Epidemiology of Inflation Expectations of Households and Internet Search- An Analysis for India

Saakshi* Sohini Sahu* Siddhartha Chattopadhyay∀

August 15, 2017

Abstract

This paper investigates how inflation expectations of individuals are formed in India. We investigate if news on inflation plays a role in formation of inflation expectations following the epidemiology-based work by Carroll (2003). The standard literature on this topic considers news coverage by the print and audio-visual media as the sources of formation of inflation expectations. Instead, we consider the Internet as a potential common source of information based on which agents form their expectations about future inflation. Based on data extracted from Google Trends, our results indicate that during the period 2008 to 2016, the Internet has indeed played a significant role in explaining how the inflation expectations of individuals have been formed in India. Additionally, we find that the Internet search sentiment, especially the negative sentiment, has a significant impact of inflation expectations.

Keywords: Inflation expectations, Epidemiology, Internet Search
JEL Classification: C53, D84, E31, E37.

* Department of Economic Sciences, Indian Institute of Technology Kanpur, India. saaki@iitk.ac.in
* Department of Economic Sciences, Indian Institute of Technology Kanpur, India. ssahu@iitk.ac.in (Corresponding Author)
∀ Department of Humanities and Social Sciences, Indian Institute of Technology Kharagpur, India. siddhartha@hss.iitkgp.ernet.in

1 Introduction

Inflation expectations are an important marker for monetary policy analysis. Price stability, one of the most sought-after macroeconomic goals, requires inflation expectations anchoring. Also, expectations about future inflation affect both the demand and supply sides of an economy, thereby making it imperative to study as to how these inflation expectations are formed by the general public.

This paper focuses on the survey-based responses on inflation expectations by the general public in India and addresses two related issues. First, we ask how inflation expectations of agents in the Indian economy, which do not follow rational expectations as per the survey data analysis (Mohanty, 2012), are formed? Second, we explore if the Internet, which is an omnipresent source of information, plays an active role in inflation expectations formation.

In the process, this paper brings together two strands in the existing literature- the first one being the epidemiological sources of inflation expectations pioneered by Carroll (2003), and the other is the use of meaningful data based on Internet search initiated by Choi and Varian (2012).

The literature on epidemiological sources of inflation expectations took off with the seminal work of Christopher Carroll in 2003. The basic idea, that draws inspiration from epidemiology, concerns how information about expected inflation emanates from a certain common source, like the forecasts made by professional forecasters, and the general public use this information from experts to update their beliefs about future inflation. Carroll's work provides a micro foundation for an aggregate expectations equation in a sticky information set-up. Since then, the epidemiological model has been explored by Lamla and Maag (2012), Pfajfar and Santoro (2013), Erhmann, Pfajfar and Santoro

(2014) and Lei et al (2015). The first few studies induct media news as the common source of information based on which the general public form their own inflation expectations. Lei et al (2015) use newspaper reports on inflation as the common source and further categorize the newspapers (general, economic, political etc.) to assess the impact of each type of newspaper report on inflation expectations formation in China.

Our work explores the epidemiological sources of inflation expectations among households in India following the general structure outlined by the studies mentioned above. However, we depart from all these studies as far as the common source of information is concerned. Instead of considering the forecasts made by the professional forecasters or those provided by the news media, we hypothesize that the Internet search on inflation-related keywords is the common source through which the general public update their future inflation expectations. Specifically, we use the inflation-related searches conducted through Google in India between 2008 and 2016, the data on which is extracted through Google Trends. Choi and Varian (2009a, 2009b, 2011) have demonstrated how the Google Trends data can be used to forecast automobile sales, forecast travel destination planning, unemployment etc. But this use of Google Trends data for epidemiological sources of inflation expectations formation, is a new addition to the existing literature, both from the point of view of a novel common source for information and also as a new application of the data.

In the process, our work analyses the Inflation Expectations Survey of Households (IESH) conducted by the Reserve Bank of India since 2006. This dataset has not yet been explored to its full potential and to the best of our knowledge, this work is the first of its kind to look at the epidemiological sources of inflation expectations using the survey-based responses on expected inflation in India.

Our findings indicate that the epidemiological sources of inflation expectations of general public in India holds for both Carroll (2003) type specification based on the inflation expectations of professional forecasters, as well as the Internet search data. Thus, we infer that the Internet is indeed a significant source from where the agents derive information and update their future inflation expectations. Additionally, we find that the Internet search "sentiment", that is the keywords used to conduct inflation-related searches, affect inflation expectations in the expected direction, especially unfavorable searches lead to higher inflation expectations in the future. All our findings are in line with the various works cited above, especially that of Lei et al. (2015).

The rest of the paper is organized as follows. Section 2 outlines the benchmark epidemiology framework pioneered by Carroll (2003), including the modifications made to accommodate our model. Section 3 discusses the data sources. Section 4 analyses the results and Section 5 concludes.

