Mind the Gap: Household Inflation Expectations and Monetary Policy in India

Pushpendu Ghosh and Aditi Singh[1]

Centre for Advanced Financial Research and Learning (CAFRAL)

Abstract

The persistent gap between household inflation expectations and measured inflation complicates the interpretation of surveys for monetary policy. Using the Reserve Bank of India's Inflation Expectations Survey of Households microdata for 2014–25, this article examines forecast bias, reporting patterns and the consequences of filtering responses. Households overpredict urban consumer price inflation by about five percentage points on average, with greater overprediction in the post-pandemic sample, particularly at the one-year horizon. Calendar patterns explain little individual variation, while rounding and the open-ended upper response category shape the reported distribution. Filters based on proximity between current inflation perceptions and official CPI bring aggregate expectations closer to realised inflation, but exclude substantial shares of responses. The three- and four-percentage-point bands retain 47.3% and 56.5% of observations, respectively. For monetary policy, closer agreement with CPI must be weighed against whose expectations remain in the measure.

Keywords: inflation expectations; household surveys; monetary policy; consumer price inflation; survey design; forecast errors; India

[1] Disclaimer: The views expressed here are solely the responsibility of the author and should not be interpreted as reflecting the views of the Centre for Advanced Financial Research and Learning (CAFRAL) or the Reserve Bank of India.

Household inflation expectations are difficult to interpret when they persistently exceed inflation measured by official statistics. For the Reserve Bank of India (RBI), this raises a practical question: what should a high reported expectation convey about households' inflation outlook? The answer matters because the same survey can be used to describe reported beliefs, assess forecasts against an official price index, or construct an indicator intended to track that index. Each use requires a different reading of the numbers.

India's Inflation Expectations Survey of Households (IESH) provides evidence on what urban households perceive and expect. In the sample examined here, households overpredict subsequent urban consumer price index (CPI) inflation by about five percentage points on average. This persistent gap complicates a literal reading of the reported level as a forecast of official inflation. It also complicates statistical adjustment: removing responses that lie far from CPI changes which households contribute to the resulting measure. An aggregate closer to official inflation may be useful for one purpose while conveying less about the range of reported household beliefs.

This article examines that trade-off. We first document the gap between expectations and realised urban CPI, including differences before and after the acute COVID-19 period. We then examine calendar patterns, rounded responses and concentration in the survey's highest category. Finally, we compare five filtering rules and assess both their effects on the aggregate and the number and composition of responses retained. Bringing these comparisons together makes the policy meaning of a statistical adjustment more explicit.

The argument builds on research showing that experience and information shape inflation expectations. Salient price histories can influence beliefs through experience-based learning (Malmendier and Nagel 2016), while dispersed and slowly updated information can produce disagreement (Mankiw, Reis and Wolfers 2003; Easaw, Golinelli and Malgarini 2013). For India, Muduli, Nadhanael and Pattanaik (2022) examine adjustment for household biases, and Singh and Bandyopadhyay (2024) study disagreement. Our contribution is to consider the interpretation of alternative measures jointly with their coverage and observed composition. The question is how much can be inferred from an apparent improvement in agreement with CPI.

The evidence supports using filtered measures as transparent supplements to the raw survey. It does not identify an optimal filter or establish that the retained respondents are better informed.

The raw distribution remains necessary for understanding the spread of reported beliefs. This distinction also defines the scope of the policy discussion: the analysis concerns how expectations should be measured and read, rather than estimating their effects on inflation or the appropriate interest rate.

1 What the survey measures

The IESH is a repeated cross-section of roughly 6,000 urban households per round and currently covers 19 cities. The survey is aligned with the bimonthly policy cycle. It elicits perceptions of current inflation and expectations three months and one year ahead. The analysis in this paper uses microdata beginning in 2014 until 2025.

We benchmark expectations against urban CPI because the survey covers urban households. We measure statistical bias by the forecast error, defined as the realised urban CPI inflation at the relevant future date minus[1] the expectation reported in the survey. A negative error therefore denotes overprediction. For example, an expectation exceeding the subsequent inflation rate by five percentage points produces an error of minus five percentage points.

There are several reasons why household expectations might be biased. A household's consumption basket can differ from the basket underlying urban CPI. Respondents may also attend to different prices, recall different periods or report beliefs with different degrees of precision. Our analysis does not identify the source of the gap. We therefore use “bias” to refer to systematic forecast error against urban CPI, without suggesting that households’ own experiences of price changes are wrong.

The repeated-cross-section design limits how changes over time can be described. Differences across periods may reflect changes in beliefs, the households sampled, interview conditions or the economic environment. We can compare distributions and conditional averages, but cannot conclude that any particular household changed its expectations in line with the overall trend. When reporting any statistic, we will report the median instead of the mean. This is because data on expectations tends to contain outliers, which distort the mean.

2. The CPI Gap And Changes After The Pandemic

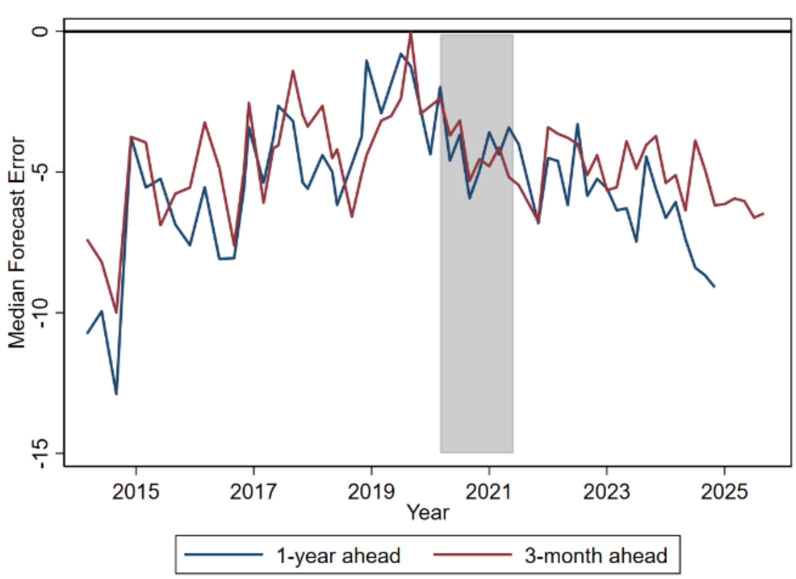
2.1 Persistent overprediction

Median forecast errors are negative at both horizons for most of the sample (Figure 1a). Ghosh and Singh (2026) show that expectations of Indian households respond significantly to major economic shocks. We find that the same is true of bias as well. During the COVID-19 period, errors temporarily became less negative as realised inflation rose towards already high expectations. The gap widened again afterwards. This illustrates why an apparent improvement in forecast accuracy needs to be carefully interpreted: the error can narrow because realised inflation changes, not because households become better at forecasting.

