Working Paper

On the accuracy of the probability method for quantifying beliefs about inflation

Author(s):
Maag, Thomas

Publication Date:
2009-07

Permanent Link:
https://doi.org/10.3929/ethz-a-005859391

Rights / License:
In Copyright - Non-Commercial Use Permitted
On the Accuracy of the Probability Method for Quantifying Beliefs about Inflation

Thomas Maag
On the Accuracy of the Probability Method for
Quantifying Beliefs about Inflation*

Thomas Maag†

July 2009

Abstract

This paper assesses the probability method for quantifying EU consumer survey data on perceived and expected inflation. Based on micro-data from the Swedish consumer survey that asks for both qualitative and quantitative responses, I find that the theoretical assumptions of the method do not hold. In particular, estimated models of response behavior indicate that qualitative inflation expectations are not ordered. Nevertheless, the probability method generates series that are highly correlated with the mean of actual quantitative beliefs. For quantifying the cross-sectional dispersion of beliefs, however, an index of qualitative variation outperforms the probability method.

JEL classification: C53, D84, E31

Keywords: quantification, inflation expectations, inflation perceptions, qualitative response data, belief formation

*I thank Simone Elmer, Sarah M. Lein, Christoph Moser, Rolf Schenker, and seminar participants at the KOF Swiss Economic Institute for helpful comments and suggestions. I am grateful to the Swedish National Institute of Economic Research for providing access to the Consumer Tendency Survey data. Financial support by the Swiss National Science Foundation is gratefully acknowledged.

†KOF Swiss Economic Institute, ETH Zurich, CH-8092 Zurich, Switzerland, E-mail: maag@kof.ethz.ch
1 Introduction

Surveys of households and firms are often qualitative. Rather than giving a quantitative estimate of a particular variable, respondents are asked to indicate their belief on a qualitative scale. In the European Union (EU), inflation perceptions and expectations of households are surveyed as part of the Joint Harmonized EU Consumer Survey programme. Within this framework, harmonized qualitative surveys are conducted in all member states, covering a national sample size of roughly 1,500 households on a monthly basis. The EU consumer survey thus provides an extensive and consistent data-set on beliefs about inflation.1 Consequently, the data has been investigated by a large literature. Only recently, the euro cash changeover and its effects on households’ inflation perceptions has given rise to a new strand of research.2 However, since the EU consumer survey is qualitative, most empirical applications rely on a method to quantify the qualitative response data in the first place. This paper assesses the validity of one particular method, the probability method for 5-category scales, and compares its accuracy to other quantification approaches.

Possibly the most widely used quantification method is the balance statistic proposed by Anderson (1952). It is originally defined as the difference between the share of respondents that perceive or expect positive inflation rates and the share of respondents that perceive or expect negative inflation rates. Theil (1952) rationalizes the balance statistic, demonstrating that it is an appropriate measure of the population mean if quantitative beliefs are uniformly distributed in the population. Furthermore, Theil (1952) suggests

---

1 Currently, the Joint Harmonized EU Consumer Survey covers a monthly sample of roughly 40,000 consumers in 27 member states. The consumer survey consists of 15 qualitative questions about the household’s financial situation, perceived economic conditions, and planned savings and spending. The questionnaire is translated into national languages and may include additional country specific questions, see European Commission (2007).

2 This literature centers on the rise in perceived inflation coinciding with the euro cash changeover, as documented in ECB (2005). Several explanations are being discussed, including increased information processing requirements due to conversion rates, overreaction to prices of frequently bought items, and anchoring of perceptions to prior expectations. See, e.g., Ehrmann (2006), Aucremanne, Collin, and Stragier (2007), Doehring and Mordonu (2007), Dziuda and Mastrobuoni (2006), Aalto-Setälä (2006), and Fluch and Stix (2007). Abstracting from the euro cash changeover, other contributions investigate belief formation in general, see, e.g., Döpke, Dovern, Fritsche, and Slacalek (2008), Forsells and Kenny (2004), Lamla and Lein (2008), and Lein and Maag (2008).
that the distributional assumption may be relaxed by imposing a normal distribution on
the quantitative beliefs instead. Combined with the assumption that respondents reply
that they perceive or expect prices to be constant if their quantitative belief is within an
indifference interval around 0% the mean and variance of the imposed distribution can be
identified. The model of Theil (1952) has been rediscovered by Carlson and Parkin (1975)
and is known today as the Carlson-Parkin method or the 3-category probability method.3
Batchelor and Orr (1988) extend the probability method to response data on 5-category
scales as it is available from the EU consumer survey. Taking into account the particular
questioning in the EU consumer survey, Berk (1999) additionally suggests an identification
scheme that links inflation expectations to inflation perceptions.

The goal of this paper is to assess the 5-category probability method and to derive
lessons for applied research. The analysis relies on joining qualitative and quantitative
response data on household level which is available from the Swedish Consumer Tendency
Survey from 01/1996 to 10/2008. A similar approach has been adopted, to the best of
my knowledge, only by Defris and Williams (1979) and Batchelor (1986). Defris and
Williams (1979) investigate a 5-year sample from an Australian consumer survey. They
document that the balance statistic as well as the 3-category probability method generate
series that are only weakly correlated with quantitative survey responses. Batchelor (1986)
In line with Defris and Williams (1979), Batchelor (1986) finds that both the balance
statistic and quantified expectations using the probability method track quantitative survey
responses only inaccurately, in particular in the short term. This result is in contrast to
my findings for Sweden.

This paper extends the existing literature in several respects. First, a detailed assessment
of the theoretical assumptions of the 5-category probability method is provided. Exist-

---

3 A less common quantification method is the regression approach of Pesaran (1987). The regression
method extends the balance statistic, allowing for a non-linear relation between response shares and
quantitative beliefs. The method is outlined in Section 4. Pesaran (1987) discusses the three-category
probability method and the regression approach in detail. For a recent survey of quantification methods
see Nardo (2003).
ing research focuses on the 3-category probability method and on testing distributional assumptions. Joining quantitative and qualitative data on household level allows to estimate unrestricted response schemes. The restrictions imposed by the 5-category probability method can then be tested using likelihood theory. Second, the accuracy of the 5-category probability method relative to the mean and cross-sectional standard deviation of quantitative survey responses is assessed in a long sample of 154 monthly observations. The discussion centers on comparing correlation coefficients employing the Fisher $z$-transformation and double block bootstrap confidence intervals. Accuracy is compared to a set of alternative quantification methods, including the 3-category probability method, the balance statistic, and the regression approach. For quantifying the cross-sectional heterogeneity of beliefs the set of alternatives includes the 3-category probability method, an index of qualitative variation, an index of ordinal variation, and the disconformity index. Third, I assess the probability method both for quantifying households’ inflation expectations as well as inflation perceptions.

To anticipate a central result, the 5-category probability method performs well for quantifying the mean of beliefs, despite that its underlying assumptions are rejected. For quantifying the cross-sectional heterogeneity, however, an index of qualitative variation is most accurate.

The paper is structured as follows. Section 2 presents the data and discusses some key statistical properties. Section 3 assesses the assumptions of the 5-category probability method. Section 4 investigates the accuracy of the approach and contrasts it with alternative methods. Section 5 draws lessons for applied research. Section 6 concludes.

2 Data

Inflation perceptions and expectations of Swedish households are being surveyed on a monthly basis since 1973. Unlike most surveys in other countries, the Swedish Consumer Tendency Survey simultaneously records qualitative and quantitative beliefs about infla-
tion. This paper uses monthly household-level response data spanning 01/1996–10/2008. During this period, the survey comprises a representative monthly sample of roughly 1,500 households which are interviewed by telephone.

The questionnaire of the Swedish Consumer Tendency Survey is based on the Joint Harmonized EU Consumer Survey questionnaire. Households’ beliefs about inflation are recorded in two steps. In a first step, the respondent is asked to report perceived inflation on a five-category ordinal scale. The question reads:

“Compared with 12 months ago, do you find that prices in general are...?” “Lower (S₁), about the same (S₂), a little higher (S₃), somewhat higher (S₄), a lot higher (S₅), don’t know”.

Following, S₁ through S₅ denote the qualitative response categories, while s₁ through s₅ are the fractions of answers in the respective qualitative response category. In a second step, the respondent is asked for a direct quantitative estimate of the current inflation rate. The question reads: “How much higher/lower in percent do you think prices are now? (In other words, the present rate of inflation)”. In a similar manner, expected inflation is first captured by asking:

“Compared to the situation today, do you think that in the next 12 months prices in general will...?” “Go down a little (S₁), stay more or less the same (S₂), go up more slowly (S₃), go up at the same rate (S₄), go up faster (S₅), don’t know”.

Quantitative estimates are recorded in a second step by asking: “Compared with today, how much in percent do you think prices will go up/down? (In other words, inflation/deflation 12 months from now)”. As a result and in contrast to the Joint Harmonized EU Consumer Survey households report qualitative as well as quantitative inflation beliefs.

---

4The sample horizon is limited by data availability. Before 1996 qualitative responses were only recorded on a 3 option ordinal scale and quantitative beliefs were only surveyed on a quarterly basis.


6Response shares are computed excluding the “don’t know” category, i.e. s₁ through s₅ sum up to 100%.
In line with the literature, the quantitative response data is adjusted for outliers. Responses outside the interval \([-30\%, 30\%]\) are omitted which reduces the sample size by 0.3\%. Moreover, only observations that simultaneously contain non-missing responses to qualitative and quantitative questions are considered. Dropping respondents that only answer the qualitative question on perceived (expected) inflation further reduces the sample size by 13\% (15\%). As will be discussed in the next section, a theoretical assumption of all quantification methods considered here is that a respondent needs to form a quantitative belief in the first place to to express a qualitative opinion at a later stage. The observed response pattern can be considered as evidence against this assumption. However, an alternative interpretation is that qualitative responses with missing quantitative responses are uninformed and should be attributed to the “don’t know” category.

The resulting sample comprises 154 monthly surveys spanning 01/1996–10/2008, which amounts to almost 200,000 observations. Throughout this paper the discussion centers on this sample. The appendix additionally presents estimation results for the shorter 01/2002–10/2008 sample. Considering this period accounts for a potential structural break due to a change in the surveying institution in 01/2002. As will be shown, results for both estimation periods are consistent, confirming the validity of the results for the 01/1996–10/2008 sample.

As a measure of actual inflation I use the year-over-year percent change in the Harmonized Index of Consumer Prices (HICP) as published by Eurostat. Over the entire sample period, 667 observations include inflation perceptions and 443 observations include inflation expectations that are classified as outliers. This view is supported by the distribution of missing quantitative answers by qualitative answer category. For inflation perceptions, about 70\% of missing quantitative responses are assigned to respondents that opt for the qualitative response category \(S_3\) (“a little higher”). For inflation expectations, 40\% of missing quantitative responses are assigned to respondents that opt for the qualitative response category \(S_4\) (“go up at the same rate”).

In 01/2002 the surveying institution has changed from Statistics Sweden to GfK Sweden. The change goes along with a decline in the share of missing quantitative responses. This might be partly due to differences in the questioning, as outlined by Palmqvist and Strömberg (2004). However, the share of missing observations rises again sharply in 2008 to levels before 2002. Hence, part of the initial decline in the share of missing quantitative responses appears to be pure coincidence.

