Maag, Thomas

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

KOF Working Papers, No. 230

Provided in Cooperation with:
KOF Swiss Economic Institute, ETH Zurich

Suggested Citation: Maag, Thomas (2009) : On the accuracy of the probability method for quantifying beliefs about inflation, KOF Working Papers, No. 230, ETH Zurich, KOF Swiss Economic Institute, Zurich, http://dx.doi.org/10.3929/ethz-a-005859391

This Version is available at:
http://hdl.handle.net/10419/50407

Terms of use:
Documents in EconStor may be saved and copied for your personal and scholarly purposes.
You are not to copy documents for public or commercial purposes, to exhibit the documents publicly, to make them publicly available on the internet, or to distribute or otherwise use the documents in public.

If the documents have been made available under an Open Content Licence (especially Creative Commons Licences), you may exercise further usage rights as specified in the indicated licence.
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†

This Version: April 2010

Abstract

This paper assesses the probability method for quantifying EU consumer survey data on perceived and expected inflation. Based on household level data from the Swedish consumer survey that asks for both qualitative and quantitative responses, it is found that the theoretical assumptions of the method do not hold. In particular, estimates of unrestricted response schemes indicate that response intervals are asymmetric and that qualitative inflation expectations are formed relative to perceptions of current inflation. 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 is most accurate.

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 beliefs on qualitative scales. In the European Union (EU), beliefs of households about inflation 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 dataset on beliefs about inflation.\footnote{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 pertaining to 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).} In particular, the EU consumer survey both asks for perceptions about current inflation and expectations about future inflation. Consequently, the response data has been investigated by a large literature. Only recently, the euro cash changeover and its effects on inflation perceptions of households has given rise to a new strand of research.\footnote{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) and Fluch and Stix (2007). Abstracting from the euro cash changeover, other contributions use the EU consumer survey data to investigate belief formation in general, see, e.g., Düpke, Doeven, Fritsche, and Slacalek (2008), Forsells and Kenny (2004), Lamla and Lein (2008) and Lein and Maag (2008).} 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. Furthermore, Theil (1952) suggests that the distributional assumption may be relaxed by imposing a normal distribution instead. Combined with the assumption that respondents perceive or expect prices to be constant in qualitative terms 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 Carl-
son and Parkin (1975) and is known today as the Carlson-Parkin method or the 3-category
data on 5-category scales as it is available from the EU consumer survey. Taking into ac-
count the particular wording of 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, taken from the Swedish Consumer Tendency Survey. To the best
of my knowledge, only two studies have investigated surveys that ask for both qualitative
and quantitative responses. Defris and Williams (1979) consider 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) investigates micro-data from the University of Michigan
Survey of Consumers. In line with Defris and Williams (1979), Batchelor (1986) finds that
both the balance statistic and quantified expectations generated with the probability method
are inaccurate, in particular in the short term. This result is in contrast to my findings for
Sweden.

This paper extends the literature in several respects. First, it provides a detailed as-
ssessment of the theoretical assumptions underlying the 5-category probability method. Ex-
isting research focuses on the 3-category probability method and on testing distributional
assumptions. Joining quantitative and qualitative responses on household-level allows to es-
timate 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
responses is assessed in a long sample of 154 monthly surveys spanning 01/1996–10/2008. The
discussion centers on comparing correlation coefficients relying on 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, the probability method is assessed for quantifying both perceptions and expectations of inflation.

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

2 Data

Inflation opinions of Swedish households are being surveyed on a monthly basis since 1973. This paper uses monthly household-level response data which is available for the period 01/1996–10/2008. Unlike most surveys in other countries, the Swedish Consumer Tendency Survey jointly asks for qualitative and quantitative beliefs about inflation. The questionnaire captures beliefs in two steps. In a first step, households are asked to report perceived inflation on a five-category ordinal scale. This qualitative question is in line with the questionnaire of the Joint Harmonized EU Consumer Survey. The question reads:

“Compared with 12 months ago, do you find that prices in general are...?” “Lower \((S_1)\), about the same \((S_2)\), a little higher \((S_3)\), somewhat higher \((S_4)\), a lot higher \((S_5)\), don’t know”.

In the following, \(S_1\) through \(S_5\) denote the qualitative response categories and \(s_1\) through \(s_5\) are the fractions of households that opt for the respective category. In a second step, households are asked for a direct quantitative estimate of the current annual inflation rate. The question reads: “How much higher/lower in percent do you think prices are now?” (In other

---

4During this period, the survey comprises a monthly sample of roughly 1,500 households which are interviewed by telephone. The 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. \(s_1\) through \(s_5\) sum up to 100%. 4
words, the present rate of inflation). As a result and in contrast to the Joint Harmonized EU Consumer Survey, households report both qualitative and quantitative beliefs about inflation.

In a similar manner, expected inflation is captured in a first step by asking:

“Compared to the situation today, do you think that in the next 12 months prices in general will . . . ?” “Go down a little ($S_1$), stay more or less the same ($S_2$), go up more slowly ($S_3$), go up at the same rate ($S_4$), go up faster ($S_5$), don’t know”.

In a second step, quantitative beliefs are captured 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)”. 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 contain non-missing responses to the qualitative and quantitative questions are considered. Regarding inflation perceptions, 13% of observations only include a qualitative but no quantitative response. Regarding inflation expectations, 15% of observations only include a qualitative but no quantitative response. As will be discussed in the next section, a theoretical assumption of quantification methods is that households form quantitative beliefs. The high shares of missing quantitative responses can therefore 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.

As Table 1 shows, the resulting sample includes almost 200,000 observations from 154 monthly surveys spanning 01/1996–10/2008. Throughout this paper, the discussion centers on this sample. The appendix additionally presents results for a shorter sample covering 01/2002–10/2008. I consider this subsample to account for a potential structural break due to a change in the surveying institution in 01/2002. As will be shown, results for both

---

7Over the entire sample period, 667 (443) observations contain quantitative inflation perceptions (expectations) that are outliers.

8This view is supported by the distribution of missing quantitative answers by qualitative response category. For inflation perceptions, about 70% of missing quantitative answers are accounted for by respondents that opt for the qualitative category $S_3$ (“a little higher”). For inflation expectations, 40% of missing quantitative answers are accounted for by respondents that opt for the qualitative category $S_4$ (“go up at the same rate”). Note that 1.3% (2.2%) of all observations only contain a quantitative but no qualitative perception (expectation).
Table 1: Summary statistics for 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>Observations</td>
<td>197,487</td>
<td>1,031</td>
<td>1,456</td>
<td>192,845</td>
</tr>
<tr>
<td>Number of months</td>
<td>154</td>
<td>154</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Mean</td>
<td>1.81</td>
<td>0.47</td>
<td>5.91</td>
<td>2.10</td>
</tr>
<tr>
<td>Median</td>
<td>0.00</td>
<td>0.00</td>
<td>5.00</td>
<td>1.00</td>
</tr>
<tr>
<td>Standard deviation</td>
<td>4.06</td>
<td>2.68</td>
<td>5.37</td>
<td>3.73</td>
</tr>
<tr>
<td>Share of integer answers</td>
<td>0.94</td>
<td>0.86</td>
<td>0.98</td>
<td>0.92</td>
</tr>
<tr>
<td>Share of focal point answers</td>
<td>0.81</td>
<td>0.50</td>
<td>0.92</td>
<td>0.72</td>
</tr>
<tr>
<td>Share of zero responses</td>
<td>0.59</td>
<td>0.08</td>
<td>0.80</td>
<td>0.44</td>
</tr>
<tr>
<td>Mean response, given $S_1$</td>
<td>-4.79</td>
<td>-20.00</td>
<td>-2.12</td>
<td>-3.89</td>
</tr>
<tr>
<td>Mean response, given $S_2$</td>
<td>0.03</td>
<td>-0.05</td>
<td>0.46</td>
<td>0.02</td>
</tr>
<tr>
<td>Mean response, given $S_3$</td>
<td>4.72</td>
<td>3.98</td>
<td>5.95</td>
<td>4.06</td>
</tr>
<tr>
<td>Mean response, given $S_4$</td>
<td>7.94</td>
<td>5.22</td>
<td>10.72</td>
<td>4.77</td>
</tr>
<tr>
<td>Mean response, given $S_5$</td>
<td>9.58</td>
<td>4.25</td>
<td>16.00</td>
<td>3.98</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.

estimation periods are consistent, confirming the validity of the results for the full sample. As a measure of actual inflation I use the year-over-year percentage change in the Harmonized Index of Consumer Prices (HICP) as published by Eurostat.

