Inflation Targeting in the Time of Supply Shocks: A Large-Language-Model Reading of India's Reserve Bank, 1997–2025[1]

Karan Singh Bagavathinathan
Independent Economist

Tandley Omprakash Sridevi
Bharathidasan Institute of Management, Tiruchirappalli, India

This draft: July 14, 2026

Abstract

India's Phillips curve looks flat across 1997–2025, spanning the 2016 move to flexible inflation targeting. We ask whether the flatness survives separating demand from supply shocks. The pooled output-gap slope is zero, but conditioning on the shock, the demand-pull slope is +2.2 before targeting and zero afterward, while the supply slope stays negative, the signature of a central bank offsetting demand before it reaches prices. A large language model reading the Reserve Bank's own assessments, blind to the gap, recovers the same regime contrast across three model families. We identify which channel flattened, scaling narrative identification with AI.

JEL classification: E31, E52, E58.
Keywords: Phillips curve; inflation targeting; supply shocks; narrative identification; large language models; India.

1 Introduction

The Phillips curve, the link between economic slack and inflation, has appeared to flatten across much of the world. The appearance is sharpest in emerging economies whose inflation is repeatedly buffeted by supply shocks. India is such a case, and an unusually clean one. The Reserve Bank of India (RBI) moved from a discretionary "multiple-indicator" approach to a formal flexible inflation-targeting (FIT) framework in 2016, with a 4% CPI target inside a 2–6% tolerance band. The decade since has been a demanding test of it: the COVID-19 disruption, the Ukraine-war commodity spike, and recurrent food-price shocks repeatedly

[1] This is the SSRN working-paper version. Corresponding author: Dr. Sridevi Tandley, Assistant Professor — Analytics & Economics, Bharathidasan Institute of Management, P. Box No. 12, MHD Campus, BHEL Complex, Tiruchirappalli — 620014, Tamil Nadu, India; email `sridevi.tandley@bim.edu`; ORCID 0000-0002-9579-5701. An online Supplementary Appendix (data sources and construction) is included at the end of this document.

pushed headline inflation through the upper band, even as average inflation fell from roughly 6.8% to 4.7% and its volatility halved. Whether the Phillips curve is genuinely absent in such a setting, or only hidden by the mix of disturbances and by a central bank that responds to them, is the question of this paper. It can now be asked with a new tool, because large language models let researchers read the policy record at a scale that was until recently infeasible by hand.

A central, long-standing challenge in inflation modelling is to separate demand from supply disturbances. The two are observationally close. A positive demand shock and a favourable supply shock both raise output, but they move prices in opposite directions: demand pushes inflation and activity the same way, while supply pushes them apart. The sign of their co-movement is thus the mechanical signature that distinguishes them. A range of mechanical methods exploit it, including structural vector autoregression (VAR) models identified by long-run or sign restrictions (Blanchard and Quah, 1989; Uhlig, 2005) and category-level decompositions of the price index (Shapiro, 2022). These richer decompositions, however, rely on data India lacks at the required frequency, such as category-level quantities or the labour-market tightness on which advanced economies gauge demand. The simplest version is a co-movement sign rule on inflation and the output gap: a quarter is read as demand-driven when the two move in the same direction and supply-driven when they move apart. It asks only for series India has, so we adopt it as our transparent mechanical benchmark. This is a classification rule applied to the data, not an estimated equation, and it is distinct from the restriction-based structural VARs above. The rule is serviceable but limited in two ways. It is dichotomous, forcing every quarter into either "demand" or "supply." And it is mechanical, with no access to the economic reasoning behind the numbers, so it cannot recognise a quarter driven by both forces at once, nor tell whether a change of regime has altered the way shocks transmit to prices.

An alternative is to read the policy record itself. Central banks, the RBI among them, document their real-time assessment of each quarter's inflation across decades of published reports. The narrative tradition of Romer and Romer (2004, 2010) reads precisely this kind of contemporaneous record, identifying monetary and fiscal shocks by their stated motivation and recovering what statistical filters miss. For India, Singh et al. (2011) take a first step in this direction. Having flagged candidate supply-shock quarters as outliers from the fitted inflation–output-gap relation, they read the RBI's monthly *Bulletin* by hand for those quarters, to confirm that a genuine supply-side factor lay behind each. That hand reading is deliberately selective, covering only the flagged quarters and only the supply side, and its limitation has always been cost: it does not scale to thousands of pages across thirty years, nor can it be applied consistently at that scale. Large language models remove the constraint, and a growing literature now reads central-bank communication at scale to quantify its content (Fernández-Fuertes, 2025; Kwon et al., 2025; Silva et al., 2025). A machine can read the entire archive and classify the demand–supply situation quarter by quarter, applying the same criteria uniformly.

This large-scale reading does two things a mechanical rule cannot. First, it provides an independent check on the mechanical classification: the central bank's own account either corroborates the sign-rule label or contradicts it, and the disagreements are themselves informative. Second, it opens a window the sign rule closes by construction. The economy frequently runs with *both* demand and supply pressures at once, a mixed state the dichotomous rule cannot represent but the text describes directly. Reading the record also lets us observe the move to inflation targeting from the inside, in how the RBI reasons about prices, rather than inferring it only from a structural break in the data.

We therefore ask whether the supply–demand attribution can be recovered from the central bank's own words, at scale, and whether doing so changes what we conclude about India's Phillips curve. We proceed in two independent steps. First, we model inflation in a reduced-form, open-economy Phillips curve, with inflation on its own lag, an output gap, the real exchange rate, and a supply-shock term. This is the workhorse specification of the Indian Phillips-curve literature (Ball et al., 2016; Behera et al., 2018; Dua and Gaur, 2010; Kapur, 2013; Patra and Kapur, 2010; Singh et al., 2011), and we classify each quarter's dominant disturbance mechanically from the sign rule. Second, and separately, we use an LLM to read the RBI's reports (1997–2025) and classify the same quarters from the text alone. The model never sees the output gap, the estimated curve, or the structural shocks, so the two classifications cannot simply restate one another, and the sign rule serves as the benchmark against which we judge the narrative read.

The narrative step is the methodological core, and reading the reports well is not a single pass. The RBI documents supply pressures with hard data, quantifying the food and fuel components of the price index contribution by contribution, whereas demand pressures rarely come with such numbers and surface instead in qualitative language. A naive reading inherits this asymmetry and over-attributes inflation to supply. We therefore iterate the prompt through several versions, each adding one correction for the bias, and run the read on the most capable language model available, with the prompts and procedure fixed so that the classification can be replicated.

The richer textual information delivers two results. It recovers the "both" state the sign rule cannot see: 29 of 115 quarters in which demand was a genuine co-driver of an inflation the RBI still framed as supply-led, against the 77 of 115 the narrative reads as supply-led overall. And it explains the policy-regime change cleanly. Under inflation targeting the demand channel is offset, so the inflation that remains is supply-driven. This reconciles a flat post-FIT Phillips curve with a live demand-pull slope before targeting, which falls from +2.2 to about zero across the regime change while the supply-quarter slope stays negative. Singh et al. (2011) first establish this contrast; we recover it by conditioning on the shock type. Following McLeay and Tenreyro (2020), a flat pooled curve is the *signature* of effective targeting, a central bank that offsets demand shocks and erases their imprint on the data, not a dead structural relationship. The flattening has been documented as an aggregate movement (Patra et al., 2021); our contribution is to identify *which* channel was responsible.

The paper makes two contributions. Substantively, it explains the change in India's empirical Phillips-curve patterns as a consequence of the change in the policy regime, triangulating the result across a mechanical model and an independent narrative read. Methodologically, it scales the narrative identification of supply and demand shocks with AI, in a deliberately hybrid design: we keep the classical reduced-form model and deploy the LLM only for the one task where human reading is the binding constraint, the large-scale, consistent classification of the demand–supply situation from policy text. The rest stays classical. Section 2 reviews the related literature; Section 3 describes the data and measurement; Section 4 sets out the specification; Section 5 reports the model-based results; Section 6 develops the LLM narrative; Section 7 examines whether the RBI looked through supply shocks; and Section 8 concludes.

2 Background and Related Literature

Separating demand- from supply-driven inflation is the measurement problem behind the Phillips curve, and many methods have been brought to it (Table 1). Each trades identification against data or assumptions: comovement rules are transparent but mechanical; structural VARs are weakly identified on few series; structural-model and category-level decompositions demand rich data or maintained assumptions.

The Indian Phillips-curve literature has relied mainly on the first row, typically augmenting the output gap with *observed* cost-push regressors (oil, food, and import prices entered directly as supply terms, in the spirit of Gordon's triangle model), so that demand is read from the gap and supply from these observables (Behera et al., 2018; Kapur, 2013). Singh et al. (2011) are the instructive exception. Although they began from the scatter of inflation against the output gap, their identification turned on cross-checking each flagged quarter against the RBI's own *Bulletin* commentary, read by hand, which places them in method with the narrative tradition. This is the logic of narrative identification (Romer and Romer, 2004, 2010): the documentary record reveals *why* inflation moved, information no

Table 1: Approaches to separating supply- from demand-driven inflation.