I have also considered the Consumer Price Index (CPI) and the Consumer Price Index excluding mortgage payments and indirect taxes (CPIX). In particular at the beginning of the sample period these indices might have obtained more attention by the Swedish public than the HICP. Employing these
Figure 1: Qualitative response shares and mean quantitative response

Notes: The lowest grey line shows the share $s_1$ of qualitative answers in the category $S_1$, the second-lowest grey line shows the cumulative share of answers in the categories $S_1$ and $S_2$, etc.

Figure 1 shows qualitative response shares, the mean of quantitative responses, and actual inflation over time. The figure indicates that quantitative inflation perceptions closely track actual inflation. The correlation coefficient between the two series is 0.78, the root mean squared error (RMSE) is 0.46%. The high correlation is particularly noteworthy as the consumer survey does not ask specifically about HICP inflation, but about the evolution of prices “in general”. Moreover, the first panel of Table 1 shows that the overall mean of inflation perceptions is 1.79%, as opposed to an average HICP inflation rate of 1.61%. Hence, inflation perceptions of the Swedish public are only slightly biased, which is in line with earlier findings of Jonung and Laidler (1988). In light of quantitative survey results for the U.S. (Bryan and Venktatu, 2001a, 2001b) and for the U.K. (Driver and Windram, 2007) inflation perceptions of the Swedish public are relatively accurate.\footnote{Relying on a monthly household survey conducted by the Federal Reserve Bank of Cleveland, Bryan and Venktatu (2001a, 2001b) find that inflation perceptions (and expectations) are biased by several percentage points. For the U.K., Driver and Windram (2007) report a correlation coefficient between actual and perceived inflation of roughly 0.5 over a similar sample period.} Figure 1 shows that inflation expectations also follow actual inflation quite closely. The contemporaneous correlation of the mean expectation with actual inflation is 0.70, the RMSE is 0.66%.

alternative indices does not significantly alter the results.
Table 1: Summary statistics of quantitative responses

<table>
<thead>
<tr>
<th></th>
<th>Perceptions</th>
<th>Expectations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall</td>
<td>Min</td>
</tr>
<tr>
<td>Observations</td>
<td>197,487</td>
<td>1,031</td>
</tr>
<tr>
<td>Number of months</td>
<td>154</td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1.81</td>
<td>0.47</td>
</tr>
<tr>
<td>Median</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>4.06</td>
<td>2.68</td>
</tr>
<tr>
<td>Share of integer answers</td>
<td>0.94</td>
<td>0.86</td>
</tr>
<tr>
<td>Share of focal point answers</td>
<td>0.81</td>
<td>0.50</td>
</tr>
<tr>
<td>Share of zero responses</td>
<td>0.59</td>
<td>0.08</td>
</tr>
<tr>
<td>Mean response, given $S_1$</td>
<td>-4.79</td>
<td>-20.00</td>
</tr>
<tr>
<td>Mean response, given $S_2$</td>
<td>0.03</td>
<td>-0.05</td>
</tr>
<tr>
<td>Mean response, given $S_3$</td>
<td>4.72</td>
<td>3.98</td>
</tr>
<tr>
<td>Mean response, given $S_4$</td>
<td>7.94</td>
<td>5.22</td>
</tr>
<tr>
<td>Mean response, given $S_5$</td>
<td>9.58</td>
<td>4.25</td>
</tr>
</tbody>
</table>

Notes: The column overall presents results for the entire sample spanning 01/1996–10/2008. Min and Max are the monthly minimum and maximum of the respective statistic. All shares are relative to the overall number of observations.

Predictive power is relatively low, as the correlation with 12 months ahead inflation is only 0.28, with the RMSE increasing to 0.97%. Table 1 documents that the sample mean of inflation expectations lies somewhat higher at 2.05%. Figure 1 shows that inflation perceptions surge in 2008, exceeding actual inflation by roughly 2%. Meanwhile, the share of qualitative responses in the lowest two categories $S_1$ and $S_2$ ("lower" and "about the same") sharply declines. Finally, Figure 1 reveals a systematic difference between qualitative inflation perceptions and expectations. Qualitative inflation perceptions are concentrated in categories $S_2$ and $S_3$ ("about the same" and "a little higher"). In 1996–2008, 87% of respondents opt for these categories. In contrast, 70% of qualitative inflation expectations fall into answer categories $S_2$ and $S_5$ ("stay more or less the same" and "go up faster").

Table 1 highlights a number of interesting properties of the quantitative response data. First, panel 1 shows that inflation perceptions and expectations are highly heterogeneous. Despite the low mean of both variables, inflation perceptions and expectations exhibit a cross-sectional standard deviation of 4.06% and 3.73% respectively. Second, panel 2
indicates that more than 90% of all quantitative answers are integers. Third, these integer answers are concentrated at a few focal points. Both for perceptions and expectations, the most frequently mentioned focal points are -5%, -2%, 0%, 2%, 5%, 10%. As Table 1 indicates, the most important focal point both for perceptions and expectations is 0%, which accounts for more than half of all focal point responses. Towards the end of the sample horizon, the share of zero responses declines significantly, attaining a minimum of 8% for perceptions and 12% for expectations in 06/2008. The high share of zero responses also explains the low median of quantitative beliefs. The median of inflation perceptions is 0%, while the median of inflation expectations is 1%. Fourth, mean quantitative inflation perceptions are generally rising in the qualitative response category. This is shown by panel 3 of Table 1 which summarizes the mean of quantitative responses conditional on the qualitative response the respondent simultaneously opts for. However, this is not the case for inflation expectations. For expectations, the mean response given $S_4$ ("go up at the same rate") is higher than the mean response given $S_5$ ("go up faster"). Also, in comparison to inflation perceptions the differences between the cross-sectional mean of expectations given qualitative answers $S_3$, $S_4$, $S_5$ are only minor. Fifth, the relation between quantitative and qualitative responses is time varying. The differences between mean, minimum, and maximum are considerable for most categories. The only exception is $S_2$ ("about the same"): given this qualitative answer the mean quantitative response is always close to 0%.

These initial results suggest that the relation between quantitative and qualitative inflation beliefs is complex. The response scheme, i.e. the formal relation between quantitative and qualitative responses, appears to be time varying. Moreover, the conditional mean of quantitative expectations is not monotonously rising in the order of the qualitative

---

12 In line with Bryan and Palmqvist (2006) focal points are defined as integers that are mentioned more often than their neighboring integers. I have not found any evidence for important non-integer focal points.

13 Of the remaining integers, 1, 3, and 4 obtain the highest response shares. This set accounts for 9% of quantitative inflation perceptions and for 15% of expectations.

14 On a monthly basis the mean of inflation perceptions is not always strictly rising too, as indicated by the minima of monthly means in panel 3 of Table 1. But this only is the case in 27 months as opposed to 136 months in which the conditional mean of inflation expectations is not monotonously rising.
response categories. While the 5-category probability method allows for a time varying response scheme, it imposes a certain symmetry on the response scheme and requires ordered qualitative data. Regarding the distributional assumptions, the above results indicate that quantitative beliefs are not normally distributed. Since the mean exceeds the median the distributions of inflation perceptions and expectations are positively skewed. The concentration of answers at focal points and in particular at 0% raises doubt whether any of the common parametric distributions adequately describes the quantitative response data. The next section thus discusses in detail whether the assumptions of the probability method are consistent with the data.

3 Validity of the Probability Method

This section tests the main theoretical assumptions of the 5-category probability method for quantifying qualitative response data. Building on contributions of Theil (1952) and Carlson and Parkin (1975), the 5-category probability method has been introduced by Batchelor and Orr (1988). To begin with, this method is briefly outlined.

Assume that previous to answering the consumer survey, respondent \( i \) forms a quantitative belief \( \pi_{it}^e \) about inflation over the upcoming 12 months\(^{15}\). Respondent \( i \) then answers the qualitative survey question on expected inflation according to the following response scheme:

\[ \begin{align*}
\pi_{it}^e < -\delta_t & : \text{ prices in general will go down a little} \ (S_1) \\
-\delta_t & \leq \pi_{it}^e < \delta_t : \text{ stay more or less the same} \ (S_2) \\
\delta_t & \leq \pi_{it}^e < \pi_r^t - \eta_t : \text{ go up more slowly} \ (S_3) \\
\pi_r^t - \eta_t & \leq \pi_{it}^e < \pi_r^t + \eta_t : \text{ go up at the same rate} \ (S_4) \\
\pi_{it}^e & \geq \pi_r^t + \eta_t : \text{ go up faster} \ (S_5)
\end{align*} \]  

(1)

The response scheme is defined by the parameters \( \delta_t, \eta_t, \pi_r^t \). In the following, \( \pi_r^t \) is

\(^{15}\)The analogous approach for quantifying perceived inflation \( \pi_{it}^p \) and detailed derivations can be found in the Appendix.
called reference inflation. It is the inflation rate that people have in mind when opting for answer $S_4$ (“prices will go up at the same rate” and, for inflation perceptions, “prices are moderately higher”). A first key assumption of the probability method restricts the response scheme to be fully defined by three parameters:

**Assumption 1:** The response intervals are symmetric around 0% and around $\pi_r^t$.

The corresponding intervals $[-\delta_t, \delta_t]$ and $[\pi_r^t - \eta_t, \pi_r^t + \eta_t]$ correspond to qualitative responses $S_2$ and $S_4$ respectively. A second assumption imposes structural homogeneity on the response scheme:

**Assumption 2:** Threshold parameters $\delta_t$ and $\eta_t$ and the reference inflation rate $\pi_r^t$ are identical across all respondents.

Quantitative inflation expectations $\pi^e_i$ will vary across respondents due to differences in information sets and information processing. To infer the mean quantitative inflation expectation from qualitative response shares the probability method imposes a distributional assumption on $\pi^e_i$. The standard assumption is that the cross-sectional distribution of quantitative beliefs is normal:

**Assumption 3:** The cross-sectional distribution of quantitative beliefs is normal, i.e.: $$\pi^e_i \sim N(\pi^e_t, \sigma^e_t).$$

The parameters of interest are the cross-sectional mean $\pi^e_t$ and standard deviation $\sigma^e_t$ of quantitative beliefs. As outlined in the Appendix, the above assumptions yield a system of 4 linearly independent equations with 5 unknowns ($\pi^e_t$, $\sigma^e_t$, $\delta_t$, $\eta_t$, $\pi_r^t$) which can be solved for $\pi^e_t$ and $\sigma^e_t$. The solution for both parameters is equal to the product of reference inflation $\pi_r^t$ and a function of the response shares $s_1^t, \ldots, s_5^t$.

The usual identification scheme restricts reference inflation $\pi_r^t$. For quantifying inflation expectations two choices of $\pi_r^t$ are apparent. First, reference inflation can be set equal to some actual rate of inflation, assuming that the respondent knows the actual rate of inflation and answers the question relative to this value. Second, reference inflation can
be set equal to previously quantified perceived inflation $\pi^p_t$ as suggested by Berk (1999). This approach is supported by empirical evidence that households are not necessarily well informed about actual inflation.\textsuperscript{16} Identifying $\pi^r_t$ is less obvious for inflation perceptions. Relying on Carlson and Parkin (1975) it is commonly assumed that inflation perceptions are unbiased over the sample horizon. This assumption can be imposed by restricting $\pi^r_t$ to a constant accordingly.\textsuperscript{17} The last assumption thus reads:

Assumption 4: The reference rate of inflation $\pi^r_t$ for quantifying inflation expectations is equal to some actual or quantified perceived rate of inflation. The reference rate of inflation for quantifying inflation perceptions is time invariant.