2 The Model

Carroll's (2003) seminal work on epidemiology of inflation expectations draws inspiration from Kermack-McKendrick (1927) model in the epidemiology literature that explains the process of transmission of disease in society from a "common source". At any given point in time, a constant fraction λ of the population gets affected by the disease that spreads from a common source, like air pollution, while the remaining $(1-\lambda)$ fraction of the population do not get affected but nevertheless remain susceptible. In the next time period, fraction λ of those who escaped the disease in the previous period get infected, while the rest do not. This is how a disease is transmitted from a common source over a period of time. Analogous to disease transmission from a common source, Carroll's (2003) model hypothesizes that news media is the "common source"

of information (transmission) based on which a section λ of the population update their expectations about future inflation. The remaining (1-λ) proportion of the population, who have not been affected by the "common source" or the news media, continue with the inflation expectations from the last period. Thus, at any given point in time, there is a combination of agents who have updated their inflation expectations and those who continue with their inflation belief from the previous period. This is how an element of "information stickiness" is introduced in the model.

$$E_t(\pi_{t+1}) = \lambda N_t(\pi_{t+1}) + (1-\lambda)\{\lambda N_{t-1}(\pi_{t+1}) + (1-\lambda)(\lambda N_{t-2}(\pi_{t+1}) + \cdots)\} \quad (1)$$

where, $E_t(\pi_{t+1})$ is the expectation of inflation rate for period t+1 conditional on the information available till time period t, π_{t+1} is the realized inflation for period $t+1$ and $N_t(\pi_{t+1})$ is the updated news at period t to form inflation expectations for period $t+1$.

The expression $\{\lambda N_{t-1}(\pi_{t+1}) + (1-\lambda)(\lambda N_{t-2}(\pi_{t+1}) + \cdots)\}$ can be expressed as the sum of an infinite series and equation (1) can be written succinctly as:

$$E_t(\pi_{t+1}) = \lambda N_t(\pi_{t+1}) + (1-\lambda)E_{t-1}(\pi_t) \quad (1a)$$

For the purpose of estimation, we use the following equation (2) that follows from equation (1a),

$$E_t(\pi_{t+1}) = \alpha_1 N_t(\pi_{t+1}) + \alpha_2 E_{t-1}(\pi_t) + \varepsilon_t \quad (2)$$

Most works on the epidemiological sources of inflation expectations, including that by Carroll (2003) use the forecasts made by the professional forecasters as the source from which the general public form their expectations (i.e. $N_t(\pi_{t+1})$). Lei et al. (2015) assume that $N_t(\pi_{t+1})$ is directly obtained from news reports. Stepping aside from these assumptions, we hypothesize that agents draw their information from the Internet search represented by the Google Trends (GT) data. Thus, we get equation (3) as follows:

$$N_t(\pi_{t+1}) = \gamma GT_t + \eta_t \quad (3)$$

Substituting equation (3) in equation (2), we get:

$$E_t(\pi_{t+1}) = \beta GT_t + \alpha E_{t-1}(\pi_t) + \mu_t \quad (4)$$

Following Lei et al (2015), we categorize the inflation-related searches into three categories- "favorable" (like low inflation, low prices), "unfavorable" (like high inflation, rising prices) and "neutral" (stable prices). The exact keywords searched for are outlined in the Data description section. Accordingly, to assess the search sentiment and its impact on inflation expectations, as in Lei (2015), we estimate the following equation:

$$E_t(\pi_{t+1}) = c + \alpha E_{t-1}(\pi_t) + \beta^{+} GT\_Unfavorable_t + \beta^{-} GT\_Favorable_t + \beta^{0} GT\_Neutral_t + \mu_t \quad (5)$$

However, due to very low number of searches in the "neutral" category, no data points are available. Thus, for search sentiment we estimate equation (6) which is a version of equation (5) after excluding the variable $GT\_Neutral_t$:

$$E_t(\pi_{t+1}) = c + \alpha E_{t-1}(\pi_t) + \beta^{+} GT\_Unfavorable_t + \beta^{-} GT\_Favorable_t + \mu_t \quad (6)$$

3 Data description

3.1 Inflation expectations

Data on inflation expectations of households in India is available from the Inflation Expectations Survey of Households (IESH) conducted quarterly by the Reserve Bank of India (the Central Bank of India) since the third quarter of 2006.

The survey covers about 5000 urban individuals in each round across 18 major cities in India at present. The sample surveyed represents both genders, nine age categories and seven different broad occupational categories in each city. The survey respondents are asked to quote quantitative numbers based on their inflation perception, one-quarter ahead inflation and one-year ahead inflation expectations. Figure 1 and Figure 2 plot the mean one-quarter ahead and mean one-year ahead inflation expectations of the general public in India against realized CPI and WPI respectively.