#	Approach	Reference	How it separates supply/demand	Limitation
1	Output-gap / comovement sign	standard AD–AS logic	sign of gap–inflation comovement	mechanical: classify and estimate from the same signal
2	Category price–quantity comovement	Shapiro (2022)	disaggregated CPI: do price and quantity move together?	needs rich disaggregated category data
3	Structural VAR (long-run)	Blanchard and Quah (1989)	demand has no permanent effect on output	weakly identified; permanent-output $\neq$ cost-push
4	Sign-restriction SVAR	Uhlig (2005)	sign pattern on impulse responses	set-identified; wide identified sets
5	Structural-model decomposition	Bernanke and Blanchard (2023)	estimated channel components	model-dependent; data-hungry
6	Narrative identification	Romer and Romer (2004, 2010); Singh et al. (2011)	read the policy record; classify shocks by their stated cause	manual; not scalable

time-series filter contains. Its cost is equally exact. Such reading is slow, demands central-bank expertise, and does not scale, amounting in practice to weeks of full-time expert work per pass.

We take up this narrative approach but remove its binding constraint. Where that hand read classified only the handful of flagged quarters by policymakers' stated reasoning, at the cost of weeks of expert labour, a large language model applies the identical criterion to every quarter, reproducibly and cross-checked against the sign rule. Romer and Romer (2023) anticipated that narrative identification would eventually be "made largely redundant by computers"; a recent literature now does so for monetary-policy shocks (Fernández-Fuertes, 2025; Silva et al., 2025). This sits within a broader turn to machine learning and text as data in financial econometrics: machine-learning methods sharpen inflation forecasts in data-rich environments (Medeiros et al., 2021), narrative and alternative-price signals carry predictive content for inflation in their own right (Hong et al., 2025; Sha and Wu, 2025), and textual tone is priced in financial markets (Chen et al., 2022). We bring the same machinery to the supply–demand decomposition of inflation, validated against expert labels, for India.

A final point cuts across these methods and shapes how we report our own. Supply and demand are not mutually exclusive: an economy can run both pressures at once, and a cost-push shock can ignite a wage–price spiral once it propagates into wages, expectations, and broader prices, the second-round effects that turn an initial supply impulse into self-sustaining inflation. This interaction is most likely when the nominal anchor is weak, as in India's decades before formal inflation targeting, when unanchored expectations let supply shocks generalise rather than fade. A clean binary classification, whether the sign rule or any single-shock attribution, cannot represent such episodes, because it must force each quarter into one bin. We therefore retain an explicit *both* category in the narrative read (Section 6), together with a generalisation index that gauges how far a supply impulse has spread into core inflation, so that the joint-pressure episodes a binary collapses remain visible.

3 Data and Measurement

No official series spans 1996–2025, so we assemble the three estimation series from primary sources following Singh et al. (2011). We summarise each here; full construction mechanics, partner-by-partner sources, and the 40-currency basket are in the online Supplementary Appendix (Supplementary Tables S1–S3), and the narrative label is described in Section 6.

Inflation is the four-quarter change in a long all-India Composite CPI (CCPI), monthly from 1996. No official all-India index runs that far back, so following Singh et al. (2011) we build it as a population-weighted composite of the four labour-population indices — rural 0.70 (Agricultural and Rural Labourers) and urban 0.30 (Industrial Workers and Urban Non-Manual Employees), reflecting India's roughly 70 per cent rural population (Census of India, 2001; National Sample Survey Office, 2006) — growth-chained onto CPI-Combined from 2011 (Supplementary Appendix, equation (S1)). This long CCPI is both our inflation series and the domestic price level in the REER below.

The real exchange rate is the trade-weighted, CPI-based 40-currency real effective exchange rate (REER) the RBI monitors under inflation targeting (2015-16 base). Its published index begins only in 2004, so we reconstruct it as a geometric, trade-weighted index of relative CPIs on the RBI's January 2021 methodology (Bhoi and Behera, 2016; Reserve Bank of India, 2021) and extend it back to 1996 (Supplementary Appendix, equation (S2)). We use genuine price series only — 36 national series in 1996 (about 85 per cent of trade weight), rising to the full 40 by 2009, the domestic price level being the CCPI above. The reconstruction tracks the published index closely (correlation 0.96 over 2004–2025; 0.995 over 2010–2021), leaving the four-quarter change Δreer_t in (1) essentially unaffected (Supplemen-

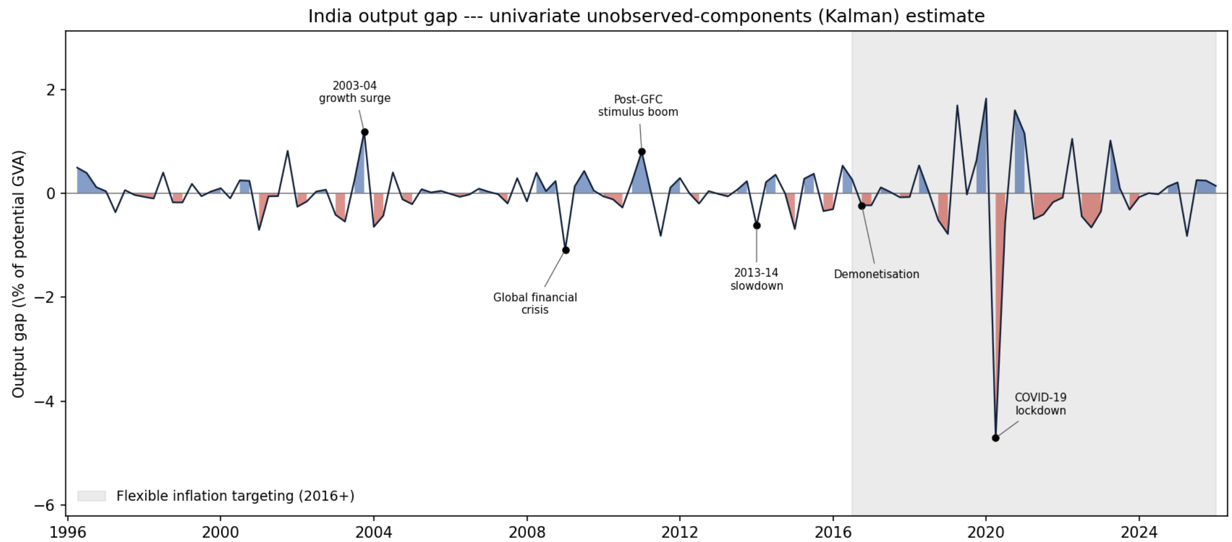

Figure 1: India's output gap, the baseline univariate unobserved-components (Kalman) estimate, in per cent of potential GVA. Gaps are small (mostly within ±1 per cent) apart from the pandemic trough (−4.7 per cent in 2020Q2). Under flexible inflation targeting (shaded, 2016 onward) the economy reads as broadly closed, consistent with the RBI's own assessment.

Alt text: Line chart of India's output gap (univariate unobserved-components estimate) in per cent of potential GVA, quarterly. The series stays mostly within plus or minus one per cent, drops to a −4.7 per cent trough in 2020Q2, and reads as broadly closed across the shaded flexible-inflation-targeting period from 2016.

tary Appendix, Figure S1).

The output gap is the deviation of real gross value added (GVA) at factor cost from trend, quarterly. We growth-chain three base-year vintages (the 1999-2000, 2004-05, and 2011-12 bases, RBI DBIE), extend the series through 2026Q1 with the latest release, seasonally adjust by an STL decomposition, and detrend. Our baseline is the univariate unobserved-components (Kalman) gap, following Singh et al. (2011), who reject the Hodrick–Prescott filter as spurious for India: it reads the post-2016 economy as broadly closed — in line with the RBI's own assessment — and is small in amplitude apart from the −4.7 per cent pandemic trough (Figure 1). Seven alternative filters are used for robustness (Appendix A).

4 Empirical Specification

We estimate a backward-looking, open-economy New Keynesian Phillips curve,

$$\pi_t = c + \beta\,\pi_{t-1} + \kappa\,\mathrm{gap}_t + \alpha\,(\mathrm{gap}_t \cdot \mathrm{sdum}_t) + \gamma\,\Delta\mathrm{reer}_t + \varepsilon_t, \tag{1}$$

where π_t is four-quarter CCPI inflation, gap_t the output gap, and Δreer_t the four-quarter change in the real effective exchange rate. The lagged term π_{t-1} captures the persistence documented in Section 3, with β measuring inertia; γ is the exchange-rate pass-through, expected negative as a real appreciation eases inflation. The coefficient of interest is the slope κ, the response of inflation to demand pressure.

