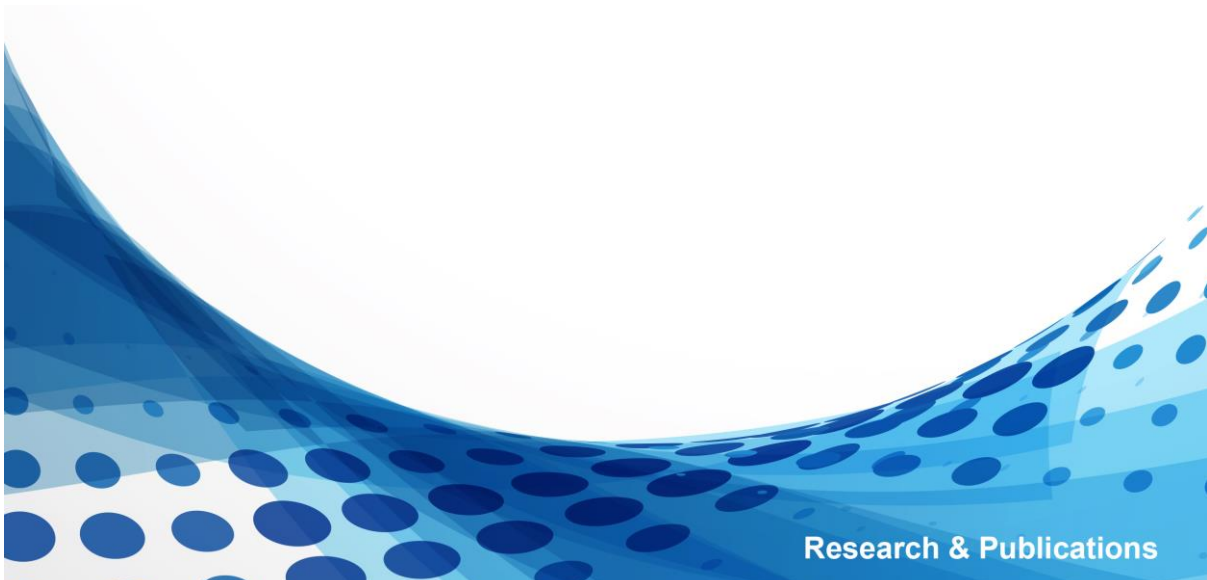INDIAN INSTITUTE OF MANAGEMENT AHMEDABAD

IIMA
Working Paper

How informative are quantified survey data? Evidence from RBI household inflation expectations survey

Gaurav Kumar Singh

Research & Publications

How informative are quantified survey data? Evidence from RBI household inflation expectations survey

Gaurav Kumar Singh

December 2020

INDIAN INSTITUTE OF MANAGEMENT AHMEDABAD

How informative are quantified survey data? Evidence from RBI household inflation expectations survey

Gaurav Kumar Singh*

ABSTRACT

Quantification[2] of the ordinal survey responses on inflation expectations is an important preliminary step for undertaking further macroeconomic analysis of the data. In this paper, we briefly describe the standard quantification methods along with the underlying assumptions. We also propose two new methods for quantification. We then apply these methods to quantify the IESH[3] data collected by the Reserve Bank of India (RBI). An interesting fact that emerges from this exercise is, simpler quantification methods are found to perform better than more complex methods for IESH data. Also, the methods with time varying weights or time varying thresholds, as the case may be, work significantly better.

KEYWORDS

Inflation expectations survey, Ordinal responses, quantified inflation expectations

1. Introduction

In the conduct of monetary policy, inflation expectations play a key role (Galí 2008, Sims 2009). Expectations drive people's economic behaviour and influence their day to day decisions related to consumption, saving, investment, and wage negotiations. These economic decisions collectively impact the real economic activity and thus actual inflation. Inflation expectation is a key driving force to the dynamics of inflation (Mishkin 2007).

As a consequence, for a central bank, it is crucial to monitor inflation expectations, making sure that they remain well anchored and consistent with policy objectives.

The standard macroeconomic models assume rational expectations (Muth 1961) of economic agents. This assumption enables an economic agent to produce a unique inflation expectation for any given horizon. However, for modeling inflation expectations in the real world, this assumption is found to suffer from several counts. The assumption lacks strong empirical support and often is found to be unrealistic (Manski 2018). Thus, for measuring inflation expectations, the central banks must rely on various direct and indirect approaches. In direct approach, some of the commonly used methods are, surveys of professionals such as economists and market participants, and surveys of firms and households. Especially, measuring inflation expectations of households over time is extremely important for the central banks since in aggregate they are important drivers of economic activity, and thus helps in assessing the impact of inflation expectations on future realized inflation (Bernanke 2007).

---

*Indian Institute of Management, Ahmedabad (IIMA)
[2]See Nardo (2003); Pesaran and Weale (2005)
[3]Inflation Expectation Survey of Households

An important survey-based source of information on urban Indian households'inflation expectations is the quarterly inflation expectation Survey of households (IESH) conducted by Reserve Bank of India since 2005 3rd quarter. It collects data from a nationally representative sample of 4000 households of urban Indians in each quarter. Among other questions, the survey solicits quarter-ahead, and year-ahead directional changes to "prices in general", i.e, whether it is going to "increase more than the current rate (a)", "increase similar to the current rate (b)", "increase less than the current rate (c)", "no change (d)" or "decline (e)", where "a", "b", "c", "d" and "e" represent the respective categories. This structure has been widely used across various countries for collecting data on inflation expectations indicating perhaps that qualitative questions of this sort lead to more reliable responses than asking more precise questions. A survey that captures inflation expectation through qualitative response is often preferable over their quantitative counterpart primarily due to their robustness to measurement errors and extreme responses and higher response rate.[4] Also, it may be noted that more precise questions yield more precise responses but not necessarily more accurate responses (Pesaran and Weale 2005). Every such survey is expected to produce responses for a specific price index, in this case, it is consumer price index (CPI) for industrial workers (CPI-IW).