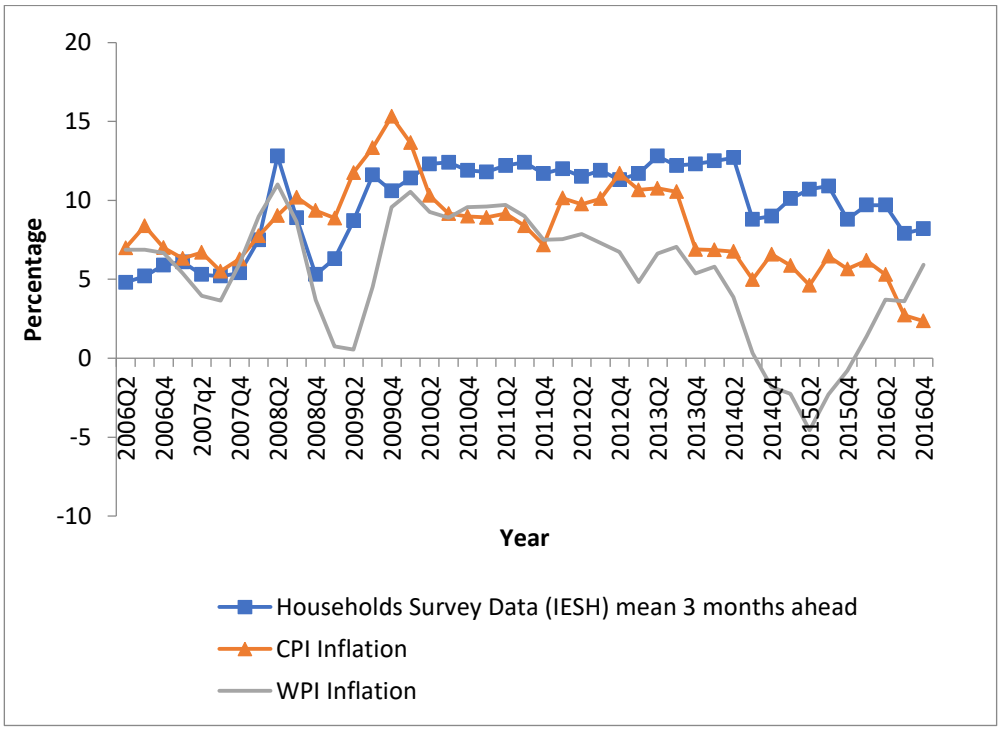


Figure 1: 3-month ahead inflation expectations of households in India

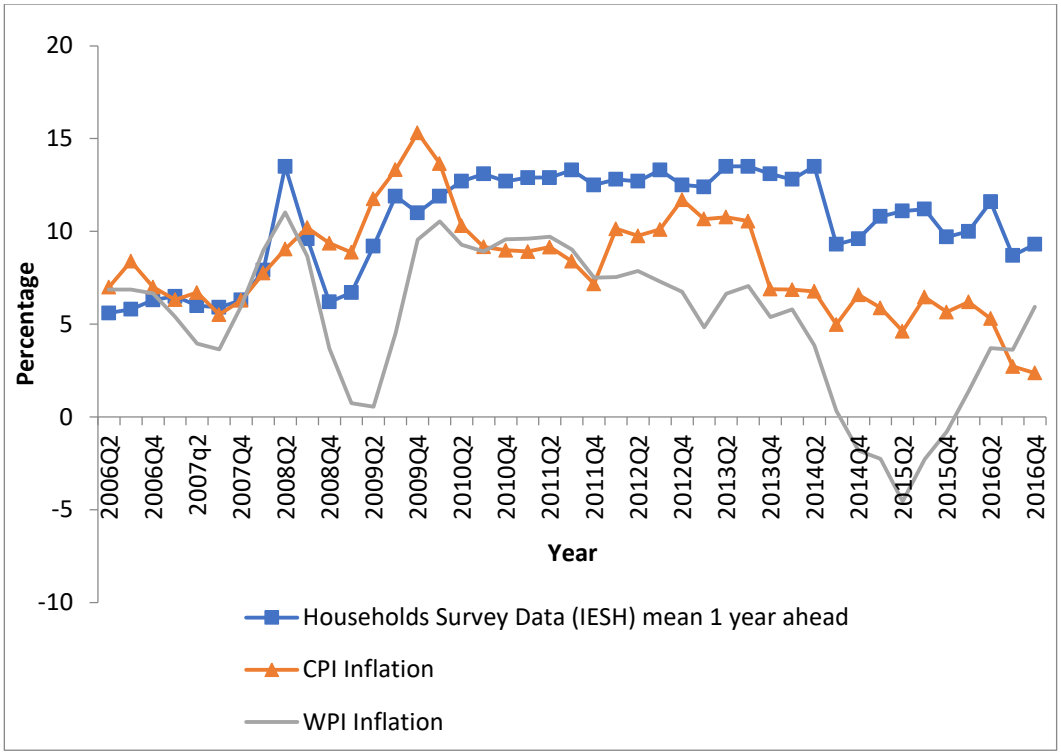
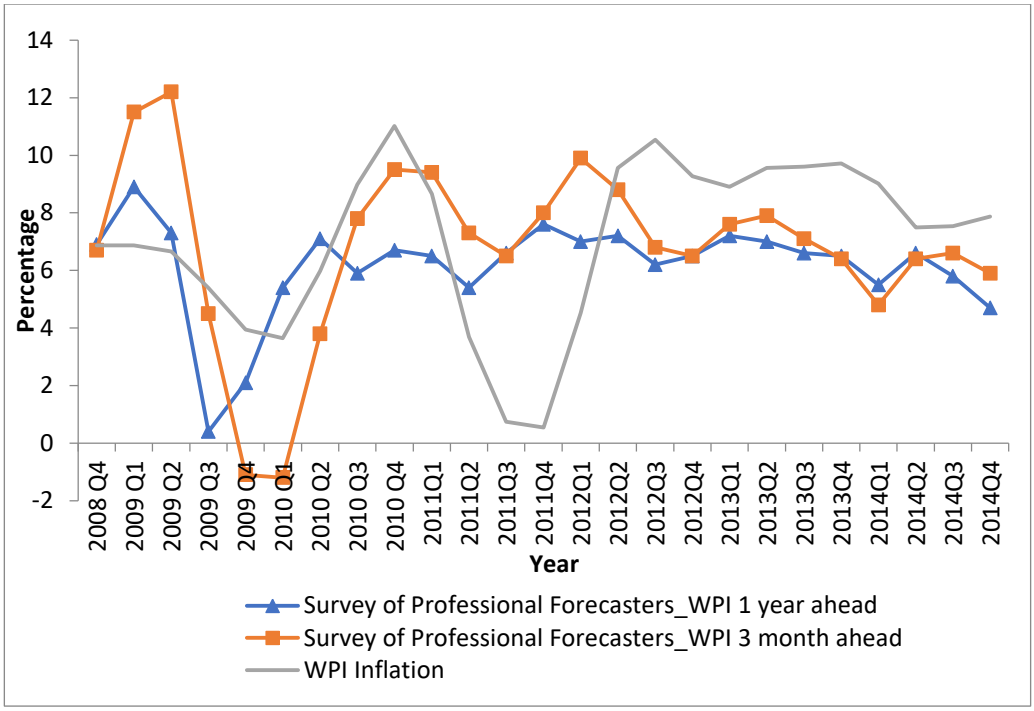