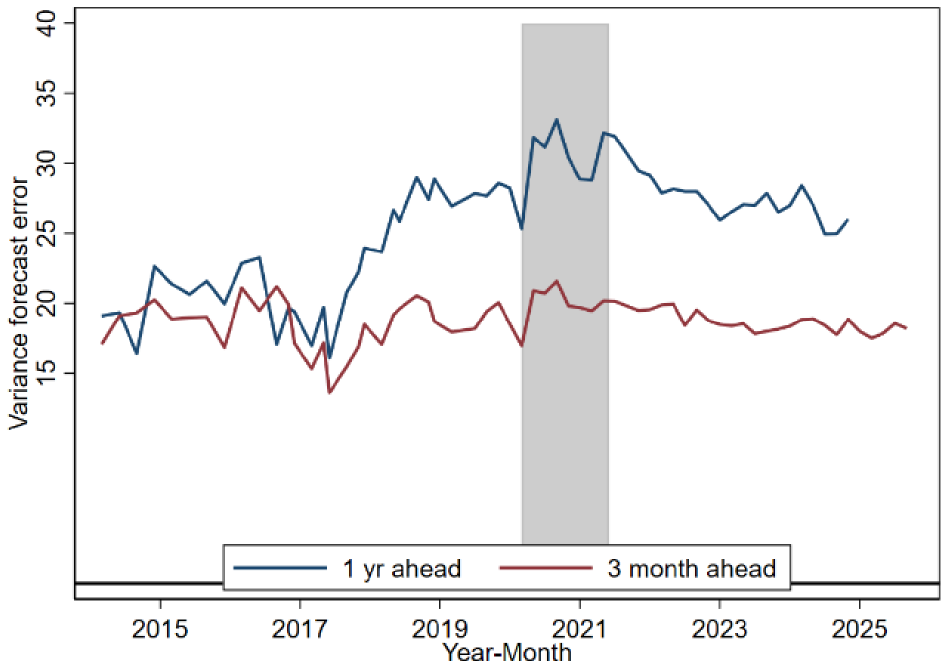
Figure 1b shows the variance of forecast errors across households. Dispersion changes sharply around the pandemic and remains elevated in parts of the subsequent sample. It is also of note that the variance of one year ahead expectations rises much more than the variance of three months ahead expectations, suggesting greater disagreement about inflation at longer horizons than in the near term. The variance consequently provides information on aggregate uncertainty that the median cannot supply. A stable median can coexist with a changing spread of responses, and policymakers should consider both when interpreting the survey.

Figure 1: Forecast errors: central tendency and dispersion

(a) Median forecast error



(b) Variance of forecast error



Notes: Forecast error is realised urban CPI inflation minus the reported expectation. Negative errors indicate overprediction. Blue: one year; red: three months. Errors are in percentage points; variances are in squared percentage points. The shaded interval marks the pandemic period. Source: Authors' calculations using RBI IESH microdata and urban CPI data, as applicable.

The persistent gap also argues against interpreting a high numerical expectation as sufficient evidence of a deterioration in monetary-policy credibility. Such an assessment requires a view about how the survey relates to the official inflation measure and how that relationship changes over time.

2.2 A larger difference at the longer horizon

To compare the pre and post pandemic environments, we exclude January 2020–December 2021 and contrast observations before January 2020 with those from January 2022 onwards. We estimate the difference in mean forecast errors first without controls and then allowing for differences across city, gender, age group and occupation. Table 1 places the two forecast horizons together so that the magnitudes can be compared directly.

Table 1: Post-pandemic differences in forecast errors (percentage points)

Variable	Three months No controls	Three months With controls	One year No controls	One year With controls
Post-period indicator	−0.456*** (0.017)	−0.512*** (0.017)	−0.881*** (0.021)	−1.082*** (0.022)
Constant	−4.980*** (0.011)	−5.107*** (0.049)	−5.004*** (0.014)	−5.092*** (0.061)
City Fixed Effects	No	Yes	No	Yes
Gender Fixed Effects	No	Yes	No	Yes
Age Fixed Effects	No	Yes	No	Yes
Occupation Fixed Effects	No	Yes	No	Yes
Observations	299,911	299,911	262,144	262,144
R-squared	0.003	0.037	0.007	0.043

Notes: The pre-period ends before January 2020; the post-period begins in January 2022. January 2020–December 2021 is excluded. Parentheses contain reported standard errors; stars are retained from the reported output. Inference does not resolve dependence at the survey-wave level. Controls are the fixed effects shown below the coefficients. Source: Authors' calculations using RBI IESH microdata and urban CPI data, as applicable.

Without controls, the post-period difference is −0.456 percentage points at three months and −0.881 percentage points at one year. Given the definition of the error, these estimates indicate greater average overprediction after the excluded pandemic period. Adding the controls makes the differences somewhat larger. With city and demographic controls, the differences are −0.512 and −1.082 percentage points. The larger difference at one year is consistent with greater difficulty judging the persistence of price pressures, although the comparison does not directly test that explanation.

2.3 Macroeconomic Information and Forecast Errors

We ask whether macroeconomic variables are associated with forecast errors. If households are using information efficiently, then public information available when households make their forecast should not systematically predict their later errors. Applying that interpretation requires correct alignment to interview dates and to the data releases then available. We assume the households' information set comprises of real GDP, urban unemployment, the policy repo rate, current urban CPI inflation and urban inflation lagged by three months.

Table 2 reports the macroeconomic regression entries for the two horizons. All five variables are associated with pre-period forecast errors in the reported specifications.

Table 2: Macroeconomic variables and forecast errors

Variable	Three months Pre-period	Three months Post-period	One year Pre-period	One year Post-period
Log GDP	8.809*** (0.597)	-10.188*** (0.993)	8.400*** (0.669)	-28.404*** (1.109)
Unemployment Rate	-0.223*** (0.040)	-1.143*** (0.268)	-1.972*** (0.044)	-7.245*** (0.299)
Interest Rate	0.964*** (0.092)	-2.133*** (0.208)	1.857*** (0.103)	-5.659*** (0.233)
Urban CPI Inflation	-0.119*** (0.016)	-0.803*** (0.053)	-0.866*** (0.018)	-1.734*** (0.059)
Lagged Urban Inflation	-0.368*** (0.016)	-0.378 (0.063)	-0.533*** (0.018)	1.628*** (0.070)

Notes: Parentheses contain reported standard errors; stars are reproduced as reported. GDP enters in logarithms; unemployment, the repo rate and inflation are rates. Urban inflation is lagged by three months in the last row. The post-period entries cannot be interpreted as total slopes without establishing whether they are interaction coefficients. Common survey-wave variation also limits conventional inference. These entries are descriptive. Source: Authors' calculations using RBI IESH microdata and urban CPI data, as applicable.

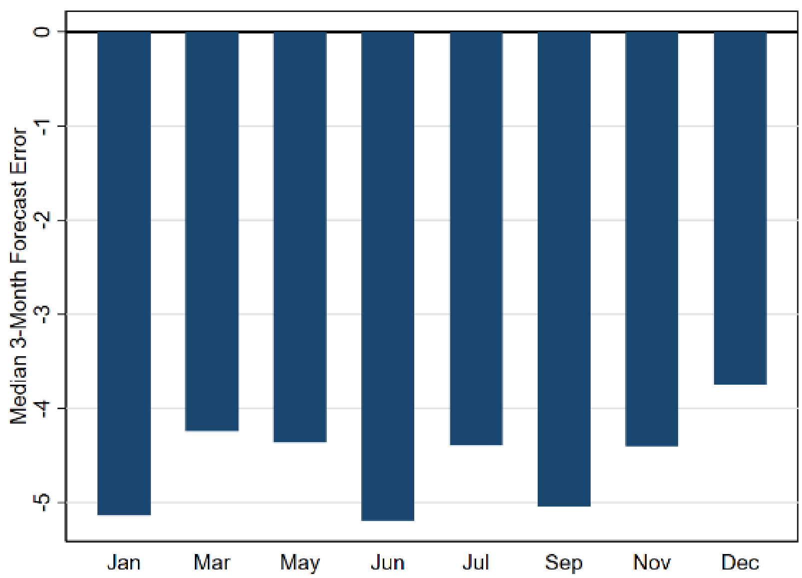
3. How Reporting Patterns Shape The Survey