Figure 1a plots the cross-sectional mean of quantitative inflation perceptions, together with qualitative response shares and the actual inflation rate. The figure indicates that quantitative inflation perceptions closely track actual inflation. The correlation between the two series is 0.78. Moreover, the first panel of Table 1 shows that the overall mean of inflation perceptions is 1.81%, as opposed to an average HICP inflation rate of 1.61%. Hence, inflation perceptions of the Swedish public are roughly unbiased during 01/1996–10/2008. This finding is in line with earlier results of Jonung and Laidler (1988) for Sweden.

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 coincidental.

I have also considered the Consumer Price Index (CPI) and the Consumer Price Index excluding mortgage payments and indirect taxes (CPIX). Particularly at the beginning of the sample period, these indices might have obtained more attention by the Swedish public than the HICP. Employing these alternative indices does not alter the conclusions in qualitative terms.

Inflation perceptions in Sweden seem relatively accurate compared to other countries. Relying on a monthly household survey conducted by the Federal Reserve Bank of Cleveland, Bryan and Venkatu (2001a, 2001b) find that inflation perceptions (and expectations) exceed actual inflation by several percentage points. For the
indicates 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.

Figure 1b shows the cross-sectional mean of inflation expectations. Again, inflation expectations exhibit pronounced comovement with actual inflation, the correlation between the two series being 0.70. Predictive power is relatively low, as the correlation with 12 months ahead inflation is only 0.28. Table 1 documents that inflation expectations average somewhat higher than perceptions at 2.05%. Moreover, the figures reveal a systematic difference between qualitative perceptions and expectations of inflation. Qualitative inflation perceptions are concentrated in categories $S_2$ and $S_3$ (“about the same” and “a little higher”). During 1996–2008, 87% of respondents opt for these categories. In contrast, 70% of qualitative inflation expectations fall into categories $S_2$ and $S_5$ (“stay more or less the same” and “go up faster”).

Table 1 highlights important properties of the quantitative response data. First, panel 1 shows that beliefs about inflation are highly heterogeneous. Despite the low cross-sectional means, inflation perceptions and expectations exhibit cross-sectional standard deviations of 4.06% and 3.73%, respectively. Second, panel 2 indicates that more than 90% of all quantitative answers are integers. Third, 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%.

Table 1 indicates, the most important focal point is 0%, which accounts for more than half of all focal point responses. Towards the end of the sample period, the share of zero responses declines significantly. It attains a minimum of 8% for perceptions and 12% for expectations in 06/2008. The high share of zero responses also explains the low medians of quantitative beliefs. Fourth, the cross-sectional mean of quantitative inflation perceptions is generally rising in the qualitative response category. This is shown in panel 3 which summarizes the conditional means of quantitative responses depending on the qualitative response. As opposed to inflation perceptions, qualitative inflation expectations are not ordered. For expectations, the mean response given $S_4$ (“go up at the same rate”) is

---

U.K., Driver and Windram (2007) report a correlation of perceived inflation with actual inflation of roughly 0.5 in a similar sample period.

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.

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.
Figure 1: Qualitative response shares and mean quantitative response

Notes: The lowest grey line shows the share $s_1$ of qualitative answers in category $S_1$, the second-lowest grey line shows the cumulative share of answers in categories $S_1$ and $S_2$, etc.
higher than the mean response given $S_5$ ("go up faster"). Also, in comparison to inflation perceptions, the differences between the cross-sectional means given qualitative responses $S_3$, $S_4$ and $S_5$ are only minor.\textsuperscript{14} Fifth, the relation between quantitative and qualitative responses is time varying. The differences between overall, minimum and maximum conditional means are considerable for most categories. The only exception is $S_2$ ("about the same"): Given this qualitative response, the mean quantitative response is always close to 0%.

These initial results suggest that the relation between quantitative and qualitative beliefs about inflation 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 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 mean and median values indicate that quantitative beliefs are positively skewed and therefore not normally distributed. The concentration of answers at focal points, in particular at 0%, raises additional 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

3.1 Theoretical Assumptions

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 proposed by Batchelor and Orr (1988). To begin with, the method is briefly outlined.

Assume that previous to answering the consumer survey, respondent $i$ forms a quantitative belief $\pi_{t,i}^e$ about inflation over the upcoming 12 months.\textsuperscript{15} Respondent $i$ then answers the

\textsuperscript{14}On a monthly basis, the mean of inflation perceptions is not always strictly rising too. This is indicated by the minima of monthly conditional means in panel 3 of Table 1. But the conditional means lack order only in 27 months, as opposed to 136 months for inflation expectations.

\textsuperscript{15}The analogous approach for quantifying perceived inflation $\pi_{t,i}^p$, and detailed derivations can be found in Appendix A.1.
qualitative survey question on expected inflation according to the following response scheme:

\[
\begin{align*}
\pi_{t,i}^e < -\delta_t & : \text{prices in general will go down a little (S}_1) \\
-\delta_t \leq \pi_{t,i}^e < \delta_t & : \text{stay more or less the same (S}_2) \\
\delta_t \leq \pi_{t,i}^e < \eta_t & : \text{go up more slowly (S}_3) \\
\eta_t - \pi_{t,i}^e < \pi_{t,i}^r + \eta_t & : \text{go up at the same rate (S}_4) \\
\pi_{t,i}^r \geq \pi_{t,i}^e + \eta_t & : \text{go up faster (S}_5) 
\end{align*}
\]

The response scheme is defined by the parameters \(\delta_t, \eta_t\) and \(\pi_{t,i}^r\). In the following, \(\pi_{t,i}^r\) is referred to as 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”). The first key assumption of the probability method restricts the response scheme to be fully defined by these three parameters:

**Assumption 1:** The response intervals are symmetric around 0\% and around \(\pi_{t,i}^r\).

The corresponding intervals \([-\delta_t, \delta_t]\) and \([\pi_{t,i}^r - \eta_t, \pi_{t,i}^r + \eta_t]\) 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 \(\pi_{t,i}^r\) are identical across respondents.

Quantitative inflation expectations \(\pi_{t,i}^e\) 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 \(\pi_{t,i}^e\). 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. \(\pi_{t,i}^e \sim N(\pi_{t,i}^r, (\sigma_{t,i}^e)^2)\).

