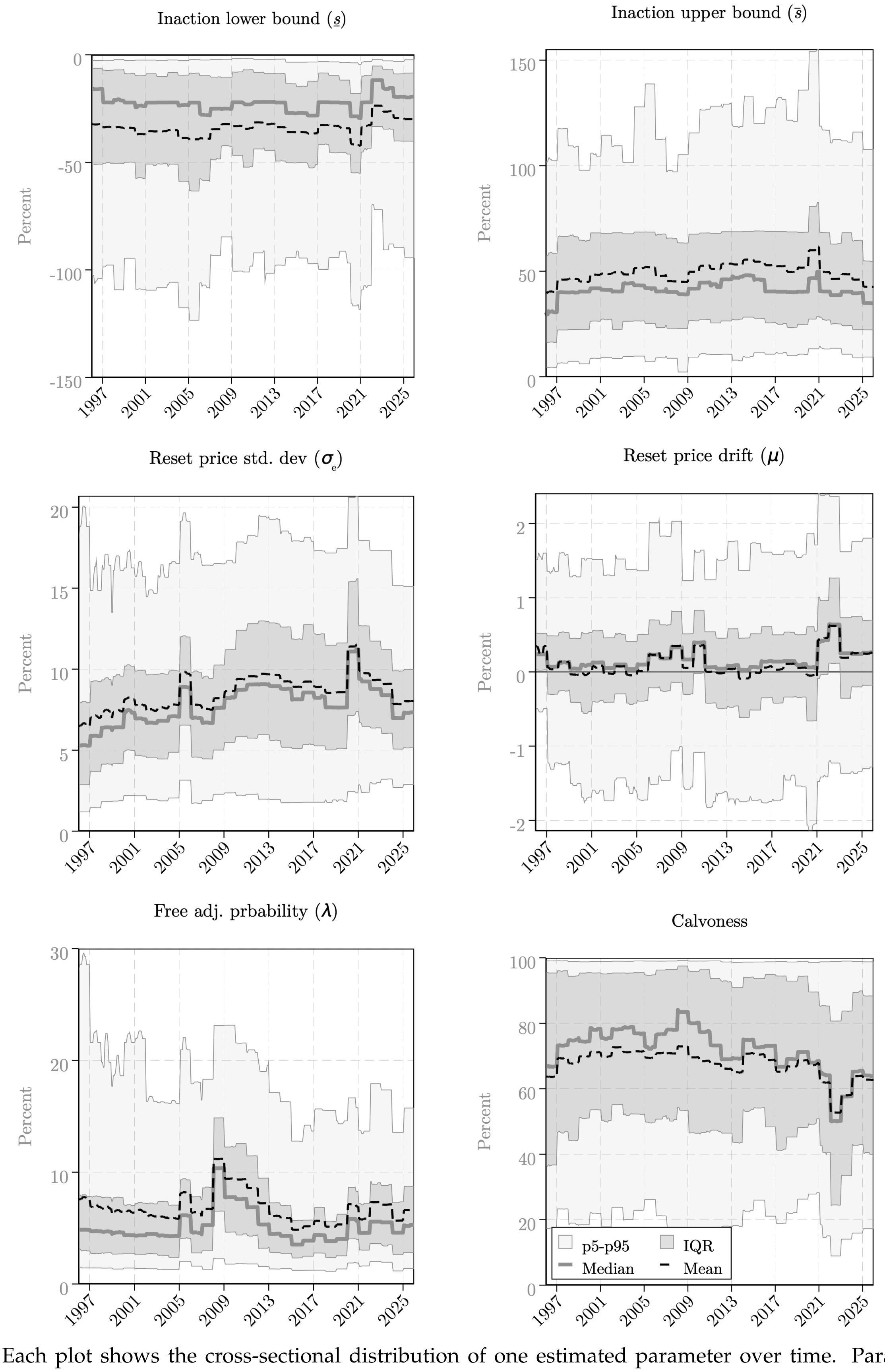
D Additional Figures

Figure A.1: Amplification ratio under volatility-shock backgrounds



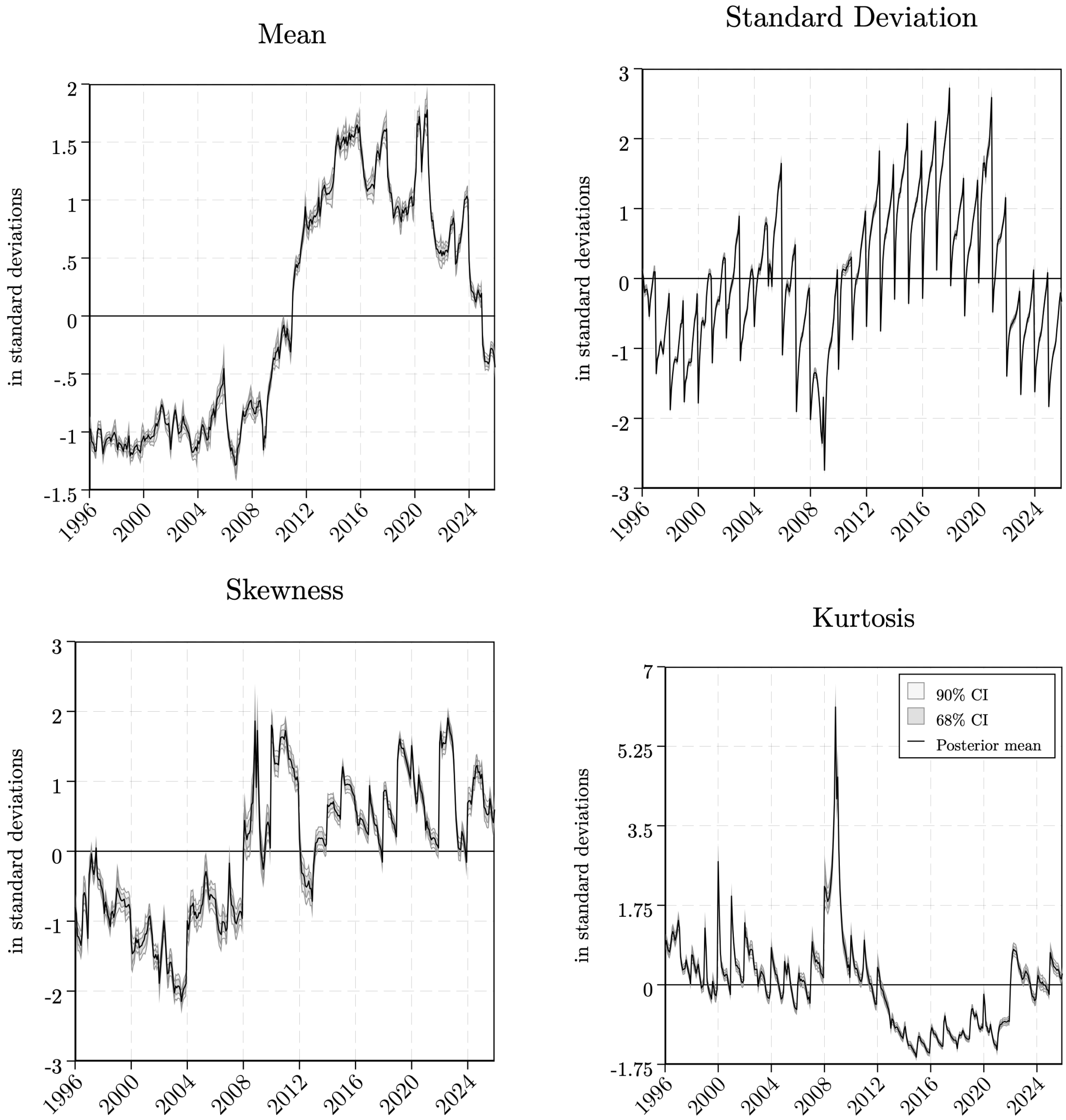
Notes: Analogous to the right panel of Figure 1, but with distributions displaced by persistent perturbations to idiosyncratic volatility σ_ε rather than to the drift. Each dot represents a distribution sampled at a different point along the transition back to the steady state. The x-axis plots the variance $\mathrm{Var}(s_t)$. The y-axis plots the net amplification ratio: impact excess inflation minus the no-shock counterfactual, normalized by the ergodic response. The star marks the ergodic distribution. Same calibration as Figure 1.

