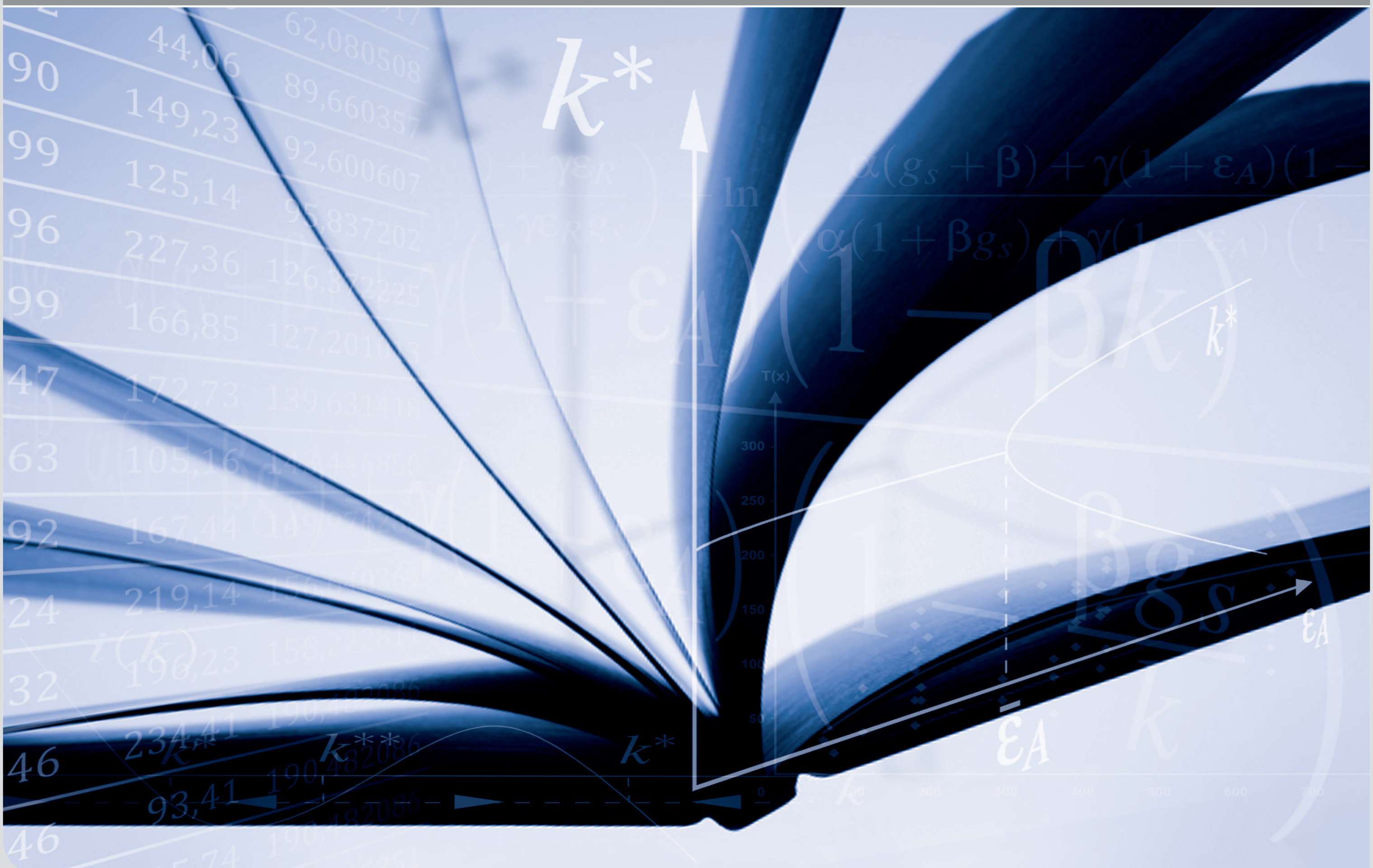
Karlsruher Institut für Technologie

Quantifying subjective uncertainty in survey expectations

by Fabian Krüger and Lora Pavlova

No. 139 | MARCH 2020

WORKING PAPER SERIES IN ECONOMICS



Impressum

Karlsruher Institut für Technologie (KIT)
Fakultät für Wirtschaftswissenschaften
Institut für Volkswirtschaftslehre (ECON)