The parameters of interest are the cross-sectional mean \(\pi_{t,i}^e\) and standard deviation \(\sigma_{t,i}^e\) of quantitative beliefs. As outlined in Appendix A.1, the above assumptions yield a system of 4 linearly independent equations with 5 unknowns \((\pi_{t,i}^e, \sigma_{t,i}^e, \delta_t, \eta_t, \pi_{t,i}^r)\) which can be solved for \(\pi_{t,i}^e\) and \(\sigma_{t,i}^e\). The solution for both parameters is equal to the product of reference inflation \(\pi_{t,i}^r\) and a function of the response shares \(s_{t,1}, \ldots, s_{t,5}\).
The usual identification scheme restricts reference inflation $\pi_t^r$. For quantifying inflation expectations two choices of $\pi_t^r$ 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 $\pi_t^p$ as suggested by Berk (1999). This approach is supported by empirical evidence that households are not necessarily well informed about actual inflation.\footnote{See, e.g., Bryan and Venkatu (2001a, 2001b) who document that inflation perceptions of U.S. households are significantly biased.} Identifying $\pi_t^r$ is less obvious for inflation perceptions. Following Carlson and Parkin (1975) it is commonly assumed that inflation perceptions are unbiased over the sample horizon. This assumption can be imposed by restricting $\pi_t^r$ to a constant accordingly.\footnote{The solution for $\pi_t^r$ is given by Equation (A.7) in the Appendix.} The last assumption thus reads:

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

### 3.2 Symmetry of the Response Scheme

Assumption 1 restricts response intervals to be symmetric around 0% and around $\pi_t^r$. To test the validity of this assumption I estimate an unrestricted response scheme defined by 4 threshold parameters. Assume that respondent $i$ answers the qualitative question according to the following scheme:\footnote{The identical scheme applies to inflation perceptions, with $\pi_t^r, i$ being replaced by $\pi_t^p, i$.}

\begin{align}
\pi_{t,i}^e + \varepsilon_{t,i} &< \mu_t^1 : prices \ in \ general \ will \ go \ down \ a \ little \ (S_1) \\
\mu_t^1 &\leq \pi_{t,i}^e + \varepsilon_{t,i} < \mu_t^2 : \ stay \ more \ or \ less \ the \ same \ (S_2) \\
\mu_t^2 &\leq \pi_{t,i}^e + \varepsilon_{t,i} < \mu_t^3 : \ go \ up \ more \ slowly \ (S_3) \\
\mu_t^3 &\leq \pi_{t,i}^e + \varepsilon_{t,i} < \mu_t^4 : \ go \ up \ at \ the \ same \ rate \ (S_4) \\
\pi_{t,i}^e + \varepsilon_{t,i} &\geq \mu_t^4 : \ go \ up \ faster \ (S_5) 
\end{align}

The idiosyncratic component $\varepsilon_{t,i}$ allows the response scheme to shift between individuals. Under the assumption that the idiosyncratic component represents the sum of independent idiosyncratic factors it is reasonable to assume that $\varepsilon_{t,i}$ is normally distributed. One thus
obtains an ordered probit model (Zavoina and McKelvey, 1975). In contrast to the usual identification scheme, I restrict the coefficient on the quantitative belief \( \pi_{t,i} \) to unity, whereas the variance of \( \varepsilon_{t,i} \) remains unrestricted. Assuming that \( \varepsilon_{t,i} \sim N(0, \sigma_\varepsilon^2) \), the following probabilities are obtained:

\[
\begin{align*}
P(S_1 | \pi_{t,i}, \mu_t, \sigma_t) &= \Phi \left( \frac{-\pi_{t,i} + \mu_1}{\sigma_t} \right) \\
P(S_2 | \pi_{t,i}, \mu_t, \sigma_t) &= \Phi \left( \frac{-\pi_{t,i} + \mu_2}{\sigma_t} \right) - \Phi \left( \frac{-\pi_{t,i} + \mu_1}{\sigma_t} \right) \\
P(S_3 | \pi_{t,i}, \mu_t, \sigma_t) &= \Phi \left( \frac{-\pi_{t,i} + \mu_3}{\sigma_t} \right) - \Phi \left( \frac{-\pi_{t,i} + \mu_2}{\sigma_t} \right) \\
P(S_4 | \pi_{t,i}, \mu_t, \sigma_t) &= \Phi \left( \frac{-\pi_{t,i} + \mu_4}{\sigma_t} \right) - \Phi \left( \frac{-\pi_{t,i} + \mu_3}{\sigma_t} \right) \\
P(S_5 | \pi_{t,i}, \mu_t, \sigma_t) &= 1 - \Phi \left( \frac{-\pi_{t,i} + \mu_4}{\sigma_t} \right)
\end{align*}
\]

where \( \mu_t = \{\mu_1, ..., \mu_4\} \). Deviating from the assumptions of the probability method, \( \mu_1, \mu_2 \) and \( \mu_3, \mu_4 \) are not required to be symmetric around 0% and the reference rate of inflation, respectively.

Table 2 presents the maximum likelihood estimation results of the unrestricted 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 point to systematic differences between perceptions and expectations. While the threshold parameters for inflation perceptions are increasing from \( \mu_1 = -6.91\% \) to \( \mu_4 = 13.89\% \), the thresholds for inflation expectations range between \( \mu_1 = -7.04\% \) and \( \mu_4 = 5.61\% \). For inflation expectations, the threshold parameters that define the response

\[^{19}\text{The test statistic is given by } LR = -2 (\log L_r - \log L_i) \rightarrow \chi^2(q), \text{ where } \log L_r \text{ is the log likelihood of the restricted model and } \log L_i \text{ is the log likelihood of the unrestricted model. The number of restrictions is given by } q = 1.\]
Table 2: Estimated response schemes 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>$\mu_5$</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>$\sigma_t$</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>N</td>
<td>197,487</td>
<td>192,845</td>
<td>110,071</td>
<td>109,782</td>
</tr>
<tr>
<td>Log L</td>
<td>-138,263</td>
<td>-220,521</td>
<td>-88,738</td>
<td>-136,562</td>
</tr>
</tbody>
</table>

Likelihood ratio tests

$H_0^1$: Symmetry such that $\mu_1 = -\mu_1^2$ ($q = 1$)

LR statistic | 20,978.94 | 35,749.40 | 17,427.61 | 25,138.60 |

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, respectively.

Intervals for $S_3$, $S_4$, $S_5$ are in a narrow range of 4 to 5 percent. 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. The second interpretation is suggested by the lack of order in conditional means documented in Section 2. The unordered nature of qualitative responses seems to be caused by the relative wording 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 gives consistent answers will always opt for qualitative response $S_4$ if the quantitative expectation corresponds to the subjective quantitative perception of inflation, irrespective of the level of expected inflation. The second interpretation is also supported by the substantially lower log likelihood of the model for expectations, despite the lower number of observations.