As stated above, the information gathered is *ordinal* in nature, and provides a directional change for "general price level". A key research issue in the use of such data is, thus, linking the qualitative data with the latent inflation expectations of the households for which it is a proxy. Quantification[5] transforms the ordinal responses on inflation expectations to cardinal measures of inflation expectation in an econometric fashion. In the literature related to macroeconomic modelling, the interest in quantification methods stems from the fact that survey results can be used as "proxies of agents' expectations (see Smith and McAleer 1995, MacDonald 2000) or as 'leading indicator' of actual changes in economic variables (see, among others, GarcÃÌa-Ferrer and Bujosa-Brun 2000)" (Nardo 2003, p.646). Therefore, as a first step, its use in econometric analysis requires its conversion into quantitative data.

There are three main approaches proposed in the literature for quantification. These are balance statistic (see Anderson Jr 1952) used by Eurostat, the probability approach (Theil 1952, Carlson and Parkin 1975, Batchelor and Orr 1988) both with time invariant and time varying parameters, and the regression approach (Pesaran 1984, Pesaran 1987). The earliest approach to quantification is the use of "balance statistic". A *balance statistic*, is defined as the difference between the proportion of economic agents that expected price level to rise and the proportion that expected a fall. It is scaled against the actual price change. As an alternative to balance statistic in the quantification of IFO-Munich Survey results (see Anderson Jr 1952), Theil (1952) first proposed a quantification approach assuming a common subjective probability distribution of expectations for all economic agents. Later, probability approach was rediscovered by Carlson and Parkin (1975) and is popularly known as **C**arlson-**P**arkin (CP) method for ordinal responses in 3-categories. Batchelor and Orr (1988) generalize the CP method for ordinal responses in 5-categories. Pesaran developed the regression approach (Pesaran 1984, Pesaran 1987) along the lines suggested by Anderson Jr (1952). The method uses the relationship between actual inflation and the response fractions of the surveyed households in the categories "a", "b", "c", "d" and "e" mentioned above as a basis for the quantification of respondents' expectations about the future.

---

[4]Batchelor (2006) highlighted interesting points in support of qualitative response such as robust estimation of mean expectation, insights into consumer expectation formation, etc.

[5]See Nardo (2003), Pesaran and Weale (2005) for a survey on quantification techniques.

Thus, quantified expectations are derived from a specific regression model rather than a specific probability distribution.

Finally, comparison of different available studies "indicates that generalization about which method of quantification works best is not possible" (Pesaran and Weale 2005). It is highly context specific.

In this paper we employ the above approaches to quantify the IESH data of RBI, and assess its performances by comparing the accuracy of quantified inflation in forecasting the actual inflation. Our empirical results show that the simpler methods viz., balance statistics (BS) and Pesaran's regression methods (PR) perform significantly better than the more complex Carlson-Parkin (CP) method. The reason is, the strong assumptions that CP method makes, do not possibly hold for IESH data. We then propose an extension of BS and PR methods with time varying weights similar to CP method with time varying thresholds. The performances of the methods with time varying weights and thresholds are found to be significantly better than the methods assuming constancy of weights and thresholds.

This paper is organized as follows. In Section 2, we give a detailed description of the IESH data collected by Reserve Bank of India (RBI), and introduce some notations that will be used throughout. In Section 3, we discuss three basic approaches mentioned above. In Section 4, we report the results of our empirical analysis and compare the performances of these approaches in quantification of IESH data. Finally, in Section 5, we give the concluding remarks.

2. Data

Inflation Expectations Survey of Households (IESH)[6] of RBI collects data from a quota sample of 4000 urban households from twelve cities in India on a quarterly basis since 2005, third Quarter. Unlike most surveys, IESH collects data on both qualitative and quantitative inflation expectations. It captures the data in two steps. First, households are asked to report their three months ahead and one year ahead expected change in the general price level on a five-category ordinal scale as stated above. Let "A", "B", "C", "D" and "E" denote the fraction of households opting for the categories "a", "b", "c", "d" and "e", respectively.

In the second step, households are asked for a direct quantitative estimate of the inflation rate in different intervals of values. The question reads: "Respondent view's on the following inflation rates: Current Inflation rate, Inflation rate after three months, Inflation rate after one year" (In other words, the perceived inflation and expected inflation, after three months and after one year). Thus, households report both qualitative and quantitative beliefs about expected inflation whereas only quantitative belief on perceived inflation.

In the following we use households data for the period 2008 : $Q3$ (IESH: round 13) to 2018 : $Q3$ (IESH: round 53). Almost 0.1% of the households have missing responses in one or more fields, and 3% of the quantitative responses are in "No idea" category. These are discarded. Also, quantitative responses for the households in category "e" are re-coded as -0.5, and in category "d" are re-coded as 0. This is to bring internal consistency to the data for the two types of responses. The anomaly in qualitative and quantitative responses may be due to an error in converting the experience of price change to an inflation value. Also households who answered no change in price do not

---

[6]RBI (2010)

necessarily mean a zero quantitative response, albeit a value in an imperceptible range around zero (Batchelor (1986)).