Figure 2: 1 year ahead inflation expectations of households in India

The Survey of Professional Forecasters (SPF) is at present conducted bi-monthly by the Reserve Bank of India. This survey was initiated in 2007 and was conducted at a quarterly frequency until April 2014. At present, 21 professional forecasters participate in this survey and give their quantitative forecasts on CPI- combined inflation, CPI-core inflation and WPI-combined inflation rates for the current and future quarters, apart from other macroeconomic forecasts. Figure 3 shows the mean one-quarter ahead and mean-one-year ahead forecasts of the professional forecasters across different categories of inflation.



Both IESH and SPF data are publicly available at the Reserve Bank of India website (https://www.rbi.org.in/).

3.2 Internet search data

The Internet search data has been collated from Google Trends (https://trends.google.co.in/trends/). Google Trends is a public web facility made available by Google Inc. that gives Google search related statistics.

Search related data in Google Trends can be filtered according to geographic zones, time period, frequency and categories. For example, we looked into the search history for the keyword "inflation" in India starting 2006: Q3 till 2016: Q4 at a quarterly frequency across "All categories". The data generated is an index known as the Google Trend Index for the searched term. It is to be noted that Google Trends does not generate data on search volume. Instead, it expresses the searched term as a fraction of the total number of searches conducted in the zone of our interest during a specified time period and assigns the number 100 against the highest fraction. The rest of the data

series is rescaled with respect to the data point that corresponds to the number 100.

We collected data on the search history of each term related to inflation, namely- "Inflation", "Core Inflation", "Headline Inflation", "CPI Inflation" and "WPI Inflation"; across each search category namely- "All Categories", "Finance Category", "Business and Industrial Category" and "News Category", that originated from India between 2006: Q3 and 2016: Q4. Within each category, searches are further divided into "Web Search" and "News Search".

For search sentiment, we collected data on the search history of terms to represent "Favorable" search, "Unfavorable" search and "Neutral" search in India during the time period 2006: Q3 and 2016: Q4.

Terms like "Low Price", "Price Low" and "Price Low", were clubbed under "Favorable" search. "High Price", "Price Rise", "Price High", "Price Increase", were clubbed under "Unfavorable" search. Not enough searches were recorded for terms related to "Neutral" search, thereby generating no data points by Google Trends.

4 Results

Table 1 reports the results[1] of the benchmark Carroll (2003) equation, represented by various specifications of equation 2, where the WPI inflation expectations of the professional forecasters (SPF) are the epidemiological source of inflation expectations for the general public. In all versions of equation 2, the coefficient of SPF is positive and significant, thereby implying that epidemiological sources of inflation expectations formation exist

[1] All equations estimated across Table 1 to Table 8, use OLS. For each equation, we conduct the Durbin-Wu-Hausman test to check for endogeneity. However, the null hypothesis of the test cannot be rejected in all cases except equation version 1(b) in Table 1 where IV estimation is done; for the rest all equations we estimate using OLS.

in the Indian economy. Past inflation expectations by the general public are also positive and significant in all versions of equation 2, thereby highlighting the contribution of agent's past expectations in forming the current expectations about future inflation. Past actual inflation (WPI inflation) matters as well for inflation expectations formation, but it has a negative relationship with the latter.