I therefore estimate a scheme for positive inflation expectations $\pi_{t,i}^e$ that accounts for the role of quantitative inflation perceptions $\pi_{t,i}^q$ at the household-level:
Table 3: Estimated relative response schemes 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>$\mu_1$</td>
<td></td>
<td>-7.172***</td>
<td>-6.014***</td>
<td>-5.056***</td>
<td>-3.928***</td>
<td>-5.316***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.184)</td>
<td>(0.116)</td>
<td>(0.0569)</td>
<td>(0.0132)</td>
<td>(0.0500)</td>
</tr>
<tr>
<td>$\mu_2$</td>
<td></td>
<td>1.322***</td>
<td>1.059***</td>
<td>0.805***</td>
<td>0.575***</td>
<td>0.814***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.0839)</td>
<td>(0.0550)</td>
<td>(0.0445)</td>
<td>(0.0104)</td>
<td>(0.0242)</td>
</tr>
<tr>
<td>$\sigma_t$</td>
<td></td>
<td>6.212***</td>
<td>4.956***</td>
<td>4.056***</td>
<td>3.182***</td>
<td>4.369***</td>
</tr>
<tr>
<td></td>
<td></td>
<td>(0.139)</td>
<td>(0.0826)</td>
<td>(0.0062)</td>
<td>(0.0039)</td>
<td>(0.0347)</td>
</tr>
<tr>
<td>N</td>
<td></td>
<td>10,046</td>
<td>15,133</td>
<td>15,584</td>
<td>19,618</td>
<td>60,381</td>
</tr>
<tr>
<td>Log $L$</td>
<td></td>
<td>-8,759</td>
<td>-12,591</td>
<td>-12,436</td>
<td>-14,921</td>
<td>-49,217</td>
</tr>
</tbody>
</table>

Likelihood ratio tests

<table>
<thead>
<tr>
<th>LR statistic</th>
<th>P-value</th>
</tr>
</thead>
<tbody>
<tr>
<td>$H_0^1$: Identical threshold parameters $\mu_1$, $\mu_2$ ($q = 6$)</td>
<td>746.60</td>
</tr>
<tr>
<td>$H_0^2$: Identical standard deviation $\sigma_t$ ($q = 3$)</td>
<td>907.96</td>
</tr>
<tr>
<td>$H_0^3$: Identical thresholds and standard deviation ($q = 9$)</td>
<td>1,020.93</td>
</tr>
</tbody>
</table>

Notes: This table shows maximum likelihood estimates of the relative response scheme (3). Monthly data, 01/2002–10/2008. Standard errors in parentheses. *, ** and *** indicate statistical significance at the 10%, 5% and 1% level, respectively.

\[
\pi_{t,i}^c + \varepsilon_{t,i} < \pi_{t,i}^p + \mu_1^1 : \text{prices in general will go up more slowly (S_3)}
\]
\[
\pi_{t,i}^p + \mu_1^1 \leq \pi_{t,i}^c + \varepsilon_{t,i} < \pi_{t,i}^p + \mu_2^1 : \text{go up at the same rate (S_4)}
\]
\[
\pi_{t,i}^c + \varepsilon_{t,i} \geq \pi_{t,i}^p + \mu_2^1 : \text{go up faster (S_5)}
\]

where $\varepsilon_{t,i} \sim N(0, \sigma_t^2)$. This scheme assumes that the respondent will opt for $S_4$ if $\pi_{t,i}^c - \pi_{t,i}^p$ lies in the range $[\mu_1^1, \mu_2^1]$. Maximum likelihood estimation results can be found in the last column of Table 3. All parameters are highly significant, confirming that the qualitative response about expected inflation is linked to the quantitative inflation perception. Moreover, the response interval is highly asymmetric: Qualitative answers are more responsive to an increase of quantitative expectations over perceptions than to a decrease. The significance of the relative response scheme suggests that despite the lack of an unambiguous relation between quantitative and qualitative expectations, the 5-category survey contains more information than a 3-category 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 normality assumption, the estimated response interval is not symmetric around 0%, both for inflation...
perceptions and expectations. Furthermore, the response interval is not symmetric around \( \pi_t^r \) for inflation expectations. 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.

3.3 Homogeneity of the Response Scheme

Assumption 2 imposes that threshold parameters \( \delta_t \) and \( \eta_t \) and reference inflation \( \pi_t^r \) are homogeneous across respondents. Since the Swedish dataset only contains one observation per individual, this assumption is tested by estimating the response scheme for different income groups. 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 show likelihood ratio tests of three restrictions. The null hypotheses state that threshold parameters are identical across income groups \((H_0^1)\), that standard deviations are identical across income groups \((H_0^2)\) and that threshold parameters and standard deviations are identical across income groups \((H_0^3)\). All three hypotheses are clearly rejected.

Table 3 presents estimation results for the relative response scheme (3) 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.

---

20This finding is consistent with Henzel and Wollmershäuser (2005) who investigate data from a special edition of the ifo World Economic Survey that directly asks respondents to 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.

21I have also considered educational groups, with unchanged qualitative results.

22Note that the mean of beliefs about inflation also depends on socioeconomic characteristics. The cross-sectional means of perceptions and expectations are declining in income. This pattern is consistent with the estimated response schemes that suggest that individuals in the highest income quartile experience deviations of inflation from zero as more relevant in qualitative terms than individuals in lower income quartiles.

15
Table 4: Estimated response schemes for perceived inflation by income groups

<table>
<thead>
<tr>
<th>Perceptions</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>9.074***</td>
<td>-8.201***</td>
<td>-6.657***</td>
<td>-5.833***</td>
<td>-7.308***</td>
<td></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>
<td></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>
<td></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>
<td></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>
<td></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>
<td></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>
<td></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>
<td></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>
<td></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>
<td></td>
</tr>
<tr>
<td>N</td>
<td>17,092</td>
<td>24,845</td>
<td>25,482</td>
<td>32,614</td>
<td>100,033</td>
<td></td>
</tr>
<tr>
<td>Log L</td>
<td>-15,584</td>
<td>-20,718</td>
<td>-19,346</td>
<td>-23,163</td>
<td>-79,851</td>
<td></td>
</tr>
</tbody>
</table>

Likelihood ratio tests

$H^0_0$: Identical threshold parameters $\mu^1_t,\ldots,\mu^4_t$ ($q = 12$) 1,868.21 0.00
$H^0_2$: Identical standard deviation $\sigma_t$ ($q = 3$) 1,115.68 0.00
$H^0_3$: Identical thresholds and standard deviation ($q = 15$) 2,080.00 0.00

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, respectively.
Table 5: Estimated response schemes for expected inflation by income groups

<table>
<thead>
<tr>
<th>Expectations</th>
<th>Income groups (1st quartile lowest)</th>
<th>Overall</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>1st quartile</td>
<td>2nd quartile</td>
</tr>
</tbody>
</table>

| $\mu_t^1$ | -9.229*** | -3.303*** | -7.031*** | -5.968*** | -7.598*** |
| $\mu_t^2$ | 0.703*** | 0.440*** | 0.591*** | 0.783*** | 0.600*** |
| $\mu_t^3$ | 2.452*** | 1.972*** | 1.816*** | 1.835*** | 1.950*** |
| $\mu_t^4$ | 7.075*** | 6.164*** | 5.136*** | 4.047*** | 5.502*** |
| $\sigma_t$ | 6.089*** | 5.448*** | 4.488*** | 3.789*** | 4.905*** |

| N           | 17,232   | 24,757   | 25,384   | 32,528   | 99,901   |
| Log $L$     | -21,816  | -31,163  | -31,206  | -39,092  | -124,179 |

**Likelihood ratio tests**

- $H_0^0$: Identical threshold parameters $\mu_t^1, ..., \mu_t^4$ ($q = 12$) LR statistic 1,102.84 P-value 0.00
- $H_0^1$: Identical standard deviation $\sigma_t$ ($q = 3$) LR statistic 1,689.36 P-value 0.00
- $H_0^2$: Identical thresholds and standard deviation ($q = 15$) LR statistic 1,804.00 P-value 0.00

**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, respectively.
3.4 Normality of Quantitative Responses

Assumption 3 requires that the cross-sectional distribution of quantitative beliefs is normal. Normality has been tested and rejected 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 patterns can also be found in the Swedish survey data, as panel 1 of Table 6 indicates. Beliefs about inflation exhibit a pronounced positive skewness and are leptokurtic. Consequently, 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 a noncentral $t$ distribution 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 this distribution using the Kolmogorov-Smirnov test. Results are summarized in panel 3 of Table 6 and show that this null hypothesis is also rejected in all months.