Kaiserstraße 12
76131 Karlsruhe

KIT – Die Forschungsuniversität in der Helmholtz-Gemeinschaft

Working Paper Series in Economics
No. 139, March 2020

ISSN 2190-9806

econpapers.wiwi.kit.edu

Quantifying Subjective Uncertainty in Survey Expectations*

Fabian Krüger
Karlsruhe Institute of Technology

Lora Pavlova
Karlsruhe Institute of Technology

March 30, 2020

Abstract

Several recent surveys ask for a person's subjective probabilities that the inflation rate falls into various outcome ranges. We provide a new measure of the uncertainty implicit in such probabilities. The measure has several advantages over existing methods: It is robust, trivial to implement, requires no functional form assumptions, and is well-defined for all logically possible probabilities. These advantages are particularly relevant when analyzing large scale consumer surveys. We illustrate the new measure using data from the Survey of Consumer Expectations.

1 Introduction

Expectations uncertainty matters in economics. Uncertain firms tend to respond less to monetary or fiscal policy (Bloom, 2009). Monitoring inflation expectations and the associated uncertainty may help recognize early signs of eroding central bank credibility or de-anchoring of inflation expectations (Grishchenko et al., 2019); central banks are paying increasing attention to consumer and firm expectations for this purpose (ECB, 2019). Subjective uncertainty also features prominently in theoretical models of expectation formation, such as rational inattention (Sims, 2003; Mackowiak and Wiederholt, 2009).

There is hence much interest in measuring uncertainty, both at the level of the aggregate economy (e.g. Baker et al., 2016; Carriero et al., 2018) and at the level of individual persons or firms. In the present paper, we propose a new measure of individual-level uncertainty based on subjective survey probabilities. Such an uncertainty measure is an important input to studies that consider either the determinants or the consequences of subjective uncertainty. See, for example, Coibion et al. (2018) for an analysis of firms' expectations, Ben-David et al. (2019) for a household finance perspective, and Rich and Tracy (2010) for an analysis of macroeconomic expert forecasts.

*Financial support from the German Research Foundation (DFG) via grant KR 5214/1-1 is gratefully acknowledged. We thank seminar and conference participants at Heidelberg University, HKMEtrics, Joint Conference on Household Expectations (Bundesbank - Banque de France), Humboldt-Universität Berlin and IWH (Halle), as well as Konstantin Görgen, Axel Lindner, Malte Knüppel, Simas Kucinskas, Sebastian Rüth and Michael Weber for helpful comments.

Manski (2004, 2018) reviews a growing number of economic surveys in which participants assess the probability of a variable falling into various outcome ranges. In macroeconomics, the Survey of Professional Forecasters (SPF; Croushore, 1993) and its European counterpart (Garcia, 2003) are popular data sources covering expert forecasts. Furthermore, several recent surveys address the probabilistic expectations of consumers and firms. Examples include the Survey of Consumer Expectations (SCE) launched by the Federal Reserve Bank of New York (Armantier et al., 2017), a similar initiative by the Bank of Canada (Gosselin and Khan, 2015), and the firm survey by Coibion et al. (2018). These data on probabilistic expectations promise to shed new light on consumers' uncertainty, complementing more traditional survey information on point expectations. The latter do not contain direct information about uncertainty. However, Binder (2017) uses information on whether numeric point expectations are a multiple of five (indicating possible rounding and higher uncertainty) in order to construct a measure of individual uncertainty.