Figure 1. Qualitative Response Fractions (A:bottom to E: top) of Next Quarter Inflation Expectation: IESH Data from 2008Q3 (Round 13) to 2018Q3 (Round 53)
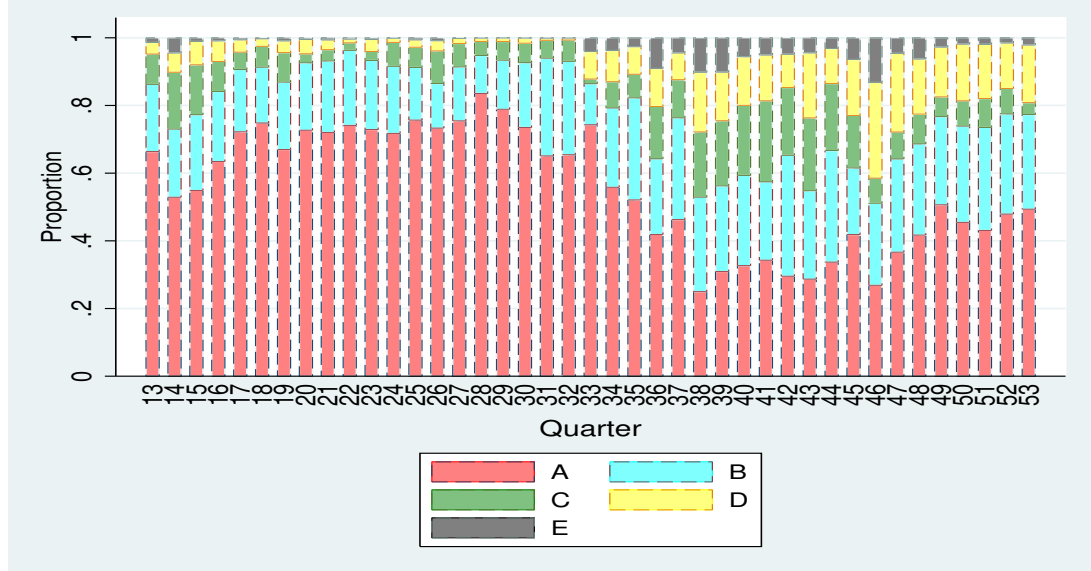

Figure 1 shows the movement of the response fractions "A" to "E" for one quarter ahead expectations across different survey rounds. Clearly, during the high inflation regime, i.e., during $2008Q3$ to $2014Q3$, the fraction "A" remains high, varying between 0.6 to 0.8. However, subsequently, with lowering of actual inflation it falls down. It is interesting to observe that during the study period $2008Q3$ - $2018Q3$, almost 90% of the responses are concentrated in the categories "a", "b" and "c", and more than 75% are in the categories "a" and "b". Figure 2 shows the plots of the mean household inflation expectations and the actual inflation as measured by CPI-IW against different rounds. It helps to explain the pattern observed in Figure 1. Clearly, the household inflation expectations are almost always higher than the actual, often by a substantial margin, and there is a tendency of inflation expectations to persist. On the other hand, it is interesting to note that the inflation forecasts by the professionals[7] follow the actual inflation closely over time.

3. Quantification Methods

We now briefly discuss the three main approaches to quantification of qualitative responses on inflation expectations.

3.1. Balance Statistic (BS) and Regression Approach of Pesaran

One of the most commonly used method for quantification of qualitative survey observations is the 'balance' statistic proposed by Anderson Jr (1952). It is simply

---

[7]RBI's professional forecasters' forecasts of quarterly inflation for the same period. RBI conducts a quarterly (now bi-monthly since 2014) survey of professional forecasters (SPF).

Figure 2. Time Series of Actual Inflation (CPI-IW), Mean Inflation Expectations of Households (IESH) and Professionals (SPF)
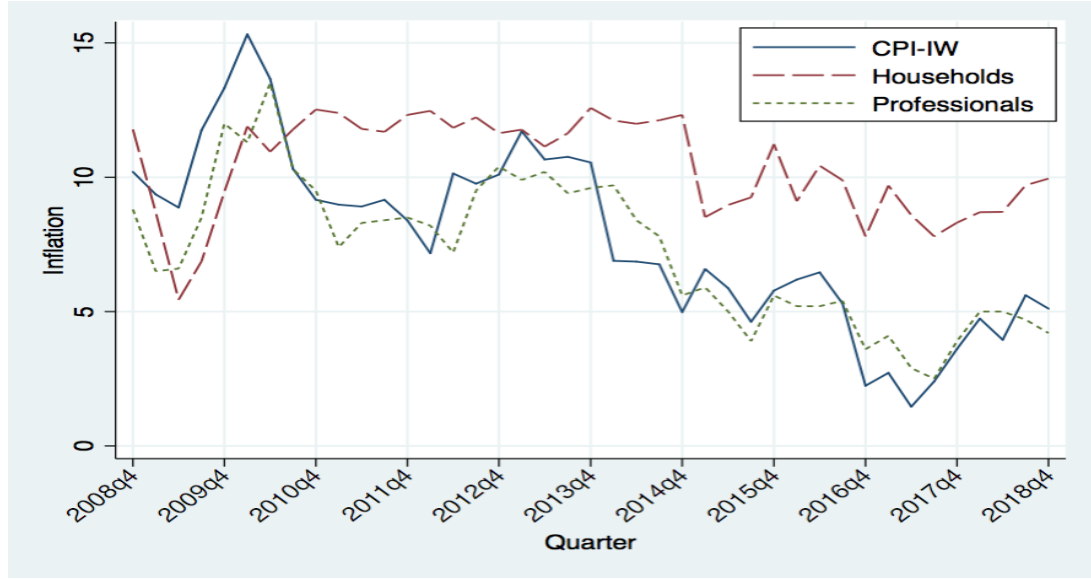
the difference between the proportion of households expecting a rise and those expecting a fall in future price level. Often, it is used as a measure to summarise the information in a survey; a positive balance being associated with rise in inflation expectations and a negative balance associated with falling inflation expectations.

