WP-2025-021

Is core inflation useful in predicting headline inflation? Evidence from a large, emerging economy

Yogeshwar Bharat (SNU), Rajeswari Sengupta (IGIDR),
Gautham Udupa (CAFRAL)

Indira Gandhi Institute of Development Research, Mumbai

August 2025

Is core inflation useful in predicting headline inflation? Evidence from a large, emerging economy

Yogeshwar Bharat (SNU), Rajeswari Sengupta (IGIDR),
Gautham Udupa (CAFRAL)

Email(corresponding author): rajeswari@igidr.ac.in

Abstract

Core inflation measure is widely tracked as a measure of trend inflation, but it does not forecast headline inflation well. In this paper, we use disaggregated state-level inflation data from India to construct a 'cleaned' core inflation measure. We do this by stripping out the passthrough of past food inflation from the raw core inflation measure. We estimate the passthrough using local projection with global supply-side instruments in order to achieve better identification. We further find that our 'cleaned' core inflation measure generates better forecasts of the headline inflation after a six-month horizon, compared to the raw core measure.

Keywords: Inflation forecasting, Core inflation, Headline inflation, State-level in- flation

JEL Code: E31, E37, E52.

Is core inflation useful in predicting headline inflation? Evidence from a large, emerging economy*

Yogeshwar Bharat† Rajeswari Sengupta‡ Gautham Udupa§

August 7, 2025

Abstract

Core inflation measure is widely tracked as a measure of trend inflation, but it does not forecast headline inflation well. In this paper, we use disaggregated, state-level inflation data from India to construct a 'cleaned' core inflation measure. We do this by stripping out the passthrough of past food inflation from the raw core inflation measure. We estimate the passthrough using local projection with global supply-side instruments in order to achieve better identification. We further find that our 'cleaned' core inflation measure generates better forecasts of the headline inflation after a six-month horizon, compared to the raw core measure.

JEL Codes: E31, E37, E52.
Keywords: Inflation forecasting, Core inflation, Headline inflation, State-level inflation

*We thank Saroj Bhattarai, VV Chari, and Kundan Kishor for helpful comments. We thank Sripad Atri and Suraj Kumar for excellent research assistance. Any remaining errors are ours.
†Shiv Nadar Institute of Eminence, Delhi-NCR, India. Email: yogeshwar.bharat@snu.edu.in
‡IGIDR, Mumbai, India. Email: rajeswari.sen@gmail.com
§CAFRAL, Mumbai, India. Email: gautham.udupa@cafral.org.in

1 Introduction

Delay in the transmission of monetary policy actions to the headline inflation rate is well documented in the literature and is an unavoidable challenge in the conduct of monetary policy. While taking a policy action, central bankers therefore have to anticipate the future path of headline inflation, and set the policy rate, targeting an inflation forecast as opposed to the actual inflation. However, forecasting headline inflation is complicated because of the different properties of its components. There is considerable ambiguity in the usefulness of different variables and econometric models in their ability to forecast the headline inflation (Atkeson and Ohanian (2001)).

In this paper, we attempt to provide a partial resolution to this challenge by constructing a 'cleaned' core inflation measure that can be used to forecast the headline inflation. We first construct the 'raw' core inflation, which is a widely-used measure used for forecasting the headline inflation, by subtracting food and fuel prices from the headline index, using disaggregated, state-level data from India. We hypothesize that the raw core inflation measure could be contaminated due to the transmission of past food and fuel prices.[1] Contemporary core inflation, therefore, contains information that may not have much impact on the future trajectory of headline inflation. Hence we construct a 'cleaned' core inflation measure by stripping out the passthrough effects of past inflation. We estimate that up to 12.3 percent of state-level food inflation passes through to local core inflation. We do not find a significant passthrough for fuel inflation. We further find that our 'cleaned' core inflation measure out-performs the raw core inflation, in terms of forecasting headline inflation, after a six-month horizon.

India provides a unique setting to study this issue. First, there is monthly data separately for headline, food, fuel, and core inflation variables at the state-level, which allows us to study whether disaggregation by geography and CPI components improves forecasting performance. Similar data is not available for the United States. Second, food and fuel form a large part (over 45 and 15 percent respectively) of the national consumption basket. Third, the Reserve Bank of India adopted inflation targeting in 2016, and tracks core inflation as a predictor of headline inflation.

We begin by establishing the properties of the data in the form of stylized facts. We show that: i. inflation varies more in the state-level data, and the median of state-level inflation tracks national data quite closely; ii. a significant part of the variation in state-month inflation is unexplained by state- or month-related factors; iii. correlations of headline, food, and fuel inflation with lags and leads of core inflation are lower in the states' data compared to the national data; and iv. the correlations are lower in the

[1] Peersman (2022) finds that a rise in global food prices increases the core harmonized index of consumer prices (HICP) of the by 0.05 percent, showing the passthrough of external prices to core inflation. We suspect that local food price increases could have a similar, if not a greater, transmission to local non-food and non-fuel prices, hence distorting the core inflation measure.

period with higher inter-state trade following the adoption of a uniform nationwide tax rate. The stylized facts are useful in understanding the features of the state-level data.

While the stylized facts (i) and (ii) are useful to understand the raw patterns in inflation, facts (iii) and (iv) are useful to identify a key factor, i.e., inter-state trade, that impacts the estimated inflation passthrough. In the absence of administrative data on inter-state trade, these two facts serve as a useful proxy. Our presumption is that states import from neighbouring states when they are affected by adverse local shocks. If that is the case, a local food shortage should lead to a smaller reduction in local food inflation compared to a counterfactual where the state does not import food. Furthermore, the food shortage should have a lower impact on core inflation assuming that the inflation expectations increase by a smaller amount.

We causally establish the passthrough of food and fuel prices to core inflation using local projections (Jordà (2005)). We adopt two novel approaches in terms of the data used and the estimation methodology. First, we use monthly state-level administrative data on food, fuel, and core inflation in India. The state-level data has greater variation across all three inflation measures compared to the respective inflation in the national data. The panel setting allows us to benefit from higher statistical power compared to using national data (Hsiao, 2022). Second, we use an instrumental variable (IV) estimation approach in which state-level food and fuel inflation is analyzed with their respective global supply-driven inflation. We take IVs for global food and oil supply shocks of Bhattarai et al. (2024) and Baumeister and Hamilton (2019), respectively. We attribute the two IVs to each state-month pair using a nonparametric first-stage regression and predict the state food and oil inflation.

We find that local or state-level food inflation has a positive and significant impact on local or state-level core inflation. In particular, a one percentage point (pp) increase in food inflation leads to a 0.09 pp increase in core inflation on impact, and this effect persists for about 11 months. The effect has a maximum coefficient estimate of 12.3 basis points in the fifth month after the food shock. In contrast, local fuel inflation, which is instrumented by global supply-driven shocks to crude oil prices, does not lead to a statistically significant impact on states' core inflation.

We then use these estimated food inflation passthrough coefficients to construct our cleaned measure of core inflation.[2] Since the passthrough coefficients are positive, the cleaned core is lower than the raw core inflation measure. The correlation between the two inflation measures is also very low at -0.075.

Next we test if the cleaned core measure is better at forecasting headline inflation compared to the raw (i.e. without eliminating the passthrough impact) core measure.

[2]The cleaned core inflation equals the current raw core inflation minus the food inflation passthrough for the previous twelve months. Passthrough of food inflation k months ago is the product of the actual food inflation k months ago and the estimated passthrough of food inflation into k-months ahead core inflation.

Using the state panel data, we calculate the root mean squared error (RMSE) for each forecasting horizon using a rolling 60-month estimation sample and a 12-month forecasting horizon. In the baseline forecasting model with the raw core inflation measure, the RMSE increases steadily from about 0.68 percentage points at the one-month horizon to about 2.83 percentage points at the 12-month horizon. In comparison, with the cleaned core measure, the one-month ahead RMSE is higher at 1.17 percentage points, but the 12-month ahead RMSE is about 2.55 percentage points. The RMSE using the raw core is higher after the sixth month implying that over a longer horizon, the cleaned core measure outperforms the raw core in forecasting headline inflation.

We also conduct robustness checks to further examine the usefulness of the cleaned core inflation measure. The results of alternate forecasting models reiterate the usefulness of the cleaned core measure over a longer horizon. The RMSE continues to be lower using the cleaned core measure and state-level data after the sixth month. We are also interested in whether past headline or core inflation from the national level time series data has better forecasting properties relative to the state-panel data. Although state-level data are useful for estimating the passthrough, national data could be better for forecasting for two possible reasons. First, if state inflation is driven by common aggregate shocks, state-level variations are less useful. In fact, we find that aggregate shocks explain about 57.1 percent of the variation in headline inflation as opposed to state-specific factors. Second, the inflation forecasting model conducted at the state-level might miss the role of inter-state trade if states offset local inflation by trading with each other.