3.1 Calendar differences explain little individual variation

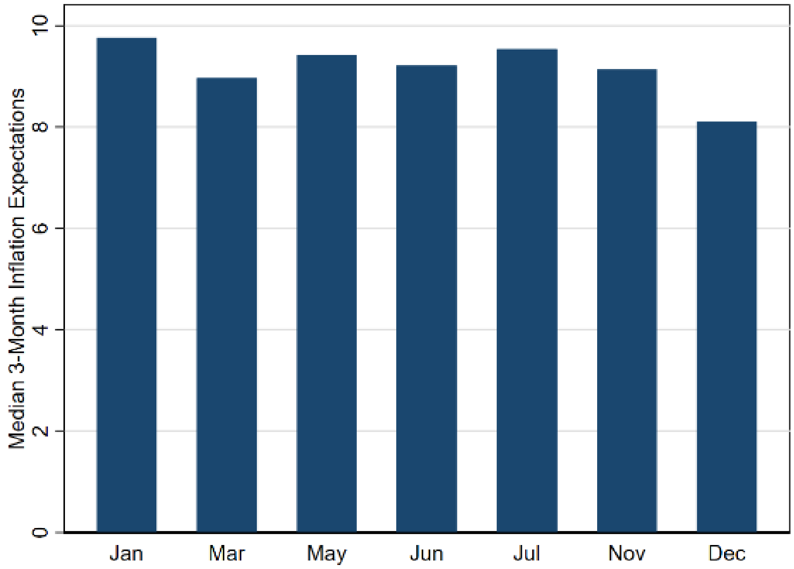
Calendar patterns can arise from seasonal price movements, the timing of survey rounds or differences in the respondents interviewed. Figure 2 pools responses by calendar month and shows expectations alongside the errors that subsequently materialise. At the three-month horizon, the median error is negative in every displayed month and varies by roughly one percentage point, from close to −4 to almost −5 percentage points. Formal tests reject equality across months, but calendar month explains only about 0.5% of individual error variation. At one year, median errors range from about −4.5 to nearly −6 percentage points, with the most negative values in June and September. This larger swing partly mirrors seasonality in realised urban CPI.

Figure 2: Calendar patterns in forecast errors and expectations

(a) Three-month median forecast error



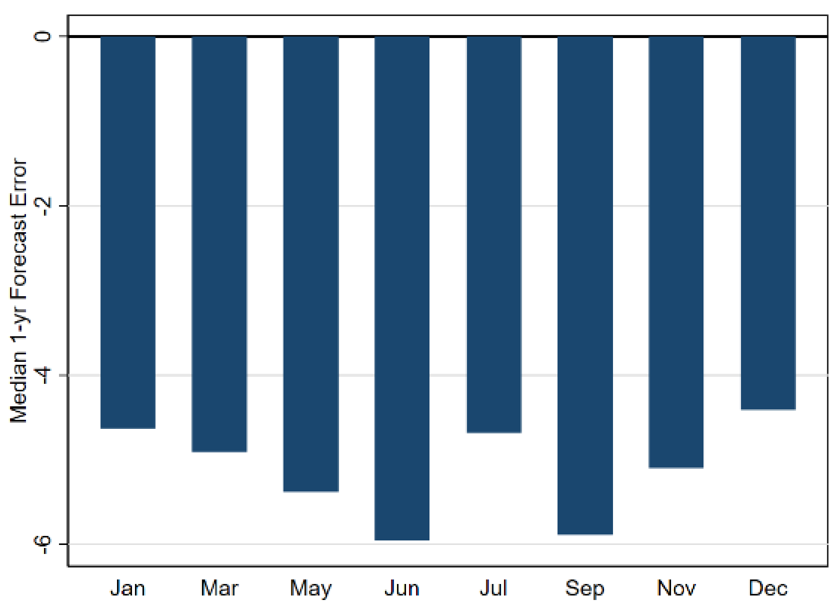
(b) Three-month median expectation



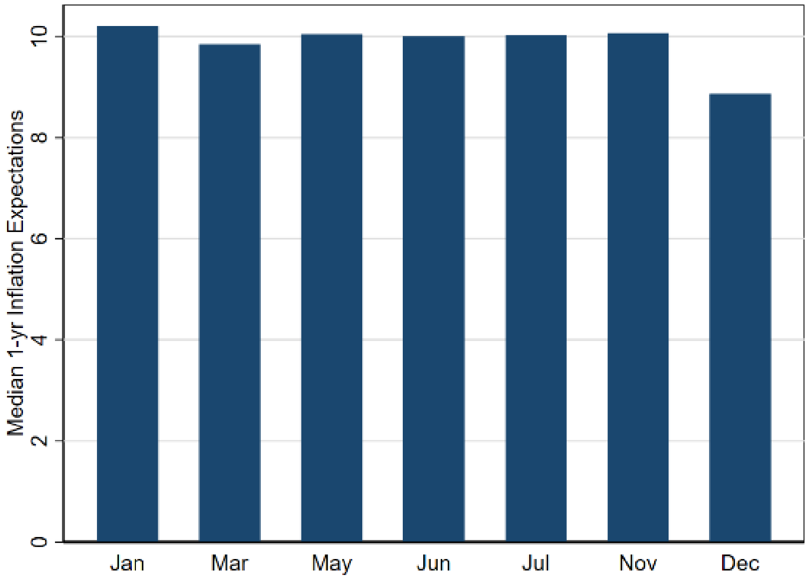
Notes: Series and units are defined in the notes following the final panel. Source: Authors' calculations using RBI IESH microdata and urban CPI data, as applicable.

Figure 2: Calendar patterns in forecast errors and expectations (continued)

(c) One-year median forecast error



(d) One-year median expectation



Notes: Observations are pooled by calendar month. Forecast errors are in percentage points; expectations are in per cent. The panels display the months available in each series. Source: Authors' calculations using RBI IESH microdata and urban CPI data, as applicable.

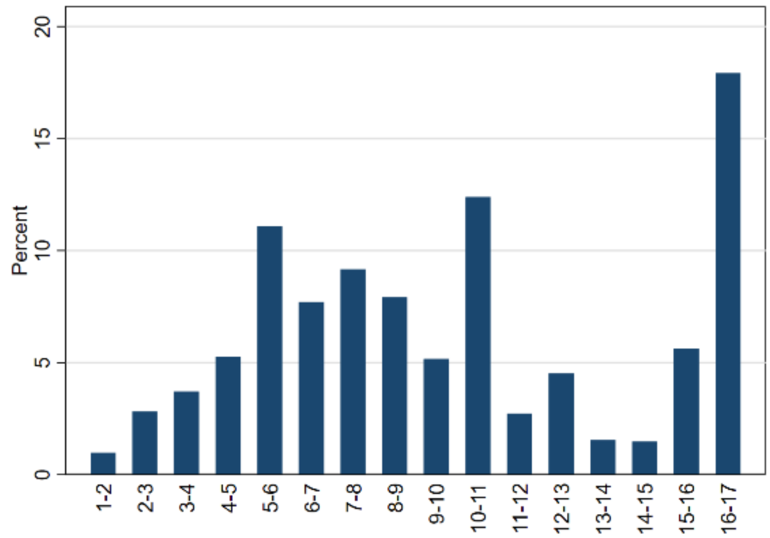
The size of the effect matters for policy interpretation. A detectable monthly pattern can be relevant when comparing releases, but it accounts for little of the persistent gap. Seasonal adjustment would not resolve the broader problem of reading reported expectations against official CPI.