Figure A.2: Cross-sectional distribution of estimated parameters over time



Notes: Each plot shows the cross-sectional distribution of one estimated parameter over time. Parameters are estimated at the item-window level and held constant within each window; only windows for which both states and parameters are estimated are considered. The solid line is the cross-sectional median and the shaded area the interquartile range (p25–p75).

Figure A.3: Aggregate price gap moments over time with posterior uncertainty

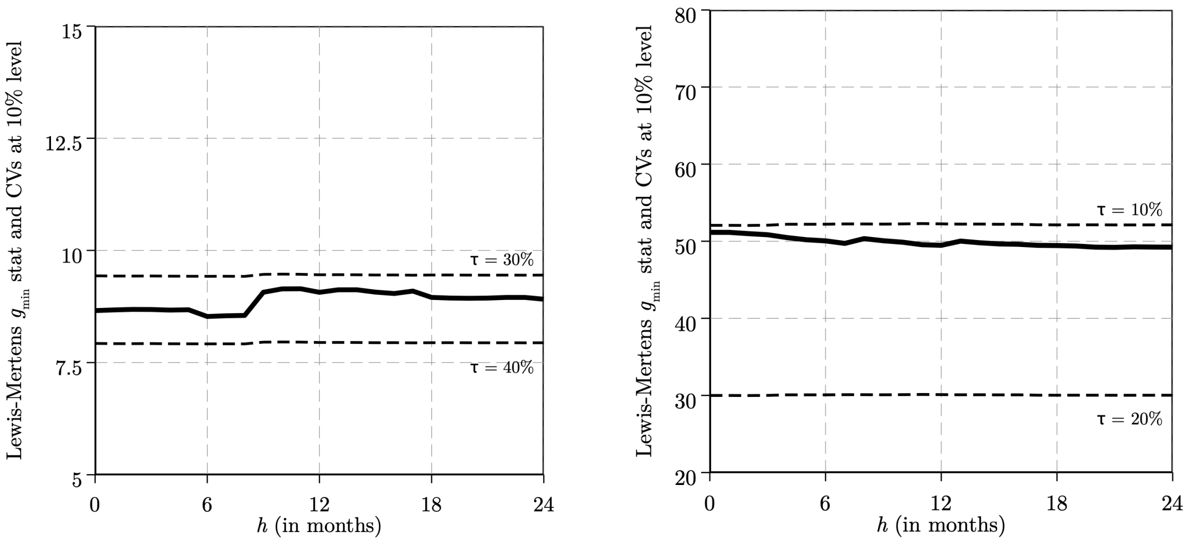


Notes: Each panel plots the CPI-weighted aggregate price gap moment over time, standardized by subtracting the draw-specific time-series mean and dividing by the draw-specific time-series standard deviation. The solid line is the posterior mean across draws. The dark shaded band is the 68% posterior credible interval and the light shaded band is the 90% posterior credible interval.