To separate demand from supply disturbances we use a simple, judgement-free rule: a demand shock moves output and prices in the *same* direction, a supply shock in *opposite* directions. We therefore set the supply indicator $\text{sdum}_t = 1$ in quarters where the changes in the gap and in inflation carry opposite signs, and 0 otherwise, and interact it with the gap in (1). The demand-quarter slope is then κ and the supply-quarter slope $\kappa + \alpha$. A live demand-pull curve implies $\kappa > 0$; cost-push dynamics in supply quarters imply $\alpha < 0$, so that the supply-quarter slope turns negative. The rule requires no judgement about which historical episodes were "supply" events; we cross-check the resulting labels against a structural decomposition and the central bank's own narrative in Sections 6 and 7.

We estimate equation (1) by OLS with HAC (Newey–West) standard errors, across all eight output-gap measures of Section 3 to confirm the result is not an artefact of any single filter. We then trace the slope over time, both by sub-period and by seven-year rolling windows. Because the output gap is endogenous to inflation, we re-estimate κ by two-stage least squares, instrumenting the gap with its own lags and twice-lagged inflation, and report the first-stage strength.

5 Results I: The Model

5.1 The pooled Phillips curve appears flat

In the pooled sample the output-gap slope is economically and statistically indistinguishable from zero, and this holds across all eight gap measures (Appendix A): for the baseline unobserved-components (Kalman) gap, $\kappa = +0.114$ ($t = 0.92$). Inflation is instead dominated by *persistence*, $\beta \approx 0.90$, alongside a small, correctly-signed exchange-rate pass-through ($\gamma \approx -0.01$: real appreciation lowers inflation). The fit is high ($R^2 \approx 0.80$) but carries essentially no demand-pull content. Taken at face value, this is the familiar "missing" or flat emerging-market Phillips curve.

5.2 Separating supply shocks revives the curve

The flat pooled slope is, however, an artefact of mixing demand and supply disturbances. Applying the sign rule of Section 4 yields 59 supply and 56 demand quarters and requires no

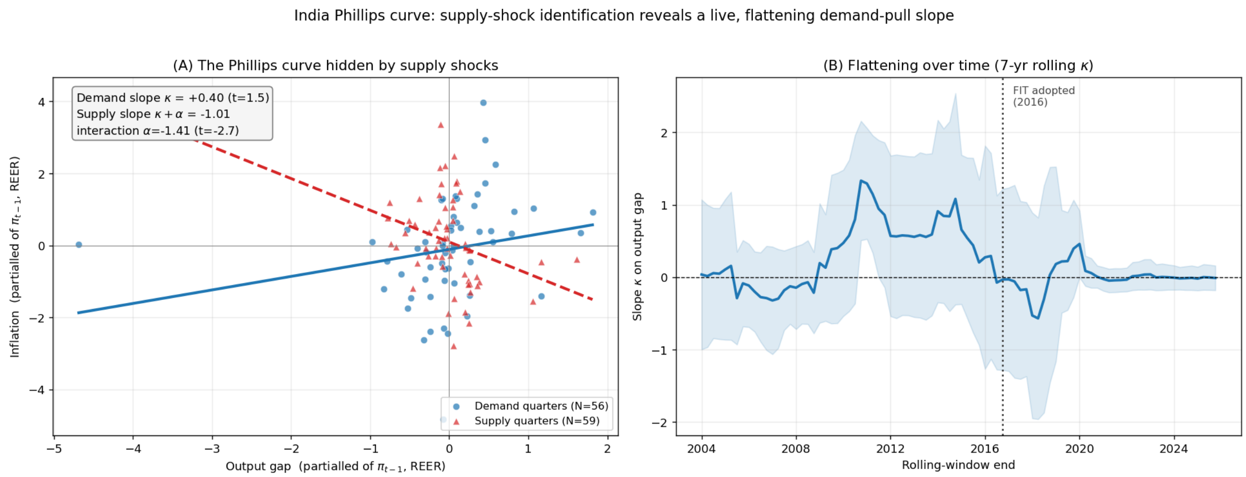

Figure 2: Supply-shock identification reveals the Phillips curve. (A) Inflation versus the output gap, both partialled of persistence and the real exchange rate, for demand versus supply quarters: the demand-pull slope is positive, the supply-quarter slope negative. (B) Seven-year rolling *pooled* slope κ^{pool} (the unconditional gap slope, no supply dummy), flattening to around zero at the adoption of flexible inflation targeting (FIT) in 2016.

Alt text: Two-panel figure. Panel A is a scatter of inflation against the output gap, both partialled of persistence and the real exchange rate, with demand quarters showing an upward-sloping (positive) fit and supply quarters a downward-sloping (negative) fit. Panel B is a time-series line of the seven-year rolling pooled output-gap slope, positive through the 2000s and flattening to around zero from the 2016 adoption of flexible inflation targeting.

judgement about which episodes are "supply" events. Conditioning on the split recovers a live Phillips curve: the demand-quarter slope is $\kappa = +0.398$ ($t = 1.51$), the supply interaction $\alpha = -1.408$ ($t = -2.68$, significant at 1%), and the supply-quarter slope $\kappa + \alpha = -1.010$, exactly the signature of cost-push shocks that raise prices while lowering output. The fit improves to $R^2 = 0.825$ (from 0.812 pooled). In short, the Phillips curve is present in demand-driven quarters but masked when supply quarters are pooled in (Figure 2A).

5.3 The slope flattened under inflation targeting

The demand-pull slope has weakened markedly. Estimating the *pooled* gap slope on a moving seven-year window (the unconditional counterpart of the conditional κ in equation (1), written κ^{pool}; Figure 2B) shows it positive and sizeable through the 2000s, averaging about +0.3 and peaking near +0.5 around 2013, before collapsing to around zero from 2016 (a post-FIT average of -0.01), precisely as India adopted flexible inflation targeting, where it has since remained. The flattening coincides with the establishment of a credible nominal anchor.

5.4 Which channel flattened

Estimating the supply/demand split *separately* in each regime isolates the mechanism. Across FIT, the demand-quarter slope collapses from $\kappa = +2.23$ ($t = 3.30$) before to $+0.09$ ($t = 0.73$) after, while the supply-quarter slope stays firmly negative in both regimes (-1.28 before, -0.48 after). Inflation targeting did not flatten the curve uniformly: it switched off the demand-pull slope while the supply channel remained an active cost-push force, so that supply becomes the dominant driver of the inflation that remains. Persistence stays high throughout ($\beta \approx 0.9$), so supply shocks move inflation around an otherwise well-anchored mean rather than dislodging it. Table 2 summarises this split.

Table 2: Inflation–output-gap slope by regime and shock type.

Sample / regime	Demand-quarter slope κ	Supply-quarter slope $\kappa + \alpha$
Full sample (1997–2025)	$+0.398$ (1.51)	-1.010^{***}
Pre-FIT (1997–2016)	$+2.225^{***}$ (3.30)	-1.283^{***}
Post-FIT (2016–2025)	$+0.090$ (0.73)	-0.476^{*}

Notes. Estimates of equation (1) on the baseline unobserved-components (Kalman) output gap; supply quarters are flagged by the comovement sign rule (59 supply, 56 demand quarters in the full sample). The demand-quarter slope is κ and the supply-quarter slope $\kappa+\alpha$; t-statistics in parentheses where reported, with $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$ (the supply-quarter slope's significance is from a Wald test of $\kappa + \alpha = 0$). Persistence is high throughout ($\beta \approx 0.89$–0.96) and exchange-rate pass-through small and correctly signed ($\gamma \approx -0.02$). The large coefficient magnitudes reflect the small amplitude of the unobserved-components gap. Under inflation targeting the demand-pull slope collapses to zero while the supply-quarter slope remains firmly negative.

5.5 Robustness

We assess two concerns in Appendix A: the choice of detrending filter and the endogeneity of the output gap. The pooled near-zero slope, the positive demand / negative supply split, and the post-2016 flattening all hold across the eight detrending methods of Section 3 (Table 6); and instrumenting the gap by two-stage least squares, though the instruments are weak, leaves the flattening intact ($\kappa_{\mathrm{IV}} \approx -0.06$ full-sample).

6 Results II: The Narrative

The model in Section 5 identifies supply quarters mechanically, from the sign of the comovement of inflation and the gap. We now ask whether the RBI's own contemporaneous account agrees. Doing so well is the methodological core of the paper, because the RBI's reports are systematically framed and the method must correct for that framing.

6.1 Reading the RBI at scale

Following the narrative tradition of Romer and Romer (2004, 2010) and its recent machine implementations (Fernández-Fuertes, 2025; Kwon et al., 2025; Silva et al., 2025), we classify each quarter's dominant inflation disturbance from the RBI's own contemporaneous assessments, read by a large language model. Four publications cover the sample: the *Macroeconomic and Monetary Developments* review (quarterly, to 2014), the *Monetary Policy Report* (half-yearly, 2014–2016), the Monetary Policy Committee *minutes* (from late 2016), and, for the earliest quarters, the macroeconomic chapters of the *Annual Report*. For each quarter we read the full set of attribution-bearing chapters, not a single section. Where several reports review the same quarter we take the freshest contemporaneous read as the headline label, since a narrative measure should capture what the RBI believed in real time, and treat later recaps as reliability re-reads. The mapping yields a reading for all 115 quarters of 1997Q2–2025Q4, the same horizon as the mechanical rule; 16 rest on the Annual Report, and the findings are unchanged when they are excluded.