3.2 Rounding and the open-ended upper category

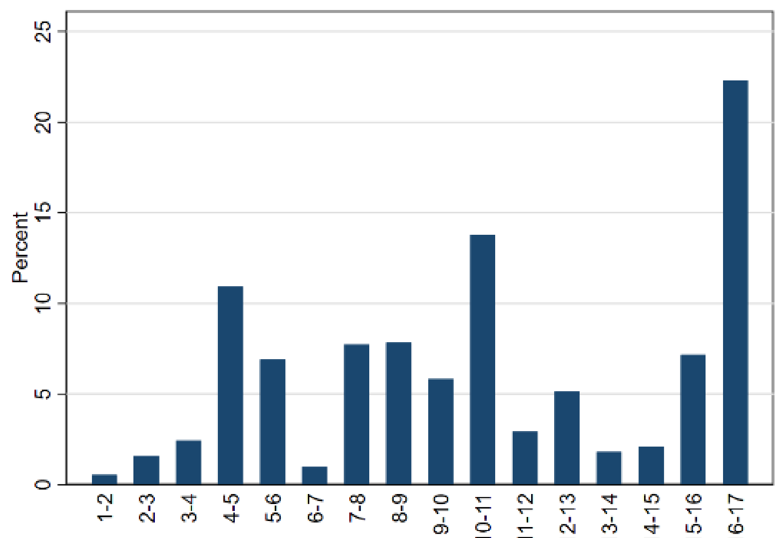
The response distributions contain substantial concentrations at particular values (Figure 3). At the three-month horizon, the bin around 10% contains about 12% of responses, while the open-ended category above 16% contains close to 18%. At one year, more than 22% of responses fall in the upper category. These concentrations affect both the shape of the distribution and the consequences of filtering it.

Figure 3: Distributions of reported inflation expectations

(a) Three-month horizon



(b) One-year horizon



Notes: Vertical axes show the percentage of responses; horizontal labels denote inflation response categories in per cent. The rightmost category represents reports above 16%, rather than a measured interval with a known upper endpoint. Source: Authors' calculations using RBI IESH microdata and urban CPI data, as applicable.

Interior rounding and concentration in the top category have different meanings. A round number can express a belief that is genuinely imprecise or simply easier to communicate. The upper category is also a feature of the questionnaire: it records that a response exceeds a threshold, without identifying the precise value. Treating both patterns as a single form of reporting error would obscure the distinction between coarseness in a respondent's answer and a boundary imposed by the survey instrument.

The share of responses at multiples of five falls from roughly 55–58% before COVID-19 to about 43–44% afterwards, even as concentration in the highest category increases. The change reinforces the need to examine the distribution over time. Less interior rounding can coexist with more observations recorded above the upper threshold, and so doesn’t necessarily imply that households are becoming more precise over time.

3.3 Broad demographic groups explain little variation

Demographic differences are statistically detectable in a large sample, but their explanatory contribution is small. Age is the strongest demographic predictor examined, yet it accounts for no more than 0.46% of variation in expectations and 0.40% of variation in forecast errors. Gender and occupation explain still less. City differences are detectable, although comparable effect sizes are not available in this exercise.

These results provide limited support for differentiating substantive policy messages solely by broad demographic labels. They leave room for differences in information access, language and communication channels that these categories do not adequately measure. They also matter for evaluating filters: a retained sample can resemble the full sample on age or gender while differing in numeracy, price exposure or other characteristics relevant to expectations.

4. Methods to Reduce Bias in Expectations

The raw IESH distribution serves two policy purposes that need not coincide. It records the beliefs and perceived cost-of-living pressures of surveyed households, including extreme and imprecise reports. It can also be used for a numerical signal about future urban CPI. Filtering may improve the second use while weakening the first by changing whose responses enter the aggregate. The goal is therefore not to cleanse the survey of supposedly bad respondents, but to construct clearly labelled supplementary measures and make their trade-offs visible. We compare five screening rules: hard bounds, proximity to measured current inflation, 25% tail trimming, 10% tail trimming, and exclusion of heaped or extreme-bin responses.

4.1 Hard Bounds

The hard-bounds rule retains reported expectations between 0% and 10%. This interval roughly encompasses realised urban CPI over the sample. One further benefit of it is that it is simple to communicate. It is nevertheless a researcher-chosen, ex post screen. Values outside the band may be unusual as CPI forecasts but can still represent genuine perceived inflation. Removing them mechanically pulls the aggregate towards the range used to define the rule. Household *i* enters the hard-bound sample $S_{\{HB\}}$ if its expectations lie between 0% and 10% for the relevant survey wave and forecast horizon:

$$S_{\{HB\}} = \{i: 0 \leq E_i \leq 10\}$$

The upper and lower bounds can be changed to accommodate periods of very high inflation or deflation. However, the current range works well for India given its past history of realised inflation.

4.2 Current-Inflation Proximity

The second method selects responses according to the distance between a respondent's perception of current inflation and measured urban CPI. We study bands from 1 to 5 percentage points. This is a transparent way to extract a CPI-aligned signal, but closeness to official CPI is not direct evidence of attention or forecasting skill. It may also exclude households whose personal consumption baskets experienced different price changes. Specifically, $S_{\{Prox\}}$ will include household *i* if and only if its perception of current inflation lies within a band around it:

$$S_{\{Prox\}} = \{i: Perception(\pi_{\{t\}}) \in \left[\pi_{\{t\}} - x, \pi_{\{t\}} + x\right]\}$$

x varies from 1 to 5 percentage points, corresponding to Method 2-A to 2-E. The retained set contains respondents whose reported interval for current inflation contains, or lies within the chosen tolerance of, the latest official urban CPI value available for the survey wave. It is selects for people who are more aware about current economic conditions.

4.3 IQR Trimming

The third method retains observations between the 25th and 75th percentiles within each survey wave. It is data-driven and removes both tails without imposing an economic range. Since the principal charts use medians, however, little movement is expected by construction: symmetric tail removal normally preserves the median. This method is more meaningfully assessed with the mean, dispersion or tail-sensitive loss measures.

$$S_{\{T25\}} = \{i: Q_{\{0.25,t\}} \leq E_{\{it\}} \leq Q_{\{0.75,t\}}\}$$

Here, the retained sample contains observations between the wave-specific 25th and 75th percentiles for the relevant horizon.

4.4 Percentile Trimming

The fourth method drops the top and bottom 10% within each wave, subject to ties at the cut-offs. It is a less aggressive version of IQR trimming. Like the 25% rule, it should be evaluated using the mean, dispersion or tail-sensitive errors because symmetric trimming will usually leave the median almost unchanged.

$$S_{\{T10\}} = \{i: Q_{\{0.10,t\}} \leq E_{\{it\}} \leq Q_{\{0.90,t\}}\}$$

Here, the retained sample contains observations between the wave-specific 10th and 90th percentiles for the relevant horizon. This allows for a larger share of households to be retained in the sample.

4.5 Heaping/Extreme Bin Filter

The fifth approach excludes interior round-number responses that are multiples of five and the open-ended extreme category. People reporting expectations in either of these categories may have more uncertainty around their estimates. This method directly tests sensitivity to response coarsening and top-coding. The remaining sample should, in principle, contain people with more certain or precise beliefs.