### 3.1 Symmetry of the Response Scheme

Assumption 1 restricts response intervals to be symmetric around 0\% and around $\pi^r_t$. To test the validity of this assumption I estimate an unrestricted response scheme with 4 threshold parameters. Assume that respondent $i$ answers the qualitative question according to the following scheme:\textsuperscript{18}

\begin{align}
\pi^e_{it} + \varepsilon_{it} < \mu^1_t & : \text{prices in general will go down a little (S}_1) \\
\mu^1_t \leq \pi^e_{it} + \varepsilon_{it} < \mu^2_t & : \text{stay more or less the same (S}_2) \\
\mu^2_t \leq \pi^e_{it} + \varepsilon_{it} < \mu^3_t & : \text{go up more slowly (S}_3) \\
\mu^3_t \leq \pi^e_{it} + \varepsilon_{it} < \mu^4_t & : \text{go up at the same rate (S}_4) \\
\pi^e_{it} + \varepsilon_{it} \geq \mu^4_t & : \text{go up faster (S}_5) \\
\end{align}

The idiosyncratic component $\varepsilon_{it}$ allows the response scheme to shift between individuals. Under the assumption that the idiosyncratic component represents the sum of various, independent individual specific factors it is reasonable to assume that $\varepsilon_{it}$ is normally distributed. This reasoning is in line with the normality assumption commonly

\textsuperscript{16}See, e.g., Bryan and Venkatu (2001a, 2001b) who document that inflation perceptions of U.S. households are significantly biased.

\textsuperscript{17}The solution for $\pi^r_t$ is given by Equation (A.8) in the Appendix.

\textsuperscript{18}The identical scheme is applied to inflation perceptions, where $\pi^e_{it}$ is replaced by $\pi^p_{it}$.
imposed by the probability method. One thus obtains an ordered probit model (Zavoina and McKelvey, 1975). Deviating from the usual identification scheme, I restrict the coefficient on the quantitative belief $\pi e_{it}$ to unity, while the variance of $\varepsilon_{it}$ remains unrestricted.

With $\varepsilon_{it} \sim N(0, \sigma_{it})$ the following probabilities are obtained:

$$P(\text{lower}|\pi e_{it}, \mu_{it}, \sigma_{it}) = P(S_1|\pi e_{it}, \mu_{it}, \sigma_{it}) = \Phi \left( \frac{-\pi e_{it} + \mu_1}{\sigma_{it}} \right)$$

$$P(\text{about the same}|\pi e_{it}, \mu_{it}, \sigma_{it}) = P(S_2|\pi e_{it}, \mu_{it}, \sigma_{it}) = \Phi \left( \frac{-\pi e_{it} + \mu_2}{\sigma_{it}} \right) - \Phi \left( \frac{-\pi e_{it} + \mu_1}{\sigma_{it}} \right)$$

$$P(\text{a little higher}|\pi e_{it}, \mu_{it}, \sigma_{it}) = P(S_3|\pi e_{it}, \mu_{it}, \sigma_{it}) = \Phi \left( \frac{-\pi e_{it} + \mu_3}{\sigma_{it}} \right) - \Phi \left( \frac{-\pi e_{it} + \mu_2}{\sigma_{it}} \right) - \Phi \left( \frac{-\pi e_{it} + \mu_1}{\sigma_{it}} \right)$$

$$P(\text{moderately higher}|\pi e_{it}, \mu_{it}, \sigma_{it}) = P(S_4|\pi e_{it}, \mu_{it}, \sigma_{it}) = \Phi \left( \frac{-\pi e_{it} + \mu_4}{\sigma_{it}} \right) - \Phi \left( \frac{-\pi e_{it} + \mu_3}{\sigma_{it}} \right) - \Phi \left( \frac{-\pi e_{it} + \mu_2}{\sigma_{it}} \right) - \Phi \left( \frac{-\pi e_{it} + \mu_1}{\sigma_{it}} \right)$$

$$P(\text{a lot higher}|\pi e_{it}, \mu_{it}, \sigma_{it}) = P(S_5|\pi e_{it}, \mu_{it}, \sigma_{it}) = 1 - \Phi \left( \frac{-\pi e_{it} + \mu_4}{\sigma_{it}} \right)$$

(3)

where $\mu_t = \{\mu_1, ..., \mu_4\}$. Unlike in the probability method, $\mu_1, \mu_2$ and $\mu_3, \mu_4$ are not required to be symmetric around 0% and a reference rate of inflation respectively.

Table 2 presents the maximum likelihood estimates of this response scheme. By construction, the threshold parameters are rising in the qualitative response category. The model is confirmed by highly significant parameter estimates which are stable across subperiods. The width of the interval $[\mu_1, \mu_2]$ corresponding to qualitative response $S_2$ ("about the same") exceeds 8% both for perceptions and expectations. The estimated parameters suggest that this interval is not symmetric around 0%. Relying on maximum likelihood theory I test the restriction that $\mu_1 = -\mu_2$ with a likelihood ratio test.\(^{19}\) The second panel of Table 2 shows that this test clearly rejects the null hypothesis of symmetry.

The estimates confirm the systematic differences between quantitative perceptions and expectations that have been noticed before. While the threshold parameters for inflation perceptions are increasing from $\mu_1 = -6.91\%$ to $\mu_4 = 13.89\%$, the thresholds for inflation

\(^{19}\)The test statistic is given by $LR = -2 (\log L_r - \log L_i) \rightarrow \chi^2(q)$, where $\log L_r$ is the log likelihood of the restricted model and $\log L_i$ is the log likelihood of the unrestricted model. The number of restrictions is given by $q = 1$. 

13
Table 2: Estimated response scheme for perceived and expected inflation

<table>
<thead>
<tr>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Perceptions</td>
<td>Expectations</td>
<td>Perceptions</td>
<td>Expectations</td>
</tr>
<tr>
<td>$\mu_1$</td>
<td>-6.909***</td>
<td>-7.404***</td>
<td>-7.595***</td>
<td>-7.883***</td>
</tr>
<tr>
<td></td>
<td>(0.0300)</td>
<td>(0.0367)</td>
<td>(0.0457)</td>
<td>(0.0556)</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td>3.194***</td>
<td>1.739***</td>
<td>2.526***</td>
<td>0.617***</td>
</tr>
<tr>
<td></td>
<td>(0.0138)</td>
<td>(0.0141)</td>
<td>(0.0188)</td>
<td>(0.0230)</td>
</tr>
<tr>
<td>$\mu_3$</td>
<td>10.95***</td>
<td>2.797***</td>
<td>10.97***</td>
<td>2.005***</td>
</tr>
<tr>
<td></td>
<td>(0.0267)</td>
<td>(0.0141)</td>
<td>(0.0362)</td>
<td>(0.0212)</td>
</tr>
<tr>
<td>$\mu_4$</td>
<td>13.89***</td>
<td>5.609***</td>
<td>14.04***</td>
<td>5.682***</td>
</tr>
<tr>
<td></td>
<td>(0.0388)</td>
<td>(0.0174)</td>
<td>(0.0520)</td>
<td>(0.0251)</td>
</tr>
<tr>
<td>$\sigma_t$</td>
<td>3.782***</td>
<td>4.427***</td>
<td>4.260***</td>
<td>5.083***</td>
</tr>
<tr>
<td></td>
<td>(0.0118)</td>
<td>(0.0163)</td>
<td>(0.0181)</td>
<td>(0.0270)</td>
</tr>
<tr>
<td>$\gamma$</td>
<td>0</td>
<td>0</td>
<td>0</td>
<td>0</td>
</tr>
<tr>
<td>$\gamma_0$</td>
<td>1</td>
<td>1</td>
<td>1</td>
<td>1</td>
</tr>
</tbody>
</table>

N                      | 197487     | 192845    | 110071     | 109782    |
Log $L$                | -138263    | -220521   | -88738     | -136562   |

Likelihood ratio tests
$H_0^1$: Symmetry such that $\mu_2 = -\mu_1$ ($q = 1$)
LR statistic          | 20978.94   | 35749.40  | 17427.61   | 25138.60  |
P-value                | 0.00       | 0.00      | 0.00       | 0.00      |

Notes: This table shows maximum likelihood estimates of the unrestricted response scheme (2). Monthly data, 01/1996–10/2008 and 01/2002–10/2008. $N$ is the number of observations, log $L$ is the log likelihood, standard errors in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level.

For inflation expectations, the threshold parameters that define the response intervals for $S_3, S_4, S_5$ are in a narrow range of 4% to 5%. This allows for two interpretations: either, the response intervals are indeed narrower for inflation expectations than for inflation perceptions, or the ordered model does not adequately describe the formation of qualitative inflation expectations in the positive region. The second interpretation is supported by the relatively high estimated standard deviation $\sigma_t$ and a substantially lower log likelihood than for inflation perceptions (despite a lower number of observations). That qualitative expectations are not necessarily ordered is indeed suggested by the relative nature of the response categories. The qualitative response $S_4$ ("prices will go up at the same rate") anchors the qualitative expectation to the perception of current inflation. A respondent who has an identical quantitative
## Table 3: Estimated relative response scheme for expected inflation

<table>
<thead>
<tr>
<th>Expectations</th>
<th>Income groups (1st quartile lowest)</th>
<th>1st quartile</th>
<th>2nd quartile</th>
<th>3rd quartile</th>
<th>4th quartile</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu^1_t$</td>
<td>$-7.172^{***}$</td>
<td>-6.014^{***}</td>
<td>-5.056^{***}</td>
<td>-3.928^{***}</td>
<td>-5.316^{***}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.184)</td>
<td>(0.116)</td>
<td>(0.0569)</td>
<td>(0.0132)</td>
<td>(0.0500)</td>
<td></td>
</tr>
<tr>
<td>$\mu^2_t$</td>
<td>1.322^{***}</td>
<td>1.059^{***}</td>
<td>0.805^{***}</td>
<td>0.575^{***}</td>
<td>0.814^{***}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.0839)</td>
<td>(0.0550)</td>
<td>(0.0445)</td>
<td>(0.0104)</td>
<td>(0.0242)</td>
<td></td>
</tr>
<tr>
<td>$\sigma_t$</td>
<td>6.212^{***}</td>
<td>4.956^{***}</td>
<td>4.056^{***}</td>
<td>3.182^{***}</td>
<td>4.369^{***}</td>
<td></td>
</tr>
<tr>
<td></td>
<td>(0.139)</td>
<td>(0.0826)</td>
<td>(0.0062)</td>
<td>(0.0039)</td>
<td>(0.0347)</td>
<td></td>
</tr>
<tr>
<td>N</td>
<td>10046</td>
<td>15133</td>
<td>15584</td>
<td>19618</td>
<td>60381</td>
<td></td>
</tr>
<tr>
<td>Log $L$</td>
<td>-8759</td>
<td>-12591</td>
<td>-12436</td>
<td>-14921</td>
<td>-49217</td>
<td></td>
</tr>
<tr>
<td>Likelihood ratio tests</td>
<td>LR statistic</td>
<td>P-value</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H^1_0$: Identical threshold parameters $\mu^1_t$, $\mu^2_t$ ($q = 6$)</td>
<td>746.60</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H^2_0$: Identical standard deviation $\sigma_t$ ($q = 3$)</td>
<td>907.96</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$H^3_0$: Identical thresholds and standard deviation ($q = 9$)</td>
<td>1020.93</td>
<td>0.00</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table shows maximum likelihood estimates of the relative response scheme (4). Monthly data, 01/2002–10/2008. Standard errors in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level.