Figure 1 illstrates subjective probability distributions ('histograms') from the April 2019 wave of the SCE.[1] Each survey participant provides probabilities for various outcome ranges ('bins') of next year's inflation rate, as represented by the horizontal axis. The SCE contains a substantial share of responses that use one or two bins only. Such responses, which we call 'sparse histograms', are made by roughly a third of the SCE participants. Sparse histograms pose a challenge for existing measures of individual uncertainty (notably Engelberg et al., 2009, EMW) which are based on fitting a parametric distribution. For sparse histograms, fitting a flexible distribution is not possible, and a simple triangular shape is commonly used instead (see the two examples in the top row of Figure 1).

Motivated by the SCE data, we propose a new uncertainty measure that is transparent, trivial to implement, and is well-defined even for sparse histograms. By contrast, existing approaches require assumptions on the support of the subjective histogram, the distribution within each bin, or the functional form of the underlying continuous distribution. Our proposed measure can be theoretically motivated as the generalized entropy function of the ranked probability score (Epstein, 1969), a strictly proper scoring rule. We therefore refer to the new measure as ERPS, for Expected Ranked Probability Score.

The remainder of this paper is structured as follows. Section 2 summarizes some stylized facts of the SCE probabilities. Section 3 describes existing methods for quantifying uncertainty. Section 4 develops the ERPS, detailing its advantages as mentioned above. Sections 5 and 6 study the behavior of the ERPS for simulated and empirical data, respectively. Section 7 concludes. The Online Appendix contains details, proofs, and additional results.

[1]Source: *Survey of Consumer Expectations*©, 2013-2020 Federal Reserve Bank of New York (FRBNY). The *SCE* data are available without charge at `http://www.newyorkfed/microeconomics/sce` and may be used subject to license terms posted there. FRBNY disclaims any responsibility for this analysis and interpretation of *Survey of Consumer Expectations* data.