4.6 Comparing the Filters

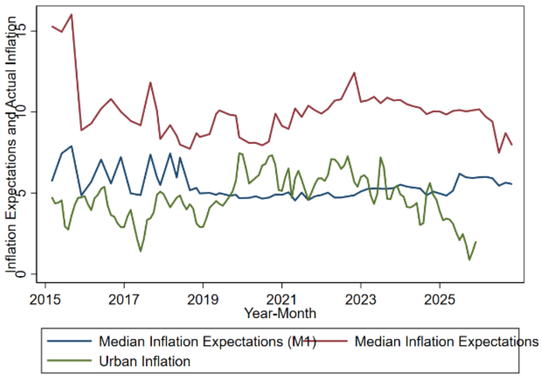


Figure 9: Hard Bounds

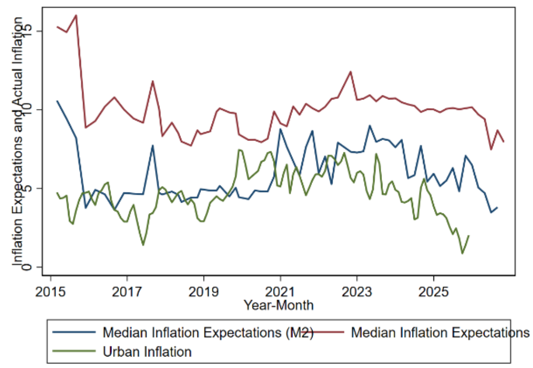


Figure 10: Proximity to Actual Inflation

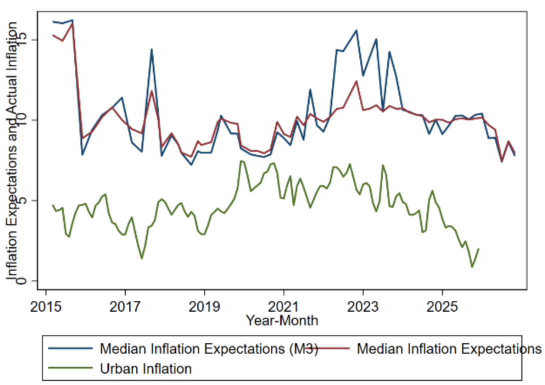


Figure 11: IQR Trimming

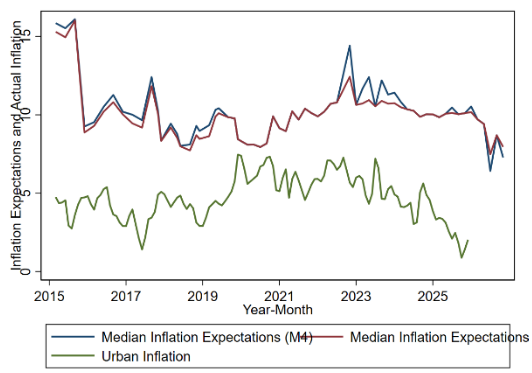


Figure 11: Percentile Trimming

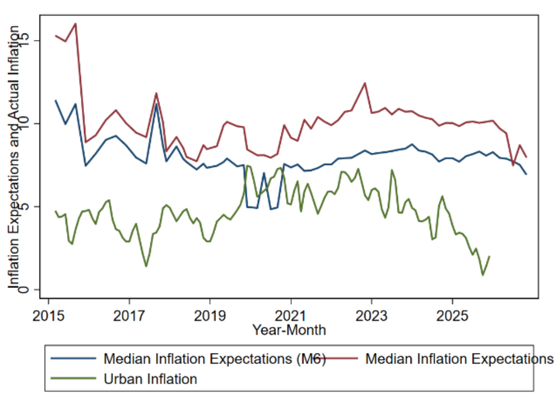


Figure 12: Heaping/Extreme Bin Filter

Figures 9-13 compare the median of the raw series, the filtered median for each method, and realised urban CPI. They show how strongly the headline measure depends on alternative screening choices. ~~Because the charts do not provide a common numerical accuracy table, the comparison below emphasises what each rule changes and the estimand it creates rather than declaring a single winner~~.

Hard bounds visibly compress the series towards the historical range of realised inflation. That makes the rule transparent and easy to reproduce, but it also explains why measured error falls. The exercise is best interpreted as a sensitivity check on extreme responses rather than evidence that the retained households are well anchored.

The current-inflation proximity filter produces a median closer to realised CPI in the displayed sample. The rule conditions on agreement with official inflation, so part of that improvement may be mechanical if current perceptions and future expectations are correlated. The resulting series should be described as the expectations of a CPI-aligned subsample, not of respondents whose attention or knowledge has been directly observed.

The limited movement under symmetric trimming is largely mechanical when the outcome is the median: removing equal shares from the two tails normally leaves the middle observation close to its original value. These filters may still matter for the mean, dispersion and tail-sensitive loss measures, but the median charts are not an informative test of their general usefulness.

The heaping and extreme-bin rule also changes the series, but it combines two conceptually different exclusions: interior round-number responses and the survey's open-ended upper category. It reduces bias more than the two trimming methods, but less than the first two methods.

Figure 14 makes this point explicit. It indicates that the filters change measured forecast errors, and the first two filters that directly restrict the range or condition on current-CPI proximity reduce bias the most. Since method 1 conditions on a fixed range, it does not allow any reposes greater than realized inflation, even though those might be the true household expectation. Our preferred measure is therefore method 2, which selects a sample of informed households, but doesn’t impose restrictions on what their inflation expectations could be.

Table 3 compares the full sample with the responses retained by the five methods. Broad demographic shares generally remain recognisable, but the rules change their relative weights. The female share is 46.62% in the full sample, compared with 45.91% under hard bounds and 48.68% under the baseline proximity filter. The young-age share rises from 34.04% to 36.29% and 36.70%, respectively. The corresponding old-age shares fall from 13.64% to 12.66% and 12.06%.

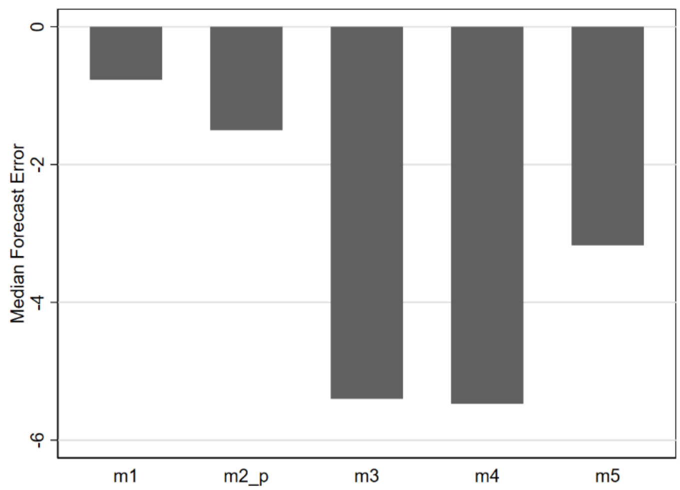


Figure 13: Median Forecast Error across methods

The proximity filter also retains a higher share of homemakers, at 32.69% compared with 31.07% in the full sample, and a lower share of self-employed respondents, at 17.59% compared with 19.04%. Among the city categories, hard bounds reduce the Tier 1 share from 67.42% to 65.95%, while increasing the Tier 3 share from 6.08% to 7.08%. The table reports all categories across all five filters, allowing these changes to be assessed together rather than relying on the retained sample size alone.

The 10% trimming rule retains the largest share of responses and has several demographic proportions close to those in the full sample. Its young-age share, for example, is 34.31%, compared with 34.04% in the full sample. This limited compositional change accompanies a median that also moves little. By contrast, the much more restrictive proximity rule retains less than a quarter of observations, even though many of its broad demographic shares are not far from the full-sample values.