We find two important but previously not highlighted properties of inflation forecasting models in India: i. national time series data performs better (i.e., has lower RMSE) compared to states data, holding everything else constant; ii. a single-variable model using past headline inflation performs better (both in national and states data) compared to a multi-variable model with the components of headline inflation.

Overall, our findings indicate that the cleaned core measure is helpful in obtaining relatively more accurate forecasts of headline inflation for longer horizons compared to the raw core measure. This suggests that the passthrough of food inflation to core inflation can play an important role in obtaining more reliable forecasts in the longer horizons especially in emerging economies like India with a very high share of food in household consumption baskets. Not adjusting for this spillover, in the longer run, diminishes the models' ability to forecast headline inflation.

Related Literature. We contribute to the literature on core and headline inflation dynamics and the passthrough of non-core prices to core prices. Anand et al. (2014) examine the passthrough of food inflation to core inflation in India, and find that fluctuations in food prices are persistent rather than transient. They find that core inflation reverts to headline, rather than headline inflation reverting back to core inflation. Peersman (2022) find evidence of second-round effects of food inflation, observing a rise in core

HICP by 0.05 percent from increases in international food prices.

We also contribute to the literature on forecasting headline inflation using core inflation. Core inflation is known to have a limited predictive power (Atkeson and Ohanian, 2001; Bullard, 2011). However more recently, there is evidence of forecasting gains when a measure of core inflation is used to forecast headline inflation (Bermingham, 2007; Pincheira-Brown et al., 2019). Relative to these papers, we argue in favour of forecasting using the cleaned core inflation measure. In addition, several papers have considered the gains from using disaggregated data in forecasting headline inflation. Marcellino et al. (2003), for instance, find that using regionally disaggregated data helps improve forecasting performance. Similarly, Chaudhuri and Bhaduri (2019) find that component-level disaggregation improves accuracy in wholesale price indices forecasts. On the contrary, however, Hubrich (2005) concludes that there are no gains to using disaggregated data, particularly when the data generating process (DGP) is not known. While we do not compare model performance from using aggregated versus disaggregated data, we do use a dataset that is disaggregated along two dimensions – region (i.e., Indian states) and components (core, food, and fuel).

The rest of the paper is organized as follows. Section 2 details the data sources and provides stylized facts; Section 3 explains the empirical methodology to estimate the passthrough of food and fuel inflation to raw core inflation; Section 4 provides the results from the passthrough estimation; Section 5 explains the construction of cleaned core measure and the results from the forecasting exercises; and Section 6 concludes.

2 Data and Stylized Facts

Our data are from three primary sources. We source the state-level and national consumer price index (CPI) data and the relevant weights from the Ministry of Statistics and Program Implementation (MoSPI). We use the oil supply shock of Baumeister and Hamilton (2019) and the food supply shock of Bhattarai et al. (2024). All data are at a monthly frequency. Our sample period is January 2011 to June 2022.

We define inflation as the year-on-year growth rate of CPI:

$$\pi^a_{s,t} = ln(CPI^a_{s,t}) - ln(CPI^a_{s,t-12}) \tag{1}$$

where $a \in \{core, food, fuel\}$ and $CPI_{s,t}$ is the consumer price index in month t for state s. We construct the core inflation measure by removing food and fuel inflation from the headline inflation.[3]

[3]The fuel CPI provided by MoSPI does not include the prices of transportation fuel. This is provided separately under the 'miscellaneous–transportation and communication' category at the state-urban and state-rural levels. We first add this component to fuel CPI at the state-urban and state-rural levels, and then construct the state-level core inflation using urban and rural area weights for each state. We detail

2.1 Stylized Facts

In this section, we establish stylized facts from India's state-level inflation data.

2.1.1 Greater Variation in the State-Level Data

Using state-level inflation data over national-level time series allows us to exploit the benefits of a richer dataset. Table 1 shows that the inflation variation is greater in the state-level data compared to the national data. For example, the standard deviation of the headline inflation is 1.35 times higher in the state data (2.80) compared to the national data (2.07). The variance differential is more pronounced in the cases of core inflation and fuel inflation, where the standard deviation ratios are, respectively, 1.5 and 2.21. We show the higher variation in the states data visually in the online appendix (Figure A1). Although there is a large variation between states, the inflation for the median state tracks the national average inflation quite closely.

2.1.2 Large Local Variation in State Inflation

We start by complementing the information in Table 1 by decomposing the variation in state-level data into those driven by permanent state-related factors and aggregate time-related factors. We achieve this by regressing the state-level inflation against state- and year-month fixed effects separately. We summarize the role played by each set of fixed effects with the respective partial R^2 number. The results are in Table 2. We report results from both unweighted (Panel A) and weighted (Panel B, the state-level consumption weights are from MoSPI) regressions.

We draw the following conclusions. First, a large component of the variation in state-level inflation is unexplained by state- or time-factors. Taking headline inflation as an example, the partial R^2 of state fixed effects is 2.7 percent and the partial R^2 for time fixed effects is 57.1 percent, implying that about 57.6 percent of the total variation is captured by the two sets of fixed effects. In other words, the unexplained component that varies at the state-month level is large – a little over 40 percent. Second, the share of the unexplained factor varies across the different components of CPI. While it is 42.4 percent for headline inflation, it is as little as 36.4 percent for food inflation and as high as 72.3 percent for fuel inflation. Interestingly, food inflation is not permanently high or low in states, as seen by low F-statistics (0.660) on the state factors. The p-values associated with this F-statistic implies that the state fixed effects are not collectively significant even at a very high threshold of 91.9 percent level.

The patterns do not change qualitatively when we weigh each state by its administrative data on consumption share in the national sample. In other words, states which have

this methodology in an online Appendix A.

relatively low consumption share in the national data do not bias our results. However, because states with high consumption share tend to have a higher time-specific variation, the average share of the unexplained component is smaller. For example, the partial-R^2 on time fixed effects for headline inflation is higher at 0.79 in the weighted regression compared to 0.57 in the unweighted regression. Nevertheless, about 20 percent of the variation in headline inflation and 51 percent of the variation in fuel inflation remain unexplained.

2.1.3 Lower Dynamic Correlations in the States Data

We now provide preliminary evidence on the correlations between headline, food, fuel, and core inflation with the help of dynamic correlation charts for both the state-level and the national-level data. Specifically, we calculate the correlation of, for example, food inflation with the leads and lags of core inflation. Similarly, we calculate the correlation between headline and core inflation as well as fuel and core inflation. Note that the goal here is not to establish a causal link between food or fuel inflation and core inflation. The causal link is established with an instrumental variable approach later in the paper.

Figure 1 provides the dynamic correlations. We report the correlations for 12 lags and 12 leads of core inflation along with the correlation with contemporaneous core inflation. The figure in the left panel uses national data, and the figure in the right panel uses state-level data. The dashed black lines correspond to correlations with headline inflation; the solid red lines are for correlations with food inflation; and the solid blue lines are for correlations with fuel inflation.

We find that the dynamic correlations are generally lower in the states data compared to the national data. For example, the correlation of food inflation with 12th lag of core inflation is close to 0.6 in national data but 0.2 in state-level data. We observe the same pattern of downward shift in correlations for all lags and leads of core inflation with both food and fuel inflation. For headline inflation, the correlations with contemporaneous and low lags and leads of core inflation are similar in the two datasets; but the correlations at higher lags and leads of core inflation are remarkably lower in the state-level data.

The correlations can be lower in the state-level data because, in principle, local food or fuel inflation shocks can be traded off more easily compared to national shocks. When food inflation starts to spike in one state, for example, due to an adverse weather shock that destroys crops, the state can import food from neighboring states that did not experience a similar weather shock. In equilibrium, the import channel reduces (increases) food inflation in the weather-(un)affected state compared to a world where trade does not happen. Crucially, therefore, local shocks likely do not have a significant effect on local inflation expectations and are less likely to pass through to local core inflation.

2.1.4 Lower Dynamic Correlations after A Tax Reform

As argued above, the ability to trade reduces dynamic correlations at the state level. In the absence of inter-state trade data, we bolster this point by comparing dynamic correlations in the state-level data before and after a nationwide tax reform. In July 2017, India replaced the value added tax (VAT) system with the Goods and Services Tax (GST). Unlike the VAT regime where the tax slabs differed by states, the taxes on a given good are uniform across Indian states. This reform is expected to have increased cross-state trade substantially (Van Leemput and Wiencek, 2017).

In line with our hypothesis, the dynamic correlations are significantly lower in the post-GST regime. The correlation of food inflation with contemporaneous core inflation reduced from 0.48 to 0.31. We find a reduction in the correlation of even fuel inflation, which is interesting given that petrol and diesel were left out of GST and follow the old VAT system even today.