Figure A.4: Lewis and Mertens (2026) weak-IV diagnostic

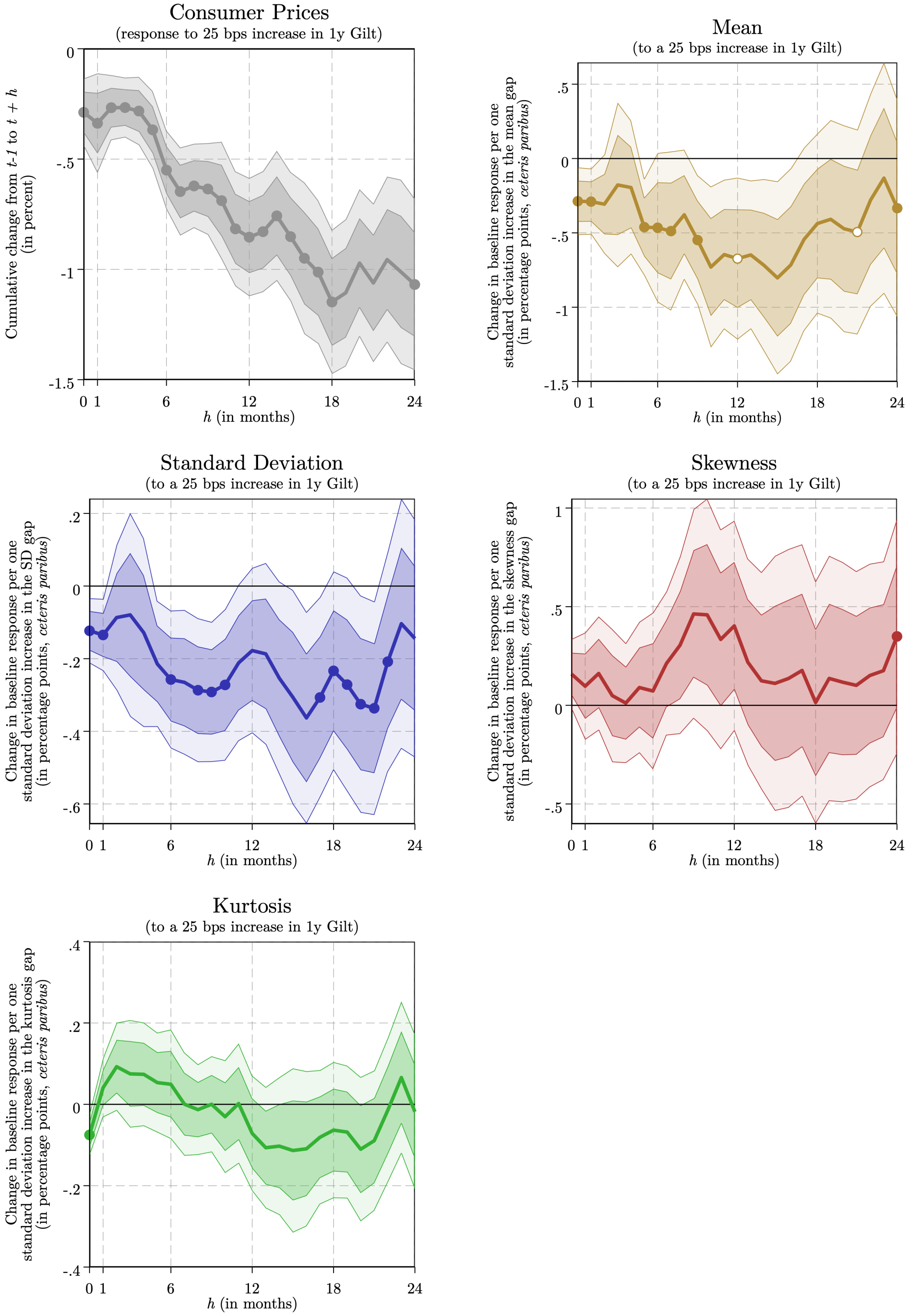
(a) Monetary policy

(b) Oil prices



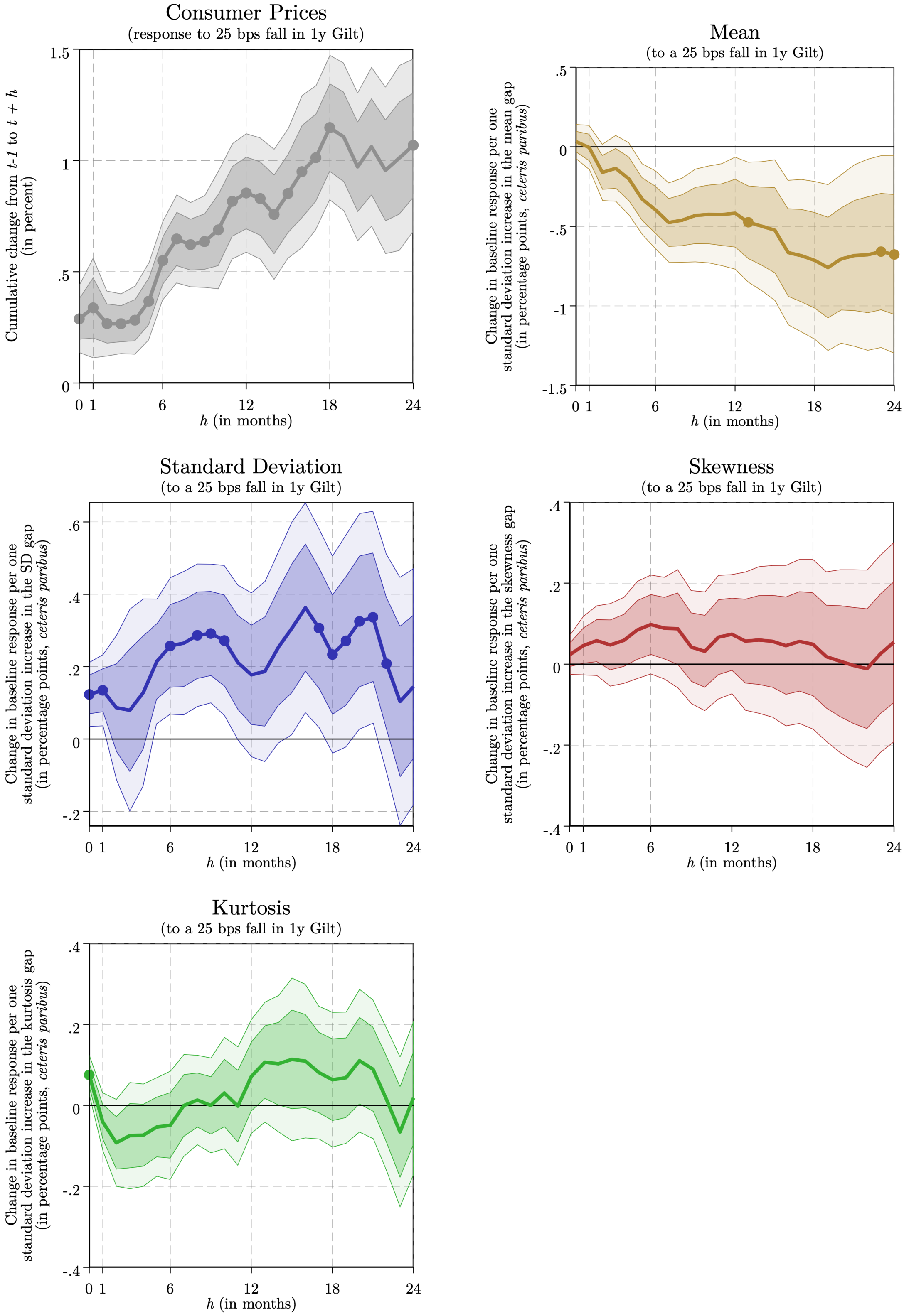
Notes: The solid line plots the Lewis and Mertens (2026) $g_{\min}$ statistic across horizons h. Dashed lines show critical values at a 10% significance level for two identification-robustness levels τ. Panel (a): changes in 1y gilts, instrumented with target, path, and QE surprises from Braun et al. (2025); dashed lines at $\tau = 30\%$ (upper) and $\tau = 40\%$ (lower). Panel (b): changes in WTI oil price, instrumented with oil supply news shocks from Känzig (2021); dashed lines at $\tau = 10\%$ (upper) and $\tau = 20\%$ (lower).

Figure A.5: Price gap moments and the transmission of a monetary tightening



Notes: Impulse response estimates from (20), with Δx_t equal to the change in the 1y Gilt and the instrument set given by the target, path, and QE factors from Braun et al. (2025). The first plot depicts the impulse response evaluated at the cross-sectional long-run means of all four moments (baseline effect). Each subsequent plot shows the incremental response when the respective moment increases by one standard deviation above its mean, holding all other moments fixed. All responses are scaled to a 25 bps increase in the 1y Gilt. Shaded areas denote 68% and 90% confidence intervals using conventional inference. Solid (hollow) circles indicate significance at the 10% (32%) level under Andrews (2018) identification-robust inference.