Two principles make the exercise a valid independent check. The model is *gap-blind*: it reads only the RBI's text and never sees the output gap, the estimated curve, or the structural shocks, so its labels cannot mechanically restate the model's. And each quarter is classified *statelessly*, in isolation, which rules out look-ahead leakage from later quarters (Fernández-Fuertes, 2025; Kwon et al., 2025). For each quarter the model reasons from both sides of the ledger, what the RBI says drove prices and what it says drove growth, since a supply–demand call is a statement about how the two moved together: cost-push prices with weakening growth indicate supply, while rising prices with strong growth indicate demand.

6.2 Prompt design is identification

A naive reading does not work, and *why* it fails is instructive. The RBI's reviews almost always lead with food and fuel, even in demand-driven episodes, so a reader that weighs the RBI's emphasis inherits its framing and over-attributes inflation to supply. The contribution of this section is not that a language model can read the reports, which it plainly can, but the *protocol* that strips out this inherited bias so that the recovered attribution reflects the economics rather than the RBI's emphasis. This is the distinctive challenge of reading a central bank's *own* narrative, as opposed to news (Kwon et al., 2025) or market prices (Fernández-Fuertes, 2025): the source is systematically framed, and the method must correct for it.

The difficulty reflects a universal asymmetry between the two disturbances. A supply shock has a home in the price index, landing on identifiable categories such as food, fuel, or freight-sensitive goods, so it can be quantified contribution by contribution. A demand shock has no such home: it raises prices broadly, in rough proportion to spending, and never appears as a line item, so it must be inferred in every economy. Data-rich economies infer it from sources India lacks at the required frequency, such as the category-level quantities that let Shapiro (2022) sign each price movement, or the labour-market tightness through which Bernanke and Blanchard (2023) gauges demand pressure. The richest structured signal about demand that India possesses is the RBI's own reasoning in its reports, which is why reading that record carefully is not a second-best curiosity here but the feasible substitute for the missing data instruments, and why recovering demand depends entirely on a protocol that can read past the RBI's supply-framing.

We build the protocol as a ladder of four prompts, each adding one principle: (P0) a naive, tone-based read; (P1) structured reasoning over prices *and* growth before labelling; (P2) a theory-grounded, country-specific list of demand drivers; and (P3) an explicit instruction to discount the RBI's food/fuel framing. The driver list in P2 is harvested from the corpus itself rather than hand-imposed, and distinguishes drivers by type (supply vs demand) and *origin* (external vs domestic). On the demand side the harvest confirms that India's channel operates largely through *fiscal* forces: Pay Commission awards (the Sixth, which financed the 2009–11 boom; the Seventh, in 2016–17), fiscal stimulus, farm-loan waivers, rural-employment programmes, and election-cycle spending. Following Kwon et al. (2025), a rate change is coded as a reaction to the underlying disturbance, not as a driver in its own right.

6.3 Measurement and validation

The read is produced by a reasoning decoder and validated across three independent large language models: Anthropic's Claude, OpenAI's GPT-4.1, and Google's Gemini 2.5 Pro. We run the headline read first on Claude (pinned `claude-opus-4-8` snapshot) and then replicate the identical protocol on the other two as an independent cross-model check. Claude, driven by prompting alone with no fine-tuning (Brown et al., 2020), reads each quarter's assessment and fills a structured record: supply and demand channels rated on ordinal severity scales, each tagged by origin, from which a transparent rule derives a four-way label (supply, demand, both, normal) and a continuous *lean* from -1 (pure supply) to $+1$ (pure demand). We use a reasoning decoder rather than a fine-tuned classifier because the task requires tracing how prices and growth move together, which embedding classifiers cannot, and because 115 quarters is far too small a sample to train one (Kwon et al., 2025). Current models expose no temperature or seed, so reproducibility rests on the pinned snapshot, a single constrained JSON-schema draw per quarter, and an archived hash of every response.

To confirm the labels are not an artefact of one model, we re-run the identical P3 protocol over all 115 quarters on the two further models. All three read India as supply-dominated with a negative mean lean; they agree on the four-way label at a Fleiss κ of 0.45 and on the binary supply-versus-demand sign in 81–87 per cent of quarters. The disagreement falls not on the supply-versus-demand axis the paper turns on but on the finer supply-versus-normal boundary in benign, low-inflation quarters (Appendix B). That three models with different training data and architectures concur on the attribution is evidence the signal lies in the RBI's text rather than in any single model's priors.

6.4 A prompt-sensitivity panel

To show that prompt design is identification, we run all four protocols over an identical set of digests for a stratified benchmark of 28 quarters, varying *only* the prompt (Table 3). As the design principles accumulate, the share of quarters read as supply-dominant falls monotonically from 79 to 50 per cent and the mean lean rises from −0.36 to −0.10. The shift is concentrated where it should be: on the ten boom quarters the mean lean climbs from near zero to +0.32, with the largest gains coming from the fiscal-driver and anti-bias steps, precisely the principles that teach the reader to look past the RBI's food/fuel framing.

Table 3: Prompt-sensitivity of the narrative supply/demand read, on a pinned `claude-opus-4-8` snapshot. Identical text, four prompts; 28 benchmark quarters. "Boom" = ten demand-candidate quarters (2006–08, 2010–11, 2017–18); "supply-classic" = six unambiguous supply shocks (2008 oil, 2013 taper, 2019 onion, 2020 COVID, 2022 Ukraine, 2023 tomato).

Protocol	% supply	mean lean	lean, boom	lean, supply-classic
P0 naive / tone	79	−0.36	+0.05	−0.71
P1 + structured (price+growth)	71	−0.29	0.00	−0.73
P2 + fiscal demand drivers	68	−0.25	+0.09	−0.67
P3 + anti-bias (full)	50	−0.10	+0.32	−0.66

Two features establish that this is genuine de-biasing rather than a thumb on the scale. First, the six supply-classic quarters stay strongly supply (−0.66 to −0.73) under *every* protocol: the de-biasing lifts the demand reading only where demand evidence exists, never on an unambiguous supply shock. Second, the de-biasing shifts quarters from supply into the *both* category (which rises from 3 of 28 under P0 to 10 under P3) rather than into pure demand, which stays at three to five quarters throughout. The protocol surfaces the joint supply–demand pressure that the RBI's framing had recorded as supply alone; it does not manufacture demand-driven quarters. We therefore adopt P3 as the headline protocol (Figure 3).

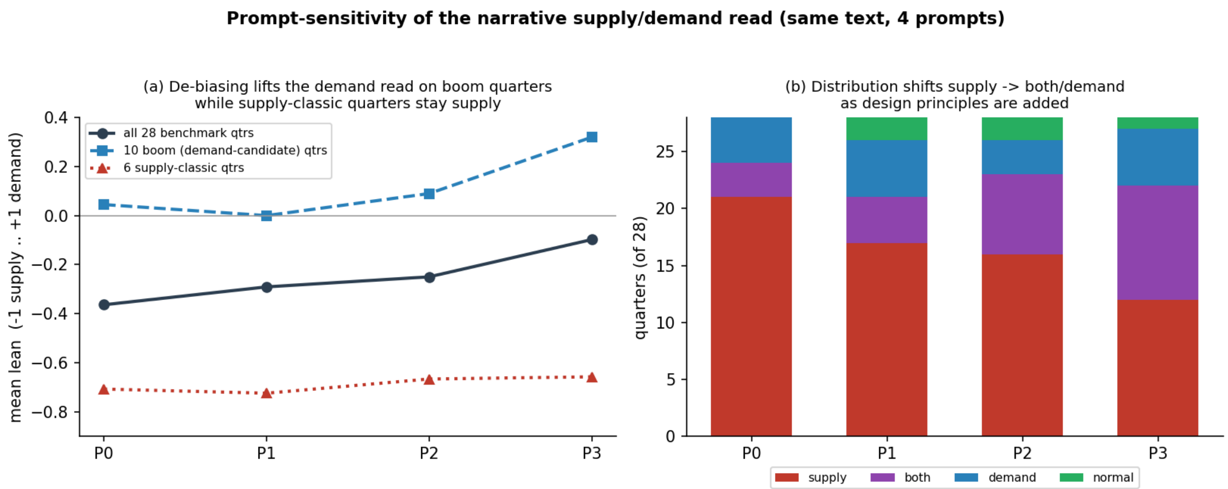

Figure 3: De-biasing lifts the demand read on boom quarters while supply-classic quarters stay supply (left), shifting the distribution from supply toward "both"/demand (right).

Alt text: Two-panel figure on prompt sensitivity. The left panel shows the mean demand–supply lean rising on boom quarters as successive de-biasing prompts are added, while the six unambiguous supply quarters stay strongly supply-leaning. The right panel shows the share of quarters classified as supply falling and the "both" and demand shares rising as the prompts accumulate from P0 to P3.