In the following we motivate the balance statistic using regression. Suppose x_t^e represents the average inflation expectations of the population, and x_{it}^e represents the inflation expectation of the i-th household. Then x_t^e is a weighted average of x_{it}^e's and is given by

$$x_t^e = \sum w_{it} x_{it}^e. \tag{1}$$

Assuming that the same general relationship (1) holds for the surveyed households we get,

$$x_t^e = \sum_{i=1,2,\ldots,N_t} w_{it} x_{it}^e = \sum_{U_t} w_{it}^+ x_{it}^{e+} + \sum_{D_t} w_{it}^- x_{it}^{e-}, \tag{2}$$

where, N_t is the sample size, w_{it} is the weight attached to the $i-$th respondent which is usually N_t^{-1}; U_t and D_t denote the sets of respondents expecting increase and decrease, respectively, and the superscripts $+$ and $-$ denote the respondent expecting increase and decrease. If we now assume

$$x_{it}^{e+} = \beta + e_i, x_{it}^{e-} = -\beta + e_i \tag{3}$$

where $\beta > 0$, and e_i's are independently distributed errors with mean 0 and variance σ_{ie}^2. Assume that σ_{ie}^2 are sufficiently small and the distributions of e_i's are appropriately truncated so that $x_{it}^{e+} > 0$ and $x_{it}^{e-} < 0$ for all i and t. Thus, for sufficiently large number of responses in U_t and D_t equation (2) entails

5

$$x_t^e = \beta \sum_{U_t} w_{it}^+ - \beta \sum_{D_t} w_{it}^-, \tag{4}$$

or simply,

$$x_t^e \approx \beta(p(U_t^e) - p(D_t^e)), \tag{5}$$

where, $p(U_t^e)$ and $p(D_t^e)$ are the proportion of households that reported an expected increase or decrease in inflation expectations, respectively.

The balance statistic $p(U_t^e) - p(D_t^e)$ is proposed by Anderson Jr (1952) and advocated by Theil (1952). It, thus, provides an accurate measure of the average change in the inflation expectations up to a scalar factor if the percentage changes of increase and decrease reported by the respondents are equal in magnitude and remain constant over time. The slope parameter β is estimated by regressing actual inflation π_t, which is chosen as a proxy for the unobservable π_t^e, on the balance statistic.

The balance statistic is motivated by assuming that the qualitative responses (on prices) are observed in three categories, "go up", "no change", and "go down". Most importantly, it assumes that if the value of the latent variable x_t lies in a symmetric interval around zero, say, $(-\delta, \delta)$, the response is "no change". However, the mean π_t^e of x_t being positive, this assumption seems to be untenable, and so the balance statistic is generally biased (Cunningham 1997). A simple method to correct for the bias is to add an intercept term to the regression model considered above. The model is,

$$\pi_t = \alpha + \beta(p(U_t^e) - p(D_t^e)) + \epsilon_t, \tag{6}$$

where α represents the bias term, and ϵ_t represents the generic error term.

Pesaran (1984) generalizes the balanced statistics by relaxing the assumption that the percentage changes of increase and decrease are constant but may have different magnitudes. In other words, the equation (3) is to be rewritten as

$$x_{it}^{e+} = \beta_u + e_i, x_{it}^{e-} = -\beta_l + e_i \tag{7}$$

where $\beta_u, \beta_l > 0$, and e_i's are independently distributed errors with mean 0 and variance σ_{ie}^2. Following similarly as above, the equation (5) reduces to

$$x_t^e \approx \beta_u p(U_t^e) - \beta_l p(D_t^e)), \tag{8}$$

which is a generalization of the balanced statistic. Arguing as above in order to estimate β_u and β_l, the following regression model is to be considered,

$$\pi_t = \beta_u p(U_t^e) - \beta_l p(D_t^e) + \epsilon_t. \tag{9}$$

In order to estimate the bias, an intercept is to be added to equation (9).

Pesaran's regression approach provides an easy generalization of the balanced statistic to the qualitative data observed in the five categories, viz., "a", "b", "c", "d" and "e" mentioned in Section 2 by assigning weights β_a, β_b, β_c, β_d and β_e to the respective fractions "A", "B", "C", "D" and "E". In other words the regression model to be considered is,

$$\pi_t = \beta_a A_t + \beta_b B_t + \beta_c C_t + \beta_e E_t + \epsilon_t, \tag{10}$$

or the bias corrected version of (9) which is given by,

$$\pi_t = \beta_0 + \beta_a A_t + \beta_b B_t + \beta_c C_t + \beta_e E_t + \epsilon_t. \tag{11}$$

The crucial assumption of Pesaran's regression method is that the model (11) also holds for future expectation x_t^e.

Pesaran further generalizes (9) by advocating asymmetrical responses of individuals in periods of increasing and decreasing inflations. He proposed to consider

$$x_{it}^{e+} = \beta_u + \lambda \pi_t + \epsilon_{it}^+, \; x_{it}^{e-} = \beta_l + \epsilon_{it}^- \tag{12}$$

where β_u, $\beta_l \geq 0$, $0 \geq \lambda \leq 1$, $\pi_t \geq 0$ and ϵ_{it}^+'s and ϵ_{it}^-'s are independently distributed random errors with mean 0 and variance σ_{it}^{+2} and σ_{it}^{-2}, respectively.

In periods of rising inflation, the relation between x_{it}^{e+} and π_t is a first approximation to an unknown, and possibly a complicated relationship that underlies changes in inflation expectation of an individual agent and the overall rate of inflation besides the individual specific factors (Pesaran 1987, p.222). Plugging in expression (12) to equation (2), while replacing π_t with mean survey expectation x_t^e we obtain

$$x_t^e = \frac{\beta_u p(U_t) - \beta_l p(D_t)}{1 - \lambda p(U_t)} + \xi_t \tag{13}$$

where, $\xi_t = \frac{\sum w_{it}^+ \epsilon_{it}^+ + \sum w_{it}^+ \epsilon_{it}^-}{1 - \lambda p(U_t)}$ are possibly heteroscedastic and autocorrelated (through the weights and through $p(U_t)$).

Clearly it is a non-linear regression model. Taking π_t as the proxy for x_t^e the estimates of β_u, β_l, and λ can be obtained. In this context it is to be mentioned that equation (13) is similar to the three category probability model with time dependent boundaries of "no change" category, say, (a_t, b_t) where, $a_t = a_0 + a_1 \pi_t + \epsilon_t$ and $b_t = b_0$. This is shown by Smith and McAleer (1995) (p.183). It needs mentioning that Pesaran presents the equation (13) merely as a link between two different sources of information, not the one providing a causal explanation for change in inflation expectation. Unfortunately, the validity of this model cannot be checked. It needs individual level data which are not available in general.