These formal tests do not answer the question which parametric distribution produces the best quantification results. The answer will also depend on the time period. During 1996–2007, the high share of zero responses cannot be reconciled with both the normal and noncentral $t$ distributions. With the rise in perceptions and expectations of inflation in 2008, the shape of the empirical distribution becomes somewhat smoother and less skewed, as the share of zero responses declines. Overall, the results indicate that differences in the relative fit of common parametric distributions are predominated by the high share of zero responses. This conjecture is consistent with Berk (1999), Dasgupta and Lahiri (1992) and Smith and McAleer (1995) who find that the accuracy of the quantified series does not significantly vary between any of the common parametric distributions.

---

23 The Kolmogorov-Smirnov statistic is given by $D_n(F) = \sup_x |F_n(x) - F(x)|$, 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.
Table 6: Tests for the distribution 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>Skewness</td>
<td>4.04</td>
<td>0.05</td>
</tr>
<tr>
<td>Kurtosis</td>
<td>40.51</td>
<td>9.96</td>
</tr>
</tbody>
</table>

*Jarque-Bera test for normal distribution*

<table>
<thead>
<tr>
<th></th>
<th>Overall</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>J-B statistic</td>
<td>6,099.24</td>
<td>771.61</td>
<td>22,361.00</td>
</tr>
<tr>
<td>P-value</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

*Kolmogorov-Smirnov test for noncentral t distribution*

<table>
<thead>
<tr>
<th></th>
<th>Overall</th>
<th>Min</th>
<th>Max</th>
</tr>
</thead>
<tbody>
<tr>
<td>K-S statistic</td>
<td>0.31</td>
<td>0.12</td>
<td>0.41</td>
</tr>
<tr>
<td>P-value</td>
<td>0.00</td>
<td>0.00</td>
<td>0.00</td>
</tr>
<tr>
<td>μ</td>
<td>0.51</td>
<td>0.10</td>
<td>4.00</td>
</tr>
<tr>
<td>df</td>
<td>1.02</td>
<td>1</td>
<td>3</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 \( \chi^2_2 \) distributed. The approximate 1% critical values for the Kolmogorov-Smirnov are given by \( 1.52N^{-0.5} \), 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.

3.5 Defined Reference Rate of Inflation

Assumption 4 requires that reference inflation \( \pi_t^r \) 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. To assess this assumption, Figure 2a shows the mean of quantitative inflation perceptions given by households that opt for qualitative response \( S_4 \). 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 reference rate of inflation is constant over time is clearly at odds with this pattern.

For expected inflation, reference inflation is commonly assumed to be equal to actual inflation or to (previously quantified) perceived inflation. Figure 2b shows that the conditional mean of inflation expectations given qualitative answer \( S_4 \) is less volatile, with a standard deviation of only 0.53%. The figure plots three alternative measures of reference inflation: The conditional mean of quantitative inflation perceptions of respondents that expect prices to “go up at the same rate” \( (S_4) \), quantified inflation perceptions and actual HICP inflation.\(^{24}\)

\(^{24}\)Inflation perceptions are quantified using the 5-category probability method with the unbiasedness assumption.
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 average level difference only 0.39%. The similarity of these series is in line with the finding that qualitative expectations are formed relative to quantitative perceptions. In contrast, the correlations of the conditional mean with quantified inflation perceptions and actual inflation are -0.15 and 0.06, respectively. In both cases, the level difference is substantial. Consequently, the assumption that reference inflation corresponds to quantified or actual inflation can be rejected. However, the correlations of the conditional mean with quantified inflation perceptions and actual inflation increase to 0.46 and 0.41 during 2002–2008. Notably, a comovement of these measures of moderate inflation with the conditional mean is apparent towards the end of the sample period, when actual inflation substantially increases.

Alternatively, Assumption 4 can be assessed based on the implied level of reference inflation. The implied level of reference inflation is obtained by combining the cross-sectional mean of quantitative responses with Assumptions 1 to 3.\textsuperscript{25} Assessing this series amounts to a joint test of Assumptions 1 to 3. Figure 2a indicates that implied reference inflation fluctuates around a similar level as the conditional mean of inflation perceptions. However, the correlation between the two series is 0.06. For inflation expectations shown in Figure 2b, the implied reference inflation averages 2% below the conditional mean. The correlation of the two series is 0.31. Provided that the true reference rate of inflation is 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 this finding, the next section assesses the joint validity of all 4 hypotheses in more detail.

3.6 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 with the 5-category probability method.\textsuperscript{26} This yields the threshold parameters $\delta_t$ and $\eta_t$ which can

\textsuperscript{25}Given the cross-sectional mean $\pi_t^e$ of quantitative inflation expectations, implied reference inflation can be obtained by rearranging Equation (A.3).

\textsuperscript{26}Inflation perceptions are quantified by imposing the unbiasedness condition. For inflation expectations it is assumed that reference inflation is equal to quantified inflation perceptions. Detailed derivations are provided in Appendix A.1.
Figure 2: Measures of reference inflation $\pi_t^r$

Notes: This figure shows alternative measures of $\pi_t^r$. 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 using the 5-category probability method. For quantifying perceptions the unbiasedness condition is imposed.
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 (“quant.”) is compared to the actual share of qualitative responses (“N”). For inflation perceptions, Figure 3a 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 positively skewed and leptokurtic. Moreover, the low fraction of quantitative responses in the implied range of $S_3$ appears to be a direct consequence of fitting the normal distribution to the high share of zero responses.

A similar pattern is obtained for inflation expectations. Figure 3b 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 range of $S_5$ exceeds the actual share of qualitative responses. This pattern also relates to the finding of the previous section, according to which the mean quantitative answer of respondents opting for qualitative answer $S_4$ is significantly higher than actual inflation or quantified inflation perceptions. Consequently, a large fraction of these quantitative answers fall into the interval of the qualitative answer $S_5$.

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 these fractions 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 $S_2$. On average, 99% of quantitative responses associated with qualitative answer $S_2$ lie within the implied response interval. $S_2$ is the most important qualitative response, accounting for roughly 59% of perceptions and 42% of expectations during 1996–2008. Regarding inflation perceptions shown in Figure A.1a, 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

Note that by construction, the actual share of qualitative responses corresponds to the predicted share of quantitative responses under the normality assumption.
opt for this qualitative category.

The pattern is different for inflation expectations. Figure A.1b 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. Moreover, the previous section has shown 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 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 for inflation perceptions. The next section assesses the implications for the accuracy of the probability method.
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.
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. Perceptions and expectations of inflation are quantified by imposing the usual restrictions. Inflation perceptions are assumed to be unbiased with respect to HICP inflation. Inflation expectations are quantified by setting reference inflation equal to HICP inflation and alternatively, following Berk (1999), to quantified perceived inflation.\(^{28}\) 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).\(^{29}\) 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 given by

\[
s_5 + 0.5s_4 - 0.5s_3 - s_1 = s_p - s_n,
\]

where \(s_p\) and \(s_n\) are the fractions of respondents that report that prices are rising and falling, respectively. The third alternative is the Pesaran (1987) regression approach for 3-category response data.\(^{30}\)

The primary measure of accuracy I consider is the (Pearson) correlation coefficient between the quantified series and the cross-sectional mean of quantitative responses. As opposed to 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. Relying on the Fisher \(z\)-transformation, the null hypothesis that two correlation coefficients are equal

\(^{28}\)Perceived inflation is quantified using the 5-category probability method under the assumption of unbiasedness with respect to HICP inflation.

\(^{29}\)Answer categories are aggregated following Berk (1999), see Appendix A.2 for details.