3 Empirical Methodology

Even though core inflation is obtained by removing food and fuel inflation from headline inflation, it is possible that core inflation itself will still be impacted by passthrough from the two volatile components. Hence, in this section, we specify a methodology to estimate the passthrough of food and fuel inflation to core inflation so that subsequently we are able to strip core inflation of this passthrough in order to obtain our cleaned core measure. We employ a panel local projection framework with instrumental variables (Jordà, 2005) and use India's state-level inflation data. The estimating equation is

$$\pi^{core}_{s,t+h} = \sum_{j=1}^{J} \alpha^h_j \pi^{core}_{s,t-j} + \sum_{k=0}^{K} \beta^h_k \pi^{food}_{s,t-k} + \sum_{l=0}^{L} \gamma^h_l \pi^{fuel}_{s,t-l} + \delta_s + \epsilon_{s,t+h} \tag{2}$$

where $\pi^{core}_{s,t+h}$ is the h periods ahead core inflation in state s, $\pi^{core}_{s,t-j}$, $\pi^{food}_{s,t-k}$ and $\pi^{fuel}_{s,t-l}$ are the j, k and l lags of core, food and fuel inflation respectively for state s. δ_s captures the state fixed-effects. We choose $J, K, L = 12$ and estimate Equation 2 for each horizon from $h = 0$ to $h = 12$. Standard errors are clustered at the state-level.

The ordinary least squares (OLS) estimates for the coefficients of interest (β^h_0 and γ^h_0) estimated using Equation 2 can be biased for two reasons. First, expectations about h-months ahead core inflation could affect contemporary food and fuel inflation (i.e., reverse causality). For example, cost-push shocks driving core inflation can drive up inflation expectations generally, which could drive up wages and prices in the agricultural sector. Second, omitted variables, such as exchange rates, could drive movements in both core inflation and food and fuel inflation. To account for such sources of endogeneity, we instrument the independent variables by the global oil supply shock from Baumeister and Hamilton (2019) (for fuel inflation) and the food price shocks from Bhattarai et al. (2024) (for food inflation). Baumeister and Hamilton (2019) develop an oil supply shock index by extracting residuals from a Bayesian structural vector autoregression (S-VAR) that models oil production, oil prices and real economic activity. Similarly, Bhattarai et al. (2024) construct their measure of a food price shock using a dynamic factor model that separates global influences into a common demand factor and two sector-specific (food-specific and non-food) factors. They then obtain the residuals of the global food price index from the estimated common factor and a food-specific factor. Both the oil-supply index and food-price shocks capture exogenous variations in global oil supply and food prices, respectively, allowing us to disentangle only the passthrough of food and fuel inflation to core inflation (and not vice versa).

We employ nonparametric first-stage regressions to generate state-month-year-specific instruments by regressing food and fuel inflation (separately) onto the two instruments interacted with a state dummy, as well as the other independent regressors in Equation 2,

$$\begin{aligned}\pi^{a}_{s,t} = &\sum_{k=0}^{K} \lambda_k \pi^{IV,food}_{s,t-k} + \sum_{k=0}^{K} \beta_k \pi^{IV,food}_{s,t-k} \times \delta_s + \sum_{l=0}^{L} \eta_l \pi^{IV,fuel}_{s,t-l} \\ &+ \sum_{l=0}^{L} \gamma_l \pi^{IV,fuel}_{s,t-l} \times \delta_s + \sum_{m=1}^{M} \theta_m \pi^{a}_{s,t-m} + \sum_{n=1}^{N} \phi_n \pi^{core}_{s,t-n} + \sum_{p=0}^{P} \alpha_{t-p} \pi^{a'}_{s,t-p} + \delta_s + \epsilon_{s,t+h} \end{aligned} \quad (3)$$

where, $a, a' \in \{\text{food}, \text{fuel}\}$ and $a \neq a'$. $\pi^{IV,food}_{s,t-k}$ is k lags of Bhattarai et al. (2024)'s food price shock, $\pi^{IV,fuel}_{s,t-l}$ is l lags of Baumeister and Hamilton (2019)'s oil-supply shock and δ_s is the state-specific dummy. We choose $K, L, M, N, P = 12$.

We then estimate Equation 2 using the state-specific instruments derived from Equation 3 as IV. From this we then obtain a series of coefficients for β^h_0 and γ^h_0, capturing the contemporaneous impact of current food and fuel inflation, respectively, on h periods ahead core inflation.

4 Food and Fuel Inflation Passthrough

We document the results of the panel local projections with state-specific IVs, estimated using Equation 2, in Figure 3. The figure plots the series of coefficients β^h_0 and γ^h_0 for $h = 0, 1, 2, ..., 12$. The F-statistic from the first stage is presented in Table A1 in Appendix B.

The left panel shows that food inflation has an immediate positive and statistically significant effect on core inflation. A one percentage point (pp) increase in food inflation leads to a 0.09 pp increase in contemporaneous core inflation. The impact remains persistent from $h = 0$ to $h = 10$, with an average passthrough of 9.75 basis points every month in this period, after which it begins to decrease. The effect peaks at the fifth horizon, where a one percentage point increase in food inflation corresponds to a 12.26 basis point increase in core inflation.

In contrast, the right panel shows that fuel inflation does not affect the dynamics of core inflation at the state level. This could indicate that fluctuations in fuel prices are generally transient and not persistent enough to pass through to influence the prices of nonfood and nonfuel commodities.

4.1 Heterogeneity – High versus Low Trading States

In this section, we estimate Equation 2 separately for states that were above and below the median trade share in financial year 2019-20.[4,5] This test is motivated by the stylized facts in Section 2.1 which showed that higher trade share is associated with lower dynamic correlations of food and fuel inflation with lags and leads of core inflation. We hypothesize that states that have higher trade shares are able to better insure local food and fuel shocks by trading with neighboring states. Local shocks, therefore, do not get transmitted to core inflation, as the production costs for core commodities may not be sensitive to local shocks in high trading states.

The results are in Figure 4. The top panel plots the passthrough estimates for food inflation and the bottom panel plots the passthrough for fuel inflation. In both these panels, the figure in the left corresponds to a group of states that are above the median in terms of trade share; the complement of this set of states is plotted on the right hand side.

The figure shows that states that trade less have higher passthrough, which is in line with our hypothesis. For food inflation, the passthrough on impact for high trading states is 0.07 pp. and is statistically significant. The passthrough for states that trade less is higher at 0.13 pp. We find this pattern up to a horizon of about 5 months, after which the passthrough falls in both the set of states. In the case of fuel inflation, there is zero passthrough in the states that trade more; however, the passthrough on impact is positive and statistically significant at 95 percent level for states that trade less. For the latter set of states, the passthrough remains statistically significant up until 4 months.

[4] A financial year runs from April to March of the following year. We are forced to use the states' trade share ranking from a year that is in our sample since such data is not available prior to the implementation of GST in 2017. We assume that the categorization of high- versus low-trading states is unaffected by changes in inter-state trade patterns following GST.

[5] States are classified into high- and low-trading categories based on their total trade value normalized by the gross state domestic product (GSDP). The total trade value is measured as the sum of a state's exports and imports for the financial year 2019-20. The export and import data is the e-way bills data from the Goods and Services Tax Network. States with a trade value relative to GSDP above the median are classified as high trading states, while those with a lower value are classified as low trading states.

5 Forecasting with Cleaned Core

We now use the estimates of passthrough from food inflation to core inflation from Section 4 to construct a cleaned core inflation measure. This measure eliminates the effects of past food inflation. We then compare the usefulness of the cleaned core inflation measure in forecasting future headline inflation with that of the raw core inflation measure.

Historically, raw core inflation has been a poor predictor of future headline inflation (see Atkeson and Ohanian (2001); Bullard (2011)). One solution to this is to construct a narrower core inflation measure that eliminates volatile components other than food and fuel inflation (Bryan and Pike, 1991; Bryan and Cecchetti, 1994; Cogley, 2002; Dolmas and Koenig, 2019). This is an extension of the logic behind the construction of the core inflation. In particular, this approach looks for narrower sets of products with relatively stable prices which capture the trend inflation and is a useful predictor of headline inflation. We differ in our approach by limiting our core inflation to a measure that strips out just the food and fuel inflation impact and do not seek to further narrow the set of products.

We proceed as follows. We first explain the construction of the cleaned core measure. We then explain the forecasting models and the results.

5.1 A Measure of Cleaned Core

Having determined the extent of food and fuel inflation passthrough to core inflation in the state-level data, we construct a measure of core inflation that excludes the effect of this passthrough. Since we previously showed that there is no significant passthrough of fuel inflation to the core, we limit ourselves to the influence of food inflation on the core inflation while constructing the cleaned core inflation.