3.2. Probability Method

Theil (1952) first advocated for quantification approach with explicit assumption on subjective distribution of expectation, terming it "probability method". Later, the same method has been independently developed by Carlson and Parkin (1975), thus giving the popular name **CP**. The theoretical underpinning of the method derives from the central limit theorem (CLT), which states that, as the number of respondents becomes larger their (mean) expectation distribution converges to a normal distribution. The connotation used in quantification is that subjective beliefs follow a normal distribution.

3.2.1. For 3-category Response Scheme

The CP method assumes that at time $t-1$ each agent predicts x_t, the inflation rate at time t, based on his/her subjective probability distribution $f_i(x_{it}/I_{t-1})$ conditional on the information available up to period $t-1$, represented by I_{t-1}. Further, CP assumes that for agent 'i', his/her expected inflation x_{it}^e for time t is the mean of $f_i(x_{it}/I_{t-1})$. Suppose now that individual expectations x_{it}^e can be viewed as independent draws from a common probability distribution $g(x_t)$ with mean x_t^e and variance σ_t^2. Then it supposes that an indifference interval, (a_{it}, b_{it}) exits in the mind of the respondent 'i' with $a_{it} < 0$, and $b_{it} > 0$, such that agent 'i' reports an increase in inflation expectation if $x_{it}^e \geq b_{it}$, and a fall if $x_{it}^e \leq a_{it}$, while he/she reports no change if $x_{it}^e \in (a_{it}, b_{it})$.

Notice that the balance statistic can be simply interpreted as the expected value of a discrete probability distribution taking three values -1, 0, 1 ('go down', 'no change' and 'go up', respectively) with positive probabilities. This transformation implicitly assumes the symmetry of the responses: the distance between 'go up' and 'no change' is the same as the distance between 'no change' and 'go down'. The CP method, on the other hand, makes two additional key assumptions to obtain an estimate of x_t^e from the knowledge of the individual x_{it}^e's. First, the indifference interval is same for all agents, that is, $a_{it} = -a_t, b_{it} = b_t$, where $a_t, b_t > 0$, and symmetric, that is, $a_{it} = -a_t, b_{it} = a_t$. Second, $g(x_t)$, the probability distribution of x_{it}^e's is normal.

The assumptions stated above can be restated as follows:

- Households are homogeneous: $a_{i,t} = -a_t$ and $b_{i,t} = b_t$
- Thresholds are symmetric: $a_t = b_t$
- Household's expectations are normally distributed: $x_{it}^e \sim \mathrm{N}(\pi_t^e, \sigma_t^2)$
- Unbiasedness of beliefs with respect to actual inflation: $\pi_t^e = \pi_t$

The survey response scheme for response fractions is thus:

$$
\begin{aligned}
p_t^D &= P(x_{it}^e < -a_t) \\
p_t^N &= P(x_{it}^e < a_t) - Pr(x_{it}^e > -a_t) \\
p_t^U &= P(x_{it}^e > a_t)
\end{aligned}
\tag{14}
$$

Denoting the cumulative distribution function of a standard normal variable by $\Phi(.)$, we can rewrite the equation (14) in the following form:

$$
p_t^D = \Phi(\frac{-a_t - \pi_t^e}{\sigma_t})
\tag{15}
$$

$$p_t^N = \Phi(\frac{a_t - \pi_t^e}{\sigma_t}) - \Phi(\frac{-a_t - \pi_t^e}{\sigma_t}) \tag{16}$$

$$p_t^U = 1 - \Phi(\frac{a_t - \pi_t^e}{\sigma_t}) \tag{17}$$

Solving equations (15)-(17) for π_t^e and σ_t we obtain,

$$\pi_t^e = a_t \frac{A_t^D + A_t^U}{A_t^D - A_t^U} \tag{18}$$

$$\sigma_t = -a_t \frac{2}{A_t^D - A_t^U} \tag{19}$$

where, $A_t^D = \Phi^{-1}(p_t^D)$ and $A_t^U = \Phi^{-1}(1 - p_t^U)$.

This leaves only a_t to be unknown. To obtain an operational expression for calculating π_t^e, CP further assumes that a_t is a time independent constant, i.e., $a_t = \tilde{a}$. To estimate $\tilde{a}$, CP further sets

$$\tilde{a} = \frac{\bar{\pi}}{\frac{1}{T}\Sigma(\frac{A_t^D + A_t^U}{A_t^D - A_t^U})} \tag{20}$$

where $\bar{\pi}$ is mean of inflations on which survey expectations are sought. Note that the role of $\tilde{a}$ here is to to scale π_t^e such that the mean of π_t's is equal to π_t^e.

3.2.2. For 5-category Response Scheme

CP method is extended to five categories by Batchelor and Orr (1988) following the argument similar to Carlson and Parkin (1975). We briefly describe the method here. In Section 1, we used the notation "a", "b", "c", "d" and "e" to solicit the survey responses categories. For brevity, we introduce the notation $a_j, j = 1, 2, ..., 5$ for the aforementioned categories to classify the survey responses for quarter-ahead, and year-ahead directional changes to the "prices in general".

Let us now define these categories explicitly,

$a_1 : x_{it}^e \geq \pi_t^r + \eta_t,$
$a_2 : \pi_t^r - \eta_t \leq x_{it}^e \leq \pi_t^r + \eta_t,$
$a_3 : \delta_t \leq x_{it}^e \leq \pi_t^r - \eta_t,$
$a_4 : -\delta_t \leq x_{it}^e \leq \delta_t,$
$a_5 : x_{it}^e \leq -\delta_t,$

and $p_t^j = P(x_{it}^e \in a_j), j = 1, 2, ..., 5$, where π_t^r is known as reference inflation perception which is assumed to be same for all the agents. Besides π_t^r the other four unknowns π_t^e, σ_t, δ_t, and η_t can be found by solving the equations

$$\Phi^{-1}(p_t^j) = A_t^j, j = 1, 2, ..., 5. \tag{21}$$

The solution is given by:

9

$$\pi_t^e = \pi_t^r \frac{A_t^4 + A_t^3}{A_t^4 + A_t^3 - A_t^2 - A_t^1} \tag{22}$$

$$\sigma_t = \pi_t^r \frac{2}{A_t^4 + A_t^3 - A_t^2 - A_t^1} \tag{23}$$

$$\delta_t = \pi_t^r \frac{A_t^4 - A_t^3}{A_t^4 + A_t^3 - A_t^2 - A_t^1} \tag{24}$$

$$\eta_t = \pi_t^r \frac{A_t^2 - A_t^1}{A_t^4 + A_t^3 - A_t^2 - A_t^1}. \tag{25}$$

For identifying π_t^r inflation perceptions are assumed to be unbiased over the sample period which leads to

$$\pi_t^r = \frac{\bar{\pi}}{\frac{1}{T}\Sigma\left(\frac{A_t^4 + A_t^3}{A_t^4 + A_t^3 - A_t^2 - A_t^1}\right)} \tag{26}$$

where, $\bar{\pi}$ is the mean of actual inflation over the sample period.

3.2.3. Asymmetric and Time-varying Thresholds

Carlson and Parkin (1975) assume that the thresholds of the indifference interval are symmetric around zero and constant across time and individuals. These are indeed strong assumptions. Here, following Seitz (1988) we consider extension of CP method with time dependent asymmetric thresholds.

Let us rewrite equation (18) with asymmetric time dependent thresholds (a_t, b_t), $a_t < 0$, $b_t > 0$ (Seitz 1988):

$$\pi_t^e = \frac{b_t A_t^D - a_t A_t^U}{A_t^D - A_t^U}. \tag{27}$$

Literature suggests that inflation level and inflation uncertainty (IU) are key determinants of inflation expectations (Seitz 1988). Though IU is not directly observed, disagreement in inflation expectations captured through survey may serve as a proxy for IU (Bomberger 1996, Barron, Stanford, and Yu 2009, Bachmann, Elstner, and Sims 2013). Following Pesaran (1984) and Batchelor (1986) we make the thresholds a_t and b_t depend on the current inflation rate (π_{t-1}) and a proxy for inflation uncertainty (ξ_{t-1}) at time $t-1$. Batchelor (1986) provides a theoretical justification for this procedure from signal detection theory.

We consider the index of qualitative variation (IQV), a measure of disagreement (Thomas 2010, Lamla and Maag 2012), as a proxy for IU. If the qualitative responses

are observed in J categories then $\xi_t = \frac{J}{J-1}(1 - \sum_{j=1,2,...,J}(p_t^j)^2)$ where, p_t^j is the proportion of responses in the j-th category at time t. Notice that $\xi_t = 1$ when $p_t^j = \frac{1}{J}, j = 1, 2, ..., J$ and $= 0$, when $p_t^j = 1$, for some j.

We assume the thresholds a_t and b_t are linearly related to (π_{t-1}) and (ξ_{t-1}). Thus a time-varying model for upper threshold b_t and lower threshold a_t is given by:

$$b_t = \beta_0^b + \beta_1^b \pi_{t-1} + \beta_2^b \xi_{t-1}, \; a_t = \beta_0^a + \beta_1^a \pi_{t-1} + \beta_2^a \xi_{t-1} \tag{28}$$

Substituting the expressions for b_t, a_t from equation (28) to equation (27) we obtain,

$$\pi_t^e = (\beta_0^b + \beta_1^b \pi_{t-1} + \beta_2^b \xi_{t-1})\left(\frac{A_t^D}{A_t^D - A_t^U}\right) - (\beta_0^a + \beta_1^a \pi_{t-1} + \beta_2^a \xi_{t-1})\left(\frac{A_t^U}{A_t^D - A_t^U}\right) \tag{29}$$

The unknown parameters $\beta_j^a, \beta_j^b, j = 0, 1, 2$ are obtained by running the following regression model:

$$\pi_t = (\beta_0^b + \beta_1^b \pi_{t-1} + \beta_2^b \xi_{t-1})\left(\frac{A_t^D}{A_t^D - A_t^U}\right) - (\beta_0^a + \beta_1^a \pi_{t-1} + \beta_2^a \xi_{t-1})\left(\frac{A_t^U}{A_t^D - A_t^U}\right) + \epsilon_t \tag{30}$$

where ϵ_t represents the error component.

Notice that the rationale for time varying thresholds discussed above may also be extended to other quantification methods, viz., balance statistic (BS) (cf. equation (5)) and Pesaran's regression method (PR) (cf. equation (9)). In (5) we model β, and in (9) we model β_l and β_u as linear functions of π_{t-1} and ξ_{t-1} like (28), and run appropriate regression models. These methods are new and not considered before in the literature. However, it needs to be mentioned in this context that we are using the time varying approach in order to get better prediction of actual inflation (π_t) rather than for explaining the time varying behaviour of the thresholds or of the β's.

4. Empirical Results and Discussion

In this section we compare the performances of nine different quantification methods discussed above and also mentioned at the foot note of Table 1 using IESH data spanning between $2008Q3$ to $2018Q3$. The results are furnished in Table 1. The performances of these methods are assessed by finding the forecast accuracy of the quantified inflation expectations (x_t^e) in forecasting the actual inflation (π_t). The forecast accuracy is measured by the root mean squared error (RMSE $= \sqrt{MSE}$). We also furnish in the table the correlation coefficient between the quantified inflation expectation x_t^e and actual inflation π_t, say, r to capture the similarity of the movements of the two series over time. Clearly higher the value of r more is the similarity of the movements of the two series over time.