These comparisons illustrate why sample size and observed composition answer different questions. A filter can remove many responses while leaving some demographic proportions broadly similar. That similarity does not recover the excluded beliefs or demonstrate that selection is unrelated to unobserved price experiences and information. Conversely, a modest change in a demographic share should not automatically disqualify a measure intended for a narrower forecasting purpose. What matters is whether its stated purpose, selection rule and interpretation remain consistent.

Table 3: Demographic composition under five filters, March 2014–December 2025 (%)

Characteristic	Full	M1	M2	M3	M4	M5
Female	46.62	45.91	48.68	47.05	46.76	46.24
Male	53.36	54.09	51.32	52.95	53.24	53.73
Young Age	34.04	36.29	36.70	35.97	34.31	35.36
Middle Age	52.30	51.04	51.23	51.41	52.30	51.42
Old Age	13.64	12.66	12.06	12.62	13.39	13.18
Homemaker	31.07	30.60	32.69	31.29	31.17	31.16
Retired	6.28	5.75	5.28	5.75	6.03	6.18
Daily Workers	10.02	9.92	9.86	9.83	9.98	10.02
Self-Employed	19.04	18.75	17.59	18.56	18.99	18.77
Other Employees	19.49	19.58	19.79	19.81	19.68	19.39
Financial Sector	4.32	4.70	4.26	4.36	4.19	4.52
Tier 1 City	67.42	65.95	67.01	67.07	67.97	65.86
Tier 2 City	26.48	26.97	26.34	26.72	26.12	27.33
Tier 3 City	6.08	7.08	6.65	6.21	5.91	6.78

Notes: Full denotes the unfiltered analysis sample. M1: hard bounds; M2: ±1-point proximity; M3: 25% tail trimming; M4: 10% tail trimming; M5: heaping and extreme-bin exclusions. Shares are reported for each characteristic; the listed categories do not always sum to 100. Denominator and category-coverage differences limit interpretation of those totals. The table describes observed composition, not population representativeness. Source: Authors' calculations using RBI IESH microdata and urban CPI data, as applicable.

5.7 Choosing the Proximity Band: Fit versus Coverage

Given that method 2 is our preferred measure, the natural next question is what the optimal band on households' current perception of inflation should be. Widening the proximity band allows for a larger sample, but trades off having more informed households. A wider band is not automatically more representative, because selection may remain correlated with unobserved numeracy, information access or the similarity of a household's consumption basket to urban CPI.

Method	Respondents Retained	Percent of Total Sample
M2a (±1 pp)	87,719	23.4%
M2b (±2 pp)	139,374	37.1%
M2c (±3 pp)	177,401	47.3%
M2d (±4 pp)	212,000	56.5%
M2e (±5 pp)	240,994	64.2%

Table 4: Coverage of Alternative Current-Inflation Proximity Bands

Table 4 reports five variants of the tolerance around measured current inflation, from 1 percentage point (M2a) to 5 percentage points (M2e). Retention rises monotonically from 23.4% to 64.2%. M2c retains 177,401 observations (47.3%), while M2d retains 212,000 (56.5%).

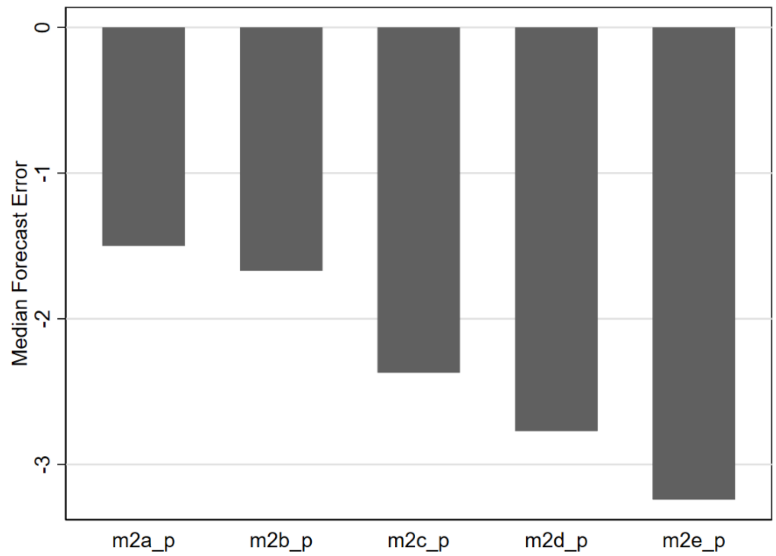


Figure 14: Median Forecast Error across different bands

Figure 15 makes the coverage trade-off visible: widening the tolerance admits more observations and moves the filtered median back towards the unfiltered series. The pattern is expected because the rule is being relaxed. What matters for policy is the incremental change in an accuracy measure for each additional share of the sample retained.

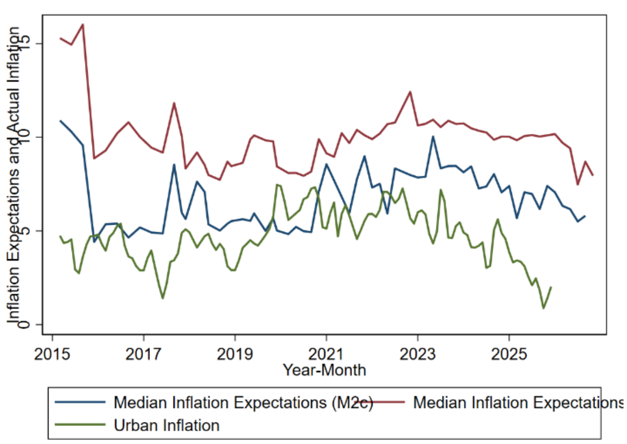


Figure 16: Method 2c

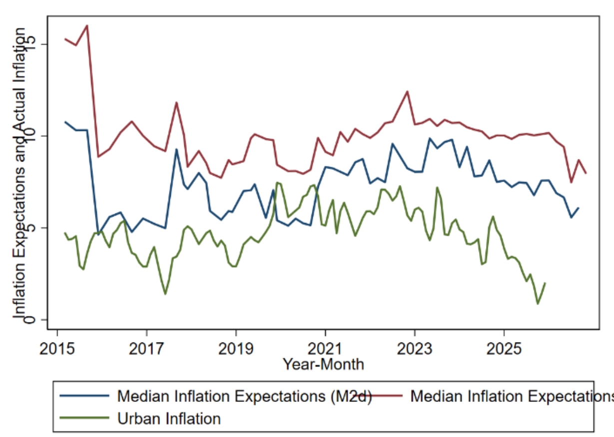


Figure 17: Method 2d

The intermediate bands appear to offer a practical coverage-fit compromise in the sample, and so M2c and M2d are our preferred filter. M2c applies the tighter screen and retains 47.3% of observations; M2d retains 56.5% and gives more scope for subgroup analysis. Figures 16 and 17 show the movements of these filtered series along with realized inflation. The limitation of these methods is equally important: they exclude many households and may favour respondents whose consumption baskets happen to resemble official urban CPI.

Transparency about the filtering method is essential. A policymaker reading a filtered series needs to know whether it changed because respondents revised their expectations or because the rule admitted a different set of responses. Reporting retention alongside the filtered aggregate would make that distinction easier to investigate. Comparisons of demographic composition provide an additional, though incomplete, description of the selection involved.