The cleaned core is obtained by removing the combined influence of current and lagged food inflation on core using the following.

$$\pi_{s,t}^{\text{Cleaned Core}} = \pi_{s,t}^{core} - \sum_{h=0}^{12} \beta_0^h \pi_{s,t-h}^{food} \tag{4}$$

where, $\pi_{s,t}^{\text{Cleaned Core}}$ is the measure of core inflation devoid of food inflation for state s at time t, $\pi_{s,t}^{core}$ is the core inflation for state s at time t and $\pi_{s,t-h}^{food}$ is h lags of food inflation for state s.

5.2 Forecasting and Evaluation Framework

In the exercise explained below, we compare the forecast performance of the cleaned and raw core inflation measures using root mean squared error (RMSE) of the headline inflation at various forecasting horizons. We choose a three-variable vector error correction model (VECM) as the baseline forecasting framework since the state-level inflation measures are not always stationary but are co-integrated.[6] The three variables in the VECM are food inflation, fuel inflation, and cleaned or raw core inflation measures.

In the state-level VECM specifications, we consider 31 states which account for 98.8 percent of the national consumption.[7] We consider the sample period prior to COVID, i.e., from January 2012 to February 2020.

5.2.1 The Vector Error Correction Model

Vector based models have been widely used to forecast inflation due to their ability to capture relationships between disaggregated components (Marcellino et al., 2003; Hubrich, 2005; Chaudhuri and Bhaduri, 2019). Such models offer flexibility that allows for disaggregation across different dimensions: across geography (for example, see Marcellino et al. (2003)), or across components (for instance, Hubrich (2005)). In our case, where we incorporate disaggregation across components, vector-based models, such as VARs and VECMs, are particularly helpful, as they are able to capture interdependencies.

We employ a VECM instead of a VAR because the inflation data are not stationary across all states and components. In such cases, the presence of unit roots renders VAR models unsuitable, making the VECM a more appropriate choice. A VECM is simply a VAR that has been re-parametrized to account for the non-stationarity of the data in levels. The VECM takes the following form:

$$\Delta y_t = \alpha\beta' y_{t-1} + \Gamma_1 \Delta y_{t-1} + \cdots + \Gamma_{p-1} \Delta y_{t-p+1} + u_t \tag{5}$$

where, y_t is a vector consisting of three variables - raw core/cleaned core, food and fuel inflation in levels. $\Gamma_1, \Gamma_2, \cdots \Gamma_{(t-p+1)}$ are 3×3 matrix of coefficients, β and α are the co-integration matrix and loading matrix, respectively, and are of the order ($3 \times r$), where r is the co-integration rank. The term $\alpha\beta' y_{t-1}$ is called the (lagged) error correction term. The optimal lag p is determined by the Bayesian Information Criteria (BIC) for each state.

To determine α and β, we need to identify the co-integration rank. To do so, we use the Johansen's co-integration test for each state. Notice that it is not always necessary for a state to have co-integrating relationships across the three variables. In cases where the co-integration rank is zero, the term $\alpha\beta$ disappears, and Equation 5 breaks down to a VAR in differences.

[6] Appendix Table A2 and A3 list the unit root test results and the co-integration ranks respectively for each of the three variables in VECM for each state.

[7] The following states are excluded due to lack of data on miscellaneous – transportation and communication CPI: Arunachal Pradesh, Andaman and Nicrobar Islands, Dadra and Nagar Haveli, Daman and Diu, and Lakshadweep.

5.2.2 Forecast Evaluation

To evaluate the performance of our forecasts, we performed a pseudo-out-of-sample rolling forecast exercise. We consider an estimation window of 60 months and a forecast horizon of 12 months. That is, starting from January 2012, we estimate the VECM on 60 months of data (i.e. January 2012 to December 2016). Using the parameters from this estimate, we obtain forecasts for the next 12 months, January 2017 to December 2017. In the second iteration, we move the estimation window by one month on either sides (February 2012 to January 2017) and obtain forecasts for the next 12 months (February 2017 to January 2018). We do this until the last estimation window, which ends in January 2020.[8]

We obtain forecasts for raw/cleaned core, food and fuel inflation for each state using VECM. Using their respective weights in the CPI basket, we obtain the forecast for headline inflation for each state. From these state-level headline inflation forecasts and the consumption weights of each state, we obtain the national-level headline inflation. We then compute the root mean squared error (RMSE) by comparing the forecasts in each month with the actual national headline inflation over different horizons.

5.3 Results

i We now discuss the results of the forecasting exercises using the raw and cleaned core inflation measures. Figure 5 plots the RMSEs across the horizons[9] We also plot RMSE from a three-variable national data VECM for comparison. "State Raw Core" (red line) represents the three-variable VECM with state-level raw (i.e., *unadjusted* for passthrough) core, food, and fuel inflation, where state-level forecasts for the three variables are aggregated to obtain forecasts for national-level headline inflation. The "State Cleaned Core" (red dotted line) refers to the three-variable VECM with state-level cleaned (i.e., *adjusted* for passthrough) core, food, and fuel inflation (i.e., *adjusted* for passthrough). The "National Raw" core (blue line) refers to the three-variable VECM directly applied to the national-level core (*unadjusted* for passthrough), food and fuel inflation.

In shorter horizons, the VECM with national-level data performs the best. For instance, the RMSE at $h = 0$ is just over 0.57 in the VECM with the national-level raw core measure, as against 0.68 in the VECM using state-level raw core and 1.17 in the VECM using the cleaned core measure. The higher accuracy of the forecasts from the national-level data indicates that there are no significant forecasting gains made from using geographically disaggregated data (in contrast to Marcellino et al. (2003)) or from using a cleaned measure of core inflation. The forecasts from raw core outperform the

[8] For the last estimation window, we obtain forecasts only for one horizon. Similarly, for the estimation window ending December 2019, we obtain forecasts only for two horizons. In general, in the last 11 estimation windows, the number of forecasts we obtain will be less than 12.

[9] We also present the full sample results in Figure A2 in Appendix E.

forecasts from cleaned core until $h = 6$ suggesting that there are no forecasting improvements made by removing the influence of food inflation in the short-run.

For longer horizons however, $h > 6$, we notice that using the cleaned core measure outperforms the other two specifications. At horizon $h = 7$, the RMSE in the VECM with raw core is 2.15, which is higher than the RMSE of 2.11 in the VECM with cleaned core. The lower RMSE of the cleaned measure of core at higher horizons indicates that the cleaned core inflation measure is useful for forecasting over the longer horizon.

Overall, the findings indicate that the cleaned core is not particularly helpful in forecasting shorter-horizons headline inflation, but is helpful in obtaining relatively more accurate forecasts of headline inflation for longer-horizons. This suggests that the passthrough of food inflation to core inflation can play an important role in obtaining more reliable forecasts in the longer horizons. Not adjusting for this spillover, in the longer run, diminishes the models' ability to forecast headline inflation.

6 Conclusions

Forecasting headline inflation is fraught with under-performing models and yet, given the lags in monetary transmission, it is crucial for central banks to be able to effectively forecast headline inflation. In fact in inflation targeting economies like India, central banks have to target the inflation forecast as opposed to actual inflation. Core inflation is generally tracked to gauge the underlying trend inflation in nonvolatile components (i.e., components other than food and fuel) of the consumer price index. However, the usefulness of this measure is also debatable.

In this paper we provide a partial resolution to this conundrum. The raw core inflation measure is potentially contaminated by the passthrough of past food and fuel inflation and may therefore have little predictive power. We estimate this passthrough in an instrumental variable setting and then remove this passthrough effect to construct a cleaned core inflation measure. We then show that this cleaned version of core inflation performs well over a longer prediction horizon in forecasting headline inflation compared to the raw core measure.

References

Anand, R., D. Ding, and M. V. Tulin (2014): *Food inflation in India: The role for monetary policy*, International Monetary Fund.

Atkeson, A. and L. E. Ohanian (2001): "Are Phillips curves useful for forecasting inflation?" *Federal Reserve Bank of Minneapolis Quarterly Review*, 25, 2–11.

Baumeister, C. and J. D. Hamilton (2019): "Structural interpretation of vector autoregressions with incomplete identification: Revisiting the role of oil supply and demand shocks," *American Economic Review*, 109, 1873–1910.

Bermingham, C. (2007): "How useful is core inflation for forecasting headline inflation?" *Economic & Social Review*, 38.

Bhattarai, S., A. Chatterjee, and G. Udupa (2024): "Food, Fuel, and Facts: Distributional Effects of External Shocks," .

Bryan, M. F. and S. G. Cecchetti (1994): "Measuring core inflation," in *Monetary policy*, The University of Chicago Press, 195–219.