Table 1 shows that among the four basic quantification methods, viz., balance-statistic (BS), regression (PR) and probability methods $(CP3$ (3-category) and $CP5$ (5-category)), BS has the maximum correlation coefficient (0.75) and second lowest $RMSE$. The probability methods $CP3$ and $CP5$ perform worse than BS and PR possibly because the assumptions underlying these methods fail to hold for IESH data. In terms of $RMSE$ regression method performs better than the others. Interestingly, as soon as we consider the time varying counterparts of BS, $CP3$ and PR,

Table 1. Performance of Different Quantification Measures

Measure	Correlation	RMSE
BS	0.75	2.51
PR	0.70	2.35
CP3	0.67	2.60
CP5	0.67	2.50
BSTV	0.85	1.72
PRTV	0.88	1.57
CP3TV	0.85	1.70
CP3AS	0.67	2.42
PRAS	0.79	1.98

Note: All the quantification measures (except CP5) are for 3-category response scheme. A translation from 5-category to 3-category is obtained by clubbing the different response categories on price increase (positive inflation) into one category. IESH data (see section 2) from $2008Q3$ to $2018Q3$ is used. Different quantification measures are; BS: balance-statistic (Anderson Jr 1952) [see equation (6) without intercept α], PR: Pesaran's regession (Pesaran 1984, Pesaran 1987) [see equation (9)], $CP3$: Carlson-Parkin (Carlson and Parkin 1975, Theil 1952) [see equation (18)], $CP5$: Batchelor (Batchelor and Orr 1988) with π_t^r is assumed unbiased over actual inflation (π_t) i.e. $\bar{\pi}$ [see equation 22], $BSTV$: BS with time varying, $PRTV$: PR with time varying coefficients, $CP3TV$: Carlson-Parkin with time-varying thresholds (Seitz 1988) [see equation (27) - (28)], $CP3AS$: Carlson-Parkin with asymmetric thresholds (Seitz 1988) [see equation (27)], $PRAS$: Pesaran's asymmetric regression [see equation (13)]. Correlation is taken with respect to actual inflation (π_t) for which the quantification is done. RMSE ($=\sqrt{\frac{1}{T}\Sigma_{t=1}^{t=T}(\hat{\pi}_t^e - \pi_t)}$) is root mean square of errors where $\hat{\pi}_t^e$ is the quantified series.

the performances of these methods enhance significantly, $PRTV$ performs the best followed by $CP3TV$ and $BSTV$. The time varying $CP3$ does not have to satisfy the strong assumptions of $CP3$. Besides these methods we also consider the Carlson-Parkin method with asymmetric thresholds ($CP3AS$), and Pesaran's regrsssion method with asymmetric response to inflation ($PRAV$); both perform better than the respective symmetric version. Finally, Figure 3 presents the plots of quantified inflation and actual inflation for nine methods. Clearly, the quantified inflation downplays the changes in the actual inflation over time.

Figure 3. Time-series of Quantified Inflation (blue, bold) and Actual Inflation (red, dotted): 2008Q3 (Round 13) to 2018Q3 (Round 53)
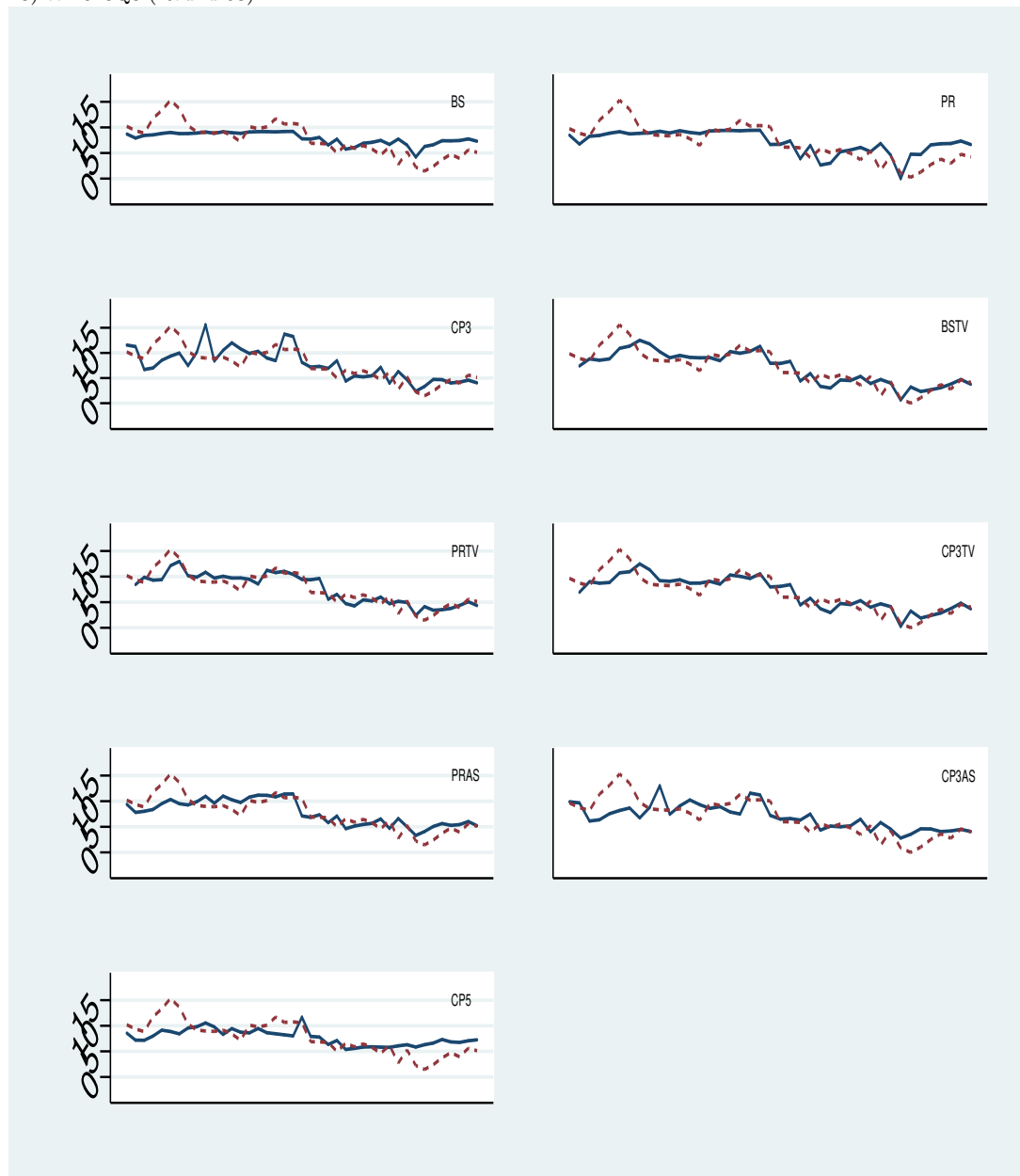
5. Conclusion