6.5 What the narrative finds

For every quarter the decoder returns a four-way label and a continuous lean. The lean merely echoes the label for the three pure categories; its information is concentrated in the *both* category, where it resolves the tilt, splitting the 29 joint-pressure quarters into 18 that lean demand and 11 that lean supply (Table 4). We use the categorical label for the distribution and as the bins for the slope analysis (Section 6.6), and the lean as a robustness split of the whole sample by net tilt; the two agree.

Applied across all 115 quarters under the headline protocol, the narrative reads India's inflation as supply-led but with demand a frequent co-driver: 77 quarters supply, 29 "both," 6 demand, and 3 normal, with a mean lean of −0.37. Demand is rarely the sole driver, only six quarters (the 2005–07 boom, the 2010 stimulus, and one quarter under targeting), but

Table 4: The decoder's two outputs, categorical label and continuous lean, across all 115 quarters. The lean is redundant with the label for supply, demand, and normal quarters; it carries information only within the "both" category, where it signs the joint pressure.

Category	n	mean lean	lean range	what the lean adds
Supply	77	-0.62	$[-0.85, -0.30]$	echoes the label (all negative)
Demand	6	$+0.56$	$[+0.40, +0.70]$	echoes the label (all positive)
Normal	3	-0.07	$[-0.10,\ 0.00]$	sits near zero
Both	29	$+0.04$	$[-0.30, +0.40]$	resolves the tilt: 18 lean demand, 11 lean supply

it is a genuine co-driver in 29, twenty-two of them pre-FIT booms and seven under inflation targeting, where strong activity generalised into core inflation even as the headline stayed food-led. In these episodes a cost-push impulse could pass into wages and broaden into core before expectations were anchored, the second-round generalisation a binary sign rule cannot register, which is why they concentrate in the pre-FIT booms. The supply side is itself substantially external: 40 per cent of the price drivers the RBI names are external in origin (world crude, global commodities, geopolitical conflict, exchange-rate pass-through) against 60 per cent domestic, the imported-supply share an inflation targeter is meant to look through (Section 7). The supply plurality is robust to the most aggressive credible demand-hunting: even after reading the full report and crediting demand through its core-inflation footprint, supply remains the leading driver. The asymmetry is itself a finding: the RBI quantifies supply contributions from the price-index components but never attributes a numerical share to demand, whose footprint is the generalisation of inflation into core. Under inflation targeting this sharpens: in 2023–25 the RBI repeatedly describes demand as firm to strong while inflation stays food-driven, the signature of a regime in which demand pressure is offset before it reaches prices.

6.6 The narrative recovers the demand-pull and its switch-off

Does the gap-blind narrative bear on the slope of the Phillips curve? The mechanical sign rule already does: conditioning on it (Section 5, Table 2) the demand-quarter slope is $+2.23$ before inflation targeting and $+0.09$ after, while the supply-quarter slope stays firmly negative, the McLeay and Tenreyro (2020) signature. Table 5 asks whether the narrative, read independently and blind to the gap, recovers it. We classify each quarter with three independent model families and also pool them into a consensus (the ensemble: majority-vote category, mean lean); each entry is the output-gap coefficient κ of the Section 5 specification within the bin.

Three findings stand out. First, the pre-FIT demand-pull is positive under every model

Table 5: The narrative Phillips slope κ (output-gap coefficient of the Section 5 specification) by shock bin, regime, model, and three-model ensemble. Significance stars are from HAC standard errors; the number of quarters is in parentheses. $^{***}p < 0.01$, $^{**}p < 0.05$, $^{*}p < 0.10$; "—" = too few ($n < 6$) to estimate. The ensemble pools the three models (majority-vote category, mean lean). Bins overlap by design ("both" is a category; the lean cuts split the whole sample). For comparison, the mechanical sign rule gives a demand-quarter slope of $+2.23^{***} \rightarrow +0.09$ (Table 2).

Shock bin	Claude	GPT-4.1	Gemini	**Ensemble**
Pre-FIT (1997Q2–2016Q3)				
"both"	+0.25 (22)	$+2.46^{*}$ (14)	$+1.72^{***}$ (22)	$+\mathbf{1.32}^{***}$ (19)
demand-leaning	$+1.06^{**}$ (22)	+1.72 (18)	$+2.19^{***}$ (18)	+2.04 (15)
supply-leaning	+0.06 (56)	−0.08 (60)	−0.20 (57)	+0.09 (63)
Post-FIT (2016Q4–2025Q4)				
"both"	-0.46^{***} (7)	— (2)	+0.20 (9)	— (4)
demand-leaning	— (2)	— (1)	+0.75 (6)	— (3)
supply-leaning	−0.01 (34)	−0.01 (34)	−0.00 (31)	−0.01 (34)

and significant in the ensemble ($\kappa = +1.32$, $p < 0.01$, in the "both" category): where a single model reads the faint, qualitatively-narrated demand signal imprecisely (Claude alone, +0.25), the consensus of three independent families renders it significant. Second, supply-leaning quarters are flat in every model and both regimes, the control: the de-biasing lifts the demand read only where demand evidence exists, never on a supply quarter. Third, the demand channel switches off under inflation targeting: post-FIT the consensus classifies only three to four quarters as demand-driven or "both," too few to estimate a slope, while the residual supply-leaning quarters stay flat. Targeting did not merely shrink the demand slope toward zero; it shrank the *number* of demand-driven quarters toward zero.

Because the split-sample table reports the two regimes as separate cells, we confirm the switch-off with a single test rather than a comparison of subsamples. Pooling the regimes on the ensemble "both" quarters (those the narrative reads as driven by demand and supply together) and interacting the gap with a post-FIT indicator FIT_t that turns on from 2016Q4, $\pi_t = c + \beta\,\pi_{t-1} + \kappa\,\widetilde{y}_t + \delta\,(\widetilde{y}_t \times \mathrm{FIT}_t) + \gamma\,\Delta\mathrm{reer}_t + \theta\,\mathrm{FIT}_t$, where δ is the coefficient of interest, the change in the demand slope across regimes, and θ is a regime intercept that absorbs the lower average inflation under targeting (persistence β and the exchange-rate coefficient γ are held common across regimes). The pre-FIT demand-pull is $\kappa = +1.32$ ($p < 0.01$) and the interaction is $\delta = -1.59$ ($p < 0.01$, HAC), so the slope falls to an insignificant $\kappa + \delta = -0.27$

under targeting: the difference itself, not just each cell, is significant.[2]

This is the same McLeay and Tenreyro (2020) signature the sign rule shows, recovered independently and gap-blind from the RBI's own text, and robust both to source (the sixteen earliest quarters, sourced from the Annual Report before the RBI's quarterly reviews began (Section 6.1), all read supply or normal, so dropping them leaves the demand panel untouched) and to the choice of model (Table 5). Because the narrative classifies by the RBI's stated *price* attribution, which quantifies only supply so that demand is read through its core-inflation footprint, the demand signal is faint, and pooling independent model families is what makes it significant. The formal estimates of Section 5 therefore rest on the mechanical sign rule; the narrative, across three models and their ensemble, supplies this independent corroboration of the demand-pull and its disappearance under targeting, the central empirical claim of the paper, now established from a completely different source once the prompting protocol corrects for the RBI's framing.

7 Results III: Central-Bank Practice

The model and the narrative both describe *what* drove inflation; this section asks what the RBI *did* about it. The prescription for a flexible inflation targeter is asymmetric: lean against demand-driven inflation, which is persistent and self-reinforcing, but look through supply-driven inflation, the first-round effect of a food or fuel shock, unless it generalises into expectations and core. India is an exacting setting for that prescription. Supply shocks are the dominant disturbance (Section 6), and a large share of them are imported: 40 per cent of the price drivers the RBI itself names are external (world crude, global commodities, exchange-rate pass-through), precisely the shocks monetary policy cannot address at source and is meant to look through.

Three pieces of evidence point to an RBI whose practice shifted in this direction at the 2016 adoption of flexible inflation targeting (FIT). First, the model. The demand-quarter Phillips slope collapses across the regime change, from $\kappa = +2.23$ before to $+0.09$ after (Table 2), while the supply-quarter slope stays firmly negative throughout. A central bank that offsets demand pressure before it reaches prices flattens the demand-pull curve while leaving the cost-push relation intact; this is the McLeay and Tenreyro (2020) signature, and it is what the estimates show.

Second, the narrative. The "both" category, the joint supply–demand episodes in which

[2] We let only the demand slope change across regimes, holding persistence (β) and the exchange-rate coefficient (γ) common. Letting them differ too is infeasible here: only four "both" quarters fall after 2016, too few to estimate separate post-FIT coefficients, a scarcity that is itself the switch-off seen from the other side.

demand was a live co-driver, carries a positive Phillips slope before targeting ($\kappa = +0.25$) that inverts to a significantly negative one under it (-0.46, $p < 0.01$, Section 6.6). Read through the category where India's demand pressure actually surfaces, the demand component was being neutralised under FIT, leaving the supply signature. Consistent with this, the RBI's post-2016 assessments increasingly describe demand as firm to strong while attributing inflation to food and fuel; the 2023–25 quarters are the clearest case, with record-low food-led inflation readings coinciding with around eight-per-cent growth, the mark of a regime that contains demand before it reaches prices and looks through the supply component that remains.