Bryan, M. F. and C. Pike (1991): "Median price changes: an alternative approach to measuring current monetary inflation," *Economic Commentary*.

Bullard, J. (2011): "Measuring inflation: the core is rotten," *Federal Reserve Bank of St. Louis Review*, 93, 223–233.

Chaudhuri, K. and S. N. Bhaduri (2019): "Inflation Forecast: Just use the Disaggregate or Combine it with the Aggregate," *Journal of Quantitative Economics*, 17, 331–343.

Cogley, T. (2002): "A simple adaptive measure of core inflation," *Journal of Money, Credit and Banking*, 94–113.

Cravino, J. and A. A. Levchenko (2017): "Multinational Firms and International Business Cycle Transmission," *The Quarterly Journal of Economics*, 132, 921–962.

Dolmas, J. and E. F. Koenig (2019): "Two Measures of Core Inflation: A Comparison," *Federal Reserve Bank of Kansas City Economic Review*, 86, 5–31.

Hsiao, C. (2022): *Analysis of Panel Data*, 64, Cambridge University Press.

Hubrich, K. (2005): "Forecasting euro area inflation: Does aggregating forecasts by HICP component improve forecast accuracy?" *International Journal of Forecasting*, 21, 119–136.

Jordà, (2005): "Estimation and Inference of Impulse Responses by Local Projections," *American Economic Review*, 95, 161–182.

Marcellino, M., J. H. Stock, and M. W. Watson (2003): "Macroeconomic forecasting in the euro area: Country specific versus area-wide information," *European Economic Review*, 47, 1–18.

Peersman, G. (2022): "International food commodity prices and missing (dis) inflation in the euro area," *Review of Economics and Statistics*, 104, 85–100.

Pincheira-Brown, P., J. Selaive, and J. L. Nolazco (2019): "Forecasting inflation in Latin America with core measures," *International Journal of Forecasting*, 35, 1060–1071.

Van Leemput, E. and E. A. Wiencek (2017): "The effect of the GST on Indian growth," *Federal Reserve International Finance Discussion Paper Note.*

	Mean	SD	Min	Max
Panel A: State-level (Unweighted)				
Headline	5.69	2.80	-2.69	18.81
Core	5.64	2.30	-1.84	19.29
Food	5.68	4.42	-8.90	21.38
Fuel	6.14	6.85	-45.68	45.63
Observations	3815			
Panel B: State-level (Weighted)				
Headline	5.70	2.56	-2.69	18.81
Core	5.69	1.86	-1.84	19.29
Food	5.71	4.28	-8.90	21.38
Fuel	5.67	4.99	-45.68	45.63
Observations	3815			
Panel C: National-level				
Headline	5.66	2.07	1.45	10.89
Core	5.67	1.53	3.35	10.29
Food	5.89	3.45	-1.70	15.40
Fuel	4.90	3.10	-1.31	11.73
Observations	152			

Table 1: Summary Statistics: State-level and National-level Inflation Data

Note: This table lists key summary statistics of state-level and national-level headline, core, food and fuel inflation in India. Panel A summarizes assuming equal weights to all states. In panel B, states are weighted according to state-level administrative consumption weights. Panel C summarizes national-level headline, core, food and fuel inflation.

	State Factors			Common Factors			Unexplained Share
	Partial R^2	F-stat	p-value	Partial R^2	F-stat	p-value	R^2
Panel A: Unweighted							
Headline Inflation	0.027	1.450	0.053	0.571	39.180	0.000	0.424
Core Inflation	0.053	4.070	0.000	0.418	20.560	0.000	0.564
Food Inflation	0.015	0.660	0.919	0.634	51.560	0.000	0.364
Fuel Inflation	0.061	6.150	0.000	0.241	9.030	0.000	0.723
Panel B: Weighted							
Headline Inflation	0.059	1.620	0.018	0.792	108.380	0.000	0.205
Core Inflation	0.094	3.670	0.000	0.716	69.220	0.000	0.275
Food Inflation	0.025	0.650	0.928	0.794	113.990	0.000	0.205
Fuel Inflation	0.083	6.060	0.000	0.457	23.330	0.000	0.518

Table 2: Role of Common and Local Shocks on State-level Inflation

Note: The decomposition follows the fixed-effects model in Cravino and Levchenko (2017) with the following specification: $\pi^a_{s,t} = \delta_s + \delta_t + \epsilon_{s,t}$ where $a \in$ (headline, core, food, fuel). The equation is estimated using OLS and the p-values correspond to the F-statistic. Results from regressions with state-level consumption weights are reported in panel B. The last column (unexplained share) captures the variation not explained by state-level time-invariant factors and aggregate time-specific factors.

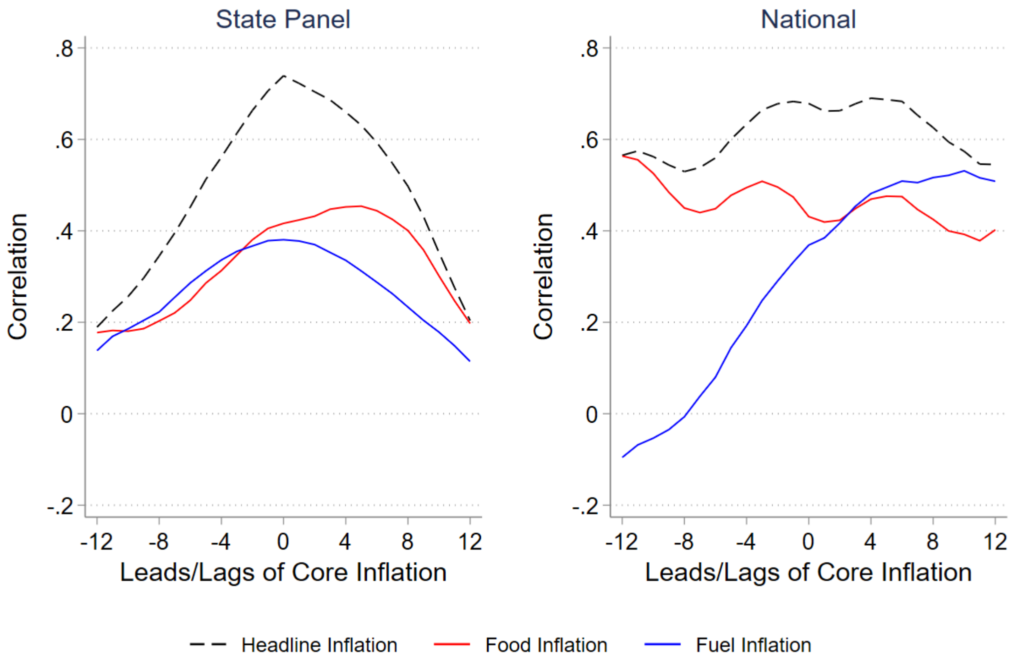

Figure 1: Dynamic Correlations – Core Inflation with Headline, Food, and Fuel Inflation

Note: This figure plots the correlation between the lags/leads of core inflation and contemporaneous headline, food, and fuel inflation. Negative values on the x-axis correspond to lags of core inflation; positive values correspond to leads of core inflation. The figure on the left plots the correlations in the state-level data. The figure on the right plots the correlations in the national-level time series data. The sample period is from January 2012 to June 2022.

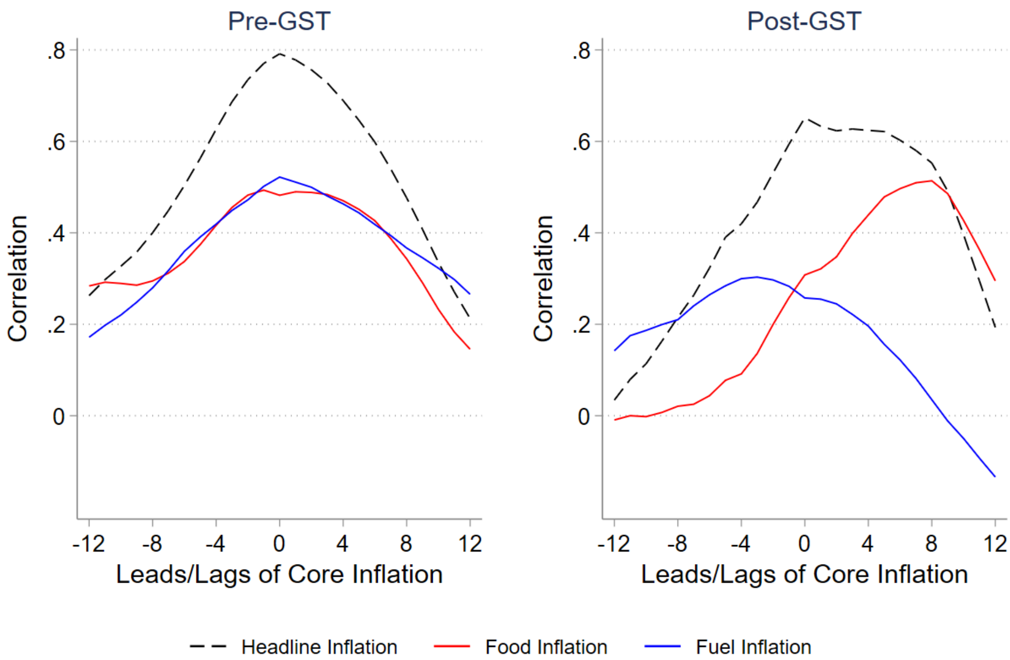

Figure 2: Dynamic Correlations – Before and After the Implementation of the Goods and Services Tax (GST)

Note: This figure plots the correlation coefficient of leads/lags of core inflation with contemporaneous headline, food, and fuel inflation in the state-level data before and after the implementation of GST. Negative values on the x-axis correspond to lags of core inflation; positive values correspond to leads of core inflation. The sample for the pre-GST period is from January 2012 to June 2017. The sample for the post-GST period is from July 2017 to June 2022.