...
are more responsive to an increase of quantitative expectations over perceptions than to a decrease. The significant estimates also reveal that, despite the lack of an unambiguous relation between quantitative and qualitative expectations, the 5-category qualitative survey contains more information than a 3-option survey that does not distinguish between $S_3$, $S_4$, and $S_5$.

In sum, the results indicate that Assumption 1 is not satisfied. Under the imposed normality assumption the estimated response interval is not symmetric around $0\%$, both for inflation perceptions and expectations.\(^{20}\) Furthermore, for expectations the response interval is not symmetric around $\pi^r_t$. The estimations confirm that qualitative inflation expectations are formed relative to perceived inflation. This result suggests that the link between expectations and perceptions should be exploited in quantifying qualitative responses. Additional evidence on the relation of reference inflation and quantitative perceptions is provided in Section 3.4.

### 3.2 Homogeneity of the Response Scheme

Assumption 2 imposes that threshold parameters $\delta_t$ and $\eta_t$ and reference inflation $\pi^r_t$ do not systematically differ across respondents. Since the Swedish data-set only contains one observation per individual this assumption is tested by estimating the response scheme for different income groups.\(^{21}\) Tables 4 and 5 summarize the estimation results for perceptions and expectations respectively. The tables show that the absolute values of threshold parameters tend to decline in income. The lower panel of Tables 4 and 5 present likelihood ratio tests of three restrictions. The null hypotheses considered are that threshold parameters are identical across income groups ($H^1_0$), that standard deviation is identical across income groups ($H^2_0$), and that both threshold parameters and standard deviation

\(^{20}\)This finding is consistent with Henzel and Wollmershäuser (2005) who investigate data from a special edition of the ifo World Economic Survey that directly asks to quantitatively indicate the indifference interval. Henzel and Wollmershäuser (2005) report that the positive threshold parameter is larger in absolute terms than the negative parameter. As opposed to the Swedish survey, however, the ifo survey queries professional forecasters and answers are given on a 3-category ordinal scale.

\(^{21}\)I have also considered education groups, with unchanged qualitative results.
Table 4: Estimated response scheme for perceived inflation by income groups

<table>
<thead>
<tr>
<th>Perceptions</th>
<th>1st quartile</th>
<th>2nd quartile</th>
<th>3rd quartile</th>
<th>4th quartile</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\mu^1_t$</td>
<td>-9.074***</td>
<td>-8.201***</td>
<td>-6.657***</td>
<td>-5.833***</td>
<td>-7.308***</td>
</tr>
<tr>
<td></td>
<td>(0.142)</td>
<td>(0.105)</td>
<td>(0.0821)</td>
<td>(0.0627)</td>
<td>(0.0459)</td>
</tr>
<tr>
<td>$\mu^2_t$</td>
<td>2.411***</td>
<td>2.487***</td>
<td>2.405***</td>
<td>2.366***</td>
<td>2.479***</td>
</tr>
<tr>
<td></td>
<td>(0.0571)</td>
<td>(0.0429)</td>
<td>(0.0352)</td>
<td>(0.0271)</td>
<td>(0.0192)</td>
</tr>
<tr>
<td>$\mu^3_t$</td>
<td>11.55***</td>
<td>11.68***</td>
<td>10.63***</td>
<td>9.405***</td>
<td>10.79***</td>
</tr>
<tr>
<td></td>
<td>(0.100)</td>
<td>(0.0824)</td>
<td>(0.0704)</td>
<td>(0.0541)</td>
<td>(0.0372)</td>
</tr>
<tr>
<td>$\mu^4_t$</td>
<td>14.60***</td>
<td>15.07***</td>
<td>13.53***</td>
<td>12.16***</td>
<td>13.77***</td>
</tr>
<tr>
<td></td>
<td>(0.136)</td>
<td>(0.119)</td>
<td>(0.104)</td>
<td>(0.0836)</td>
<td>(0.0535)</td>
</tr>
<tr>
<td>$\sigma_t$</td>
<td>5.232***</td>
<td>4.682***</td>
<td>3.759***</td>
<td>3.179***</td>
<td>4.120***</td>
</tr>
<tr>
<td></td>
<td>(0.0583)</td>
<td>(0.0425)</td>
<td>(0.0326)</td>
<td>(0.0238)</td>
<td>(0.0183)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Likelihood ratio tests</th>
<th>LR statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0^1$: Identical threshold parameters $\mu^1_t, \ldots, \mu^4_t$ ($q = 12$)</td>
<td>1868.21</td>
<td>0.00</td>
</tr>
<tr>
<td>$H_0^2$: Identical standard deviation $\sigma_t$ ($q = 3$)</td>
<td>1115.68</td>
<td>0.00</td>
</tr>
<tr>
<td>$H_0^3$: Identical thresholds and standard deviation ($q = 15$)</td>
<td>2080.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Notes: This table shows maximum likelihood estimates of the unrestricted response scheme (2). Monthly data, 01/2002–10/2008. Standard errors in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level.

are identical across income groups ($H_0^3$). All three hypotheses are clearly rejected.

Table 3 presents estimation results for the relative response scheme (4) that links expected inflation to perceived inflation. Again, all three hypotheses are clearly rejected. The estimates show the same pattern as above: the absolute values of threshold parameters are declining in income. Overall, these results suggest that the response scheme systematically differs across income groups, which implies that Assumption 2 is violated.\(^{22}\)

\(^{22}\)Note that the mean perceptions and expectations of inflation are also systematically linked to socioeconomic characteristics. The cross-sectional mean of quantitative perceptions and expectations is declining in income. This is consistent with the response scheme estimates. These suggest that individuals in the highest income quartile experience any deviation of inflation from zero as more relevant in qualitative terms than individuals in lower income quartiles.
Table 5: Estimated response scheme for expected inflation by income groups

<table>
<thead>
<tr>
<th>Expectations</th>
<th>Income groups (1st quartile lowest)</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st quartile</td>
</tr>
<tr>
<td>$\mu_1^t$</td>
<td>-9.229***</td>
</tr>
<tr>
<td></td>
<td>(0.165)</td>
</tr>
<tr>
<td>$\mu_2^t$</td>
<td>0.703***</td>
</tr>
<tr>
<td></td>
<td>(0.0686)</td>
</tr>
<tr>
<td>$\mu_3^t$</td>
<td>2.452***</td>
</tr>
<tr>
<td></td>
<td>(0.0633)</td>
</tr>
<tr>
<td>$\mu_4^t$</td>
<td>7.075***</td>
</tr>
<tr>
<td></td>
<td>(0.0805)</td>
</tr>
<tr>
<td>$\sigma_t$</td>
<td>6.089***</td>
</tr>
<tr>
<td></td>
<td>(0.0827)</td>
</tr>
</tbody>
</table>

N  17232  24757  25384  32528  99901  
Log $L$  -21816  -31163  -31206  -39092  -124179

Likelihood ratio tests

<table>
<thead>
<tr>
<th>$H_1^0$: Identical threshold parameters $\mu_1^t,\ldots,\mu_4^t$ ($q = 12$)</th>
<th>LR statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1102.84</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$H_2^0$: Identical standard deviation $\sigma_t$ ($q = 3$)</th>
<th>LR statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1689.36</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>$H_3^0$: Identical thresholds and standard deviation ($q = 15$)</th>
<th>LR statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>1804.00</td>
<td>0.00</td>
<td></td>
</tr>
</tbody>
</table>

Notes: This table shows maximum likelihood estimates of the unrestricted response scheme (2). Monthly data, 01/2002–10/2008. Standard errors in parentheses. *, **, and *** indicate statistical significance at the 10%, 5%, and 1% level.

3.3 Normality of Quantitative Responses

Assumption 3 requires that the cross-sectional distribution of quantitative beliefs is normal. Normality has been tested and rejected both for inflation expectations of consumers (Batchelor and Dua, 1987) and professional forecasters (Carlson, 1975, Lahiri and Teigland, 1987). These studies generally find that quantitative beliefs are positively skewed and leptokurtic. Both properties are also present in the Swedish survey data, as panel 1 of Table 6 indicates. Inflation perceptions as well as inflation expectations exhibit a pronounced positive skewness and are leptokurtic. Not surprisingly, the Jarque-Bera test rejects the null hypothesis of normality in every single survey month, as panel 2 of Table 6 shows.

More generally, the probability method requires that beliefs follow some identifiable parametric distribution. Lahiri and Teigland (1987) suggest that a noncentral $t$ distri-
Table 6: Tests for the distribution of quantitative responses

<table>
<thead>
<tr>
<th></th>
<th>Perceptions</th>
<th></th>
<th>Expectations</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Overall</td>
<td>Min</td>
<td>Max</td>
<td>Overall</td>
</tr>
<tr>
<td>Skewness</td>
<td>4.04</td>
<td>0.05</td>
<td>8.50</td>
<td>3.64</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>40.51</td>
<td>9.96</td>
<td>117.20</td>
<td>43.96</td>
</tr>
<tr>
<td>Jarque-Bera test for normal distribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>J-B statistic</td>
<td>6,099.24</td>
<td>771.61</td>
<td>22,361.00</td>
<td>5,837.78</td>
</tr>
<tr>
<td>P-value</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>Kolmogorov-Smirnov test for noncentral $t$ distribution</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>K-S statistic</td>
<td>1.03</td>
<td>0.12</td>
<td>0.41</td>
<td>0.24</td>
</tr>
<tr>
<td>P-value</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>$\mu$</td>
<td>0.51</td>
<td>0.10</td>
<td>4.00</td>
<td>0.74</td>
</tr>
<tr>
<td>$df$</td>
<td>1.02</td>
<td>1</td>
<td>3</td>
<td>1.01</td>
</tr>
</tbody>
</table>

Notes: Monthly data, 01/1996–10/2008. Overall denotes the mean of monthly statistics, Min and Max are the monthly minimum and maximum of the respective statistic. The Jarque-Bera statistic is asymptotically $\chi^2$ distributed. The approximate 1% critical values for the Kolmogorov-Smirnov are given by $1.52\sqrt{N}$, where $N$ is the number of observations. The noncentral $t$ distribution is defined by the noncentrality parameter $\mu$ and the degrees of freedom $df$. The table shows the parameters that minimize the Kolmogorov-Smirnov statistic.

bution as an alternative to the normal distribution. The noncentral $t$ distribution allows for positive skewness and fat tails. I formally test whether quantitative responses follow a noncentral $t$ distribution employing the Kolmogorov-Smirnov test. These formal tests do not answer the question which parametric distribution produces the best results. The answer will also depend on the time period. In 1996–2007, the high share of zero responses cannot be reconciled with both the normal and noncentral $t$ distribution. With the rise in inflation perceptions and expectations in 2008, the shape of the distribution becomes somewhat smoother and less skewed as the share of zero responses declines. In general, however, the results indicate that differences between the commonly considered parametric distributions are predominated by the high share of zero responses, which cannot be accounted for by any of these distributions. This conjecture is in line

\[ D_n(F) = \sup_x |F_n(x) - F(x)|, \]

where $F_n(.)$ is the empirical distribution function. Note that the noncentral $t$ distribution is equal to the $t$ distribution if the noncentrality parameter $\mu$ is zero.

\[^{23}\text{The Kolmogorov-Smirnov statistic is given by } D_n(F) = \sup_x |F_n(x) - F(x)|, \text{ where } F_n(.) \text{ is the empirical distribution function. Note that the noncentral } t \text{ distribution is equal to the } t \text{ distribution if the noncentrality parameter } \mu \text{ is zero.}\]
with previous literature which finds that the accuracy of the probability method does not significantly vary between any of the common parametric distributions, see Berk (1999), Dasgupta and Lahiri (1992), and Smith and McAleer (1995).