4.4 Widening the proximity band

Relaxing the proximity rule increases coverage substantially (Table 5). The ±1-percentage-point band retains 23.4% of responses, the ±2 band 37.1%, the ±3 band 47.3%, the ±4 band 56.5%, and the ±5 band 64.2%. In counts, these correspond to 87,719, 139,374, 177,401, 212,000 and 240,994 observations. A wider band includes more responses whose current perceptions differ from official CPI, moving the definition of the retained group closer to the full sample.

Table 5: Sample retention as the proximity band widens

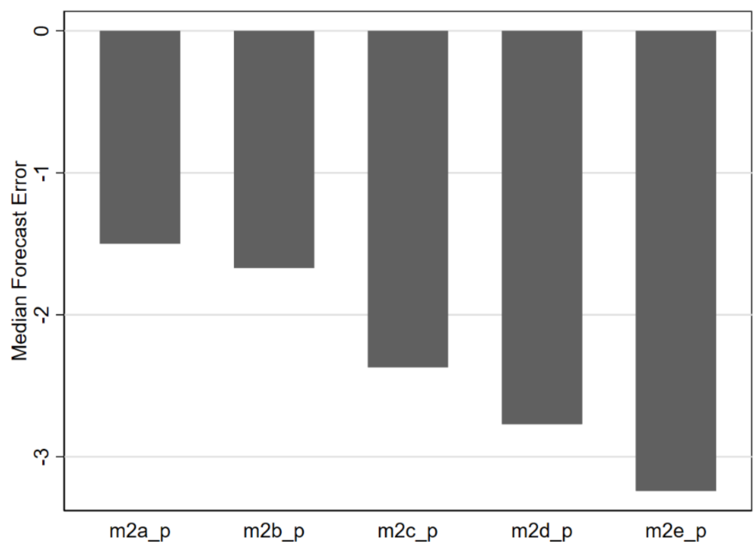
Current-inflation proximity band	Observations retained	Retention (%)
M2a (±1 percentage points)	87,719	23.4
M2b (±2 percentage points)	139,374	37.1
M2c (±3 percentage points)	177,401	47.3
M2d (±4 percentage points)	212,000	56.5
M2e (±5 percentage points)	240,994	64.2

Notes: M2A–M2E widen the tolerance around measured current urban inflation from ±1 to ±5 percentage points. Counts and percentages have the same interpretation as in Table 3a. Source: Authors' calculations using RBI IESH microdata and urban CPI data, as applicable.

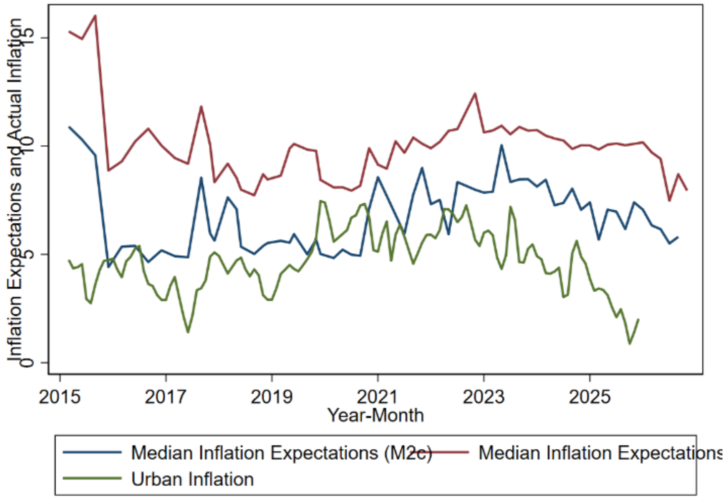
Figure 5 shows the associated changes in median forecast errors and the series under the ±3 and ±4 bands. Widening the tolerance moves the filtered median towards the unfiltered series. This is the expected consequence of relaxing the screen, rather than independent evidence that one threshold is correct. The intermediate bands offer candidate compromises between agreement with CPI and coverage in this sample. Choosing between them for regular use would require an explicit objective and validation on later observations.

Figure 5: Alternative proximity bands: errors and selected series

(a) Median forecast errors across bands from ±1 to ±5 percentage points



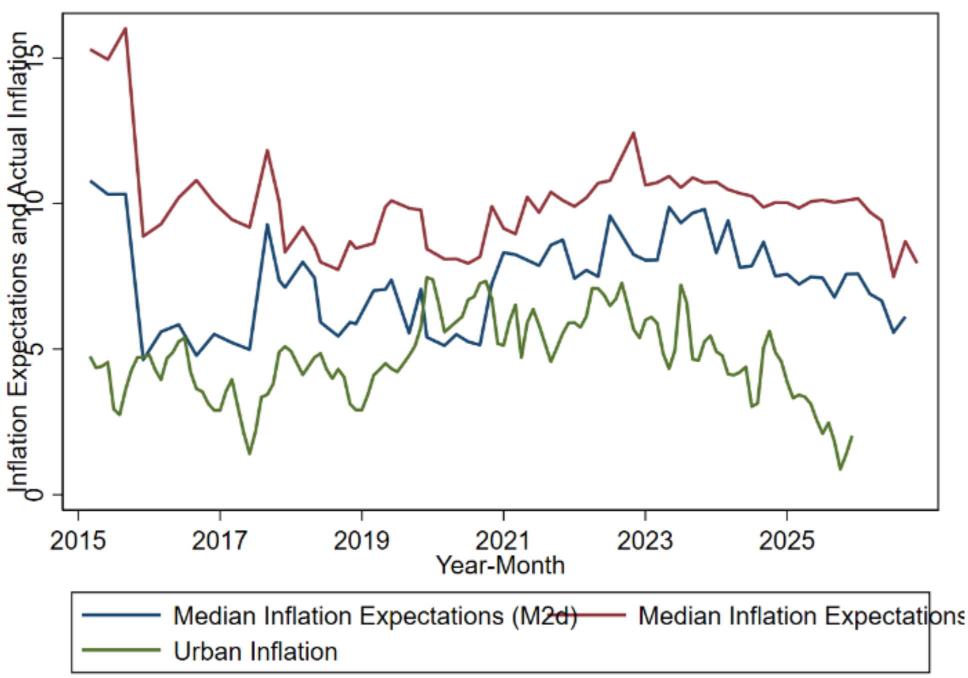
(b) Proximity band of ±3 percentage points (M2C)



Notes: Series and units are defined in the notes following the final panel. Source: Authors' calculations using RBI IESH microdata and urban CPI data, as applicable.

Figure 5: Alternative proximity bands: errors and selected series (continued)

(c) Proximity band of ±4 percentage points (M2D)



Notes: Panel (a) reports errors in percentage points; m2a_p–m2e_p denote the five bands. In panels (b) and (c), blue is the filtered median, red the unfiltered median and green urban CPI inflation (all in per cent). Method codes are equivalent irrespective of letter case. Source: Authors' calculations using RBI IESH microdata and urban CPI data, as applicable.

Table 6 examines demographic composition across all five proximity bands. The female share is 48.68% under the narrowest rule and 47.64% under the widest, compared with 46.62% in the full sample. The young-age share remains above the full-sample share under every band: it is 37.16% under ±3 and 36.92% under ±4, compared with 34.04% in the full sample. Corresponding old-age shares are 12.28% and 12.30%, compared with 13.64%.

The occupational and city shares also help describe what widening the rule changes. Self-employed respondents account for 17.67% under ±3 and 17.96% under ±4, below their 19.04% full-sample share. The Tier 1 shares are 66.73% and 66.77%, compared with 67.42%. These similarities and differences are relevant when interpreting a supplementary series. They are insufficient to identify the retained households as more attentive or to conclude that a wider band resolves all selection concerns.