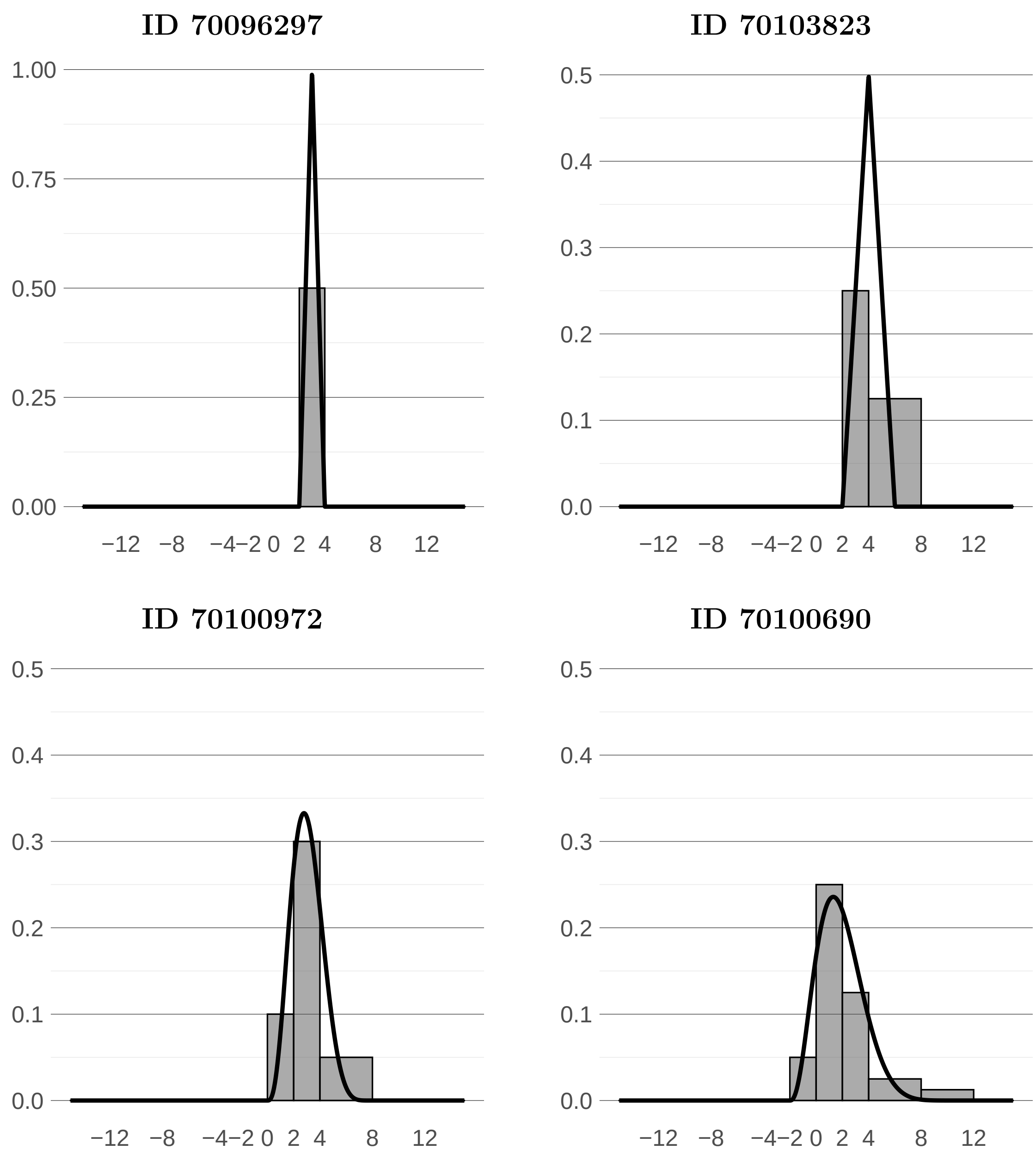


Figure 1: Illustration of probabilistic inflation expectations from the April 2019 wave of the SCE. The area of a rectangle corresponds to the subjective probability of the corresponding outcome range. For example, in the bottom left panel, the probability for an outcome between 4 and 8 equals $4 \times 1/20 = 1/5$. Solid lines indicate fitted probability density functions via the EMW method.

2 Subjective probabilities in the SCE data

The SCE is conducted at a monthly frequency with a sample size of about 1,200-1,300 respondents per month. The core module of the SCE asks, among others, for subjective probabilities of various outcome ranges, covering three variables: The Consumer Price Index (CPI) at two different horizons, real estate prices, and the respondent's personal earnings. In the SCE questionnaire made available by Federal Reserve Bank of New York (2020), the relevant question codes are Q9 and Q9c (CPI inflation rate), C1 (growth rate of the average home price nationwide) and Q24 (growth rate of the respondent's personal earnings). The relevant outcome ranges (in percent), which are the same for all variables, can be represented by the intervals

$$(-\infty, -12]; (-12, -8]; (-8, -4]; (-4, -2]; (-2, 0); [0, 2); [2, 4); [4, 8); [8, 12); [12, \infty).$$

These outcome ranges are reflected in the horizontal axis labels of Figure 1. In the case of inflation, for example, the two rightmost intervals refer to an inflation rate between 8% and 12% and to an inflation rate of 12% or more.[2]

Table 1 compares the SCE to expert forecasts in the SPF, in terms of participants' probability reponses. The table's upper panel presents summary statistics on the number of histogram bins used by SCE participants (that is, the number of bins containing strictly positive probability mass). We focus on the time period from January 2014 to March 2019 for comparability to the SPF (see below). For inflation and the average home price, around 30% of the participants uses one or two bins ('sparse histograms'). For personal earnings, roughly half of the participants use one or two bins. The mean number of bins used is somewhat higher for inflation and the average home price (4.2 – 4.4), compared to personal earnings (3.3). Finally, more than a quarter of the participants use one or both of the outer bins that correspond to the intervals $(-\infty, -12]$ and $[12, \infty)$.

The lower panel of Table 1 presents analogous statistics for the SPF. The SPF histograms are similar in design to those of the SCE, except that the two surveys use different numerical ranges for the histogram bins. While the SPF's bin definitions have been adapted over time (Federal Reserve Bank of Philadelphia, 2020), they are constant over the time period reported in Table 1. The number of bins (ten) is the same as in the SCE, except for GDP (eleven). While the share of participants using two bins and the mean number of bins used are comparable to the SCE, there are some major differences to the SCE: First, the SPF features a much smaller share of participants who use a single bin. For example, this share is about ten percentage points lower for the inflation variables. Second, the share of participants using at least one outer bin is much smaller in the SPF. For example, this share is more than 20 percentage points lower for the inflation variables.