3.4 Defined Reference Rate of Inflation

Assumption 4 requires that the reference rate of inflation $\pi_r^t$ that people have in mind when opting for the answer that prices are “somewhat/moderately higher” or are expected to “go up at the same rate” is equal to some defined value. This assumption is required to identify the system of equations that is generated by Assumptions 1 to 3. As outlined above, for identifying perceived inflation it is typically assumed that reference inflation is a constant such that perceptions are unbiased. For expected inflation, reference inflation is commonly assumed to be equal to actual inflation or to (previously quantified) perceived inflation.

To assess these assumptions, Figure 2 shows the mean of quantitative responses given by households that opt for qualitative answer $S_4$. The conditional mean of perceptions is highly volatile with a standard deviation of 1.15%. It averages at 7.93% but shows a declining tendency over time. The assumption that the moderate rate of inflation is constant over time is clearly at odds with this pattern.

The conditional mean given qualitative answer $S_4$ is less volatile for inflation expectations, with a standard deviation of only 0.53%. Figure 2 includes three candidate series that might correspond to the conditional mean: the mean quantitative inflation perception of those survey participants that expect prices to “go up at the same rate” ($S_4$), quantified inflation perceptions (based on the 5-category probability method under the unbiasedness assumption), and actual HICP inflation. Clearly, the conditional mean of inflation perceptions closely follows the conditional mean of inflation expectations. The correlation coefficient of the two series is 0.94, the mean difference is only 0.39%. This confirms the result of Section 3.2 that qualitative inflation expectations are linked to quantitative perceptions. Correlation of the conditional mean with quantified inflation perceptions and
Figure 2: Measures of reference inflation $\pi^r_t$

Notes: This figure shows alternative measures of $\pi^r_t$. Mean response $S_4$ is the mean quantitative belief of respondents that opt for qualitative answer $S_4$. Conditional mean perception is the mean quantitative inflation perception of respondents that opt for $S_4$ in the question about expected inflation. Implied reference inflation and quantified inflation perceptions are derived under the 5-category probability method. In quantifying perceptions the HICP unbiasedness condition is imposed.

actual HICP inflation is -0.15 and 0.06 respectively. Both series are nearly uncorrelated with the conditional mean of inflation perceptions. Hence, Assumption 4 that the reference rate of inflation is equal to actual inflation or quantified perceived inflation can be rejected. This of course raises doubt about the usual procedure to employ these measures in quantifying inflation expectations.

An alternative route to test Assumption 4 is to assess the implied level of moderate inflation by joining the mean of actual quantitative beliefs with Assumptions 1 to 3.\(^{24}\) Comparing the implied level of moderate inflation with the conditional mean given qualitative answer $S_4$ amounts to a joint test of Assumptions 1 to 3 of the probability method. Figure 2 indicates that for inflation perceptions the implied reference inflation fluctuates around a similar level as the conditional mean. However, the correlation with the conditional mean is only 0.06. For inflation expectations, the implied reference inflation lies

\(^{24}\)Given the mean $\pi^r_t$ of quantitative inflation expectations, implied reference inflation can be obtained by rearranging Equation (A.4) to $\pi^r_t = \pi^e_t - \frac{1}{T} \sum_{t=1}^{T} \frac{\pi^r_t - \pi^e_t \sigma^2 + \sigma^2}{\sigma^2 + \sigma^2 - \sigma^2 + \sigma^2}$.
about 2% below the conditional mean. The correlation of the two series is 0.31. Provided that the true reference rate of inflation is indeed equal to the conditional mean given qualitative answer $S_4$ these results suggests that Assumptions 1 to 3 can be jointly rejected. In light of these findings the next section assesses the joint validity of all 4 hypotheses in more detail.

3.5 Joint Assessment

While all four hypotheses can be individually rejected, this section investigates the joint validity of the assumptions. The focus does not lie on rejection/non-rejection but rather on the degree of overall validity. I proceed by quantifying the qualitative survey data employing the 5-category probability method.\(^{25}\) This yields the threshold parameters $\delta_t$ and $\eta_t$ which can be used to construct the implied response scheme on a monthly basis.

Figure 3 shows box plots of the distribution of monthly response shares. For each answer category, the fraction of quantitative beliefs that lie within the implied response interval is compared to the actual share of qualitative beliefs. Note that, by construction, the actual share of qualitative responses corresponds to the the predicted share of quantitative responses under the normality assumption. For inflation perceptions, the figure signals pronounced deviations of implied from actual response fractions in categories $S_3$ and $S_5$. The high share of quantitative responses in the implied range of $S_5$ is consistent with the previous finding that the distribution of responses is asymmetrical and leptokurtic. More importantly, the low fraction of quantitative responses in the implied range of $S_3$ appears to be a direct consequence from fitting the normal distribution to the high share of zero responses.

A similar pattern can be observed for inflation expectations. Figure 3 illustrates that the deviation of the implied from the actual response share is highest for categories $S_3$, $S_4$, and $S_5$. Similar to perceptions, the fraction of quantitative responses in the implied

\(^{25}\)Inflation perceptions are quantified by imposing the HICP unbiasedness condition. For inflation expectations it is assumed that the reference rate of inflation is equal to quantified inflation perceptions. Detailed derivations are provided in the Appendix.
Figure 3: Actual and theoretical response fractions

Notes: These figures show the fraction of quantitative answers within the implied response interval (quant.) and the actual share of qualitative responses (N). S1 through S5 are the qualitative response categories. Sample period 01/1996–10/2008. Perceptions are quantified using the 5-category probability method unbiased with respect to HICP inflation. Expectations are quantified using the 5-category probability method with quantified perceptions as reference inflation. Each box covers the range between the 25th and 75th percentile of monthly fractions and contains a median line. Upper (lower) adjacent values are given by the highest value not greater than the 75th (25th) percentile + (-) 3/2 of the interquartile range.

range of S5 exceeds the actual share of qualitative responses. This pattern also relates to the finding of Section 3.4 that the mean quantitative answer of respondents opting for qualitative answer S4 is significantly higher than actual inflation or quantified inflation perceptions. By consequence, a large fraction of these quantitative answers fall into the interval of the qualitative answer S5.

Further insights can be gained by looking at the fraction of quantitative responses that lie below or above the implied response interval. Figure A.1 in the Appendix shows this fraction, relative to the number of responses in the respective qualitative response category. Both for perceptions and expectations, the figure reveals that the 5-category probability method best accommodates qualitative answer S2. On average, 99% of quantitative responses associated with qualitative answer S2 lie within the implied response interval. S2 also is the most important qualitative response, accounting for roughly 59% of perceptions and 42% of expectations in 1996–2008. Regarding inflation perceptions, coverage for the
second most important category $S_3$, which obtains 30% of responses, is lower. Only about 30% of quantitative responses are within the implied response interval. A relatively large share of quantitative responses lies below the implied response interval, indicating that the interval around 0% is too wide. The worst coverage results for $S_4$, but only 4% of respondents opt for this qualitative category.

The pattern is different for inflation expectations. Figure A.1 indicates that only about 10% of quantitative beliefs fall into the implied response intervals for $S_3$ and $S_4$. Most quantitative responses are above the implied interval. This can be explained by the high share of on average 27% of responses in category $S_5$. Fitting this share leads to a downward shift of the lower response intervals. Also, Section 3.4 shows that quantified perceptions are significantly lower than reference inflation $\pi^r$. Hence, the implied response intervals linked to quantified perceptions will be too low. The resulting distortion might be notable as categories $S_3$ and $S_4$ account for 25% of qualitative responses.

The above findings also hold in the 01/2002–10/2008 subperiod, as Figures A.2 and A.3 in the Appendix confirm. In sum, the results suggest that Assumptions 1 through 4 are invalid. This leads to significant distortions primarily concerning the incorporation of information from positive categories $S_3$, $S_4$, $S_5$, which seem more pronounced for inflation expectations than perceptions. The next section assesses the implications for the accuracy of the probability method.

4 Accuracy of the Probability Method

4.1 Level and Dynamics of Beliefs

This section assesses the accuracy of the 5-category probability method relative to the mean of actual quantitative survey responses. Inflation perceptions and expectations are identified by imposing the usual restrictions. Inflation perceptions are assumed to be unbiased with respect to HICP inflation. Reference inflation for quantifying inflation expectations is set equal to HICP inflation or, following Berk (1999), to quantified perceived
inflation. The 5-category probability method is compared to a set of alternative quantification methods. The first alternative is the 3-category probability method of Carlson and Parkin (1975). The second alternative is the scaled balance statistic with mean and variance of actual inflation. In line with the literature, the 5-category balance statistic is computed as $s_5 + 0.5s_4 - 0.5s_2 - s_1$, while the 3-category balance statistic is given by $s_5 + s_4 + s_3 - s_1 = s_p^n - s_f^n$. Here, $s_p^n$ and $s_f^n$ are the fractions of respondents that report that inflation is rising or falling. The third alternative is the Pesaran (1987) regression approach for 3-category response data.

The primary measure of accuracy I consider is the (Pearson) correlation coefficient between the quantified series and the mean quantitative answer. Unlike the mean absolute error (MAE) or root mean squared error (RMSE), the correlation coefficient is robust to a constant scaling of the involved series. In particular, the correlation coefficient is unaffected by the average level of reference inflation $\pi^r$. Another advantage of employing the correlation coefficient is that its distributional properties have been explored. The Fisher $z$-transformation of the correlation coefficient results in an approximately normal random variable, provided the underlying data follows a bivariate normal distribution (Fisher, 1925). Relying on the Fisher $z$-transformation one can test the null hypothesis that two correlation coefficients are equal using the following statistic:

$$z = \tanh^{-1}(\rho_1) - \tanh^{-1}(\rho_2)$$  \hspace{1cm} (5)

where $\tanh^{-1}(\rho_i) = 0.5\ln\left(\frac{1+\rho_i}{1-\rho_i}\right)$. This statistic is approximately normal with variance

---

26Perceived inflation is quantified using the 5-category probability method under the assumption of unbiasedness with respect to HICP inflation.

27Answer categories are aggregated following Berk (1999), see the Appendix for details.

28Unlike the early regression approaches suggested by Theil (1952) and Anderson (1952), the Pesaran (1987) approach allows for asymmetric response behavior in periods of rising and falling inflation. The Pesaran approach is based on nonlinear least squares estimation of the model $\pi_t = \beta_1 s_p^n - 2\beta_2 s_f^n + \varepsilon_t$, where $\pi_t$ denotes actual HICP inflation. Expected inflation is generated in a second step as a prediction of this model based on answering fractions about inflation expectations (where coefficient estimates are obtained in the first step using perceptions data). A measure of perceived inflation is computed as the prediction of the model using the perceptions data it has been estimated with.

29The squared correlation coefficient is equal to the coefficient of determination in a simple linear regression.
$\frac{1}{T_{1-3}} + \frac{1}{T_{2-3}}$. However, the approximation may be inaccurate in the present case because $|\rho_1|$ is high (Mudholkar, 2006) and the underlying series are serially dependent. I therefore assess significance based on double block bootstrap confidence intervals for the $z$ statistic.\textsuperscript{30}

Table 7 summarizes the results. The underlying series are plotted in Figures A.4 and A.5 in the Appendix. All statistics are provided for levels and first differences of the series. Since the quantified series and the mean quantitative beliefs are highly persistent the discussion focusses on results for first differences.\textsuperscript{31} These results are not subject to spurious regression problems as the first differences are stationary. However, the results on correlation are broadly consistent among levels and first differences because the block bootstrap accounts for serial dependence.