The ±3 and ±4 measures thus have a specific potential role: they show whether movements among responses relatively close to current official CPI corroborate movements in the raw survey. The tighter measure retains 47.3% of observations; the wider one retains 56.5% and leaves more observations for subgroup comparisons. Neither is a corrected estimate of the beliefs of all surveyed households. Publishing both without their definitions and retention rates would conceal a substantial part of what the filtering exercise changes.

5. Implications For Monetary Policy And Survey Design

The first implication concerns how the level of household expectations is read. A reported expectation should be interpreted alongside its historical relationship with the inflation benchmark and the distribution of responses. The persistent gap in this sample means that the numerical level cannot be treated as an unbiased forecast of urban CPI. Equally, overprediction alone cannot establish that households have lost confidence in monetary policy. Such an interpretation would require evidence separating changes in that relationship from changes in price experiences, information and reporting.

Table 6: Demographic composition across proximity bands, March 2014–December 2025 (%)

Characteristic	Full	M2A	M2B	M2C	M2D	M2E
Female	46.62	48.68	48.56	48.07	47.72	47.64
Male	53.36	51.32	51.44	51.93	52.28	52.36
Young Age	34.04	36.70	37.11	37.16	36.92	36.75
Middle Age	52.30	51.23	50.63	50.56	50.78	50.93
Old Age	13.64	12.06	12.25	12.28	12.30	12.32
Homemaker	31.07	32.69	32.26	31.96	31.73	31.74
Retired	6.28	5.28	5.66	5.69	5.72	5.69
Daily Workers	10.02	9.86	9.74	9.64	9.67	9.91
Self-Employed	19.04	17.59	17.46	17.67	17.96	18.05
Other Employees	19.49	19.79	19.33	19.54	19.64	19.60
Financial Sector	4.32	4.26	4.39	4.51	4.48	4.42
Tier 1 City	67.42	67.01	66.84	66.73	66.77	67.00
Tier 2 City	26.48	26.34	26.32	26.48	26.55	26.45
Tier 3 City	6.08	6.65	6.84	6.79	6.68	6.55

Notes: Full denotes the unfiltered analysis sample. M2A, M2B, M2C, M2D and M2E correspond to tolerances of ±1, ±2, ±3, ±4 and ±5 percentage points, respectively. Shares are reported for each characteristic; the listed categories do not always sum to 100. Denominator and category-coverage differences limit interpretation of those totals. The table describes observed composition, not population representativeness. Source: Authors' calculations using RBI IESH microdata and urban CPI data, as applicable.

A useful presentation would place the unfiltered median beside selected percentiles, current official inflation and historical forecast errors at the relevant horizons. The percentiles would show whether a movement in the median is accompanied by a wider shift in responses or by changing dispersion. Historical errors would provide context for the size of the gap without suggesting that an average adjustment is valid in every period. The post-pandemic differences show why the stability of such an adjustment cannot simply be assumed.

Filtered measures could then be presented as supplementary indicators with explicit labels. A hard-bounds series would describe sensitivity to the selected numerical range. Proximity measures would describe expectations among respondents whose current perceptions lie within stated distances of urban CPI. A reporting framework could include the ±3 and ±4 variants alongside the raw median, but should also show retention for each wave. Agreement across these measures would provide descriptive corroboration; divergence would identify a reason to inspect the distribution and the composition of the retained sample.

Operational use requires more validation than the present comparisons supply. The filter and its threshold should be specified before evaluating future performance. Comparisons should use consistent horizons and missing-data rules, and report signed errors alongside measures such as mean absolute or root mean squared error that do not allow positive and negative errors to offset each other. Retention and demographic composition should accompany those measures. These requirements follow from the different purposes of the indicators; a rule suited to predicting CPI need not be best for describing household beliefs.

The reporting patterns also motivate changes that can be tested in the questionnaire. A pilot could ask for a central estimate together with a plausible range, or use probability questions to allow respondents to express uncertainty more directly. Retaining the existing question alongside an alternative for several rounds would help distinguish a change in elicitation from a change in expectations. The open-ended upper category deserves separate attention because it limits what can be learned about the magnitude and dispersion of high reports.

Survey administration can help interpret changes over time as well. Information on contact attempts, interview mode and duration, item non-response and refusals would provide a basis for examining how collection conditions relate to response patterns. Such information would be particularly valuable when comparing periods with different interview conditions. It would not remove all measurement problems, but would make some of the competing explanations for changes in the distribution observable.

Communication is another area for testing. The larger increase in overprediction at the one-year horizon makes the expected persistence of inflation pressures a relevant subject for communication. Explaining why particular price pressures may be temporary or persistent, the risks around the outlook and the timing of policy effects could address questions that a current inflation release alone cannot answer. The regressions do not establish that such messages would improve forecasts, so their effects should be evaluated rather than assumed.

Finally, small explanatory contributions from broad demographic categories favour attention to specific information barriers and communication channels. Different languages or delivery formats may improve access without presuming that an age or occupational group holds a single shared view. The demographic tables can inform monitoring of who remains in a filtered sample, while the raw distribution continues to record the diversity that any aggregate necessarily compresses.

6. Conclusion

Household inflation expectations are useful for policy when their interpretation matches the question being asked. The raw survey describes reported beliefs among surveyed urban households. Forecast errors assess those beliefs against subsequent urban CPI. Filtering constructs aggregates for selected responses and changes both their numerical properties and their coverage. The distinction is consequential in this sample: rules that move expectations closer to CPI can exclude a large share of observations, while similarity on broad demographics does not establish the absence of selection.

The evidence supports transparent supplementary measures, fuller presentation of the raw distribution and testing of survey and communication changes. Retaining the connection between each indicator, its construction and the responses it represents would make the IESH more informative for monetary-policy discussion without asking a single headline number to serve every purpose.

Endnotes

1. Since 3 month CPI is not published, so for 3 month ahead forecast error, we treat the actual Y-o-Y CPI for that month as the realized CPI.

References

Easaw, Joshy Z, Roberto Golinelli and Marco Malgarini (2013): "What Determines Households' Inflation Expectations? Theory and Evidence from a Household Survey," European Economic Review, Vol 61, pp 1–13.

Ghosh, Pushpendu and Aditi Singh (2026): "Anchored for Whom? Household Inflation Expectations after a Decade of Inflation Targeting" Available at SSRN: https://ssrn.com/abstract=7229058 or http://dx.doi.org/10.2139/ssrn.7229058

Malmendier, Ulrike and Stefan Nagel (2016): "Learning from Inflation Experiences," Quarterly Journal of Economics, Vol 131, No 1, pp 53–87.

Mankiw, N Gregory, Ricardo Reis and Justin Wolfers (2003): "Disagreement about Inflation Expectations," NBER Macroeconomics Annual, Vol 18, pp 209–248.

Muduli, Silu, G V Nadhanael and Sitikantha Pattanaik (2022): "Assessing Inflation Expectations Adjusting for Households' Biases," RBI Bulletin, Vol 76, No 12, pp 97–107.

Reserve Bank of India (various years): Inflation Expectations Survey of Households, Database on Indian Economy, Mumbai.

Singh, Gaurav Kumar and Tathagata Bandyopadhyay (2024): "Determinants of Disagreement: Learning from Inflation Expectations Survey of Households," Journal of Forecasting, Vol 43, No 2, pp 326–343.