Estimating Equation:
$$E_t[\pi_{t+1}] = \alpha_0 + \alpha_1 N_t[\pi_{t+1}] + \alpha_2 E_{t-1}[\pi_t] + \alpha_3 P_t[\pi_{t-1}] + \epsilon_t$$

Equation	α_0	α_1	α_2	α_3	$\overline{R^2}$	Durbin-Watson
1a	0.23 (3.06)***	0.88 (19.80)***	--	--	0.44	1.95
1b	0.14 (0.94)	0.92 (12.63)***	--	--	0.27	2.02
2	3.61 (3.14)***	0.20 (3.01)***	0.57 (5.28)***	--	0.56	1.83
3	--	0.53 (3.78)***	0.89 (21.47)***	-0.32 (-2.49)**	0.52	2.22
4	3.00 (2.60)**	0.42 (3.09)***	0.63 (5.79)***	-0.23 (-1.86)*	0.59	2.05
5	--	--	0.94 (20.26)***	0.10 (1.42)	0.31	1.92

NOTE: Columns 2-5 report the coefficient values for different versions of equation with t-statistics at parenthesis. Columns 6-7 report diagnostics for each model and column 8 reports the methodology that is used for estimation. $E_t[\pi_{t+1}]$ is 3 month ahead mean inflation expectation survey of household at time period t. $N_t[\pi_{t+1}]$ is mean 3 month ahead professional forecast of WPI at time period t. $P_t[\pi_{t-1}]$ is one period lag of actual WPI inflation. The time period for estimation is 2008Q1 to 2016Q2. *, ** , *** indicates 10, 5 and 1 per cent significance level respectively.

Table 1: Results of baseline epidemiology equation using SPF_WPI

We repeat the exercise as done in Table 1, to check for epidemiological sources of expectation formation when the CPI forecasts of the professional forecasters are considered.

Estimating Equation:

$$E_t[\pi_{t+1}] = \alpha_0 + \alpha_1 N_t[\pi_{t+1}] + \alpha_2 E_{t-1}[\pi_t] + \alpha_3 P_t[\pi_{t-1}] + \epsilon_t$$

Equation	α_0	α_1	α_2	α_3	$\overline{R^2}$	Durbin-Watson
1	--	0.37 (3.98)***	0.72 (9.83)***	--	0.51	2.03
2	1.92 (1.30)	0.28 (2.45)**	0.61 (5.54)***	--	0.52	1.95
3	--	0.68 (3.06)***	0.74 (10.16)***	-0.31 (-1.54)	0.53	2.25
4	2.02 (1.40)	0.60 (2.65)**	0.62 (5.79)***	-0.32 (-1.62)	0.55	2.17
5	--	--	0.79 (9.93)***	0.25 (2.74)**	0.41	1.94

NOTE: Columns 2-5 report the coefficient values for different versions of equation with t-statistics at parenthesis. Columns 6-7 report diagnostics for each model and column 8 reports the methodology that is used for estimation. $E_t[\pi_{t+1}]$ is 3 month ahead mean inflation expectation survey of household at time period t. $N_t[\pi_{t+1}]$ is mean 3 month ahead professional forecast of CPI at time period t. $P_t[\pi_{t-1}]$ is one period lag of actual CPI inflation. The time period for estimation is 2008Q1 to 2016Q2. *, ** , *** indicates 10, 5 and 1 per cent significance level respectively.

Table 2: Results of baseline epidemiology equation using SPF_CPI

Similar to the results obtained in Table 1, we find that the coefficient on the CPI forecasts made by the professional forecasters is positive and significant, thereby implying Carroll-type the epidemiological sources of inflation expectations formation by general public. Past inflation expectation by agents are also positive and significant, thereby indicating that not all agents update their expectations based on current information.

Thus overall, using the two types of inflation numbers (WPI and CPI) provided by the professional forecasters as the source of news for the general public, we do find evidence of Carroll-like epidemiological sources of inflation expectations formation in India.

Next, we depart from the SPF-based Carroll type specification of the epidemiology equation and substitute it by the "inflation-related" searches of the Internet. Our contention is that, some fraction of the population updates their beliefs based on the inflation-related information that they obtain from the Internet, while the rest of the population continue with their previous beliefs about expected inflation.

In this context, as outlined in section 3.2, we summarize the results for each of the following searched terms- "Inflation", "Core Inflation", "Headline Inflation", "CPI Inflation" and "WPI Inflation"; across each search category namely- "All Categories", "Finance Category", "Business and Industrial Category" and "News Category", that originated from India between 2006: Q3 and 2016: Q4.

Table 3 reports the estimation results for five searched terms, namely, "Inflation", "Core Inflation", "Headline Inflation", "CPI Inflation" and "WPI Inflation", under the section "All Categories". Within "All Categories', there are two sub-categories, namely "Web search" and "News search", the results of both are presented for each search word. While web search is general in nature (suppose a student is searching for "what is inflation"), news search caters to searches specific within the news section. For our context, news search is more relevant, although we present the result for web search for robustness check.