Figure A.6: Price gap moments and the transmission of a monetary easing

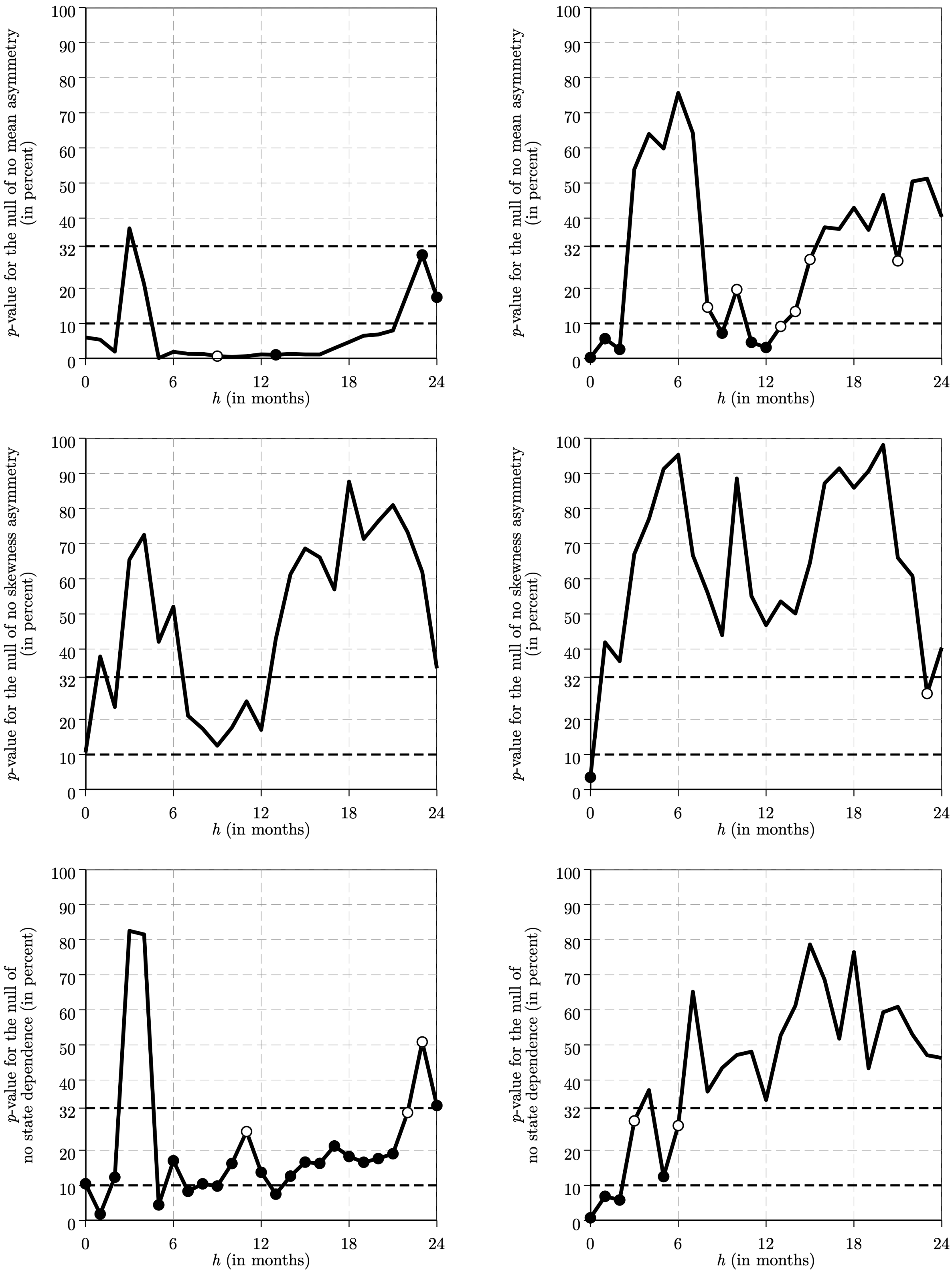


Notes: Impulse response estimates from (20), with Δx_t equal to the change in the 1y Gilt and the instrument set given by the target, path, and QE factors from Braun et al. (2025). The first plot depicts the impulse response evaluated at the cross-sectional long-run means of all four moments (baseline effect). Each subsequent plot shows the incremental response when the respective moment increases by one standard deviation above its mean, holding all other moments fixed. All responses are scaled to a 25 bps decrease in the 1y Gilt. Shaded areas denote 68% and 90% confidence intervals using conventional inference. Solid (hollow) circles indicate significance at the 10% (32%) level under Andrews (2018) identification-robust inference.