Given its large sample size and the empirical patterns just reported, the SCE necessarily contains some histograms with non-standard shapes that are hard to capture

[2]The inclusion (or exclusion) of interval limits is not specified by the SCE survey questions. For example, the survey question leaves it unspecified whether an inflation rate of exactly 12% belongs to the last or penultimate bin. Our choice of half-open intervals (with the exception of the $(-2, 0)$ interval) are arbitrary – as is any choice in that regard – but seem unlikely to be of empirical relevance.

	n	Share of respondents using one bin	two bins	outer bin(s)	Mean nr. of bins
		SCE			
Average Home Price	71477	16.1	15.9	39.8	4.2
Inflation (one-year)	81483	12.4	17.1	39.3	4.4
Inflation (three-year)	81640	13.0	17.6	39.0	4.4
Personal Wage	54933	26.7	23.9	28.5	3.3
		SPF			
Inflation (GDP def.)	719	2.2	14.0	17.0	4.5
GDP	748	3.2	19.1	7.2	4.5
Inflation (CPI)	724	1.4	14.1	13.5	4.6
Inflation (PCE)	687	0.9	16.3	13.5	4.6
Unemployment	713	8.3	30.3	44.2	3.2

Table 1: Summary statistics on the number of bins used in the SCE (January 2014 to March 2019 waves) and SPF (2014:Q1 to 2019:Q1 waves); n denotes the total number of responses. We exclude histograms that do not sum to one (less than 0.3% of responses in both surveys).

by parametric distributions. Examples include positive probability on exactly two non-adjacent bins, or substantial probability mass in one or both outer bins. These features call for simple and robust methods that allow to quantify the uncertainty in any possible histogram.

3 Existing uncertainty measures

Survey probabilities as in Figure 1 do not specify a full probability distribution since the endpoints of the histogram's support as well as the distribution within each bin are unknown. Based on the raw probabilities alone, it is hence impossible to compute each participant's subjective mean or variance. Here we briefly review two methods that use parametric assumptions in order to account for missing information.

3.1 Distribution fitting

Following earlier work by Dominitz and Manski (1997), Engelberg et al. (2009, EMW) propose to fit a continuous distribution to the histogram probabilities. Their choice of continuous distribution depends on the number of histogram bins being used: EMW propose to fit a simple triangular distribution if the histogram is sparse, and to fit a flexible generalized Beta distribution if the forecaster uses three or more bins. In case the forecaster uses the leftmost bin (left limit of $-\infty$) or rightmost bin (right limit of $+\infty$), EMW propose to treat the limits of the distribution's support as a free parameter. We provide formal details on the EMW method in Online Appendix A. The method is used to derive uncertainty measures that are reported in official SCE publications such as Armantier et al. (2017), and are made available for download by Federal Reserve Bank of New York (2020).

The EMW method provides a full analytical distribution from which any feature of interest (such as subjective measures of location, spread, or tail risk) can be computed. However, this wealth of information comes at a cost: First, the choice of a particular parametric distribution is potentially restrictive, and seems hard to justify if the histogram is sparse. Second, the approach entails a discontinuity when moving from two bins (approximated via a triangular distribution) to three bins (approximated via a generalized beta distribution). Finally, practical implementation requires judgmental choices pertaining, e.g., to parameter limits imposed in numerical optimization, or to the handling of certain 'undefined' cases that are not covered by EMW's original proposal, because they did not occur in their SPF data. These choices may reasonably be made differently by different authors. Full reproducibility hence requires careful documentation of all implementation choices.

For the SPF data, the drawbacks of the EMW method arguably play a minor role since both the share of sparse histograms and the share of 'undefined' cases is small. This observation explains the widespread and successful use of the EMW method for the SPF and similar data sets. By contrast, given the properties of the SCE discussed above, the EMW method seems less well adapted to larger-scale consumer surveys.