For all search terms across web search and news search, we find that the coefficients are both positive and significant. This indicates that people indeed use the Internet search engine to query about inflation and update their expectations accordingly. However, as the coefficients for past expectations of current inflation, across all specifications, are also positive and

significant, it implies that not all people are updating their expectations about future inflation.

Estimating Equation: $E_t[\pi_{t+1}] = \alpha_0 + \alpha_1 GT_t[\pi_{t+1}] + \alpha_2 E_{t-1}[\pi_t] + \epsilon_t$					
Search term	Version	α_1	α_2	$\overline{R^2}$	Durbin-Watson
Inflation	1	0.48 (2.40)**	0.86 (14.39)***	0.64	1.92
	2	0.40 (1.88)**	0.92 (20.08)***	0.45	1.96
Core Inflation	1	0.36 (1.82)*	0.88 (12.93)***	0.62	1.94
	2	0.57 (2.40)**	0.87 (15.23)***	0.38	2.03
Headline Inflation	1	0.43 (2.27)**	0.87 (14.68)***	0.64	1.86
	2	0.57 (2.56)**	0.83 (12.31)***	0.49	1.90
CPI Inflation	1	0.36 (1.88)*	0.87 (12.50)***	0.62	1.88
	2	0.66 (2.17)**	0.78 (7.74)***	0.37	1.82
WPI Inflation	1	0.44 (2.11)**	0.84 (10.85)***	0.63	1.93
	2	0.57 (2.27)**	0.83 (10.52)***	0.37	1.94

NOTE: Columns 3 and 4 reports the coefficient values with t-statistics at parenthesis. Columns 5 and 6 report diagnostics for each model and column 7 reports the methodology that is used for estimation. $GT_t[\pi_{t+1}]$ is data extracted from Google Trend for word **Inflation, "Core Inflation", "Headline Inflation", "CPI Inflation" and "WPI Inflation"**. $E_t[\pi_{t+1}]$ is 3 month ahead mean inflation expectation survey of household at time period t. Version 1 indicates web search and version 2 indicates the result of news search under all category. The time period for estimation is 2007Q4 to 2016Q3. *, ** , *** indicates 10, 5 and 1 per cent significance level respectively.

Table 3: All Category Results

For robustness check, we repeat the estimation exercise as done in Table 3, for the various inflation-related searches under specific search categories.

Tables 4, 5 and 6 present the results for searches for five inflation-related keywords as in Table 3, under the "Business and Industrial Category", "Finance Category" and "News Category" respectively. To ensure that people were searching for inflation-related news that might be different from

inflation-related other queries, we present two versions of the searches for each keyword which is web search and news search.

Similar to the results in Table 3, we find that irrespective of the search category, the epidemiological source of inflation expectations based on Internet search, is positive in all cases and significant in all but five cases. Inflation expectations formed in the previous period also continue to be positive and significant, thereby implying the agents' stickiness in updating expectations.

As in the previous cases, all versions of the epidemiology equations are checked for possible endogeneity by conducting the Durbin-Wu-Hausman test. In all cases, the null hypothesis that "OLS estimators are consistent" cannot be rejected, hence OLS estimators are adopted. To address the concern of auto-correlation, we check the Durbin-Watson statistic for each regression and for almost all equations there is no indication of auto-correlation.

Estimating Equation:

$$E_t[\pi_{t+1}] = \alpha_0 + \alpha_1 GT_t[\pi_{t+1}] + \alpha_2 E_{t-1}[\pi_t] + \epsilon_t$$

Search	Version	α_1	α_2	$\overline{R^2}$	Durbin-Watson
Inflation	1	0.42 (2.20)**	0.89 (16.80)***	0.63	1.94
	2	0.34 (1.56)	0.94 (22.61)***	0.43	1.99
Core Inflation	1	0.38 (2.06)**	0.89 (15.34)***	0.63	1.91
	2	0.80 (3.00)***	0.73 (8.00)***	0.30	1.85
Headline Inflation	1	0.50 (2.29)**	0.83 (10.98)***	0.47	1.88
	2	--	--	--	--
CPI Inflation	1	0.31 (1.55)	0.90 (13.97)***	0.61	1.93
	2	0.78 (3.58)***	0.80 (13.39)***	0.55	2.26
WPI Inflation	1	0.46 (2.65)**	0.85 (13.91)***	0.65	1.89
	2	0.34 (1.31)	0.92 (14.01)***	0.31	1.95

NOTE: Columns 3 and 4 reports the coefficient values with t-statistics at parenthesis. Columns 5 and 6 report diagnostics for each model and column 7 reports the methodology that is used for estimation. $GT_t[\pi_{t+1}]$ is data extracted from Google Trend for word **Inflation, "Core Inflation", "Headline Inflation", "CPI Inflation" and "WPI Inflation"**. $E_t[\pi_{t+1}]$ is 3 month ahead mean inflation expectation survey of household at time period t. Version 1 indicates web search and version 2 indicates the result of news search under business and industrial category. The time period for estimation is 2007Q4 to 2016Q3. *, ** , *** indicates 10, 5 and 1 per cent significance level respectively.