\(^{30}\)Unlike 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

\[
\pi_t = \frac{\beta_1 s_p^t - \beta_2 s_n^t}{1 - \beta_3 s_p^t} + \varepsilon_t,
\]

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.
\( (\rho_1 = \rho_2) \) can be tested using the following statistic:

\[
z = \tanh^{-1}(\rho_1) - \tanh^{-1}(\rho_2)
\]

where \( \tanh^{-1}(\rho_i) = 0.5\ln \left( \frac{1+\rho_i}{1-\rho_i} \right) \). The \( z \) statistic is approximately normal with variance \( \frac{1}{T_1-3} + \frac{1}{T_2-3} \), where \( T_1 \) and \( T_2 \) are the sample sizes underlying correlation coefficients \( \rho_1 \) and \( \rho_2 \), respectively. However, the normal approximation may be inaccurate in the present case because \( |\rho_i| \) is high and the underlying series are serially dependent (Mudholkar, 2006). I therefore assess significance based on double block bootstrap confidence intervals for the \( z \) statistic.\(^{31}\)

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. 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 respective row (\( \rho_2 \)) is zero. Since the quantified series and the mean quantitative beliefs are highly persistent, the discussion focuses on results for first differences.\(^{32}\) These results are not subject to spurious regression problems as the first differences are stationary. However, the results on the significance of correlation are mostly consistent for levels and first differences.

Panel 1 of Table 7 indicates that in terms of correlation with the mean of quantitative perceptions, all quantification methods perform well. The correlation in first differences is 0.86 for the series generated with the 5-category probability method. The 3-category probability method generates virtually identical results. Interestingly, the 5-category and 3-category balance statistics are more accurate, with correlation coefficients of 0.91 and 0.89, respectively. 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.

Regarding expectations, panel 2 of Table 7 shows that the accuracy of the 5-category probability method depends on the imposed reference inflation. Employing quantified per-

---

\(^{31}\)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.

\(^{32}\)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>First differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias</td>
<td>MAE</td>
</tr>
<tr>
<td>P.HICP (5 cat.)</td>
<td>-0.17</td>
<td>0.22</td>
</tr>
<tr>
<td>P.HICP (3 cat.)</td>
<td>-0.17</td>
<td>0.19</td>
</tr>
<tr>
<td>Balance (5 cat.)</td>
<td>-0.20</td>
<td>0.20</td>
</tr>
<tr>
<td>Balance (3 cat.)</td>
<td>-0.20</td>
<td>0.22</td>
</tr>
<tr>
<td>Pesaran (3 cat.)</td>
<td>-0.21</td>
<td>0.25</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>P.Perc. (5 cat.)</td>
<td>-0.85</td>
<td>0.85</td>
</tr>
<tr>
<td>P.HICP (5 cat.)</td>
<td>-0.87</td>
<td>0.87</td>
</tr>
<tr>
<td>P.HICP (3 cat.)</td>
<td>-0.49</td>
<td>0.49</td>
</tr>
<tr>
<td>Balance (5 cat.)</td>
<td>-0.49</td>
<td>0.56</td>
</tr>
<tr>
<td>Balance (3 cat.)</td>
<td>-0.49</td>
<td>0.51</td>
</tr>
<tr>
<td>Pesaran (3 cat.)</td>
<td>0.12</td>
<td>0.22</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 assumption 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 assumption that expectations are unbiased. *Bias* is defined as the difference between the mean of quantified beliefs and the mean of quantitative responses. *ρ* is the correlation coefficient with the mean of quantitative responses. *z* is the Fisher z statistic for testing the null hypothesis that the difference between the correlation coefficient in the first row of each panel and the correlation coefficient in the respective row is zero. *, ** and *** indicate statistical significance of *z* at the 10%, 5% and 1% level based on double block bootstrap percentile confidence intervals.
ceptions generates significantly better results than employing actual HICP inflation, as the
z-statistic indicates. The correlation coefficients are 0.80 and 0.50, respectively. 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 statistically significant, however. Results for the sample
01/2002–10/2008 in Table A.1 confirm these findings.

In sum, all quantification methods generate series that are highly correlated with the cross-
sectional mean of quantitative inflation perceptions. The 5-category balance statistic tracks
actual quantitative perceptions most accurately. 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. However, the 5-category
probability method may perform weakly to quantify expectations, depending on the chosen
reference inflation. Moreover, the similar accuracy of the 5-category probability method and
the 3-category methods signals that the 5-category probability method does not efficiently
use information from positive response categories.

4.2 Heterogeneity of Beliefs

The cross-sectional heterogeneity of beliefs is subject to increasing research in macroeco-
nomics. This section investigates how to best infer cross-sectional heterogeneity from qual-
itative survey data. Cross-sectional heterogeneity is measured by the standard deviation of
quantitative beliefs.

The 5-category probability method not only allows to identify the mean but also the
standard deviation of the fitted normal distribution given by Equation (A.4). 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.9). The second

33Batchelor (1986) documents that the quantified series do not predict the direction of change in mean quan-
titative 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.
alternative is an index of qualitative variation (IQV) based on response shares \( s_1 \) through \( s_5 \):

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

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 \( 0 \leq IQV \leq 1 \). 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).\(^{34}\) This index is given by:

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

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 \( DIS = s_p + s_n - (s_p - s_n)^2 \).\(^{35}\)

Table 8 summarizes the results. Since the time series of heterogeneity are less persistent than the series of means, the discussion centers on results in levels.\(^{36}\) The alternative measures of heterogeneity 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 actual 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 weakly. The correlation coefficient is 0.30. The 3-category probability method performs significantly better, as the \( z \) statistic indicates. Qualitative measures of variation are even more highly

\(^{35}\)The disconformity index relies on the same theoretical assumptions as the 3-category balance statistic, see Batchelor (1986).
\(^{36}\)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 heterogeneity

<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>
</tbody>
</table>
| Implied SD (P.HICP, 3 cat.)    | -1.83| 1.83| 1.88 | 0.68 | -0.52*| 0.00 | 0.28| 0.36 | 0.36 | -0.06
| IQV (5 cat.)                   | 0.83 | -0.90**|       |      | 0.48| -0.19**|      |
| DSQ (5 cat.)                   | 0.87 | -1.01**|       |      | 0.47| -0.19**|      |
| DIS (3 cat.)                   | 0.35 | -0.05 |       |      | 0.30| 0.01  |      |

<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>0.80</td>
<td>-0.29**</td>
</tr>
<tr>
<td>DSQ (5 cat.)</td>
<td>0.45</td>
<td>0.32</td>
</tr>
<tr>
<td>DIS (3 cat.)</td>
<td>0.30</td>
<td>0.49**</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. Bias is defined as the difference between the mean of quantified beliefs and the mean of quantitative responses. ρ is the correlation coefficient with the standard deviation of quantitative responses. z is the Fisher z statistic for testing the null hypothesis that the difference between the correlation coefficient in the first row of each panel and the correlation coefficient in the respective row is zero. *, ** and *** indicate statistical significance of z at the 10%, 5% and 1% level based on double block bootstrap percentile confidence intervals.
correlated with the standard deviation of quantitative perceptions. The correlation coefficients of the IQV and the DSQ are 0.83 and 0.87, respectively. 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 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 of the 3-category approach is about as accurate as the 5-category probability method with reference inflation given by quantified perceptions. The IQV most closely tracks actual heterogeneity of quantitative responses. Its 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 is only moderately correlated with actual heterogeneity. Even lower is the correlation for the disconformity index, which is in line with earlier findings of Batchelor (1986). Results for first differences are broadly consistent, with the exception that for quantifying expectations the IQV does not outperform other methods anymore. Results for the shorter sample spanning 01/2002–10/2008 in Table A.2 confirm the relatively high accuracy of the IQV compared to other methods.37

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 cross-sectional heterogeneity. Moreover, the 3-category probability method performs better than the 5-category method, which is consistent with findings of the previous section.38 The IQV generally 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 heterogeneity of inflation perceptions. A possible interpretation of this result is that the IQV is less distorted by the lack of order in qualitative inflation expectations than the DSQ index.