3.2 Mass-at-midpoint method

The mass-at-midpoint (MAM) method (see Glas, 2019, and the references therein) assumes that the subjective distribution is discrete, with point mass at $\{m_k\}_{k:p_k>0}$, where m_k denotes the midpoint of bin $k = 1, \ldots, K$. Hence the method assumes point mass at the subset of bins that receive nonzero probability. Under this assumption, the subjective mean and standard deviation can easily be computed. An advantage of this method is that it can be applied irrespectively of the number of bins used. In particular, it avoids the discontinuity inherent in the EMW method. A disadvantage of the MAM method arises whenever the participants uses one of the outer bins (i.e., whenever $p_1 > 0$ or $p_{10} > 0$). In this case, the subjective mean and standard deviation depend on the endpoints of the outer bins, which are not specified by the survey design and for which assumptions seem hard to justify. This disadvantage is especially relevant for the SCE, where about one third of the participants uses at least one outer bin (see Appendix A for details). In the following, we therefore focus on the more widely used EMW method as a benchmark for our proposed method.

4 A new approach to quantifying uncertainty in survey histograms

4.1 General idea: Quantifying uncertainty via entropy

We treat each survey response as a vector of probabilities $\underline{p} := \big(p_1, \quad p_2, \quad \ldots, \quad p_K\big)'$, where p_k denotes the subjective probability that the inflation rate is within the interval r_k that defines the range of bin k. In practice, the intervals $\{r_k\}_{k=1}^K$ are disjoint and their union is the real line. Hence the probabilities $\underline{p}$ form a subjective survey histogram as in Figure 1.

Our proposed measure of uncertainty is based on the concept of entropy. Informally, if the entropy of distribution $\underline{p}$ is large, then a forecaster with subjective distribution $\underline{p}$ places a high probability on making large forecast errors. In that sense, $\underline{p}$ corresponds to high uncertainty. Vice versa, under a low-entropy distribution $\underline{p}$, large forecast errors are unlikely, and hence low entropy corresponds to low uncertainty.

More formally, entropy relates to strictly proper scoring rules (Gneiting and Raftery, 2007). In economics, scoring rules are commonly used for eliciting beliefs in experiments (Schotter and Trevino, 2014) and for evaluating probabilistic forecasts (e.g. Boero et al., 2011). In a discrete setup, scoring rules are functions of the form $\mathrm{S}(\underline{p}, k^*)$ that measure the performance of the probabilistic forecast $\underline{p}$ if the outcome k^* realizes. The integer $k^* \in \{1, 2, \ldots, K\}$ indicates the histogram bin that contains the realization. We consider specific choices of S below. For each of these choices, a smaller value of S indicates a better forecast. A scoring rule S is called strictly proper if a forecaster minimizes their expected score by stating what they think is the true probability distribution $\underline{p}$ (conditional on their information set); see Gneiting and Katzfuss (2014, Section 3.1.1) for a formal definition. The function

$$\mathrm{ES}(\underline{p}) \quad = \quad \sum_{k=1}^{K} p_k \ \mathrm{S}(\underline{p}, k)$$

is called the entropy function associated with the scoring rule S (e.g. Gneiting and Raftery, 2007, Section 2.2). We propose to use this function in order to measure the subjective uncertainty in a probabilistic survey forecast $\underline{p}$.