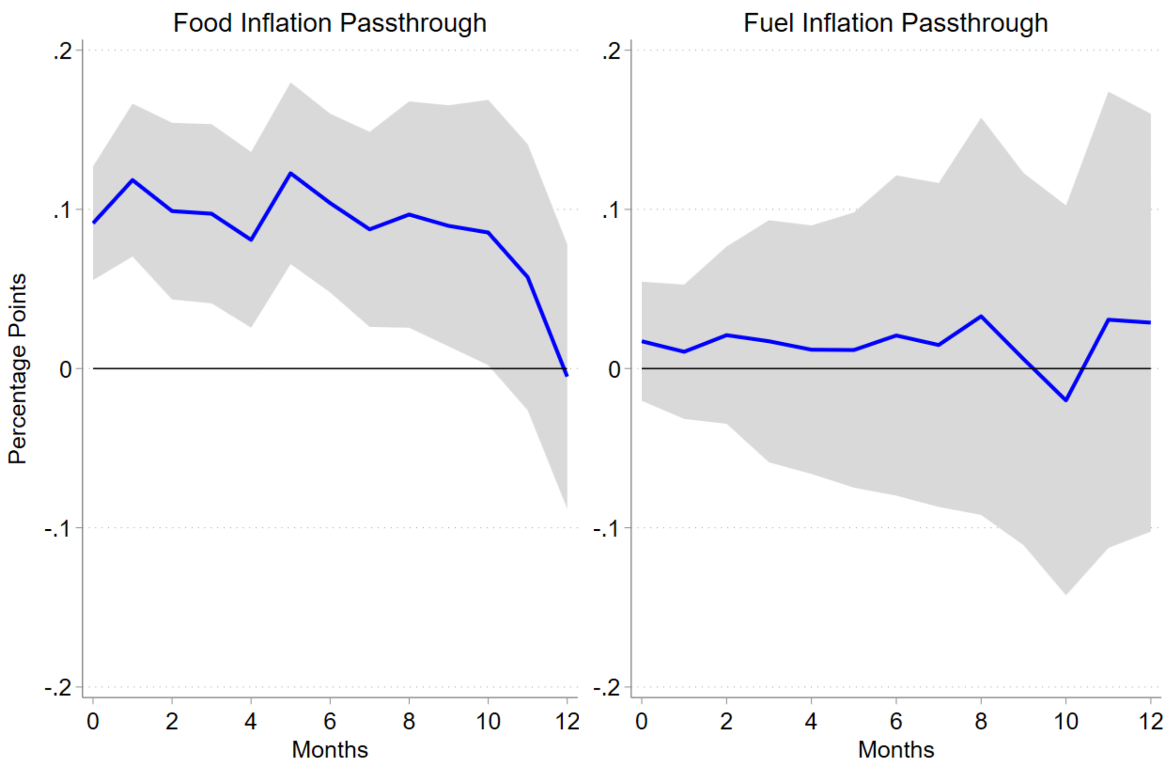

Figure 3: Passthrough of Food and Fuel Inflation to Core Inflation from Panel Local Projections

Note: This figure plots the estimated passthrough of food and fuel inflation to core inflation using panel local projections with instrument variables (Equation 2). We use Bhattarai et al. (2024)'s food shock index as an instrument for food inflation and Baumeister and Hamilton (2019)'s oil supply shock as an instrument for fuel inflation. The figure on the left captures the estimated passthrough of food inflation to core inflation, and the figure on the right captures the estimated passthrough of fuel inflation to core inflation. The blue line is the estimated impulse responses (point estimates), and the shaded region represents 95 percent confidence intervals.

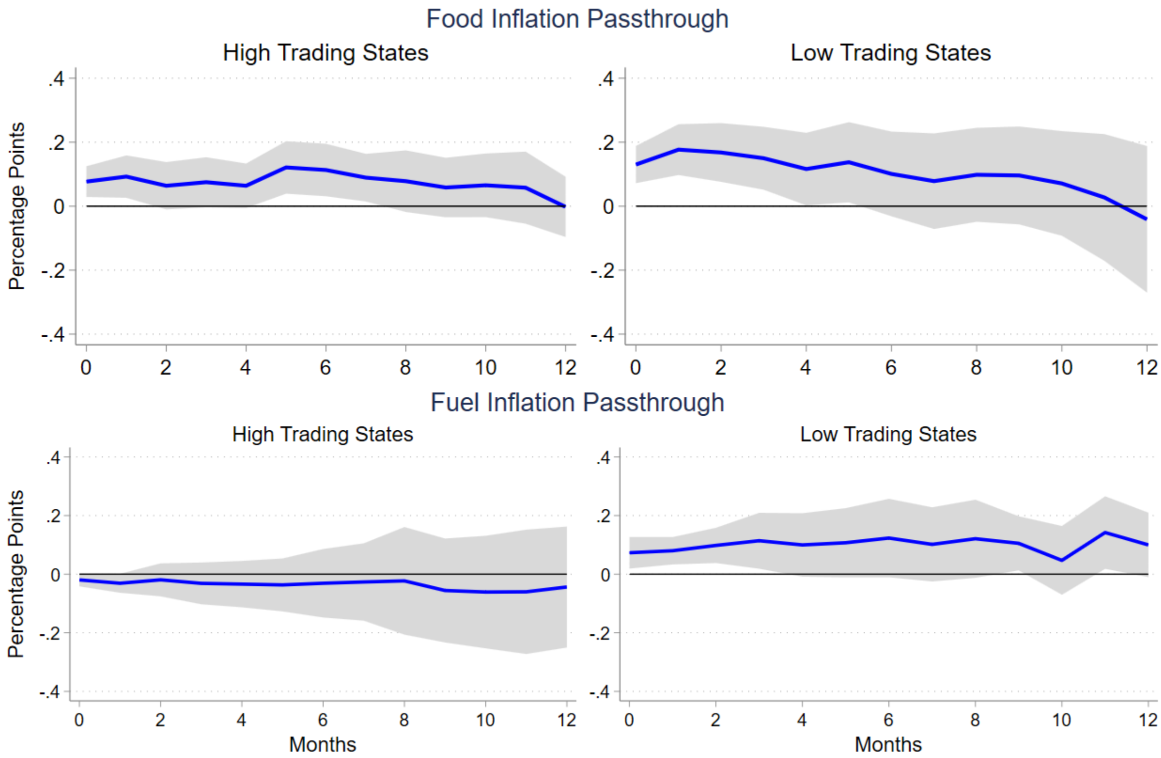

Figure 4: Passthrough of Food and Fuel Inflation to Core Inflation by Trade Heterogeneity

Note: This figure plots the estimated passthrough of food and fuel inflation to core inflation by states depending on the trade share. States with a trade value (defined as the sum of exports and imports) relative to GSDP above the median are classified as high trading states, while those with a lower value are classified as low trading states. The top panel reports the estimated passthrough of food inflation to core inflation, and the bottom panel reports the estimated passthrough of fuel inflation to core inflation. The figures on the left correspond to high trading states, and the figures on the right correspond to low trading states. The blue line is the estimated impulse responses (point estimates) and the shaded region represents the 95 percent confidence interval.