37 Note that in this sample 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.

38 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.
5 Which Quantification Method Should Be Used?

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

Despite the violation of theoretical assumptions, the accuracy of the 5-category probability method for quantifying the cross-sectional 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 use the additional information from three positive response categories. The most accurate method is the 5-category balance statistic.

For quantifying the cross-sectional mean of inflation expectations, the accuracy of the 5-category probability method largely depends on the chosen reference inflation $\pi^*_t$. Employing quantified inflation perceptions yields significantly more accurate results than employing actual HICP inflation. At first sight this finding is inconsistent with the zero correlation of quantified perceptions and reference inflation during the full sample as documented in Section 3.5. 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 increases to 0.46 during 2002–2008. In particular, quantified perceptions surge in 2007 and 2008, while reference inflation increases to about 6%. This suggests that the 5-category probability method with quantified perceptions as reference inflation will gain relative accuracy once inflation perceptions substantially diverge from actual inflation.\(^{39}\)

It remains the question why the simple balance statistic tends to trace mean beliefs even more closely than 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

\(^{39}\text{This is the case, e.g., around the euro cash changeover in 2002 (ECB, 2005).}\)
Table 9: Actual and imposed conditional means

<table>
<thead>
<tr>
<th>Mean response, given $S_1$</th>
<th>Perceptions</th>
<th>Expectations</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Actual</td>
<td>Balance</td>
</tr>
<tr>
<td></td>
<td>-4.79</td>
<td>-3.96</td>
</tr>
<tr>
<td>Mean response, given $S_2$</td>
<td>0.03</td>
<td>-0.23</td>
</tr>
<tr>
<td>Mean response, given $S_3$</td>
<td>4.72</td>
<td>3.50</td>
</tr>
<tr>
<td>Mean response, given $S_4$</td>
<td>7.94</td>
<td>7.22</td>
</tr>
<tr>
<td>Mean response, given $S_5$</td>
<td>9.58</td>
<td>10.95</td>
</tr>
</tbody>
</table>

Notes: The column Actual presents results for actual quantitative survey responses, 01/1996–10/2008. 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. Conditional means are scaled to match the mean and standard deviation of actual conditional means.

Table 9 reveals that both methods impose (scaled) conditional means that match actual conditional means quite well. 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.

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.

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. In particular, the IQV seems less distorted by a lack of order in qualitative inflation expectations than other quantification methods. The findings also indicate that the 5-category probability method with quantified perceptions as reference inflation might gain relative accuracy once inflation perceptions deviate from actual inflation. This also points to a general limitation of the results: Findings on the quantification of expectations might differ from other countries as Swedish consumers are relatively well

\[\text{The (unscaled) conditional means of the 5-category balance statistic are -1, 0.5, 0, 0.5, 1. The implied conditional means of the probability method are computed by numerical integration using the quantified parameters } \pi_t, \sigma, \delta, \eta, \pi_t. \text{ Perceptions are quantified employing the 5-category probability method unbiased with respect to HICP inflation. Expectations are quantified using the 5-category probability method with reference inflation given by quantified perceptions.}\]
informed about actual HICP inflation.

6 Conclusion

This paper assesses the validity and accuracy of the 5-category probability method for quantifying household perceptions and expectations of inflation. The analysis capitalizes on jointly available qualitative and quantitative response data from the Swedish Consumer Tendency Survey. Relying on monthly household-level data covering 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 found that qualitative inflation expectations are formed relative to inflation perceptions, which is a direct result of survey design. An important consequence is that the conditional mean of quantitative inflation expectations is not monotonously rising in qualitative response categories. Furthermore, quantitative beliefs are not normally distributed and cannot be reconciled with a noncentral $t$ distribution either. The usual assumptions on the reference rate of inflation (the “moderate” rate of inflation) are shown to be at odds with the data.

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 inflation. Nevertheless, 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 use information from positive qualitative response categories. In sum, the results are still 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 beliefs is the index of qualitative variation.

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 identify the reference rate of inflation. Moreover, a non-parametric analysis should 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). All derivations equally hold for quantifying inflation perceptions, in which case \( \pi_t^e \) is substituted with \( \pi_t^P \).

The response scheme for inflation perceptions is given by:

- \( \pi_{t,i}^P < -\delta_t \): prices in general are lower (S1)
- \( -\delta_t \leq \pi_{t,i}^P < \delta_t \): about the same (S2)
- \( \delta_t \leq \pi_{t,i}^P < \pi_t^r - \eta_t \): a little higher (S3)
- \( \pi_t^r - \eta_t \leq \pi_{t,i}^P < \pi_t^r + \eta_t \): moderately higher (S4)
- \( \pi_{t,i}^P \geq \pi_t^r + \eta_t \): a lot higher (S5)

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

\[
\begin{align*}
    s_t^1 &= P(\pi_{t,i}^e < -\delta_t) = \Phi \left( \frac{-\delta_t - \pi_t^e}{\sigma_t} \right) \\
    s_t^2 &= P(-\delta_t \leq \pi_{t,i}^e < \delta_t) = \Phi \left( \frac{\delta_t - \pi_t^e}{\sigma_t} \right) - \Phi \left( \frac{-\delta_t - \pi_t^e}{\sigma_t} \right) \\
    s_t^3 &= P(\delta_t \leq \pi_{t,i}^e < \pi_t^r - \eta_t) = \Phi \left( \frac{\pi_t^r - \eta_t - \pi_t^e}{\sigma_t} \right) - \Phi \left( \frac{\delta_t - \pi_t^e}{\sigma_t} \right) \\
    s_t^4 &= P(\pi_t^r - \eta_t \leq \pi_{t,i}^e < \pi_t^r + \eta_t) = \Phi \left( \frac{\pi_t^r + \eta_t - \pi_t^e}{\sigma_t} \right) - \Phi \left( \frac{\pi_t^r - \eta_t - \pi_t^e}{\sigma_t} \right) \\
    s_t^5 &= P(\pi_{t,i}^e \geq \pi_t^r + \eta_t) = 1 - \Phi \left( \frac{\pi_t^r + \eta_t - \pi_t^e}{\sigma_t} \right)
\end{align*}
\] (A.1)

where \( \Phi(.) \) is the standard normal cumulative distribution function. The system of equations (A.1) can be rewritten to obtain a system of 4 linearly independent equations with 5 unknowns (\( \pi_t^e, \sigma_t, \delta_t, \eta_t, \pi_t^r \)).
System (A.2) can be solved for the mean $\pi_t^e$ of expected inflation:

$$
\pi_t^e = \frac{G_t^2 + G_t^3}{G_t^2 + G_t^3 - G_t^4 - G_t^5}
$$

(A.3)

In the following, $\pi_t^e$ is referred to as “expected inflation” (rather than “mean of expected inflation”). The remaining unknowns are given by:

$$
\begin{align*}
\sigma_t &= \frac{\pi_t^r}{G_t^2 + G_t^3 - G_t^4 - G_t^5} \\
\delta_t &= \frac{G_t^4}{G_t^2 + G_t^3 - G_t^4 - G_t^5} \\
\eta_t &= \frac{G_t^1}{G_t^2 + G_t^3 - G_t^4 - G_t^5}
\end{align*}
$$