Third, the episodes are consistent with look-through subject to a guard rail. The 2020–22 supply surge (pandemic disruptions, then the Ukraine commodity spike) was initially accommodated and met with tightening only as the shock threatened to generalise into core and expectations, not on impact. As those shocks faded over 2023–25, inflation returned toward the four-per-cent target without a demand-driven overshoot, even as activity stayed strong. We read this practice from the regime change in the estimated slopes and from the RBI's own contemporaneous reasoning rather than from a separately estimated policy-reaction function, for which the post-2016 sample is still short; a direct reaction-function test is a natural extension as that sample lengthens. On the evidence available, across model and narrative, the RBI under flexible inflation targeting leaned against demand and looked through, without ignoring, the supply shocks that dominate Indian inflation.

8 Conclusion

India since the late 1990s offers a natural experiment in inflation management: a switch to flexible inflation targeting in 2016, set against an inflation process dominated by recurrent supply shocks. The pooled Phillips curve is flat, the familiar "missing" relationship, but the flatness is an artefact of mixing disturbances. Conditioning on the shock recovers a live demand-pull slope before targeting that switches off after it, while the cost-push relation in supply quarters stays negative throughout: the signature of a regime that neutralises demand pressure before it reaches prices. Read independently and blind to the gap, the RBI's own narrative recovers the same regime contrast: inflation supply-dominated (77 of 115 quarters), demand rarely the sole driver but a frequent co-driver, and the inflation–gap slope in the "both" category flipping from $+0.25$ before targeting to -0.46 after.

Methodologically, the paper scales the narrative identification of Romer and Romer (2004, 2010) with a large language model, in a deliberately hybrid design: we keep the classical reduced-form model and deploy the machine only where human reading is the binding constraint: consistent, large-scale classification of the demand–supply situation from decades of central-bank text. The identifying content is in the prompt, which must read past the RBI's own supply-framing; the read is reproducible from a pinned model snapshot and robust across three independent model families (Fleiss $\kappa = 0.45$).

Three caveats mark out future work: the narrative leaves demand to be inferred from its core-inflation footprint, so the mechanical rule carries the formal estimates while the narrative corroborates; the targeting sample is still short for a direct test of the policy-reaction function; and the method, validated here against India, should be applied to other inflation-targeting economies with different shock mixes. The broader lesson is that the apparent disappearance of the Phillips curve can be a measurement problem, and, increasingly, one that machine reading of the policy record can help to solve, the more so as language models improve with each generation.

Appendix A Robustness of the Phillips-curve estimates

This appendix supports two claims made in Section 5: that the results do not depend on the choice of detrending filter, and that they survive instrumenting the (endogenous) output gap. Both are robustness checks; the baseline in the main text is the univariate unobserved-components (Kalman) gap, following Singh et al. (2011).

Across detrending methods

Table 6 re-estimates equation (1) — both the pooled slope and the sign-rule split into demand- and supply-quarter slopes — under each of the eight detrending methods of Section 3. Three features are invariant across all eight. The *pooled* slope is economically and statistically near zero everywhere (κ between -0.04 and $+0.11$). The *split* always recovers the expected pattern: a positive demand-quarter slope κ and a negative cost-push interaction α, so the supply-quarter slope $\kappa + \alpha$ is negative. And the demand slope *falls under inflation targeting* in every method — κ pre-FIT exceeds κ post-FIT throughout; for the Kalman baseline this harness gives $+2.17 \rightarrow +0.10$, matching the headline interaction estimate of Table 2 ($+2.23 \rightarrow +0.09$). The methods differ only in amplitude: the band-pass and quadratic gaps are highly compressed, so their slopes are small in level, while the unobserved-components (Kalman) gap, with the smallest variance, gives the largest and clearest contrast (the slope scales inversely with gap variance). The signs, significance of α, and the regime flattening do not depend on the filter.

Table 6: Phillips-curve slope across eight detrending methods. Columns: pooled gap slope κ; the sign-rule split into the demand-quarter slope κ and the cost-push interaction α (HAC t-statistics in parentheses), with the implied supply-quarter slope $\kappa+\alpha$; and the demand-quarter slope estimated separately pre- and post-FIT (2016Q4). The baseline (Kalman) gap is in bold. The pooled near-zero slope, the positive demand / negative supply split, and the post-FIT flattening hold under every method.

Detrending method	κ pooled	κ (dem.) (split)	α (split)	$\kappa+\alpha$ (supply)	κ pre-FIT	κ post-FIT	R^2
HP filter ($\lambda = 400$)	0.00	0.05 (1.06)	−0.35 (−2.61)	−0.30	0.40	−0.00	0.82
HP filter ($\lambda = 1600$)	−0.00	0.04 (0.98)	−0.27 (−2.45)	−0.22	0.37	−0.03	0.82
HP filter ($\lambda = 6400$)	−0.00	0.04 (0.92)	−0.17 (−1.90)	−0.14	0.29	−0.05	0.81
Hamilton (2018) regression	0.02	0.05 (1.26)	−0.09 (−1.73)	−0.03	0.20	−0.02	0.85
Christiano–Fitzgerald band-pass	−0.04	0.03 (0.54)	−0.16 (−1.55)	−0.13	0.25	−0.03	0.81
Baxter–King band-pass	−0.03	0.03 (0.74)	−0.13 (−1.36)	−0.10	0.24	−0.10	0.88
Quadratic trend	−0.00	−0.00 (0.00)	−0.01 (−0.14)	−0.01	0.06	−0.07	0.81
Unobs. components (Kalman)	**0.11**	**0.40** (1.51)	**−1.41** (−2.68)	**−1.01**	**2.17**	**0.10**	**0.83**

Instrumenting the output gap

Because the output gap is endogenous to inflation, we re-estimate the demand-quarter slope κ by two-stage least squares, instrumenting the gap with its own lags and twice-lagged inflation. The instruments are weak — the first-stage $F \approx 2$–11, below the Stock–Yogo threshold — so the IV estimates are imprecise and we report them only as a robustness check, not as the headline. With that caveat, OLS and IV agree on the central qualitative result: the full-sample IV slope is itself near zero ($\kappa_{\text{IV}} \approx -0.06$) and the post-2016 flattening is preserved.

Appendix B Cross-model validation of the narrative read

The headline narrative labels are a single pinned snapshot of the decoder (`claude-opus-4-8`) under protocol P3. To confirm that the supply–demand attribution reflects the RBI's text rather than one model's priors, we re-run the *identical* P3 protocol — same digests, same schema — over all 115 quarters on two further flagship families: OpenAI's `gpt-4.1` and Google's `gemini-2.5-pro`. Table 7 reports each model's distribution and the cross-model agreement.

All three families read India as supply-dominated with a negative mean lean. They agree on the four-way categorical label at a Fleiss κ of 0.45 and on the binary supply-versus-demand *sign* in 81–87 per cent of quarters. The disagreement is concentrated not on the supply-versus-demand axis the paper turns on but on the finer supply-versus-normal boundary in benign, low-inflation quarters: GPT-4.1 is the most supply-leaning, reading more quiet quarters as "supply" where Claude and Gemini read "both" or "normal." That three models with different training data and architectures concur on the attribution is the convergent-validity check behind the headline read; every raw response is archived with its hash for audit.

Table 7: Cross-model agreement on the P3 narrative read, all 115 quarters (1997Q2–2025Q4). Each model runs the identical protocol on the identical digests.

Panel A. Distribution of the four-way label, by model

Model	supply	both	demand	normal	mean lean
`claude-opus-4-8` (headline)	77	29	6	3	−0.37
`gpt-4.1`	94	16	3	2	−0.54
`gemini-2.5-pro`	71	31	8	5	−0.45

Panel B. Pairwise agreement

Pair	category κ	exact agree	sign agree
Claude – GPT	0.41	76%	87%
Claude – Gemini	0.54	77%	81%
GPT – Gemini	0.39	72%	85%
Fleiss κ (three models, four-way label): **0.45**			

The per-model and ensemble *slope* results — whether the demand-pull and its switch-off hold under each model's classification — are reported in the main text, Table 5 of Section 6.6.

References

Ball, L., A. Chari, and P. Mishra (2016). Understanding inflation in India. NBER Working Paper 22948, National Bureau of Economic Research.

Behera, H. K., G. Wahi, and M. Kapur (2018). Phillips curve relationship in an emerging economy: Evidence from India. *Economic Analysis and Policy 59*, 116–126.

Bernanke, B. S. and O. J. Blanchard (2023). What caused the U.S. pandemic-era inflation? NBER Working Paper 31417, National Bureau of Economic Research.

Bhoi, B. K. and H. K. Behera (2016). India's potential output revisited. RBI Working Paper WPS (DEPR) 05/2016, Reserve Bank of India.