Panel 1 of Table 7 indicates that, in terms of correlation with actual perceptions, all quantification methods perform well. The 5-category probability method and the 3-category method produce virtually identical results. Interestingly however, the simple balance statistics are more accurate, with correlation coefficients of 0.91 and 0.89 in first differences. The $z$-statistic indicates that the correlation coefficient for the 5-category balance statistic is significantly higher than the correlation coefficient for the 5-category probability method. The regression approach shows a similar performance as the probability methods.

For expectations, panel 2 of Table 7 shows that the accuracy of the 5-category probability method depends on the chosen reference rate of inflation. Employing quantified perceptions generates significantly better results in first differences than employing actual HICP inflation. The correlation coefficients are 0.80 and 0.50 respectively and the $z$-statistic is highly significant. The most accurate method for quantifying inflation expectations is again a balance statistic. Moreover, the 3-category regression approach is slightly more accurate than the 5-category probability method. These differences are not

\textsuperscript{30}Matlab codes are available from the author. The double moving block bootstrap of the percentile confidence interval is based on 1,000 first level replications and 2,500 second level replications and a block size of 5. See Efron and Tibshirani (1993) for a description of the method.

\textsuperscript{31}Employing an Augmented Dickey-Fuller test the null hypothesis of a unit root cannot be rejected on the 10% level for all actual and quantified mean series.
Table 7: Accuracy of quantified inflation perceptions and expectations

<table>
<thead>
<tr>
<th>Perceptions</th>
<th>Level</th>
<th></th>
<th>First differences</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias</td>
<td>MAE</td>
<td>RMSE</td>
<td>ρ</td>
</tr>
<tr>
<td>P.HICP (5 cat.)</td>
<td>-0.17</td>
<td>0.22</td>
<td>0.32</td>
<td>0.97</td>
</tr>
<tr>
<td>P.HICP (3 cat.)</td>
<td>-0.17</td>
<td>0.19</td>
<td>0.24</td>
<td>0.98</td>
</tr>
<tr>
<td>Balance (5 cat.)</td>
<td>-0.20</td>
<td>0.20</td>
<td>0.25</td>
<td>0.99</td>
</tr>
<tr>
<td>Balance (3 cat.)</td>
<td>-0.20</td>
<td>0.22</td>
<td>0.31</td>
<td>0.97</td>
</tr>
<tr>
<td>Pesaran (3 cat.)</td>
<td>-0.21</td>
<td>0.25</td>
<td>0.47</td>
<td>0.94</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Expectations</th>
<th>Level</th>
<th></th>
<th>First differences</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias</td>
<td>MAE</td>
<td>RMSE</td>
<td>ρ</td>
</tr>
<tr>
<td>P.Perc. (5 cat.)</td>
<td>-0.85</td>
<td>0.85</td>
<td>0.89</td>
<td>0.94</td>
</tr>
<tr>
<td>P.HICP (5 cat.)</td>
<td>-0.87</td>
<td>0.87</td>
<td>0.96</td>
<td>0.86</td>
</tr>
<tr>
<td>P.HICP (3 cat.)</td>
<td>-0.49</td>
<td>0.49</td>
<td>0.24</td>
<td>0.95</td>
</tr>
<tr>
<td>Balance (5 cat.)</td>
<td>-0.49</td>
<td>0.56</td>
<td>0.25</td>
<td>0.89</td>
</tr>
<tr>
<td>Balance (3 cat.)</td>
<td>-0.49</td>
<td>0.51</td>
<td>0.31</td>
<td>0.96</td>
</tr>
<tr>
<td>Pesaran (3 cat.)</td>
<td>0.12</td>
<td>0.22</td>
<td>0.29</td>
<td>0.95</td>
</tr>
</tbody>
</table>

Notes: Accuracy is measured relative to the mean of quantitative responses, 01/1996–10/2008. For perceptions, P.HICP (5 cat.) and P.HICP (3 cat.) denote the 5-category and 3-category probability method under the condition that perceptions are unbiased with respect to HICP inflation. For expectations, P.Perc. (5 cat.) and P.HICP (5 cat.) are the 5-category probability methods with reference inflation given by quantified perceptions and HICP inflation, respectively. P.HICP (3 cat.) is the 3-category probability method under the condition that expectations are unbiased. ρ is the correlation coefficient with the mean of quantitative responses. The last column in each panel shows the Fisher z statistic for testing the null hypothesis that the difference between the correlation coefficient in the first row of each panel (ρ₁) and the correlation coefficient in the corresponding row (ρ₂) is zero. *, **, and *** indicate statistical significance of z at the 10%, 5%, and 1% level based on double block bootstrap percentile confidence intervals.
statistically significant, however. Results for the subsample 01/2002–10/2008 in Table A.1 confirm these findings.

In sum, the 5-category balance statistic tracks actual quantitative perceptions most accurately. Still, the 5-category probability method as well as the other methods are highly correlated with mean quantitative perceptions. For expectations, none of the alternative methods performs significantly better than the 5-category probability method with reference inflation given by quantified perceptions. The reasonable performance of the probability method is in contrast to findings of Batchelor (1986) for the U.S.\textsuperscript{32} However, the 5-category probability method may perform weakly to quantify expectations, depending on the chosen reference rate of inflation. Also, the similar accuracy of the 5-category probability method and the 3-category methods signals that the 5-category method is not fully efficient at incorporating information from positive answer categories.

4.2 Dispersion of Beliefs

The cross-sectional heterogeneity of beliefs is subject to increasing research in macroeconomics. This section investigates how to best infer cross-sectional heterogeneity, measured by the standard deviation of quantitative beliefs, from qualitative survey data. The 5-category probability method not only allows to identify the mean but also the implied standard deviation of the fitted normal distribution given by Equation (A.5). In addition, I consider four alternative measures of heterogeneity. The first alternative is implied standard deviation from the 3-category probability method given by Equation (A.11). The second alternative is an index of qualitative variation (IQV) based on the response shares $s_1$ through $s_5$:

$$IQV = \frac{K}{K-1} \left( 1 - \sum_{i=1}^{K} s_i^2 \right)$$

\textsuperscript{32}Batchelor (1986) documents that the quantified series do not predict the direction of change in mean quantitative responses. In the present case, a comparison of signs confirms the high correlation in first differences. For inflation perceptions, the balance statistic and the probability method indicate the correct direction of change of the mean quantitative response in 131 and 129 out of 153 months, for expectations in 121 and 120 months.
where \( K = 5 \) is the number of response categories and \( s_i \) the fraction of answers in category \( i \). The scaling factor \( \frac{K}{K-1} \) ensures that \( 0 \leq q^2 \leq 1 \). Unlike the probability method, the IQV does not account for the ordered nature of the data. The third alternative is the \( d^2 \)-index of ordinal variation proposed by Lacy (2006).\(^{33}\) This index is given by:

\[
DSQ = \sum_{i=1}^{K-1} F_i (1 - F_i)
\]

where \( K = 5 \) is the number of response categories and \( F_i \) the cumulative response share in category \( i \), e.g., \( F_3 = s_1 + s_2 + s_3 \). As the IQV, the DSQ statistic attains its minimum of 0 if all answers lie in the same response category. But while the IQV is maximal when answers are uniformly distributed, the DSQ attains its maximum of 1 if the distribution is polarized, i.e. if \( s_1 = s_5 = 0.5 \). The fourth alternative is the disconformity index of Theil (1955) defined as \( DIS = s_p + s_n - (s_p - s_n)^2 \).\(^{34}\)

Table 8 summarizes the results on accuracy. Since the null hypothesis of a unit-root process can be rejected for the standard deviation of quantitative responses the discussion centers on results in levels.\(^{35}\) The alternative measures of dispersion are plotted in Figures A.6 and A.7. Table 8 shows that both the 5-category probability method and the 3-category probability method considerably underestimate the heterogeneity of quantitative beliefs. The implied standard deviation lies 1.7% to 2.4% below the actual standard deviation of quantitative responses. This finding is consistent with earlier results of Defris and Williams (1979) and Batchelor (1986).

Regarding inflation perceptions, the first panel of Table 8 shows that implied standard deviation from the 5-category probability method traces actual heterogeneity only poorly. The correlation coefficient is 0.30. The 3-category probability method performs significantly better, as the \( z \)-statistic indicates. However, the qualitative measures of variation

---


\(^{34}\)The disconformity index relies on the same theoretical assumptions as the 3-category balance statistic, see Batchelor (1986).

\(^{35}\)Both for perceptions and expectations, the Augmented Dickey-Fuller test rejects the null hypothesis of a unit root for the standard deviation of actual quantitative responses and for the quantified series using the 5-category probability method linked to HICP inflation.
Table 8: Accuracy of measures of cross-sectional dispersion

<table>
<thead>
<tr>
<th>Perceptions</th>
<th>Level</th>
<th>First differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias</td>
<td>MAE</td>
</tr>
<tr>
<td>Implied SD (P.HICP, 5 cat.)</td>
<td>-1.76</td>
<td>1.76</td>
</tr>
<tr>
<td>Implied SD (P.HICP, 3 cat.)</td>
<td>-1.83</td>
<td>1.83</td>
</tr>
<tr>
<td>IQV (5 cat.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSQ (5 cat.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIS (3 cat.)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Expectations</th>
<th>Level</th>
<th>First differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias</td>
<td>MAE</td>
</tr>
<tr>
<td>Implied SD (P.Perc., 5 cat.)</td>
<td>-2.42</td>
<td>2.42</td>
</tr>
<tr>
<td>Implied SD (P.HICP, 5 cat.)</td>
<td>-2.44</td>
<td>2.44</td>
</tr>
<tr>
<td>Implied SD (P.HICP, 3 cat.)</td>
<td>-1.95</td>
<td>1.95</td>
</tr>
<tr>
<td>IQV (5 cat.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSQ (5 cat.)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DIS (3 cat.)</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

Notes: Accuracy is measured relative to the cross-sectional standard deviation of quantitative responses, 01/1996–10/2008. For perceptions, Implied SD (P.HICP, 5 cat.) is the implied standard deviation from the 5-category probability method under the HICP unbiasedness condition. For expectations, Implied SD (P.Perc., 5 cat.) and Implied SD (P.HICP, 5 cat.) are implied standard deviations from the 5-category probability method with reference inflation given by quantified perceptions and actual HICP inflation, respectively. Both for perceptions and expectations, Implied SD (P.HICP, 3 cat.) is the implied standard deviation from the 3-category probability method under the HICP unbiasedness condition. IQV is the index of qualitative variation, DSQ is the $d^2$-index of ordinal variation proposed by Lacy (2006), and DIS is the disconformity statistic. $\rho$ is the correlation coefficient with the standard deviation of quantitative responses. The last column in each panel shows the Fisher $z$ statistic for testing the null hypothesis that the difference between the correlation coefficient in the first row of each panel ($\rho_1$) and the correlation coefficient in the corresponding row ($\rho_2$) is zero. *, **, and *** indicate statistical significance of $z$ at the 10%, 5%, and 1% level based on double block bootstrap percentile confidence intervals.
are even more highly correlated with the standard deviation of quantitative beliefs. The correlation coefficients of the IQV and the DSQ are 0.83 and 0.87. The performance of the 3-category disconformity index is substantially lower, its correlation with actual standard deviation of quantitative responses is similar to the 5-category probability method.