(A.4) (A.5) (A.6)

For quantifying inflation perceptions, $\pi_t^r$ is commonly identified by restricting inflation perceptions to be unbiased over the sample period. Rearranging Equation (A.3) and imposing unbiasedness yields:

$$
\pi_t^r = \frac{\pi}{t \sum_{t=1}^T \frac{G_t^1 + G_t^3}{G_t^2 + G_t^3 - G_t^4 - G_t^5}}
$$

(A.7)

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 given by:

\[
\begin{align*}
    s_n^t &= P(\pi_{t,i}^e < -\delta) = \Phi \left( \frac{-\delta - \pi_{t,i}^e}{\sigma_t} \right) \\
    s_s^t &= P(-\delta < \pi_{t,i}^e < \delta) = \Phi \left( \frac{\delta - \pi_{t,i}^e}{\sigma_t} \right) - \Phi \left( \frac{-\delta - \pi_{t,i}^e}{\sigma_t} \right) \\
    s_p^t &= P(\delta < \pi_{t,i}^e) = 1 - \Phi \left( \frac{\delta - \pi_{t,i}^e}{\sigma_t} \right)
\end{align*}
\]

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

\[
\begin{align*}
    \pi_t^e &= \delta \frac{\Phi^{-1}(s_n^t) + \Phi^{-1}(1 - s_p^t)}{\Phi^{-1}(s_n^t) - \Phi^{-1}(1 - s_p^t)} \quad (A.8) \\
    \sigma_t &= \delta \frac{2}{\Phi^{-1}(1 - s_p^t) - \Phi^{-1}(s_n^t)} \quad (A.9)
\end{align*}
\]

The identical equations hold for perceived inflation \( \pi_t^p \). In line with the existing literature, the system is identified by imposing unbiasedness of beliefs with respect to actual inflation during the sample period. The parameter \( \delta \) is restricted accordingly:

\[
\delta = \frac{\pi}{\frac{1}{T} \sum_{t=1}^{T} \left[ \frac{\Phi^{-1}(s_n^t) + \Phi^{-1}(1 - s_p^t)}{\Phi^{-1}(s_n^t) - \Phi^{-1}(1 - s_p^t)} \right]} \]

42
A.3 Further Results
Figure A.1: Response fractions below and above the implied interval

Notes: These figures show the fractions of quantitative responses that lie below (low) and above (high) the implied response interval defined by the simultaneous qualitative response. Fractions are relative to the sum of qualitative responses in the respective 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, 2002–2008

Notes: Sample period 01/2002–10/2008. See footnote of Figure 3 for a detailed description.
Figure A.3: Response fractions below and above the implied interval, 2002–2008

Notes: Sample period 01/2002–10/2008. See footnote of Figure A.1 for a detailed description.
Figure A.4: Cross-sectional mean of inflation perceptions
Figure A.5: Cross-sectional mean of inflation expectations
Table A.1: Accuracy of quantified inflation perceptions and expectations, 2002–2008

<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>P.HICP (5 cat.)</td>
<td>-0.26</td>
<td>0.27</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>Expectations</th>
<th>Level</th>
<th>First differences</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Bias</td>
<td>MAE</td>
</tr>
<tr>
<td>P.Perc. (5 cat.)</td>
<td>-0.76</td>
<td>0.76</td>
</tr>
<tr>
<td>P.HICP (5 cat.)</td>
<td>-0.92</td>
<td>0.92</td>
</tr>
<tr>
<td>P.HICP (3 cat.)</td>
<td>-0.39</td>
<td>0.39</td>
</tr>
<tr>
<td>Balance (5 cat.)</td>
<td>-0.22</td>
<td>0.35</td>
</tr>
<tr>
<td>Balance (3 cat.)</td>
<td>-0.27</td>
<td>0.31</td>
</tr>
<tr>
<td>Pesaran (3 cat.)</td>
<td>0.05</td>
<td>0.22</td>
</tr>
</tbody>
</table>

*Notes:* Sample period 01/2002–10/2008. See Table 7 for a detailed description.
Notes: SD quantitative is the cross-sectional standard deviation of quantitative survey responses. Implied SD (P.HICP, 5 cat.) is the implied standard deviation from the 5-category probability method under the HICP unbiasedness condition. 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.
Figure A.7: Heterogeneity of inflation expectations

Notes: SD quantitative is the cross-sectional standard deviation of quantitative survey responses. 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. 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.
## Table A.2: Accuracy of measures of cross-sectional heterogeneity, 2002–2008

<table>
<thead>
<tr>
<th></th>
<th>Perceptions</th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Level</td>
<td>First differences</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td>Bias</td>
<td>MAE</td>
<td>RMSE</td>
<td>ρ</td>
<td>z</td>
<td>Bias</td>
<td>MAE</td>
<td>RMSE</td>
<td>ρ</td>
</tr>
<tr>
<td>Implied SD (P.HICP, 5 cat.)</td>
<td>-2.00</td>
<td>2.00</td>
<td>2.04</td>
<td>0.27</td>
<td></td>
<td></td>
<td>0.00</td>
<td>0.28</td>
<td>0.34</td>
<td>0.20</td>
</tr>
<tr>
<td>Implied SD (P.HICP, 3 cat.)</td>
<td>-1.86</td>
<td>1.86</td>
<td>1.88</td>
<td>0.65</td>
<td>-0.51</td>
<td></td>
<td>0.00</td>
<td>0.24</td>
<td>0.31</td>
<td>0.45</td>
</tr>
<tr>
<td>IQV (5 cat.)</td>
<td></td>
<td>0.80</td>
<td>-0.82*</td>
<td></td>
<td></td>
<td></td>
<td>0.47</td>
<td>-0.26**</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DSQ (5 cat.)</td>
<td>-0.15</td>
<td>0.43</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.33</td>
<td>0.13</td>
<td></td>
<td></td>
</tr>
<tr>
<td>DC (3 cat.)</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.29</td>
<td>0.09</td>
<td></td>
<td></td>
</tr>
</tbody>
</table>

|                  | Expectations |                           |                           |                         |                      |                         |                      |                      |                      |                      |
|                  | Level        | First differences          |                           |                         |                      |                         |                      |                      |                      |                      |
|                  |             | Bias  | MAE   | RMSE  | ρ      | z     | Bias | MAE   | RMSE  | ρ     | z     |
| Implied SD (P.Perc, 5 cat.) | -2.35 | 2.35  | 2.40  | 0.30   |        |      | 0.00 | 0.28  | 0.36  | 0.27  |        |
| Implied SD (P.HICP, 5 cat.)  | -2.51  | 2.51  | 2.57  | 0.30   | 0.01   |      | 0.00 | 0.33  | 0.42  | 0.27  | -0.01 |
| Implied SD (P.HICP, 3 cat.)  | -1.88  | 1.88  | 1.90  | 0.42   | -0.13  |      | 0.00 | 0.29  | 0.36  | 0.26  | 0.01  |
| IQV (5 cat.)      |           | 0.33  | -0.03 |        |        |      | 0.05 | 0.22  |        |        |        |
| DSQ (5 cat.)      | -0.09  | 0.41  |        |        |        |      | 0.10 | 0.17  |        |        |        |
| DC (3 cat.)       |           | 0.09  | 0.23  |        |        |      | 0.13 | 0.14  |        |        |        |

**Notes:** Sample period 01/2002–10/2008. See Table 8 for a detailed description.