Figure A.7: Hypothesis tests for state-dependent transmission

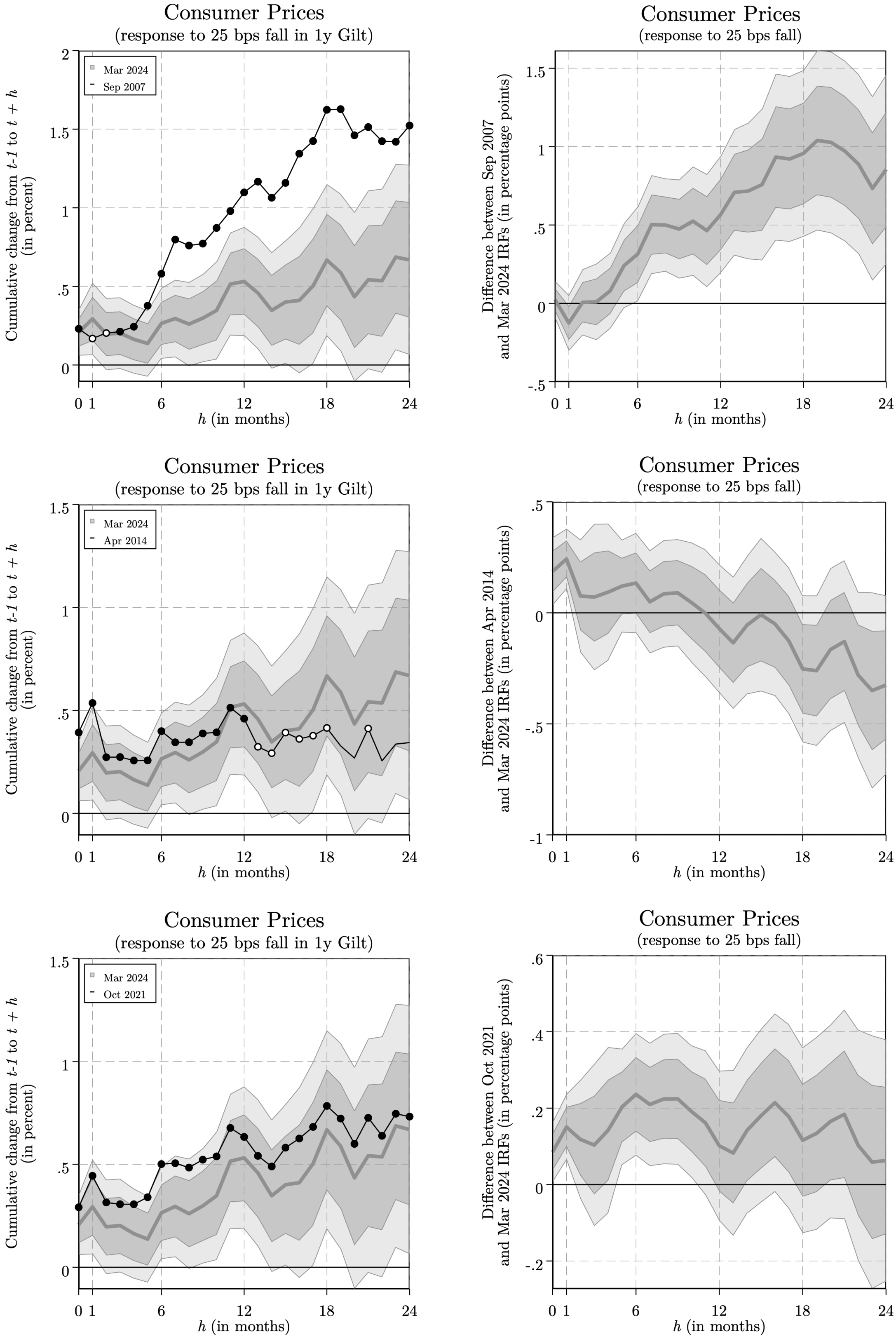
(a) Monetary policy

(b) Oil prices



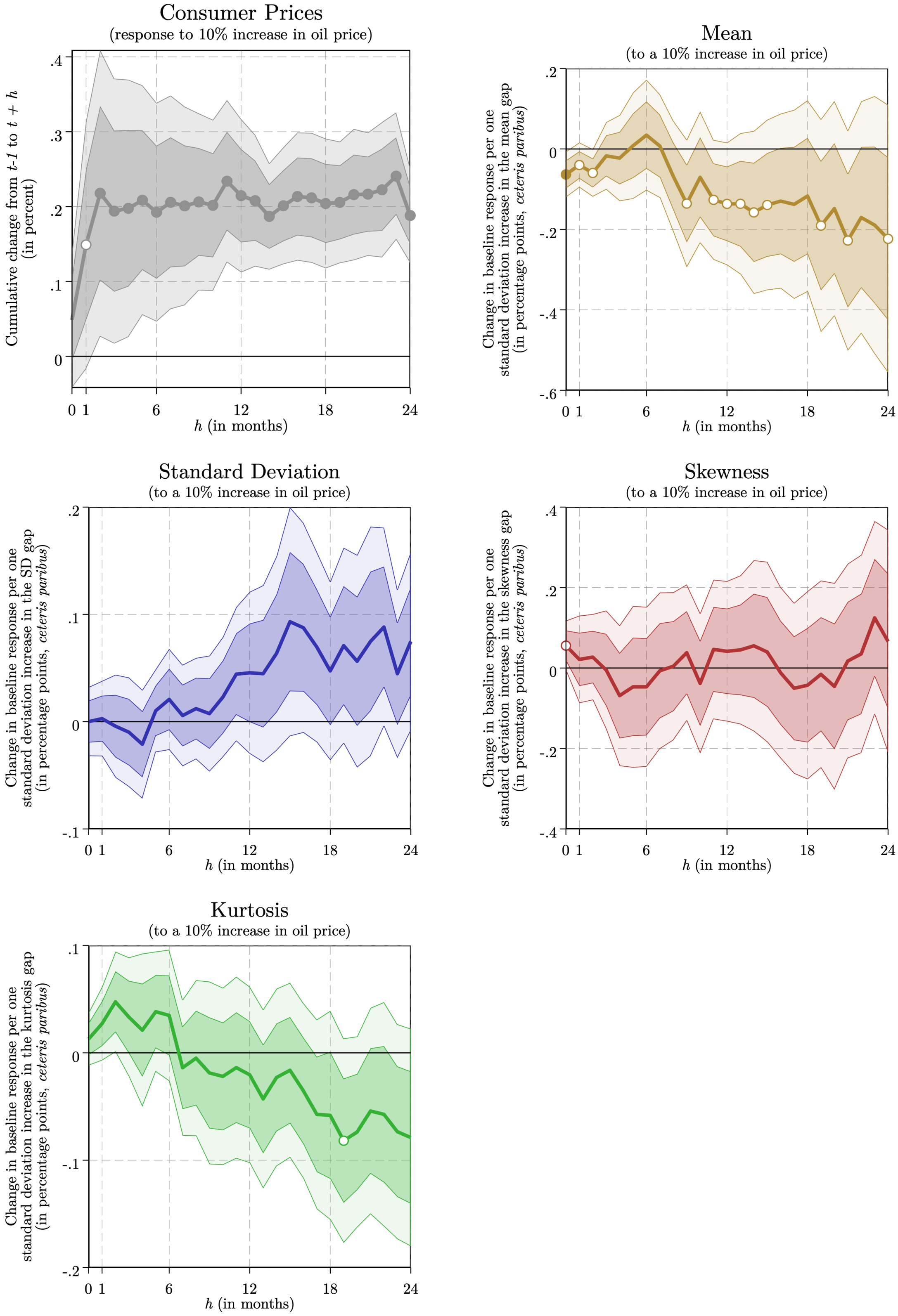
Notes: Each row reports p-values (in percent) across horizons h for one of three tests from (20): first row for no mean asymmetry ($\text{H}_0\colon \beta_h^+ = \beta_h^-$), second row for no skewness asymmetry ($\text{H}_0\colon \zeta_h^+ = \zeta_h^-$), and third row for no state dependence (H_0: all six interaction coefficients equal to zero). Dashed lines mark the 32% and 10% significance levels. Panel (a) uses the 1y Gilt instrumented with target, path, and QE factors from Braun et al. (2025). Panel (b) uses Oil prices instrumented with oil supply news shocks from Känzig (2021). Solid (hollow) circles indicate significance at the 10% (32%) level under Andrews (2018) identification-robust inference.

Figure A.8: Responses to a monetary easing selected dates



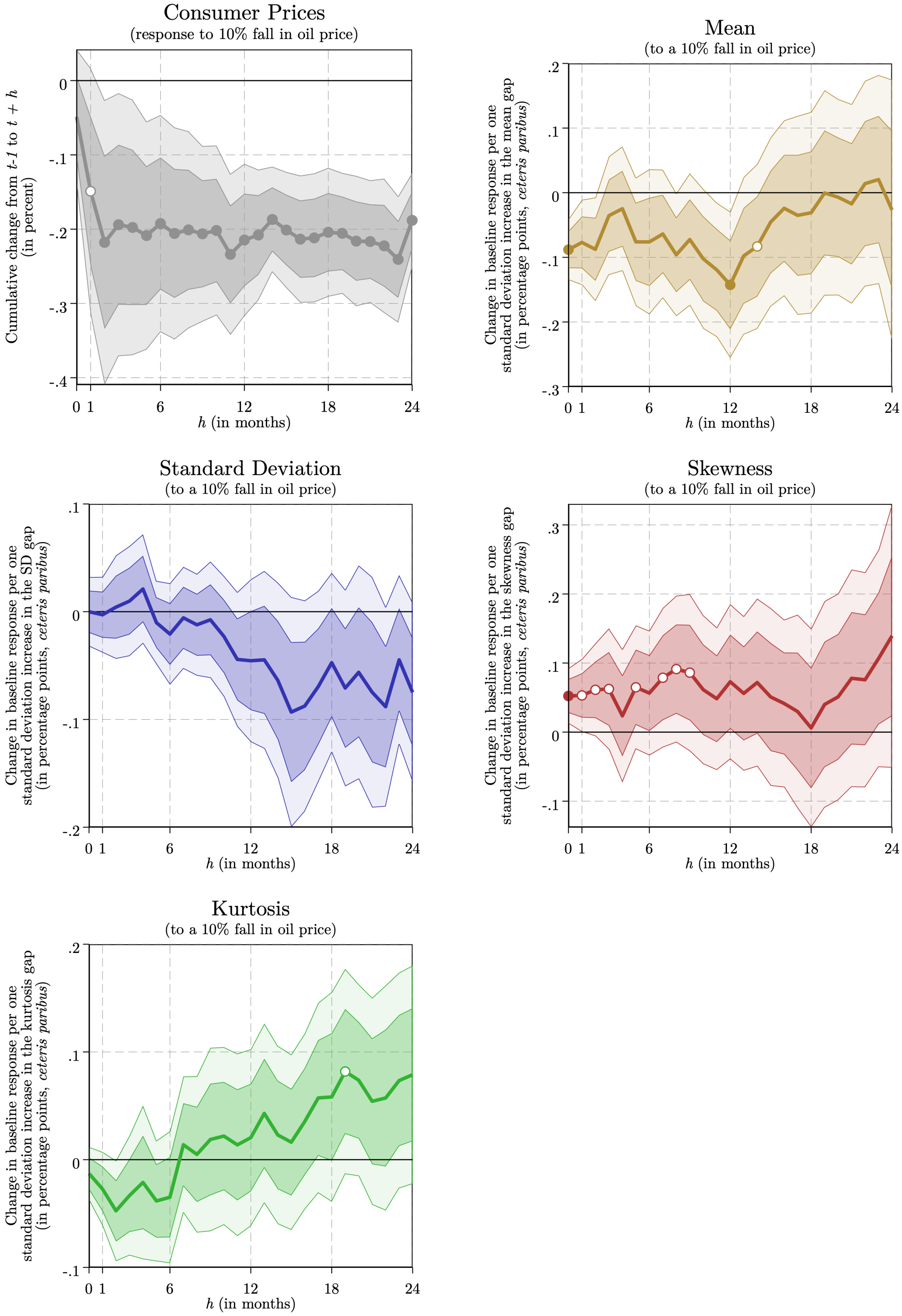
Notes: The left column plots the estimated impulse response of cumulative inflation to a 25 bps decrease in the 1y Gilt rate from (20), with Δx_t equal to the change in the 1y Gilt instrumented by the target, path, and QE factors from Braun et al. (2025). The grey line with shaded bands is the response evaluated at the benchmark distribution (March 2024); the black line is the response evaluated at the alternative date (September 2007, April 2014, and October 2021 in rows 1, 2, and 3, respectively). The right column plots the difference between the impulse response at the alternative date and at the benchmark date, with shaded areas denoting the 68% and 90% confidence bands. A white (black) filled circle indicates that the difference is statistically significant at the 32% (10%) level.