The implied standard deviation from the 5-category probability method performs better for quantifying heterogeneity of inflation expectations, as the second panel of Table 8 indicates. Again, the correlation depends on the choice of the reference inflation. Employing actual HICP inflation instead of quantified perceptions reduces the correlation coefficient significantly from 0.67 to 0.52. The implied standard deviation from the 3-category approach is about as accurate as the 5-category probability method with quantified perceptions as reference inflation. The IQV most closely tracks actual heterogeneity of quantitative responses. The correlation with actual standard deviation is 0.80, which is significantly higher than the correlation of the 5-category probability method. Unlike for perceptions, the DSQ-statistic performs substantially worse with a correlation of only 0.45. The disconformity index is also only weakly correlated with actual heterogeneity, which is in line with earlier findings of Batchelor (1986). The results for first differences are broadly consistent, although the stationary series are over-differentiated. Moreover, results for the subsample 01/2002–10/2008 in Table A.2 confirm the above findings.\(^{36}\)

In sum, these results suggest that while the probability method is relatively accurate in describing the central tendency, it is considerably less accurate in capturing the level and dynamics of cross-sectional heterogeneity. Consistent with findings of the previous section, the 3-category probability method performs better than the 5-category method.\(^{37}\) The IQV, however, dominates the other methods in terms of correlation with the cross-sectional standard deviation of quantitative beliefs. The DSQ is only accurate for quantifying the

\(^{36}\)Note, however, that the probability method gains relative accuracy for quantifying inflation expectations. I have also assessed accuracy of the square root of the index of qualitative variation, the square root of the DSQ-statistic, and the square root of the disconformity index. The correlations with actual standard deviation of survey responses do not significantly change, both in their absolute level and relative ordering.

\(^{37}\)The reasonable performance of the 3-category probability method also reinforces the results of Dasgupta and Lahiri (1993) who show that implied dispersion from the 3-category method is useful for predicting business cycle turning points.
heterogeneity of inflation perceptions. This suggests that the IQV is less distorted by the unordered qualitative answers on expected inflation than the DSQ. Finally, the assessment of accuracy relies on correlation coefficients. Bias, MAE, and RMSE seem less informative since units for qualitative indices are not directly. Moreover, these measures are driven by the identifying assumptions.

5 Which Quantification Method Should Be Used?

The assumptions of the 5-category probability method can be individually and jointly rejected, as Section 3 shows. Moreover, the response scheme estimates indicate that three separate positive categories $S_3$, $S_4$, $S_5$ contain additional information over just one positive category. This is the case for inflation perceptions as well as for inflation expectations. For inflation expectations, however, the qualitative answers are not strictly increasing in quantitative expectations since positive responses are given relative to perceived inflation.

Despite that the theoretical assumptions are violated, the accuracy of the 5-category probability method for quantifying the mean of inflation perceptions is high. The relative performance of the 3-category probability method indicates, however, that the 5-category method does not efficiently incorporate the additional information from three positive response categories. The most accurate method is the 5-category balance statistic.

Regarding the mean of inflation expectations, the accuracy of the 5-category probability method largely depends on the chosen reference inflation $\pi_r$. Employing quantified inflation perceptions yields significantly more accurate results than employing actual HICP inflation. This appears to be inconsistent with the finding of Section 3.4 that quantified perceptions are uncorrelated with reference inflation. But as Figure 2 indicates, the relation between quantified perceptions and reference inflation becomes stronger towards the end of the sample period. The correlation coefficient of reference inflation and quantified perceptions is 0.46 in 2002–2008.\footnote{As in Section 3.4 reference inflation is measured by the conditional mean of quantitative perceptions.} In particular, quantified perceptions surge in 2007 and 2008, while reference inflation increases to about 6%. This suggests that the 5-category
Table 9: Actual and imposed conditional means

<table>
<thead>
<tr>
<th></th>
<th>Perceptions</th>
<th></th>
<th>Expectations</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actual</td>
<td>Balance</td>
<td>P.HICP</td>
<td>Actual</td>
</tr>
<tr>
<td>Mean response, given $S_1$</td>
<td>-4.79</td>
<td>-3.96</td>
<td>-4.94</td>
<td>-3.89</td>
</tr>
<tr>
<td>Mean response, given $S_2$</td>
<td>0.03</td>
<td>-0.23</td>
<td>0.37</td>
<td>0.02</td>
</tr>
<tr>
<td>Mean response, given $S_3$</td>
<td>4.72</td>
<td>3.50</td>
<td>4.62</td>
<td>4.06</td>
</tr>
<tr>
<td>Mean response, given $S_4$</td>
<td>7.94</td>
<td>7.22</td>
<td>7.70</td>
<td>4.77</td>
</tr>
<tr>
<td>Mean response, given $S_5$</td>
<td>9.58</td>
<td>10.95</td>
<td>9.75</td>
<td>3.98</td>
</tr>
</tbody>
</table>

Notes: The column Actual presents results for actual quantitative survey responses, 01/1996–10/2008. The column Balance shows the mean values that the balance statistic attributes to qualitative survey categories. P.HICP denotes implied conditional means of the 5-category probability method unbiased with respect to HICP inflation. P.Perc. are the implied conditional means of the 5-category probability method with reference inflation given by quantified perceptions. The conditional means are scaled to match the mean and standard deviation of actual conditional means.

probability method with quantified perceptions as reference inflation gains relative accuracy once inflation perceptions diverge from actual inflation.\(^{39}\)

It remains the question why the simple balance statistic traces mean beliefs at least as closely as the 5-category probability method. Table 9 provides a possible explanation. The table shows the mean of quantitative responses conditional on the qualitative response. Conditional means of actual quantitative responses are compared to the imposed conditional means of the balance statistic and of the 5-category probability method. The imposed conditional means are scaled to match the mean and variance of actual conditional means.\(^{40}\) Table 9 reveals that both methods impose (scaled) conditional means that match actual conditional means quite closely. In other words, the conditional means imposed by the balance statistic are roughly proportional to the actual conditional means in the data. Not surprisingly, the fit is better for perceptions than for expectations. This also suggests that the balance statistic, unlike the probability method, will lose relative accuracy once the ratio of conditional means changes.

\(^{39}\)This is the case, e.g., around the euro cash changeover in 2002 (ECB, 2005).

\(^{40}\)The (unscaled) conditional means of the 5-category balance statistic are -1, 0.5, 0, 0.5, 1. The implied conditional means of the probability method are computed by numerical integration using the quantified parameters $\pi_t^e$, $\sigma_t$, $\delta_t$, $\eta_t$, $\pi_t^r$. Perceptions are quantified employing the 5-category probability method unbiased with respect to HICP inflation. Expectations are quantified using the 5-category probability method with reference inflation given by quantified perceptions.
The accuracy of the 5-category probability method is low for quantifying the cross-sectional standard deviation of beliefs. Here, an index of qualitative variation is preferable. The index of qualitative variation also dominates the DSQ-index of ordinal variation for quantifying the heterogeneity of expectations. Similar to the findings on the choice of reference inflation, the IQV seems less distorted by the unordered qualitative answers than the DSQ.

In sum, these results are in favor of the 5-category probability method for quantifying the mean of beliefs. The index of qualitative variation is preferable for quantifying the cross-sectional standard deviation of beliefs. The findings also indicate, that the 5-category probability method with quantified perceptions as reference inflation might gain relative accuracy if inflation perceptions deviate from actual inflation. This, however, points to a general limitation of this study’s results: findings on the quantification of expectations might differ from other countries as Swedish consumers are relatively well informed about actual HICP inflation.

6 Conclusion

This paper assesses the validity and accuracy of the 5-category probability method for quantifying inflation perceptions and expectations. The analysis capitalizes on jointly available qualitative and quantitative response data from the Swedish Consumer Tendency Survey. Relying on monthly data on household level, 01/1996–10/2008, the theoretical assumptions of the 5-category probability method are individually and jointly tested and rejected. Maximum likelihood estimations of unrestricted response schemes indicate that the actual response scheme is neither symmetric nor homogeneous across individuals. Moreover, it is shown that qualitative inflation expectations are formed relative to inflation perceptions, which is a direct result of the survey design. By consequence, qualitative inflation expectations are not monotonously rising in quantitative inflation expectations. Furthermore, quantitative beliefs are not normally distributed and cannot be reconciled with a noncentral t distribution either. Finally, the reference rate of inflation (the “moderate”
rate of inflation) is shown to be time varying for perceptions and significantly different from actual inflation for expectations. The joint assessment indicates that, by consequence, the 5-category method does not adequately accommodate the positive qualitative response categories $S_3$, $S_4$, $S_5$.

Nevertheless, the accuracy of the 5-category probability method in terms of correlation with the mean of actual quantitative beliefs is high. For quantifying inflation expectations the accuracy of the method strongly depends on the identifying restriction imposed by the choice of reference inflation. Relying on double block bootstrap confidence intervals for Fisher’s $z$ statistic it is shown that setting reference inflation equal to previously quantified inflation perceptions yields significantly better results than setting reference inflation equal to actual HICP inflation. However, the 5-category probability method is not more accurate than the balance statistic and the 3-category probability method. This suggests that the 5-category probability method does not efficiently incorporate information from positive qualitative response categories. In sum, however, the results are in favor of the 5-category probability method. In particular, the 5-category probability method with quantified perceptions as reference inflation might gain relative accuracy once inflation perceptions substantially deviate from actual inflation. The most accurate measure of the cross-sectional standard deviation of quantitative beliefs is the index of qualitative variation. This index performs significantly better than the 5-category probability method and other common approaches.

The findings for Sweden suggest a number of avenues for further research. To exploit the additional information from three positive response categories research is needed on how to choose the reference rate of inflation. Moreover, a non-parametric analysis will generate further insights that will help to improve the imposed response scheme and distribution. Looking ahead, implications for survey design should also be discussed in more depth. The results for Sweden indicate that it is difficult to efficiently handle the relative nature of the positive qualitative responses about inflation expectations. Obvious alternative survey designs include to adopt the same response scheme for expectations as currently in use for perceptions or to directly ask for quantitative responses in the first place.
References


Appendix

A.1 Derivation of the 5-Category Probability Method

This section derives the 5 category probability method based on the assumptions introduced in Section 3. The method has been originally proposed by Batchelor and Orr (1988) and relies on earlier contributions of Theil (1952) and Carlson and Parkin (1975). Again all derivations equally hold for quantifying inflation perceptions as well as inflation expectations. The response scheme for inflation perceptions is given by:

\[
\begin{align*}
\pi^p_{it} < -\delta_t : \text{ prices in general are lower (S_1)} \\
-\delta_t \leq \pi^p_{it} < \delta_t : \text{ about the same (S_2)} \\
\delta_t \leq \pi^p_{it} < \pi^r_t - \eta_t : \text{ a little higher (S_3)} \\
\pi^r_t - \eta_t \leq \pi^p_{it} < \pi^r_t + \eta_t : \text{ moderately higher (S_4)} \\
\pi^p_{it} \geq \pi^r_t + \eta_t : \text{ a lot higher (S_5)}
\end{align*}
\]

(A.1)