4.2 Expected Ranked Probability Score (ERPS)

As our preferred choice of scoring rule S, we consider the ranked probability score (RPS; Epstein, 1969):

$$\mathrm{RPS}(\underline{p}, k^*) \quad = \quad \begin{cases} \sum_{k=1}^{K}(1-P_k)^2 & \text{if } k^*=1 \\ \sum_{k=1}^{k^*-1}(P_k)^2 + \sum_{k=k^*}^{K}(1-P_k)^2 & \text{if } k^* \in \{2,3,\ldots,K\}, \end{cases}$$

where $P_k = \sum_{j=1}^{k} p_j$ is the cumulative probability of the first k bins. As its name suggests, the RPS is designed for ranked categorical variables. That is, the RPS treats the realizing bin $k^* \in \{1, \ldots, K\}$ as an ordinal variable, with $k^* = 1$ representing a smaller outcome than $k^* = 2$. Thus, the RPS rewards forecasters who put much probability mass into bins that are equal or close to the realizing bin k^*. For example, if a forecaster places unit probability mass on the first bin, then $k^* = 2$ yields a lower (i.e., better) RPS than $k^* = 3$. Boero et al. (2011) pervasively argue that this feature of the RPS is well in line with survey histograms, and propose to use it for evaluating the histograms' predictive accuracy.

The entropy function of the RPS is given by

$$\begin{aligned}
\mathrm{ERPS}(\underline{p}) &= \sum_{k=1}^{K} p_k \ \mathrm{RPS}(\underline{p}, k) \\
&= \sum_{k=1}^{K} P_k(1 - P_k). \qquad (1)
\end{aligned}$$

The latter equation, which is our proposed uncertainty measure, is trivial to compute from the histogram probabilities.

Since it attaches only an ordinal but not a numerical interpretation to the bins, the ERPS at (1) does not depend on the bins' outcome ranges or the (unknown) distribution of probability mass within each bin. The ordinal interpretation hence renders parametric assumptions obsolete, and explains the simplicity and robustness of the ERPS. A drawback of the ordinal interpretation is that the ERPS is not comparable across different bin definitions, such as design A involving ten bins of length one and design B involving five bins of length two. This concern may be relevant if the bin definitions must be adapted over time in order to account for changes in the distribution of the predictand. Such redefinitions occurred for the SPF which started in 1968 (see Federal Reserve Bank of Philadelphia, 2020). The concern seems largely irrelevant for the SCE, where the bins cover a very wide range of outcome values (see Section 2). Hence redefinitions have not occurred until now, and seem unlikely in the future.

4.3 Comparison to other entropy-based measures

Here we relate the ERPS to entropy functions for two other popular scoring rules. The logarithmic score (LS; Good, 1952) and Brier score (BS; Brier, 1950) are given by

$$\begin{aligned}
\mathrm{LS}(\underline{p}, k^*) &= -\log p_{k^*} \\
\mathrm{BS}(\underline{p}, k^*) &= \sum_{k=1}^{K} (\mathbb{I}_{k=k^*} - p_k)^2,
\end{aligned}$$

where $\mathbb{I}_{k=k^*}$ is an indicator function that equals one if $k = k^*$, and equals zero otherwise. Their respective entropy functions are given by

$$\begin{aligned}
\mathrm{ELS}(\underline{p}) &= -\sum_{k=1}^{K} p_k \log p_k. \\
\mathrm{EBS}(\underline{p}) &= \sum_{k=1}^{K} p_k(1 - p_k).
\end{aligned}$$

The ELS was famously developed by Shannon (1948) and is typically called 'Shannon Entropy'. In economics, it plays a key role in the theory of rational inattention (Sims, 2003). The EBS is much less widely used, with the interesting exception of López-Menéndez and Pérez-Suárez (2019) who quantify uncertainty in (aggregate) tendency surveys.

The BS are LS are both designed for multinomial random variables, that is, the outcome categories $k^* \in \{1, \dots, K\}$ are viewed as interchangeable. Hence the EBS and ELS are invariant to permutations of the histogram probabilities $p_1, \dots, p_K$. For example, for a hypothetical three-bin histogram, the probabilities $\underline{p}_a = (1/4, 1/2, 1/4)'$ yield the same EBS as the probabilities $\underline{p}_b = (1/2, 1/4, 1/4)'$. This assessment seems implausible, given that $\underline{p}_b$ is obtained from $\underline{p}_a$ by shifting probability mass from the central bin to the more extreme leftmost bin. Under the ERPS, which utilizes an ordinal interpretation, $\underline{p}_b$ is considered more uncertain than $\underline{p}_a$.