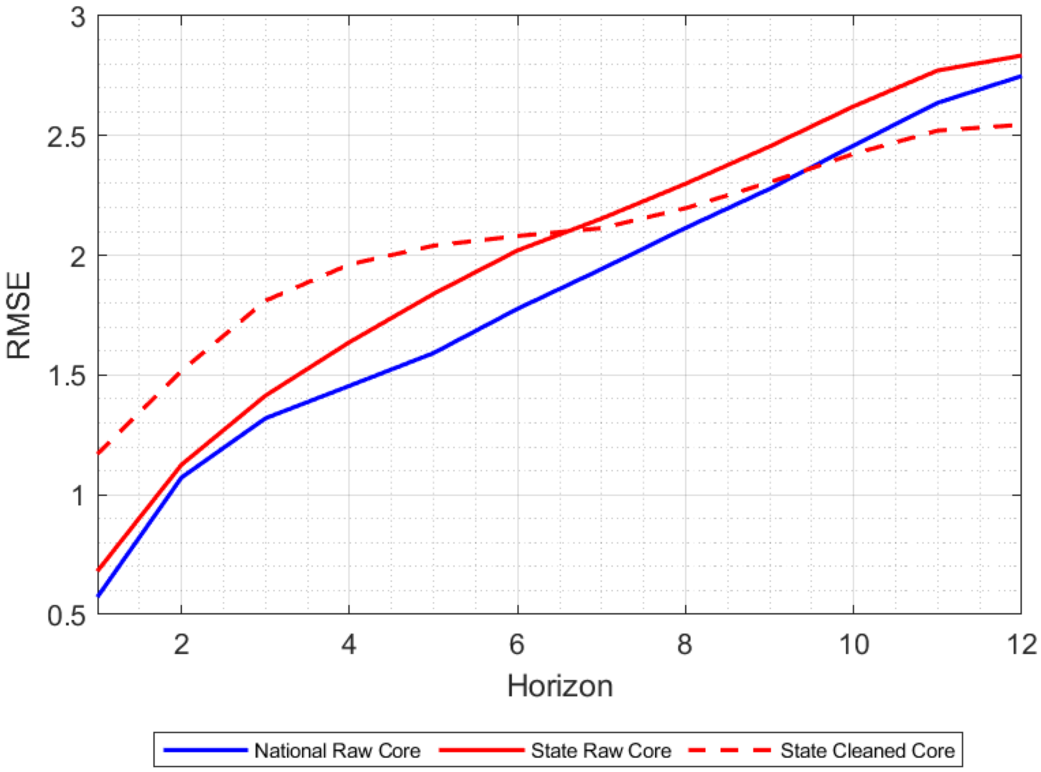

Figure 5: Comparison of Forecasting Performance with Raw and Cleaned Core Inflation Measures

Note: The figure plots the root mean squared errors (RMSEs) generated from rolling-window pseudo out-of-sample forecasts and actual data. The RMSEs are computed against headline inflation and scaled by 100. State Raw Core represents a three-variable VECM with state-level raw (i.e.,*unadjusted* for passthrough) core, food, and fuel inflation, where state-level forecasts for the three variables are aggregated to get forecasts for national-level headline inflation. State Cleaned Core refers to a three-variable VECM with state-level cleaned (i.e., *adjusted* for passthrough) core, food, and fuel inflation. National Raw Core refers to the three-variable VECM directly applied to the national-level core (*unadjusted* for passthrough), food, and fuel inflation.

Online Appendix

A Constructing Core CPI

The state-level core CPI is not calculated by MoSPI. We construct this variable by computing the core CPIs at the state-urban and state-rural levels and then aggregating them to the state-level using each state's administrative data on consumption weights in the urban and rural areas.

The construction of core CPI is not straight-forward, for example by just subtracting food and fuel CPIs from the headline CPI at the state-level, for two reasons. First, the main 'fuel and light' CPI excludes consumption of transportation fuel, which is an important component that is instead included as part of a separate CPI category (Miscellaneous – transportation and communication).[10] Second, the miscellaneous – transportation and communication CPI is only available at the state-urban and state-rural levels, but not at the state-level. We construct the state-urban and state-rural core CPI as,

$$\mathrm{CPI}_{it}^{\mathrm{core}} = \frac{\mathrm{CPI}_{it}^{\mathrm{headline}} - \mathrm{CPI}_{it}^{\mathrm{food}} \times w_i^{\mathrm{food}} - \mathrm{CPI}_{it}^{\mathrm{fuel}} \times w_i^{\mathrm{fuel}} - \mathrm{CPI}_{it}^{\mathrm{TC}} \times w_i^{\mathrm{TC}}}{1 - w_i^{\mathrm{food}} - w_i^{\mathrm{fuel}} - w_i^{\mathrm{TC}}} \tag{6}$$

where i indicates state-geography (i.e, urban and rural), $\mathrm{CPI}^{\mathrm{food}}$ is the 'food and beverages' CPI, $\mathrm{CPI}^{\mathrm{fuel}}$ is the 'fuel and light' CPI (does not include price of transportation fuel), $\mathrm{CPI}^{\mathrm{TC}}$ is the 'transportation and communication' CPI; and w^{food}, w^{fuel}, and w^{TC} are, respectively, the weights of food and beverages, fuel and light, and transportation and communication in the state-geography consumption basket.

Next, we aggregate the state-geography level core CPI to the state-level using the urban-rural weights for each state:

$$\mathrm{CPI}_{s,t}^{\mathrm{core}} = w_{urban(s)} \times \mathrm{CPI}_{urban(s),t}^{\mathrm{core}} + w_{rural(s)} \times \mathrm{CPI}_{rural(s),t}^{\mathrm{core}} \tag{7}$$

where $w_{urban(s)}$ and $w_{rural(s)}$ are, respectively, the urban and rural consumption weights for state s, such that $w_{urban(s)} + w_{rural(s)} = 1$ for each s.

A.0.1 Adjusted Fuel CPI

As the 'fuel and light' CPI index does not account for the prices of transportation fuel, we construct an adjusted measure to incorporate this element. This is done by computing a weighted-average of 'fuel and light' CPI and the 'miscellaneous – transport and com-

[10] The CPI indices are available for the following categories: i. headline (representing the weighted-average of all the prices); ii. food and beverages; iii. pan, tobacco, and intoxicants; iv. clothing and footwear; v. fuel and light; and vi. miscellaneous. Within the miscellaneous category, the CPI indices are further divided into household goods and services, health, transport and communication, recreation, education, and personal care and effects.

munication' at the state-geography level and then aggregating to the state-level (using urban-rural weights, similar to Equation 7).

B First Stage

	(1)	(2)	(3)
		Core Inflation	
Food Inflation	0.0913***	0.0917***	
	(0.0182)	(0.0188)	
Fuel Inflation	0.0172		0.0185
	(0.0191)		(0.0193)
Cragg-Donald Wald F statistic)	542.9	1262.0	1338.3
Kleibergen-Paap rk Wald F statistic	2833.2	13554.5	9059.9
State Fixed Effects	Yes	Yes	Yes
State Cluster	Yes	Yes	Yes
Food Inflation IV	Yes	Yes	No
Fuel Inflation IV	Yes	No	Yes
Observations	2914	2914	2914

Standard errors in parentheses

* $p < 0.10$, ** $p < 0.05$, *** $p < 0.01$

Table A1: First-Stage Regression for $h = 0$.

Note: This table documents the results from the first-stage regressions of the panel local projections with horizon $h = 0$. Column one contains the results of the passthrough estimates by instrumenting both food and fuel inflation with the Bhattarai et al. (2024)'s food shock index and Baumeister and Hamilton (2019)'s oil supply shock index, respectively. In columns two and three, we instrument only for food and fuel inflation, respectively. We control for state-level time-invariant factors and cluster the standard error at the state level.