Figure A.9: Price gap moments and the transmission of an oil price increase



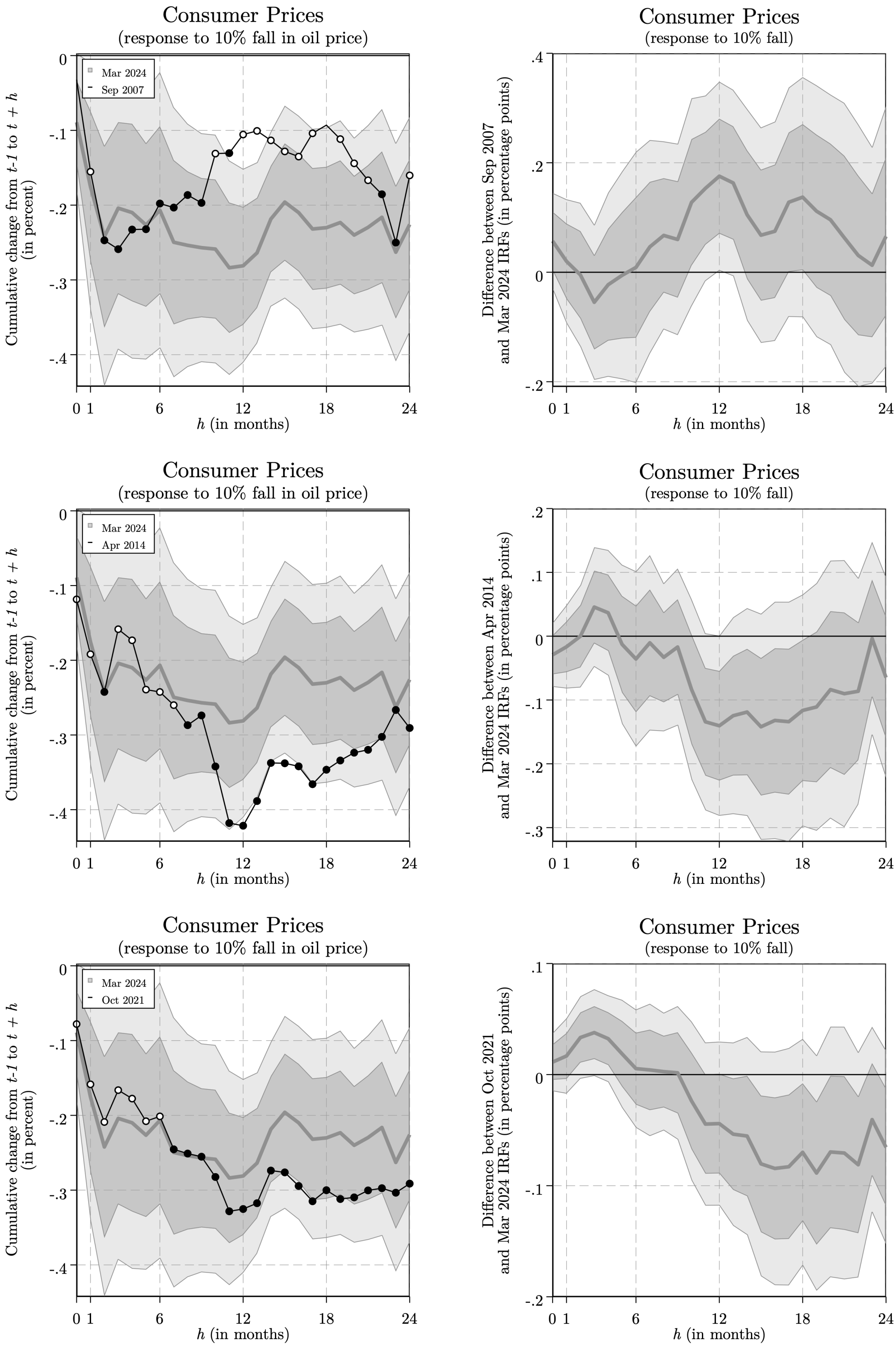
Notes: Impulse response estimates from (20), with Δx_t equal to the log change in oil prices and the instrument given by oil supply news shocks from Känzig (2021). The first plot depicts the impulse response evaluated at the cross-sectional long-run means of all four moments (baseline effect). Each subsequent plot shows the incremental response when the respective moment increases by one standard deviation above its mean, holding all other moments fixed. All responses are scaled to a 10% increase in oil prices. Shaded areas denote 68% and 90% confidence intervals using conventional inference. Solid (hollow) circles indicate significance at the 10% (32%) level under Andrews (2018) identification-robust inference.

Figure A.10: Price gap moments and the transmission of an oil price decrease



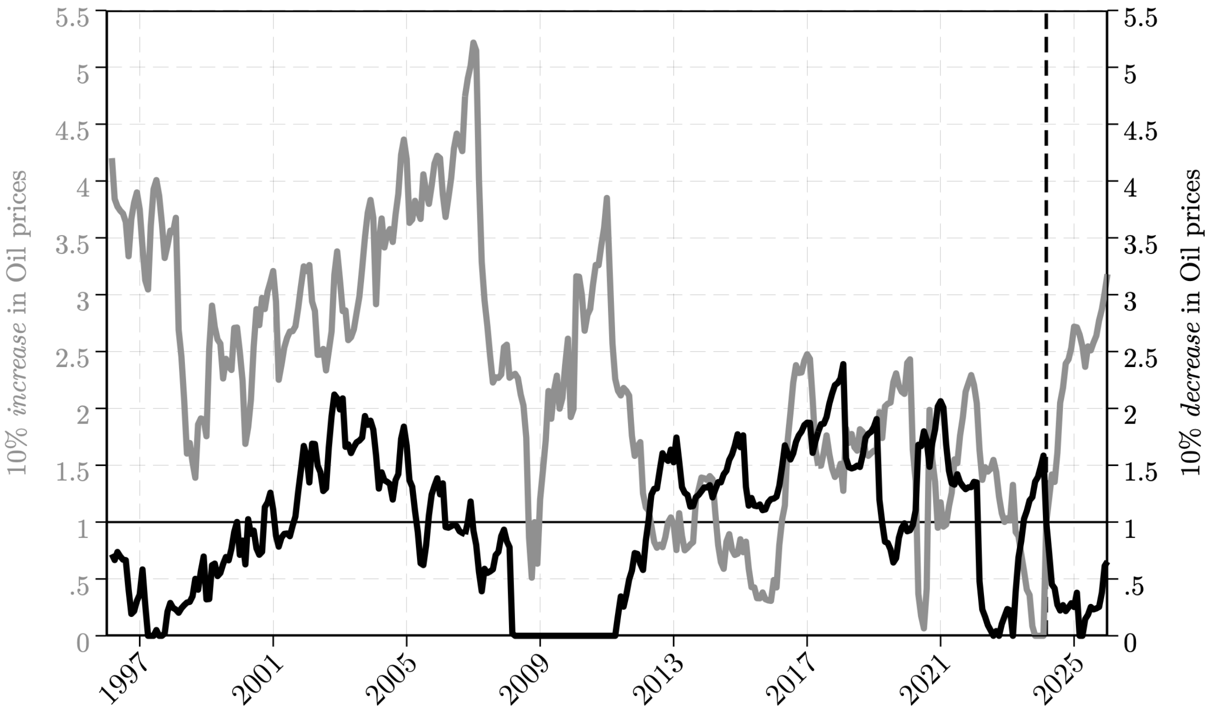
Notes: Impulse response estimates from (20), with Δx_t equal to the log change in oil prices and the instrument given by oil supply news shocks from Känzig (2021). The first plot depicts the impulse response evaluated at the cross-sectional long-run means of all four moments (baseline effect). Each subsequent plot shows the incremental response when the respective moment increases by one standard deviation above its mean, holding all other moments fixed. All responses are scaled to a 10% decrease in oil prices. Shaded areas denote 68% and 90% confidence intervals using conventional inference. Solid (hollow) circles indicate significance at the 10% (32%) level under Andrews (2018) identification-robust inference.