Table 4: Business and Industrial Category Results

Estimating Equation:

$$E_t[\pi_{t+1}] = \alpha_0 + \alpha_1 GT_t[\pi_{t+1}] + \alpha_2 E_{t-1}[\pi_t] + \epsilon_t$$

Search	Version	α_1	α_2	$\overline{R^2}$	Durbin-Watson
Inflation	1	0.49 (2.45)**	0.85 (12.87)***	0.64	1.90
	2	0.35 (1.50)	0.92 (17.08)***	0.43	1.99
Core Inflation	1	0.30 (1.53)	0.93 (18.97)***	0.61	1.94
	2	0.42 (1.89)*	0.90 (15.54)***	0.35	1.95
Headline Inflation	1	0.38 (2.09)**	0.88 (14.79)***	0.63	1.88
	2	0.36 (1.84)*	0.91 (16.34)***	0.45	1.96
CPI Inflation	1	0.39 (2.06)**	0.88 (13.82)***	0.59	1.87
	2	--	--	--	--
WPI Inflation	1	0.31 (1.71)*	0.91 (15.71)***	0.58	1.92
	2	--	--	--	--

NOTE: Columns 3 and 4 reports the coefficient values with t-statistics at parenthesis. Columns 5 and 6 report diagnostics for each model and column 7 reports the methodology that is used for estimation. $GT_t[\pi_{t+1}]$ is data extracted from Google Trend for word **Inflation, "Core Inflation", "Headline Inflation", "CPI Inflation" and "WPI Inflation"**. $E_t[\pi_{t+1}]$ is 3 month ahead mean inflation expectation survey of household at time period t. Version 1 indicates web search and version 2 indicates the result of news search under finance category. The time period for estimation is 2007Q4 to 2016Q3. *, ** , *** indicates 10, 5 and 1 per cent significance level respectively.

Table 5: Finance Category Results

To allay doubts that how do we distinguish between a general search on any inflation-related keyword and a specific search inflation-related news, we focus specifically on the findings of Table 6 that relate to searches in the "News Category" exclusively. The overall conclusion for news-specific searches do not differ much from the other categories, thereby indicating that irrespective of the category in which inflation-related keywords have been searched for, the epidemiological sources of inflation expectations are robust.

Estimating Equation:

$$E_t[\pi_{t+1}] = \alpha_0 + \alpha_1 GT_t[\pi_{t+1}] + \alpha_2 E_{t-1}[\pi_t] + \epsilon_t$$

Search	Version	α_1	α_2	$\overline{R^2}$	Durbin-Watson
Inflation	1	0.41 (2.15)**	0.89 (16.96)***	0.63	1.94
	2	0.35 (1.62)	0.94 (22.89)***	0.44	2.00
Core Inflation	1	0.28 (1.52)	0.91 (14.91)***	0.61	1.92
	2	0.70 (2.75)***	0.78 (9.31)***	0.28	1.87
Headline Inflation	1	0.54 (2.23)**	0.85 (11.71)***	0.37	1.93
	2	--	--	--	--
CPI Inflation	1	0.29 (1.48)	0.92 (15.75)***	0.61	1.96
	2	1.10 (4.60)***	0.69 (9.82)***	0.46	2.33
WPI Inflation	1	0.43 (2.14)**	0.87 (13.52)***	0.60	1.92
	2	0.49 (2.07)**	0.87 (12.96)***	0.36	1.99

NOTE: Columns 3 and 4 reports the coefficient values with t-statistics at parenthesis. Columns 5 and 6 report diagnostics for each model and column 7 reports the methodology that is used for estimation. $GT_t[\pi_{t+1}]$ is data extracted from Google Trend for word **Inflation, "Core Inflation", "Headline Inflation", "CPI Inflation" and "WPI Inflation"**. $E_t[\pi_{t+1}]$ is 3 month ahead mean inflation expectation survey of household at time period t. Version 1 indicates web search and version 2 indicates the result of news search under news category. The time period for estimation is 2007Q4 to 2016Q3. *, ** , *** indicates 10, 5 and 1 per cent significance level respectively.

Table 6: News Category Results

After establishing the role played by Internet search in inflation expectations formation, we next check if the Internet search "sentiments" affect inflation expectations by both the households and the professional forecasters. Following Lei et al. (2015) who had estimated how news sentiment (how news on inflation are reported) affect expectations formed about future inflations, we try to find out if search sentiments affect inflation expectations. It is expected that favorable searches should lead to lower inflation expectations while unfavorable searches are associated with higher inflation expectations. The direction of neutral search can be ambivalent, although Lei et al (2015) find neutral news lead to higher inflation expectations. Search sentiments are "favorable" when searches involve terms like "low inflation", "low prices" etc. "Unfavorable" search sentiment reflects searches like "high price", "price rise", "price high", "price increase", etc. Not enough searches were recorded for terms related to "neutral" search, thereby generating no data points by Google Trends.