C Unit Root and Cointegration Tests

State	Core Inflation		Food Inflation		Fuel Inflation	
	DF Statistic	p-Value	DF Statistic	p-Value	DF Statistic	p-Value
Andhra Pradesh	-2.065**	0.038	-1.624*	0.098	-2.643***	0.009
Assam	-1.409	0.147	-1.196	0.212	-1.868*	0.059
Bihar	-1.480	0.129	-1.653*	0.093	-1.984**	0.046
Chandigarh	-1.454	0.136	-1.601	0.103	-2.628***	0.009
Chhattisgarh	-1.690*	0.086	-2.010**	0.043	-1.447	0.138
Goa	-1.215	0.205	-0.923	0.312	-1.560	0.112
Gujarat	-1.952**	0.049	-1.116	0.241	-2.352**	0.019
Haryana	-0.996	0.285	-1.442	0.139	-1.571	0.109
Himachal Pradesh	-0.944	0.304	-1.365	0.159	-2.036**	0.041
Jammu and Kashmir	-1.131	0.235	-1.294	0.179	-2.396**	0.017
Jharkhand	-1.606	0.102	-1.427	0.143	-1.753*	0.075
Karnataka	-1.388	0.152	-1.295	0.179	-1.195	0.212
Kerala	-1.609	0.101	-0.941	0.305	-1.299	0.178
Madhya Pradesh	-1.779*	0.072	-0.949	0.302	-1.908*	0.054
Maharashtra	-1.655*	0.092	-1.077	0.255	-1.466	0.133
Manipur	-1.589	0.105	-1.089	0.251	-2.131**	0.033
Meghalaya	-2.388**	0.017	-2.117**	0.034	-2.538**	0.012
Mizoram	-1.557	0.112	-0.654	0.410	-2.085**	0.036
NCT of Delhi	-1.285	0.182	-1.654*	0.092	-3.113***	0.003
Nagaland	-1.366	0.159	-0.940	0.305	-1.582	0.107
Odisha	-2.086**	0.036	-1.604	0.102	-1.287	0.182
Puducherry	-0.979	0.291	-1.605	0.102	-2.695***	0.008
Punjab	-1.301	0.177	-1.363	0.160	-1.597	0.104
Rajasthan	-1.936*	0.051	-1.090	0.251	-1.782*	0.071
Sikkim	-1.166	0.223	-1.177	0.219	-1.747*	0.076
Tamil Nadu	-1.478	0.130	-0.859	0.335	-2.281**	0.023
Telangana	-1.642*	0.095	-1.493	0.126	-2.208**	0.027
Tripura	-0.842	0.341	-1.657*	0.092	-2.515**	0.012
Uttar Pradesh	-2.125**	0.033	-1.358	0.161	-1.175	0.219
Uttarakhand	-1.706*	0.083	-1.135	0.234	-2.572**	0.011
West Bengal	-1.864*	0.060	-1.302	0.177	-1.809*	0.067
National-level	-1.836*	0.063	-0.896	0.321	-1.458	0.135

Table A2: Augmented Dickey Fuller Test by State for (Raw) Core, Food and Fuel Inflation

Note: In this table, we report results from the Augmented Dickey-Fuller (ADF) test to check for the presence of unit root under the null. We fit a standard autoregressive model without drift for each state. The lag order is determined through Bayesian information criterion (BIC). *, ** and *** indicate statistical significance at 10, 5, and 1 percent level respectively.

State	**Cointegration Rank**
Andhra Pradesh	2
Assam	0
Bihar	1
Chandigarh	0
Chhattisgarh	1
Goa	0
Gujarat	0
Haryana	1
Himachal Pradesh	0
Jammu and Kashmir	0
Jharkhand	1
Karnataka	0
Kerala	0
Madhya Pradesh	0
Maharashtra	0
Manipur	0
Meghalaya	0
Mizoram	0
NCT of Delhi	2
Nagaland	0
Odisha	0
Puducherry	0
Punjab	0
Rajasthan	1
Sikkim	0
Tamil Nadu	0
Telangana	0
Tripura	0
Uttar Pradesh	1
Uttarakhand	1
West Bengal	0
National-level	1

Table A3: Cointegration Rank by State for (Raw) Core, Food and Fuel Inflation

Note: The cointegration rank has been computed using the Johansen's Cointegration (JCI) Trace statistic test for (raw) core, food, and fuel inflation at the state level. We employ the version of the model without deterministic trends and include an intercept. The lag order is chosen using Bayesian information criterion (BIC).

D Inflation – State Versus National Data

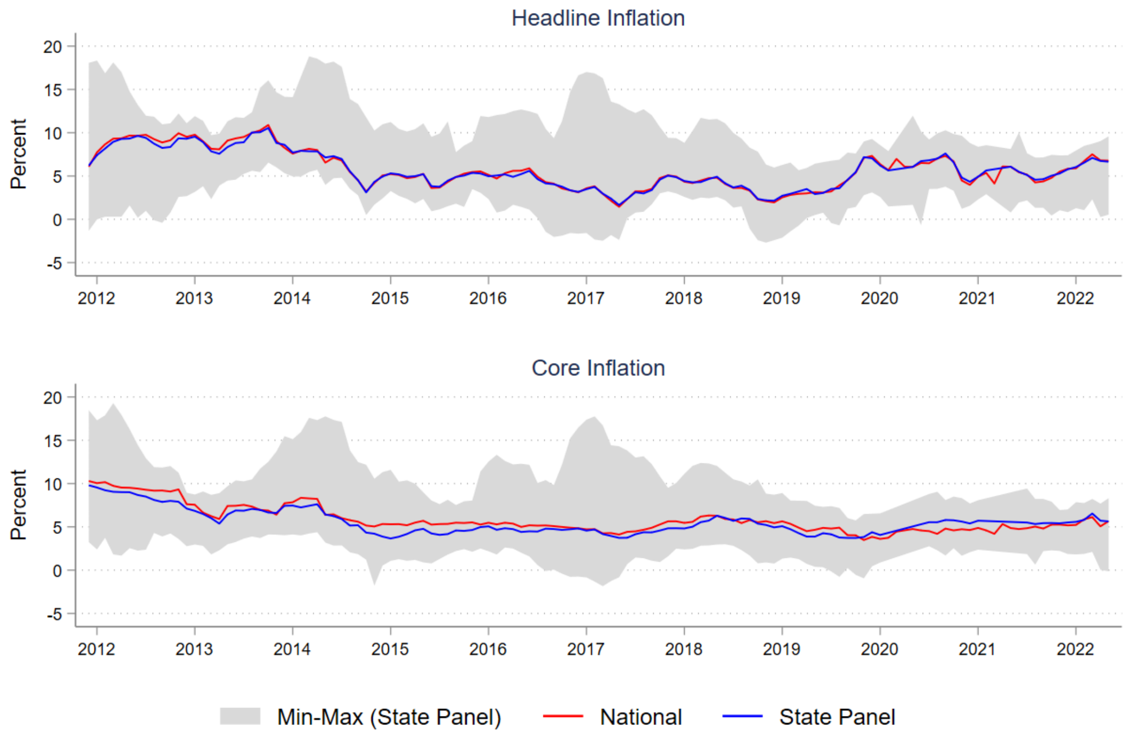

Figure A1: Inflation in the States Versus National Data

Note: This figure shows the link between inflation in the states versus national data. The top panel is for headline inflation whereas the bottom panel is for core inflation. The gray shaded region in each chart depicts the minimum and maximum inflation in the states data; the red solid line depicts national inflation; the blue solid line depicts the inflation for the median state.

E Full-Sample Forecasting Results

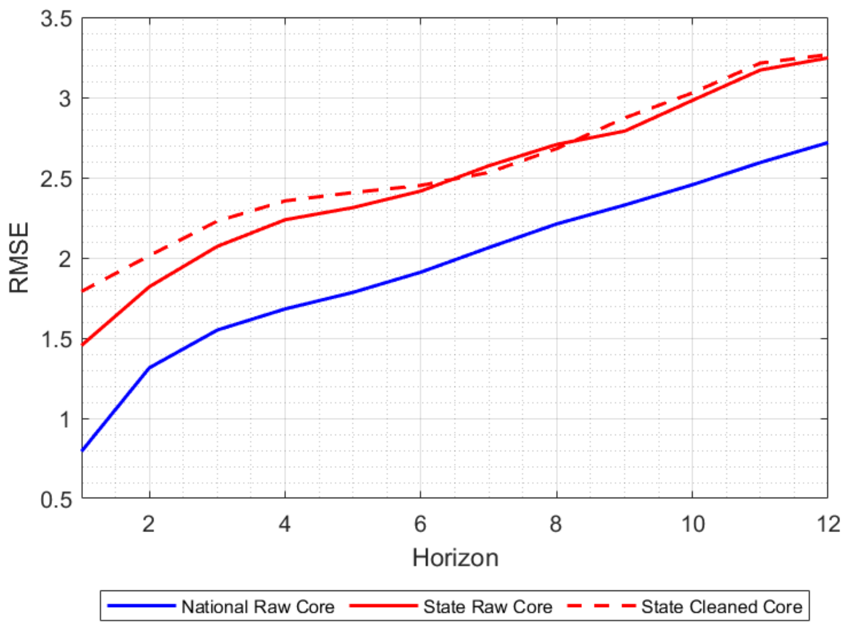

Figure A2: RMSE from the Pseudo Out of Sample Rolling Forecasts of Headline Inflation for the Sample Period from January 2012 to June 2022.

Note: The figure plots the root mean squared errors (RMSEs) generated from rolling-window pseudo out-of-sample forecasts and actual data. The RMSEs are computed against headline inflation and scaled by 100. State VECM Raw Core represents the three-variable VECM with state-level core (*unadjusted* for passthrough), food and fuel inflation where state-level forecasts for the three variables are aggregated to get forecasts for national-level headline inflation. State VECM Cleaned Core refers to the three-variable VECM with state-level core (*adjusted* for passthrough), food and fuel inflation. India VECM Raw Core refers to the three-variable VECM directly applied to the national-level core (*unadjusted* for passthrough), food and fuel inflation.