Blanchard, O. J. and D. Quah (1989). The dynamic effects of aggregate demand and supply disturbances. *American Economic Review 79*(4), 655–673.

Brown, T. B., B. Mann, N. Ryder, et al. (2020). Language models are few-shot learners. In *Advances in Neural Information Processing Systems (NeurIPS)*, Volume 33, pp. 1877–1901.

Census of India (2001). Census of india 2001: Rural–urban distribution of population. Office of the Registrar General and Census Commissioner, Ministry of Home Affairs, Government of India.

Chen, C. Y.-H., M. R. Fengler, W. K. Härdle, and Y. Liu (2022). Media-expressed tone, option characteristics, and stock return predictability. *Journal of Economic Dynamics and Control 134*, 104290.

Dua, P. and U. Gaur (2010). Determination of inflation in an open economy Phillips curve framework: The case of developed and developing asian countries. *Macroeconomics and Finance in Emerging Market Economies 3*(1), 33–51.

Fernández-Fuertes, R. (2025, November). Monetary policy shocks: A new hope — large language models and central bank communication. Job market paper, Bocconi University.

Hamilton, J. D. (2018). Why you should never use the Hodrick-Prescott filter. *Review of Economics and Statistics 100*(5), 831–843.

Hong, Y., F. Jiang, L. Meng, and B. Xue (2025). Forecasting inflation using economic narratives. *Journal of Business & Economic Statistics 43*(1), 216–231.

Kapur, M. (2013). Revisiting the Phillips curve for India and inflation forecasting. *Journal of Asian Economics 25*, 17–27.

Kwon, B., T. Park, P. Rungcharoenkitkul, and F. Smets (2025, October). Parsing the pulse: Decomposing macroeconomic sentiment with LLMs. BIS Working Paper 1294, Bank for International Settlements.

McLeay, M. and S. Tenreyro (2020). Optimal inflation and the identification of the Phillips curve. *NBER Macroeconomics Annual 34*, 199–255.

Medeiros, M. C., G. F. R. Vasconcelos, A. Veiga, and E. Zilberman (2021). Forecasting inflation in a data-rich environment: The benefits of machine learning methods. *Journal of Business & Economic Statistics 39*(1), 98–119.

National Sample Survey Office (2006). Employment and unemployment situation in India, 2004–05. NSS Report No. 515 (61st Round), Ministry of Statistics and Programme Implementation, Government of India.

Patra, M. D., H. Behera, and J. John (2021, November). Is the Phillips curve in India dead, inert and stirring to life or alive and well? Rbi bulletin, Reserve Bank of India.

Patra, M. D. and M. Kapur (2010). A monetary policy model for India. *Macroeconomics and Finance in Emerging Market Economies 3*(1), 23–41.

Reserve Bank of India (2021, January). Revision of the nominal and real effective exchange rate indices. Technical report, RBI Bulletin.

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

Romer, C. D. and D. H. Romer (2010). The macroeconomic effects of tax changes: Estimates based on a new measure of fiscal shocks. *American Economic Review 100*(3), 763–801.

Romer, C. D. and D. H. Romer (2023). Does monetary policy matter? the narrative approach after 35 years. *American Economic Review*.

Sha, Y. and X. Wu (2025). Black market prices as inflation predictor: Evidence from China's hyperinflation. *Finance Research Letters 84*, 107805.

Shapiro, A. H. (2022). Decomposing supply and demand driven inflation. Working Paper 2022-18, Federal Reserve Bank of San Francisco.

Silva, T. C., K. Moriya, and R. M. Veyrune (2025, June). From text to quantified insights: A large-scale LLM analysis of central bank communication. IMF Working Paper WP/25/109, International Monetary Fund.

Singh, B. K., A. Kanakaraj, and T. O. Sridevi (2011, June). Revisiting the empirical existence of the Phillips curve for India. *Journal of Asian Economics 22*(3), 247–258.

Uhlig, H. (2005). What are the effects of monetary policy on output? results from an agnostic identification procedure. *Journal of Monetary Economics 52*(2), 381–419.

Online Supplementary Appendix

Inflation Targeting in the Time of Supply Shocks:
A Large-Language-Model Reading of India's Reserve Bank, 1997–2025

This supplementary appendix documents the data sources and construction for the paper. Cross-references of the form "Section 3" and "Appendix A" refer to the main text above.

S1 Data sources and construction

This appendix documents every series used in the estimation: its definition and coverage, its primary source, and how it was constructed. Three of the four series are assembled by us from primary sources to obtain consistent 1996–2025 coverage, because no official series spans the sample. Table S1 summarises the series; Table S2 details the inputs to the reconstructed real effective exchange rate, and Table S3 lists all 40 partner currencies individually with their trade weight and CPI source; Table S4 reports summary statistics for the estimation variables over the full sample. All series are at their latest vintage as of the June 2026 data release.

Construction details

Composite CPI. The long all-India Composite CPI is the population-weighted average of the four labour-population indices,

$$\mathrm{CCPI}_t = 0.35\,\mathrm{AL}_t + 0.35\,\mathrm{RL}_t + 0.15\,\mathrm{IW}_t + 0.15\,\mathrm{UNME}_t, \tag{S1}$$

the rural weight (0.70) split between the CPIs for Agricultural and Rural Labourers (AL, RL) and the urban weight (0.30) between those for Industrial Workers and Urban Non-Manual Employees (IW, UNME), reflecting India's roughly 70 per cent rural population (Census of India, 2001; National Sample Survey Office, 2006). The four labour indices are each rebased to 2004 = 100 and spliced by *growth-chaining* rather than fixed linking factors: for a spine x and source s, $x_t = x_{t+12} \cdot s_t / s_{t+12}$, so the year-on-year inflation of the spliced series equals that of the underlying source at every month and leaves no splice artefact. We chain the CCPI back from CPI-Combined (2011 onward) through the CPI-IW 2001 and 1982 bases, and bridge the single no-overlap join (1982↔2001) by extrapolating CPI-IW across 2006 with CPI-UNME's seasonal growth — an implied link of 4.64, within 0.4 per cent of the official 4.63, confirming the join is pure rebasing.

Real effective exchange rate. The REER is a geometric, trade-weighted index of relative consumer prices on the Reserve Bank's January 2021 methodology (Bhoi and Behera, 2016; Reserve Bank of India, 2021),

$$\mathrm{REER}_t = 100 \prod_i \left[\frac{R_{i,t}}{R_{i,b}} \cdot \frac{P_t^{\mathrm{in}}/P_b^{\mathrm{in}}}{P_{i,t}/P_{i,b}} \right]^{w_i}, \tag{S2}$$

with $R_{i,t}$ the rupee price of partner currency i (a rise being a rupee appreciation), P^{in} the domestic price level (the CCPI above), P_i partner i's consumer prices, w_i its share in India's 2015-16 trade (the 40 weights normalised to sum to one and renormalised over the currencies present whenever a partner CPI is unavailable), and b the base period (Indian fiscal year 2015-16). Bilateral rates are IMF period averages, the pre-1999 euro spliced from the Deutsche mark at its irrevocable conversion rate. The effective weight is about 0.85 of the basket in the late 1990s, reaches the full basket by 2009, and dips below one again only in the closing months as foreign CPI releases catch up. Within the estimation sample no synthetic or interpolated price data are used.[3] Over the 261-month overlap the reconstruction matches the published index at correlation 0.96 (mean absolute difference 1.9 index points), and 0.995 over 2010–2021 (difference 0.3). Figure S1 plots the reconstructed index against the published series and its backward extension to 1996.

Output gap. The three GVA vintages span 1996Q2–2009Q3 (1999-2000 base), 2004Q2–2014Q3 (2004-05), and 2011Q2–2025Q3 (2011-12, the level anchor); the chained series is extended through 2026Q1 with the MOSPI June 2026 release (new 2022-23 base), which reports year-on-year real-GVA growth of 8.0 per cent for 2025Q4 and 7.9 per cent for 2026Q1. A single fixed seasonal factor is inadequate because the vintages' seasonal patterns differ — it leaves a spurious quarterly sawtooth — so we use a time-varying STL seasonal. The eight detrending methods are the Hodrick–Prescott filter ($\lambda = 400, 1600, 6400$), the Hamilton (2018) regression filter, the Christiano–Fitzgerald and Baxter–King band-pass filters, a quadratic trend, and univariate and bivariate unobserved-components (Kalman) models, the latter identifying the gap jointly with a Phillips-curve inflation equation (Appendix A).

[3]The one interpolated value is the United States CPI for October 2025, set to the geometric mean of the adjacent September and November levels (a constant-growth fill); the Bureau of Labor Statistics release for that month was delayed by the late-2025 federal government shutdown. It falls in 2025Q4 — beyond the narrative horizon of Section 6 — and enters no estimate.