In this paper we present an empirical study comparing nine different methods of quantification for converting qualitative survey responses on future price changes to measures of inflation expectations. For the empirical study, we use IESH quarterly survey data of RBI from 2008 Q3 to 2018 Q3. Given that the IESH survey data are noisy, biased upwards (cf. Figure 1) reflecting the idiosyncrasies of the respondents, one of our important research questions is to investigate whether the survey data can still provide useful information to the central bank, in this case RBI, about the future inflation. Our empirical study suggests, the answer to this question is in affirmative.

Notice that all the methods discussed here are derived from one of the three basic approaches viz., Balance Statistic (BS) approach, probability based approach of Carlson-Perkin (CP) and Pesaran's Regression (PR) based approach. In addition to the methods available in the literature, we propose two methods for quantification viz., BS and PR with time varying weights. Our empirical study indicates that the simpler methods like BS and PR perform better than the more complicated CP method. Possibly, the assumptions underlying the CP method do not hold for IESH data. The PR method with time varying weights is found to be the best followed by the BS with time varying weights. Thus, for quantification of IESH data, our analysis suggests, simpler methods would yield better forecast for the future inflation.

References

Anderson Jr O (1952) The business test of the ifo-institute for economic research, munich, and its theoretical model. Revue de l'Institut international de statistique pp 1–17

Bachmann R, Elstner S, Sims ER (2013) Uncertainty and economic activity: Evidence from business survey data. American Economic Journal: Macroeconomics 5(2):217–49, , URL http://www.aeaweb.org/articles?id=10.1257/mac.5.2.217

Barron OE, Stanford MH, Yu Y (2009) Further evidence on the relation between analysts' forecast dispersion and stock returns*. Contemporary Accounting Research 26(2):329–357, , URL https://onlinelibrary.wiley.com/doi/abs/10.1506/car.26.2.1, https://onlinelibrary.wiley.com/doi/pdf/10.1506/car.26.2.1

Batchelor R (2006) How robust are quantified survey data? evidence from the united states

Batchelor RA (1986) The psychophysics of inflation. Journal of Economic Psychology 7(3):269–290, URL https://ideas.repec.org/a/eee/joepsy/v7y1986i3p269-290.html

Batchelor RA, Orr AB (1988) Inflation expectations revisited. Economica pp 317–331

Bernanke BS (2007) Inflation expectations and inflation forecasting. Tech. rep.

Bomberger WA (1996) Disagreement as a measure of uncertainty. Journal of Money, Credit and Banking 28(3):381–92, URL https://EconPapers.repec.org/RePEc:mcb:jmoncb:v:28:y:1996:i:3:p:381-92

Carlson JA, Parkin M (1975) Inflation expectations. Economica 42(166):123–138

Cunningham A (1997) Quantifying survey data. Quarterly bulletin, Bank of England

Galí J (2008) Monetary Policy, Inflation and the Business Cycle: An Introduction to the New Keynesian Framework. Princeton. Princeton Univ. Press

GarcÃIa-Ferrer A, Bujosa-Brun M (2000) Forecasting oecd industrial turning points using unobserved components models with business survey data. International Journal of Forecasting 16(2):207 – 227, , URL http://www.sciencedirect.com/science/article/pii/S0169207099000497

Lamla MJ, Maag T (2012) The role of media for inflation forecast disagreement of households and professional forecasters. Journal of Money, Credit and Banking 44(7):1325–1350, , URL https://onlinelibrary.wiley.com/doi/abs/10.1111/j.1538-4616.2012.00534.x, https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1538-4616.2012.00534.x

MacDonald R (2000) Expectations formation and risk in three financial markets: Surveying what the surveys say. Journal of Economic Surveys 14(1):69–100, , URL https://onlinelibrary.wiley.com/doi/abs/10.1111/1467-6419.00105, https://onlinelibrary.wiley.com/doi/pdf/10.1111/1467-6419.00105

Mishkin FS (2007) Inflation dynamics. Working Paper 13147, National Bureau of Economic Research, , URL http://www.nber.org/papers/w13147

Muth JF (1961) Rational expectations and the theory of price movements. Econometrica: Journal of the Econometric Society pp 315–335

Nardo M (2003) The quantification of qualitative survey data: A critical assessment. Journal of Economic Surveys 17(5):645–668, URL https://EconPapers.repec.org/RePEc:bla:jecsur:v:17:y:2003:i:5:p:645-668

Pesaran M (1984) Expectations formation and macroeconomic modelling'in p. malgrange and p. muet (eds), contemporary macroeconomic modelling

Pesaran M, Weale M (2005) Survey Expectations. Cambridge Working Papers in Economics 0536, Faculty of Economics, University of Cambridge, URL https://ideas.repec.org/p/cam/camdae/0536.html

Pesaran MH (1987) The limits to rational expectations. 339.3 PESl, Blackwell Oxford

RBI (2010) Inflation expectations survey of households

Seitz H (1988) The estimation of inflation forecasts from business survey data. Applied Economics 20(4):427–438

Sims CA (2009) Inflation expectations, uncertainty and monetary policy. Working Paper 275, Bank for International Settlements

Smith J, McAleer M (1995) Alternative procedures for converting qualitative response data to

quantitative expectations: An application to australian manufacturing. Journal of Applied Econometrics 10(2):165–185, URL http://www.jstor.org/stable/2284971

Theil H (1952) On the time shape of economic microvariables and the munich business test. Revue de l'Institute International de Statistique 20

Thomas M (2010) Essays on inflation expectation formation