Under the assumptions introduced in Section 3 the relation between aggregate response shares and expected inflation \(\pi^e_t\) in period \(t\) is given by:

\[
\begin{align*}
s^1_t &= P(\pi^e_{it} < -\delta_t) = \Phi\left(\frac{-\delta_t - \pi^e_t}{\sigma_t}\right) \\
s^2_t &= P(-\delta_t \leq \pi^e_{it} < \delta_t) = \Phi\left(\frac{\delta_t - \pi^e_t}{\sigma_t}\right) - \Phi\left(\frac{-\delta_t - \pi^e_t}{\sigma_t}\right) \\
s^3_t &= P(\delta_t \leq \pi^e_{it} < \pi^r_t - \eta_t) = \Phi\left(\frac{\pi^r_t - \eta_t - \pi^e_t}{\sigma_t}\right) - \Phi\left(\frac{\delta_t - \pi^e_t}{\sigma_t}\right) \\
s^4_t &= P(\pi^r_t - \eta_t \leq \pi^e_{it} < \pi^r_t + \eta_t) = \Phi\left(\frac{\pi^r_t + \eta_t - \pi^e_t}{\sigma_t}\right) - \Phi\left(\frac{\pi^r_t - \eta_t - \pi^e_t}{\sigma_t}\right) \\
s^5_t &= P(\pi^r_t + \eta_t \leq \pi^e_{it}) = 1 - \Phi\left(\frac{\pi^r_t + \eta_t - \pi^e_t}{\sigma_t}\right)
\end{align*}
\]

(A.2)

where \(\Phi(.)\) is the standard normal cumulative distribution function. The same holds for inflation perceptions, in which case \(\pi^e_t\) is substituted with \(\pi^p_t\). System (A.2) can be rewritten
to get a system of 4 linearly independent equations with 5 unknowns ($\pi^e_t$, $\sigma_t$, $\delta_t$, $\eta_t$, $\pi^r_t$):

\[
G^1_t = \Phi^{-1}(s^1_t) = \frac{-\delta_t - \pi^e_t}{\sigma_t}
\]
\[
G^2_t = \Phi^{-1}(1 - s^5_t - s^4_t - s^3_t - s^2_t) = \frac{-\delta_t - \pi^e_t}{\sigma_t}
\]
\[
G^3_t = \Phi^{-1}(1 - s^5_t - s^4_t) = \frac{\delta_t - \pi^e_t}{\sigma_t}
\]
\[
G^4_t = \Phi^{-1}(1 - s^5_t - s^4_t) = \frac{\pi^r_t - \eta_t - \pi^e_t}{\sigma_t}
\]
\[
G^5_t = \Phi^{-1}(1 - s^5_t) = \frac{\pi^r_t + \eta_t - \pi^e_t}{\sigma_t}
\]

(A.3)

System (A.3) can be solved to get the following expression for the mean $\pi^e_t$ of expected inflation:

\[
\pi^e_t = \pi^r_t \frac{G^2_t + G^3_t}{G^2_t + G^3_t - G^4_t - G^5_t}
\]

(A.4)

For simplicity $\pi^e_t$ is called expected inflation (rather than mean of expected inflation). The remaining unknowns are given by:

\[
\sigma_t = \pi^r_t \frac{-2}{G^2_t + G^3_t - G^4_t - G^5_t}
\]

(A.5)
\[
\delta_t = \pi^r_t \frac{G^2_t - G^3_t}{G^2_t + G^3_t - G^4_t - G^5_t}
\]

(A.6)
\[
\eta_t = \pi^r_t \frac{G^4_t - G^5_t}{G^2_t + G^3_t - G^4_t - G^5_t}
\]

(A.7)

For quantifying inflation perceptions, $\pi^r_t$ is commonly identified by restricting inflation perceptions to be unbiased over the sample period. Rearranging Equation (A.4) and imposing unbiasedness yields:

\[
\pi^r_t = \frac{\pi}{\frac{1}{T} \sum_{t=1}^{T} \frac{G^2_t + G^3_t}{G^2_t + G^3_t - G^4_t - G^5_t}}
\]

(A.8)

where $T$ is the number of periods and $\pi$ the average actual rate of inflation.
A.2 Derivation of the 3-Category Probability Method

Theil (1952) and Carlson and Parkin (1975) have originally developed the probability method for three-option ordinal scales. In line with Berk (1999), the EU consumer survey responses are aggregated to three categories by defining \( s_n = s_1 \), \( s_s = s_2 \), and \( s_p = s_3 + s_4 + s_5 \). The relation between response shares and expected inflation is then given by:

\[
\begin{align*}
    s^n_t &= P(\pi_e^{<\delta}) = \Phi\left(\frac{-\delta - \pi_e^t}{\sigma_t}\right) \\
    s^s_t &= P(-\delta < \pi_e^{<\delta}) = \Phi\left(\frac{\delta - \pi_e^t}{\sigma_t}\right) - \Phi\left(\frac{-\delta - \pi_e^t}{\sigma_t}\right) \\
    s^p_t &= P(\delta < \pi_e^{<\delta}) = 1 - \Phi\left(\frac{\delta - \pi_e^t}{\sigma_t}\right)
\end{align*}
\]

(A.9)

where \( \Phi(.) \) is a standard normal cumulative distribution function. Hence at any \( t \) we have a system of 2 linearly independent equations with 3 unknowns \( (\delta, \mu_t, \sigma_t) \). Solving yields:

\[
\begin{align*}
    \pi_e^t &= \delta \frac{\Phi^{-1}(s^n_t) + \Phi^{-1}(1-s^p_t)}{\Phi^{-1}(s^p_t) - \Phi^{-1}(1-s^n_t)} \\
    \sigma_t &= \frac{2}{\Phi^{-1}(1-s^p_t) - \Phi^{-1}(s^n_t)}
\end{align*}
\]

(A.10) (A.11)

Again the identical equations hold for perceived inflation \( \pi_p^t \). I follow the existing literature and identify the system by imposing an unbiasedness assumption. \( \delta \) can be recovered by equating the mean of expected inflation (perceived inflation) to average actual inflation in the sample period:

\[
\delta = \frac{\sum_{t=1}^{T} \pi_t}{T} \frac{\Phi^{-1}(s^n_t) + \Phi^{-1}(1-s^p_t)}{\Phi^{-1}(s^p_t) - \Phi^{-1}(1-s^n_t)}
\]

(A.12)
A.3 Additional Results

Figure A.1: Response fractions below (low) and above (high) the implied response interval

Notes: These figures show the fraction of quantitative responses that lie below or above the implied response interval (defined by the simultaneous qualitative response) relative to the sum of responses in the respective qualitative response category. Sample period 01/1996–10/2008. Quantified perceptions are based on the 5 category probability method unbiased with respect to HICP inflation. Expectations are quantified using the probability method with quantified perceptions as reference inflation. Each box covers the range between the 25th and 75th percentile of monthly fractions and contains a median line.
Figure A.2: Actual and theoretical response fractions, 01/2002–10/2008

Notes: See footnote of Figure 3 for a detailed description.

Figure A.3: Response fractions below (low) and above (high) the implied response interval, 01/2002–10/2008

Notes: See footnote of Figure A.1 for a detailed description.
Figure A.4: Mean of inflation perceptions

Figure A.5: Mean of inflation expectations
Table A.1: Accuracy of quantified inflation perceptions and expectations, 01/2002–10/2008

<table>
<thead>
<tr>
<th>Perceptions</th>
<th>Level</th>
<th>First differences</th>
</tr>
</thead>
<tbody>
<tr>
<td>P.HICP (5 cat.)</td>
<td>Bias</td>
<td>MAE</td>
</tr>
<tr>
<td>-0.26</td>
<td>0.27</td>
<td>0.40</td>
</tr>
<tr>
<td>P.HICP (3 cat.)</td>
<td>-0.11</td>
<td>0.14</td>
</tr>
<tr>
<td>Balance (5 cat.)</td>
<td>-0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>Balance (3 cat.)</td>
<td>-0.15</td>
<td>0.19</td>
</tr>
<tr>
<td>Pesaran (3 cat.)</td>
<td>-0.27</td>
<td>0.31</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Pesaran (3 cat.)</th>
<th>Bias</th>
<th>MAE</th>
<th>RMSE</th>
<th>ρ</th>
<th>z</th>
</tr>
</thead>
<tbody>
<tr>
<td>P.Perc. (5 cat.)</td>
<td>-0.76</td>
<td>0.76</td>
<td>0.80</td>
<td>0.97</td>
<td></td>
</tr>
<tr>
<td>P.HICP (5 cat.)</td>
<td>-0.92</td>
<td>0.92</td>
<td>1.00</td>
<td>0.89</td>
<td>0.70**</td>
</tr>
<tr>
<td>P.HICP (3 cat.)</td>
<td>-0.39</td>
<td>0.39</td>
<td>0.18</td>
<td>0.97</td>
<td>0.09</td>
</tr>
<tr>
<td>Balance (5 cat.)</td>
<td>-0.22</td>
<td>0.35</td>
<td>0.25</td>
<td>0.82</td>
<td>0.99</td>
</tr>
<tr>
<td>Balance (3 cat.)</td>
<td>-0.27</td>
<td>0.31</td>
<td>0.33</td>
<td>0.95</td>
<td>0.31</td>
</tr>
<tr>
<td>Pesaran (3 cat.)</td>
<td>0.05</td>
<td>0.22</td>
<td>0.32</td>
<td>0.92</td>
<td>0.53</td>
</tr>
</tbody>
</table>

Notes: See Table 7 for a detailed description.
Figure A.6: Dispersion of inflation perceptions

Notes: SD quantitative is the cross-sectional standard deviation of quantitative survey responses. Perceptions and 3-category probability method are restricted to be unbiased with respect to HICP inflation, reference inflation for quantifying expectations is given by quantified perceptions. Index of ord. var. is the $d^2$ statistic.

Figure A.7: Dispersion of inflation expectations

Notes: SD quantitative is the cross-sectional standard deviation of quantitative survey responses. Perceptions and 3-category probability method are restricted to be unbiased with respect to HICP inflation, reference inflation for quantifying expectations is given by quantified perceptions. Index of ord. var. is the $d^2$ statistic.
Table A.2: Accuracy of measures of cross-sectional dispersion, 01/2002–10/2008

<table>
<thead>
<tr>
<th>Perceptions</th>
<th>Level</th>
<th>First differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias</td>
<td>MAE</td>
</tr>
<tr>
<td>Implied SD (P.HICP, 5 cat.)</td>
<td>-2.00</td>
<td>2.00</td>
</tr>
<tr>
<td>Implied SD (P.HICP, 3 cat.)</td>
<td>-1.86</td>
<td>1.86</td>
</tr>
<tr>
<td>IQV (5 cat.)</td>
<td>0.80</td>
<td>-0.82*</td>
</tr>
<tr>
<td>DSQ (5 cat.)</td>
<td>0.79</td>
<td>-0.82</td>
</tr>
<tr>
<td>DC (3 cat.)</td>
<td>-0.15</td>
<td>0.43</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Expectations</th>
<th>Level</th>
<th>First differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias</td>
<td>MAE</td>
</tr>
<tr>
<td>Implied SD (P.Perc, 5 cat.)</td>
<td>-2.35</td>
<td>2.35</td>
</tr>
<tr>
<td>Implied SD (P.HICP, 5 cat.)</td>
<td>-2.51</td>
<td>2.51</td>
</tr>
<tr>
<td>Implied SD (P.HICP, 3 cat.)</td>
<td>-1.88</td>
<td>1.88</td>
</tr>
<tr>
<td>IQV (5 cat.)</td>
<td>0.33</td>
<td>-0.03</td>
</tr>
<tr>
<td>DSQ (5 cat.)</td>
<td>-0.09</td>
<td>0.41</td>
</tr>
<tr>
<td>DC (3 cat.)</td>
<td>0.09</td>
<td>0.23</td>
</tr>
</tbody>
</table>

Notes: See table 8 for a detailed description.