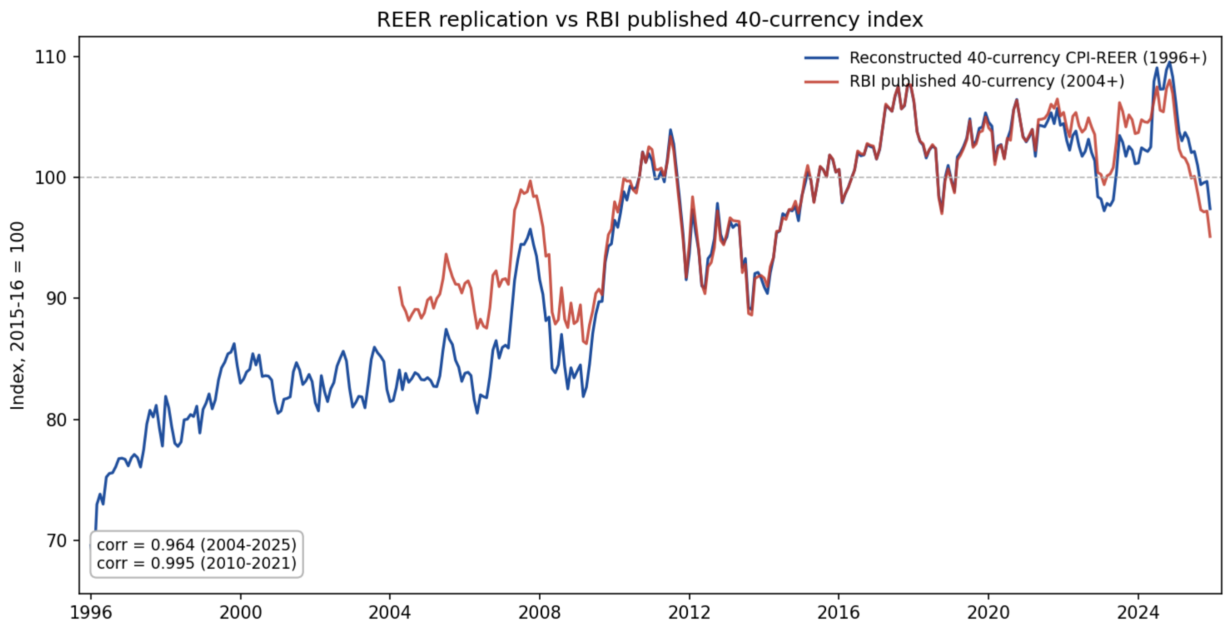

Figure S1: The reconstructed 40-currency CPI-REER (2015-16 = 100) against the Reserve Bank's published index. The two track closely over their common span (correlation 0.96 over 2004–2025; 0.995 over 2010–2021); the reconstruction extends the series eight years before the official index begins, back to 1996.

Alt text: Line chart of the reconstructed 40-currency CPI-based real effective exchange rate index (2015-16 = 100) plotted against the Reserve Bank of India's published index from 2004; the two lines lie almost on top of one another over their common span, and the reconstructed line extends back to 1996 where no official series exists.

Table S1: Data series: definitions, coverage, sources, and construction.

Series	Definition & coverage	Source	Construction
Inflation, π (Composite CPI)	Four-quarter change of a long all-India Composite CPI (CCPI); monthly, 1996–2025	Labour Bureau (CPI-IW, AL, RL) and NSO/MOSPI (CPI-UNME, CPI-Combined), via the Reserve Bank's DBIE	Population-weighted composite (rural 0.70: AL, RL; urban 0.30: IW, UNME) of the four labour indices, growth-chained onto CPI-Combined from 2011; method of Singh et al. (2011) (Section 3)
REER, reer	40-currency CPI-based real effective exchange rate, 2015-16=100; monthly, 1996–2025	Partner CPI: IMF-IFS, BIS, national agencies; bilateral rates: IMF-IFS; weights: Reserve Bank (see Table S2)	Geometric trade-weighted index on the Reserve Bank's January 2021 methodology (Bhoi and Behera, 2016; Reserve Bank of India, 2021); genuine series only, weights renormalised over 36 partners (1996) rising to the full 40 (2009); domestic price is the CCPI; validated against the published index (correlation 0.96 over 2004–2025) (Section 3)
Output gap, gap	Deviation of real gross value added (GVA) at factor cost from trend; quarterly, 1996Q2–2026Q1	Reserve Bank DBIE (1999-2000, 2004-05, 2011-12 bases); MOSPI/CSO press release of June 2026 (new 2022-23-base series, latest growth)	Three base-year vintages growth-chained and extended through 2026Q1 with the latest release's GVA growth, STL seasonal adjustment, then detrended; baseline univariate unobserved-components (Kalman) gap (Singh et al., 2011), with seven alternative filters (HP at $\lambda = 400/1600/6400$, Hamilton, Christiano–Fitzgerald, Baxter–King, quadratic) for robustness (Section 3; Appendix A)
Narrative label	Quarter-level supply / demand / both classification; quarterly, 1997Q2–2025Q4	Reserve Bank publications: *Macroeconomic and Monetary Developments*, *Monetary Policy Report*, MPC minutes, and the *Annual Report*	Large-language-model read of the Bank's contemporaneous assessments under a fixed, pre-specified protocol (P3); headline labels from a pinned `claude-opus-4-8` snapshot, cross-validated against GPT-4.1 and Gemini 2.5 Pro (Fleiss $\kappa = 0.45$) (Section 6)

Table S2: Inputs to the reconstructed 40-currency CPI-REER.

Input	Source	Notes
Partner CPI — IMF-IFS	IMF *International Financial Statistics*, monthly all-items CPI	The majority of the 40 partners; where a country's series begins late, the longer IFS vintage is growth-chained backward so early movements are real, not interpolated
Partner CPI — BIS	BIS long consumer-price series	Hong Kong, Malaysia, Thailand
Partner CPI — national	ECB (euro-area HICP), SingStat (Singapore), Nepal Rastra Bank, DGBAS (Taiwan)	Used where the national series extends further or more cleanly than IFS
Excluded (early sample)	—	United Arab Emirates (until 2007), Qatar and Iraq (until 2009), Oman (until 2001) have no genuine monthly CPI early on; they are dropped until their own series begin and the weights renormalised — no synthetic data is used
Bilateral exchange rates	IMF-IFS, national currency per US dollar, period average	Pre-1999 euro spliced from the Deutsche mark at its irrevocable conversion rate
Trade weights	Reserve Bank of India, January 2021 Bulletin	40-currency basket, 2015-16 trade shares, normalised to sum to one

Table S3: The 40-currency basket: 2015-16 trade weight and CPI source for each partner. "From" gives the entry year where later than 1996 (blank = available from 1996); the four Gulf economies enter when a continuous genuine monthly CPI begins. Source codes: IMF-IFS = IMF *International Financial Statistics*; BIS = BIS long CPI; ECB = euro-area HICP; SingStat = Singapore Dept. of Statistics; DGBAS = Taiwan DGBAS; NRB = Nepal Rastra Bank.

Country	Wt%	CPI source	From	Country	Wt%	CPI source	From
Euro area	11.4	ECB		Brazil	1.5	IMF-IFS	
China	10.0	IMF-IFS		Thailand	1.3	BIS	
UAE	9.4	IMF-IFS	2007	Vietnam	1.1	IMF-IFS	
US	9.1	IMF-IFS		Bangladesh	0.9	IMF-IFS	
Saudi Arabia	6.4	IMF-IFS		Taiwan	0.9	DGBAS	
Switzerland	3.7	IMF-IFS		Angola	0.9	IMF-IFS	
Hong Kong	2.9	BIS		Russia	0.9	IMF-IFS	
Indonesia	2.9	IMF-IFS		Turkey	0.9	IMF-IFS	
Singapore	2.8	SingStat		Mexico	0.9	IMF-IFS	
Iraq	2.7	IMF-IFS	2009	Israel	0.9	IMF-IFS	
Japan	2.5	IMF-IFS		Canada	0.8	IMF-IFS	
Korea	2.5	IMF-IFS		Sri Lanka	0.8	IMF-IFS	
Kuwait	2.5	IMF-IFS		Egypt	0.7	IMF-IFS	
Qatar	2.4	IMF-IFS	2009	Oman	0.7	IMF-IFS	2001
Nigeria	2.3	IMF-IFS		Nepal	0.6	NRB	
UK	2.2	IMF-IFS		Kenya	0.6	IMF-IFS	
Malaysia	2.2	BIS		Tanzania	0.5	IMF-IFS	
Iran	2.1	IMF-IFS		Chile	0.5	IMF-IFS	
Australia	2.0	IMF-IFS		Ukraine	0.4	IMF-IFS	
South Africa	1.8	IMF-IFS		Ghana	0.2	IMF-IFS	

Table S4: Summary statistics of the estimation variables, full sample 1997Q2–2025Q4 (N = 115).

Variable	N	Mean	Std. dev.	Min	Max
Inflation, π (four-quarter CCPI change, %)	115	6.00	3.07	0.76	15.30
Output gap (UC/Kalman, %)	115	−0.00	0.65	−4.70	1.82
Δreer (four-quarter change, %)	115	1.16	5.10	−11.80	16.49