Figure A.11: Responses to an oil price decrease at selected dates



Notes: The left column plots the estimated impulse response of cumulative inflation to a 10% decrease in oil prices from (20), with Δx_t equal to the log change in oil prices instrumented by oil supply news shocks from Känzig (2021). The grey line with shaded bands is the response evaluated at the benchmark distribution (March 2024); the black line is the response evaluated at the alternative date (September 2007, April 2014, and October 2021 in rows 1, 2, and 3, respectively). The right column plots the difference between the impulse response at the alternative date and at the benchmark date, with shaded areas denoting the 68% and 90% confidence bands. A white (black) filled circle indicates that the difference is statistically significant at the 32% (10%) level.

Figure A.12: Time-varying oil price passthrough at the 2-year horizon



Notes: The figure plots the estimated cumulative inflation passthrough to oil price shocks 24 months ahead, relative to the benchmark distribution (2024m3). The grey line (left axis) and the black line (right axis) show the passthrough for a 10% increase and decrease in oil prices, respectively. Both series are normalised to 1 at 2024m3 (vertical dashed line): values above 1 indicate a larger-than-reference passthrough, values below 1 a smaller one. The passthrough is set to zero before normalisation if the estimated response to a positive oil price increase is negative, or the response to a negative oil price increase is positive. The passthrough is estimated based on coefficients from (20) with Δx_t equal to the log change in oil prices and the instrument given by oil supply news shocks from Känzig (2021) and the standardized moments of the aggregate price gaps distribution over time.

E Additional Tables

Table A.1: Calibration for Simulated Illustrations

Parameter	Description	Value	Source	Data	Model
(a) Parameters set externally					
β	Discount factor (monthly)	0.9966	Nakamura and Steinsson (2010)		
μ	Steady-state drift (monthly)	0.0028	Nakamura and Steinsson (2010)		
ϵ	Demand elasticity	4	Nakamura and Steinsson (2010)		
(b) Parameters calibrated internally					
σ_ε	Idiosyncratic shock s.d.	0.039	Freq. of price change	8.7%	8.71%
λ	Free adjustment probability	0.065	Share of free adjustments	75%	74.95%
$\bar{c}$	Menu cost (normalized)	0.090	Median $\lvert\Delta p\rvert$	8.5%	8.49%
(c) Implied *Ss* rule					
$\underline{s}$	Lower threshold	-0.202			
$\bar{s}$	Upper threshold	0.189			

Notes: Target moments are from Nakamura and Steinsson (2010). The free adjustment probability λ is pinned directly by the free adjustment share, $\lambda = 0.75 \times 0.087$; σ_ε and $\bar{c}$ are jointly calibrated to match the remaining two target moments. The *Ss* rule parameters are obtained from the firm's optimal policy given the structural parameters. Thresholds are expressed in terms of the reset price gap.

Table A.2: Monte Carlo Post-Estimation Diagnostics

	$T = 12$			$T = 24$		
	Average	Minimum	Maximum	Average	Minimum	Maximum
(a) Post-burn-in acceptance rate						
Accept. rate	34.08	27.24	41.62	34.28	28.14	41.88
(b) MCSE of Posterior Mean						
$\underline{s}$	0.01	0.00	0.11	0.00	0.00	0.04
$\bar{s}$	0.01	0.00	0.13	0.01	0.00	0.05
μ	0.00	0.00	0.01	0.00	0.00	0.00
σ_e	0.02	0.01	0.02	0.01	0.01	0.01
λ	0.01	0.01	0.01	0.00	0.00	0.01

Notes: Monte Carlo averages over 1,000 replications. Each Monte Carlo replication is a balanced panel of 300 quote-lines and $T = 12$ or $T = 24$ periods. Panel (a) reports the post-burn-in Metropolis–Hastings (MH) acceptance rate in percent; the algorithm targets a rate of 33%. Panel (b) reports the Monte Carlo standard error (MCSE) of the posterior mean for each structural parameter, defined as the posterior standard deviation divided by the square root of the effective sample size (ESS). All values in Panel (b) are in percentage points.

Table A.3: Monte Carlo Results: Low Calvoness vs. High Calvoness

	Low Calvoness			High Calvoness		
	True	Full	Fixed	True	Full	Fixed
(a) Parameters						
$\underline{s}$	-20.18	-0.00	0.00	-20.18	-0.28	0.00
$\bar{s}$	18.90	0.01	0.00	18.90	0.14	0.00
μ	0.28	-0.06	-0.06	0.28	-0.06	-0.05
σ_e	3.85	0.03	0.04	3.85	0.11	0.12
λ	0.25	0.20	0.20	32.63	-0.93	-1.00
(b) Cross-sectional moments of price gaps*						
Mean	-0.89	-0.04	-0.05	-0.42	0.00	0.01
Std Dev	9.15	-0.03	-0.04	6.04	-0.15	-0.18
Skewness	0.02	-0.02	-0.02	-0.16	-0.01	-0.00
Kurtosis	2.29	0.03	0.03	4.47	0.24	0.20
Calvoness	4.90	0.08	-0.00	94.90	0.25	0.00

Notes: The "True" column reports DGP parameter values in Panel (a) and cross-sectional price gap moments (including zeros) and Calvoness averaged over time periods and across replications in Panel (b). The Full and Fixed columns report the mean bias of the posterior mean across 1,000 replications, averaging first over time periods within each replication for Panel (b). "Full" refers to estimation with all parameters free; "Fixed" refers to estimation with inaction bands ($\underline{s}$ and $\bar{s}$) fixed at their true DGP values. The Low Calvoness calibration sets $\lambda = 0.25\%$ to yield an approximate Calvoness of 5%; the High Calvoness calibration sets $\lambda = 32.63\%$ (five times the baseline calibration value) to yield an approximate Calvoness of 95%, where Calvoness is the share of price changes with size within the inaction band. The remaining parameters are kept at their baseline calibration values. True values are in percent, except skewness and kurtosis which are scale-free. All Full and Fixed columns are in percentage points, except for skewness and kurtosis which are scale-free. Each replication is a balanced panel of 300 quote-lines with $T = 12$ periods. * The average correlation between estimated (posterior mean) and true individual price gaps is 0.58 (Full) and 0.58 (Fixed) for the Low Calvoness DGP, and 0.79 (Full) and 0.79 (Fixed) for the High Calvoness DGP.