$$E_t(\pi_{t+1}) = c + \alpha E_{t-1}(\pi_t) + \beta^{+} GT\_Unfavorable\ Search_t^{+} + \beta^{-} GT\_Favorable\ Search_t^{-} + \beta^{0} GT\_Netural\ Search_t^{0} + \mu_t$$

	α	β^{+}	β^{-}	$\overline{R^2}$	D-W Stats
Household Survey Data	0.91 (37.66)***	3.94 (1.78)*	-2.94 (-1.34)*	0.65	2.03
SPF_WPI	0.75 (12.12)***	12.04 (2.88)**	-11.04 (-2.64)**	0.66	1.28
SPF_CPI	0.87 (32.77)***	6.08 (2.57)**	-5.08 (-2.15)**	0.69	1.69

NOTE: T-stats are mentioned at parenthesis. Sample spans from 2006Q3 to 2016Q3. *, ** , *** indicates 10, 5 and 1 per cent significance level respectively.

Table 7: Search sentiment and inflation expectations formation

Table 7 indicates that unfavorable searches increase inflation expectations for both the general public and the professional forecasters. Coefficient of favorable search for the general public has the right sign but is not

significant. In case of professional forecasters, the favorable search has huge impact on its WPI forecasts.

For robustness check, we re-estimate the search sentiment equation by including the past realized inflation as in Lei et al. (2015). Table 8 indicates that when past inflation is included, the search sentiment does not have significant coefficients, although for the professional forecasters, the search sentiments are significant and have the expected signs.

$$E_t(\pi_{t+1}) = c + \alpha E_{t-1}(\pi_t) + \beta^{+} GT\_Unfavorable\ Search_t^{+} + \beta^{-} GT\_Favorable\ Search_t^{-} + \beta^{0} GT\_Netural\ Search_t^{0} + \varphi\pi_{t-1} + \mu_t$$

	α	β^{+}	β^{-}	Φ	$\overline{R^2}$	D-W Stats
Household Survey Data	0.77 (9.32)***	1.12 (0.42)	-0.12 (-0.04)	0.18 (1.79)*	0.67	1.93
	0.88 (18.35)***	2.32 (0.76)	-1.32 (-0.43)	0.07 (0.77)	0.64	1.98
SPF_CPI	0.29 (2.25)**	4.57 (2.40)**	-3.57 (-1.88)**	0.54 (4.48)***	0.80	1.42
	0.93 (16.48)***	7.94 (2.77)*	-6.94 (-2.42)*	-0.09 (-1.14)	0.69	1.93
SP_WPI	0.59 (5.19)***	12.54 (3.05)**	-11.54 (-2.81)**	0.12 (1.56)	0.68	1.24
	0.04 (0.20)	2.80 (0.68)	-1.80 (-0.44)	0.76 (4.11)***	0.77	1.08

NOTE: T-stats are mentioned at parenthesis. Sample spans from 2006Q3 to 2016Q3. The first row of each section comprises of lagged CPI inflation and the second row of each section is WPI inflation. *, ** , *** indicates 10, 5 and 1 per cent significance level respectively.

Table 8: Search sentiment and inflation expectations formation (robustness check)

Overall, the inferences drawn here match with Lei et al. (2015) that finds similar impact of media reports on inflation expectations.

5 Conclusion

This paper explores the epidemiological sources of inflation expectations formation of general public in India. While Carroll's (2003) benchmark study considers the inflation expectations of professional forecasters as the "common source" based on which people form their expectations, our work uses

the Internet search data (Google Trends) as the "common source" of information. Our results indicate that Internet search is a significant source of information for expectations updating. Additionally, we find that the Internet search sentiment, especially the negative sentiment, has a significant impact of inflation expectations formed by public. Overall, our findings conform to the findings of the epidemiology literature and conclude that the epidemiological sources of inflation expectations formation do exist for the Indian economy.

6 Reference

Carroll, Christopher D., (2003), Macroeconomic Expectations of Households and Professional Forecasters, The Quarterly Journal of Economics, 118, issue 1, p. 269-298.

Choi, Hyunyoung and Varian, Hal, (2012), Predicting the Present with Google Trends, The Economic Record, 88, p. 2-9,

Ehrmann, Michael, Pfajfar, Damjan and Santoro, Emiliano, (2014), Consumer Attitudes and the Epidemiology of Inflation Expectations, No 2014-029, Discussion Paper, Tilburg University, Center for Economic Research,

Lamla, Michael and Maag, Thomas, (2012), The Role of Media for Inflation Forecast Disagreement of Households and Professional Forecasters, Journal of Money, Credit and Banking, 44, issue 7, p. 1325-1350.

Lei, Chengyao, Lu, Zhe and Zhang, Chengsi, (2015), News on inflation and the epidemiology of inflation expectations in China, Economic Systems, 39, issue 4, p. 644-653.

Mohanty, D., (2012), The Importance of Inflation Expectations, Speech at the S P Jain Institute of Management & Research, (November 9), Mumbai, India.

Pfajfar, Damjan and Santoro, Emiliano, (2013), News on Inflation and the Epidemiology of Inflation Expectations, Journal of Money, Credit and Banking, 45, issue 6, p. 1045-1067.