Table A.4: Description, Sources, and Coverage of Macro Data

Variable	Source	Sample
(a) Price variables		
1-year Gilt yield (percent)[a]	BoE	1996m1–2026m1
WTI crude oil spot price (WTISPLC)	FRED	1996m1–2026m1
Class level COICOP price indices	ONS	1996m1–2026m1
Aggregate UK CPI index	ONS	1996m1–2026m1
(b) Other macroeconomic variables		
Real GDP index	ONS	1997m1–2026m1
Sterling effective exchange rate index	ONS	1997m1–2026m1
Unemployment rate (percent)	ONS	1996m1–2026m1
Corporate bond option-adjusted spreads[b]	BoE & Datastream	1997m1–2026m1
FTSE All-Share index	Datastream	1997m1–2026m1
(c) Shocks		
UK monetary policy target, path, and QE surprises[c]	Braun et al. (2025)	1997m6–2026m1
Oil supply news shock	Känzig (2021)	1975m1–2025m12

Notes: The Sample column reports the observations used in our analysis after merging each series with the relevant estimation sample.

[a] The 1-year Gilt yield is constructed as the monthly average of the daily 1-year nominal spot rate from the Bank of England government liability yield curve.

[b] The corporate bond spread series equals the BoE Millennium M11 option-adjusted spread on sterling industrial corporates (AAA–BBB) through November 2010, and the iBoxx GBP Non-Financials OAS (Refinitiv Datastream: .IBBGB0188OAS) from December 2010 onwards, with the join date and level shifts chosen to minimise the squared level discontinuity between the two series at the splice point, weighted by the number of observations on each side.

[c] The target, path, and QE surprises are aggregated from event to monthly frequency using triangular weights that account for the timing of events within the month. All events in the Braun et al. (2025) Factors dataset are included.

Table A.5: Price Gap Moments: Summary Statistics and Selected Episodes

	Mean	Std. dev.	Skewness	Kurtosis
(a) Descriptive statistics				
Mean	7.61	19.00	0.49	5.59
Median	6.53	18.96	0.54	5.57
Standard deviation	3.94	2.87	0.39	1.66
Minimum	2.53	11.18	-0.38	3.17
Maximum	14.63	26.50	1.73	13.20
(b) Raw price gap moments at selected dates				
Sep 2007	4.38	17.45	0.15	7.29
Apr 2014	13.36	22.18	0.73	3.87
Oct 2021	9.89	20.45	0.56	3.93
Mar 2024 (benchmark)	8.81	17.34	0.51	5.14
(c) Standardized price gap moments at selected dates				
Sep 2007	-0.82	-0.54	-0.87	1.02
Apr 2014	1.46	1.11	0.62	-1.04
Oct 2021	0.58	0.51	0.18	-1.00
Mar 2024 (benchmark)	0.30	-0.58	0.06	-0.27

Notes: Mean, standard deviation, skewness, and kurtosis of the aggregate price gap distribution. Panel (a) reports summary statistics over the full sample. Panels (b) and (c) report raw and standardized moments, respectively, at four selected dates. Three dates fall one year before inflation extremes: September 2007 precedes the pre-pandemic peak, October 2021 precedes the overall maximum, and April 2014 precedes the sample minimum in year-on-year inflation. The benchmark (March 2024) is the date that minimizes the sum of squares of standardized moments subject to 12 months ahead being within the sample. In panel (c), each moment is standardized by subtracting its full-sample mean and dividing by its full-sample standard deviation.

Table A.6: Price Gaps *vs.* Random-Walk: Alternative Inflation Metrics

Horizon	OOS Forecast Evaluation Sample		
	Full sample	Low inflation	High inflation
(a) Average monthly inflation rate, t to $t+h$			
3	1.17*	1.13*	1.17*
6	1.15	1.10	1.16
9	1.03	1.10	1.02
12	0.92	1.13	0.91
15	0.84*	1.18	0.81*
18	0.78*	1.22	0.73**
21	0.73**	1.17	0.68**
24	0.69**	1.11	0.64**
(b) Quarterly inflation rate at $t+h$			
3	1.17*	1.13*	1.17*
6	1.03	0.99	1.03
9	0.90	0.88	0.91
12	0.84*	0.78*	0.84*
15	0.77**	0.80*	0.77**
18	0.73**	0.86	0.71**
21	0.72**	0.74**	0.72**
24	0.75**	0.70**	0.75**

Notes: RMSPE ratio of Price Gaps relative to the random-walk benchmark, adapted to match each inflation metric. Panel (a) forecasts the average monthly inflation rate between t and $t+h$; the random-walk benchmark is the average monthly inflation rate over the previous 12 months. Panel (b) forecasts the quarterly inflation rate at $t+h$; the random-walk benchmark is the average quarterly inflation rate over the previous four non-overlapping quarters. Values below 1 indicate the competing model outperforms the benchmark. Full sample: 2016m2–2026m1; Low inflation: 2016m2–2021m1; High inflation: 2021m2–2026m1. Diebold-Mariano test of equal predictive accuracy with Newey-West HAC standard errors (lag $= h$). * (**) denotes rejection at the 32% (10%) significance level.

Table A.7: Price Gaps *vs.* Alternative Benchmarks

Horizon	OOS Forecast Evaluation Sample		
	Full sample	Low inflation	High inflation
(a) Benchmark: fixed coefficients AR			
12	1.09	1.49*	1.06
15	0.96	1.38*	0.92
18	0.85*	1.20	0.82**
21	0.81**	1.01	0.80**
24	0.82**	0.92	0.81**
(b) Benchmark: recursively estimated AR			
12	1.04	1.09	1.04
15	1.02	0.95	1.02
18	1.01	0.89	1.02
21	0.98	0.86*	0.99
24	0.95*	0.91*	0.95*

Notes: RMSPE ratio of Price Gaps relative to alternative benchmarks. Forecasts are of year-on-year inflation h months ahead. AR benchmark $\pi^y_{t+h} = \alpha + \beta\,\pi^y_t + \varepsilon_t$; panel (a) uses fixed coefficients, panel (b) recursively re-estimated at each forecast origin. Values below 1 indicate the competing model outperforms the benchmark. Full sample: 2016m2–2026m1; Low inflation: 2016m2–2021m1; High inflation: 2021m2–2026m1. Diebold-Mariano test of equal predictive accuracy with Newey-West HAC standard errors (lag = h). * (**) denotes rejection at the 32% (10%) significance level.

Table A.8: Alternative Predictors *vs.* Random-Walk

Horizon	OOS Forecast Evaluation Sample		
	Full sample	Low inflation	High inflation
(a) Reset inflation			
12	0.88	0.74*	0.89
15	0.80*	0.69*	0.81*
18	0.76**	0.67**	0.76**
21	0.73**	0.65**	0.73**
24	0.70**	0.63**	0.71**
(b) Pricing-pressure indices			
12	0.88	0.79	0.89
15	0.76*	0.72*	0.76*
18	0.69**	0.67*	0.69**
21	0.67**	0.61**	0.67**
24	0.67**	0.56**	0.67**

Notes: RMSPE ratios relative to the random-walk benchmark. Panel (a) uses reset-price inflation as the predictor, while Panel (b) uses the pricing-pressure indices. Both measures are described in Appendix C.4. Forecasts are of year-on-year inflation h months ahead. The random-walk benchmark is the year-on-year inflation rate at t. Values below 1 indicate that the competing model outperforms the benchmark. Full sample: 2016m2–2026m1; Low inflation: 2016m2–2021m1; High inflation: 2021m2–2026m1. Diebold-Mariano test of equal predictive accuracy with Newey-West HAC standard errors (lag = h). * (**) denotes rejection at the 32% (10%) significance level.