ELS and EBS are both maximized by the vector

$$\underline{p}^{**} = \tau \times (1/K),$$

where τ is a $K \times 1$ vector of ones (see Shannon 1948 and López-Menéndez and Pérez-Suárez 2019). Hence flat probabilities represent maximal uncertainty, as seems natural in a multinomial setup. By contrast, we show in Online Appendix B that the maximal ERPS is attained for the vector

$$\underline{p}^* = \begin{pmatrix}1/2, & 0, & \dots, & 0, & 1/2\end{pmatrix}'$$

that places probability one half on each of the two outer bins. The intuition for this solution is that under $\underline{p}^*$, it is certain that one of the two outer bins will materialize. Both outcomes produce a large score $\text{RPS}(\underline{p}^*, k)$, since $\underline{p}^*$ places no probability mass on the neighboring bins.

5 Simulation study on sparse histograms

As we have argued, a key advantage of the ERPS over the EMW method is that the former requires no case distinction when moving from a sparse histogram (using two bins) to a histogram using three bins. We demonstrate the quantitative relevance of this point in a simulation study based on the June 2013 to April 2019 waves of the SCE. We focus on participants who use two adjacent bins, none of which is an outer bin in the SCE's histogram design shown in Section 2. We further require that the histogram probabilities sum to one and exceed one percent, which is the magnitude of the perturbation we consider. These selection criteria leave us with 13518 two-bin histograms. For each of these histograms, we consider two simple perturbations: First, we move one percentage point of probability mass from the left bin to its left neighboring bin. For example, suppose that the original histogram allocates 50% probability to the two bins $[0, 2)$ and $[2, 4)$. The perturbed histogram then places probability $1\%, 49\%$ and 50% to the three bins $[-2, 0)$, $[0, 2)$ and $[2, 4)$ respectively. Second, we apply an analogous perturbation to the right histogram bin, such that the perturbed histogram contains one percent of probability mass in a third bin located to the right of the original histogram. We choose a perturbation size of one percentage point since it is the smallest size that seems empirically plausible.

For each setup (no perturbation, left perturbation, and right perturbation), we consider the ERPS as well as the standard deviation (EMW-SD) and interquartile range (EMW-IQR) of the distribution produced via the EMW method. Given the small perturbation size, we contend that an uncertainty measure should be robust across the three setups. As a scale-free measure of robustness, we consider the rank correlation of uncertainty across the three setups. For EMW-SD, the rank correlation between the baseline setup and two perturbed versions is at 0.704 and 0.465 for left and right perturbation, respectively. The results for EMW-IQR are similar, with rank correlations of 0.711 and 0.632. Given that these numbers are considerably smaller than one, both EMW-SD and EMW-IQR are sensitive to small shifts of probability mass.

The reported correlations further indicate that the impact of right perturbation is larger than the impact of left perturbation. This effect is due to the empirical pattern that many of the two-bin histograms focus on the bins $[2, 4)$ and $[4, 8)$. According to the SCE's bin design shown in Section 2, the left neighbor of these bins is at $[0, 2)$, whereas the right neighbor is at $[4, 8)$. Hence left perturbation expands the support of the histogram by two units, whereas right perturbation expands the support by four units. This asymmetry matters here since the Engelberg et al. (2009) algorithm adopts the support of the histogram if only interior bins are used.

For ERPS, the impact of the perturbation can be described analytically. Let $\underline{p}$ denote a two-bin histogram, and $\tilde{\underline{p}}_L$ and $\tilde{\underline{p}}_R$ its perturbed version with probability mass shifted to the left and right neighboring bin. Let δ denote the size of the perturbation (with $\delta = 0.01$ in our simulation study). Then Equation (1) yields that

$$\text{ERPS}(\tilde{\underline{p}}_L) = \text{ERPS}(\tilde{\underline{p}}_R) = \text{ERPS}(\underline{p}) + \delta\ (1 - \delta),$$

i.e. both perturbations lead to an additive increase in ERPS by $\delta\ (1 - \delta)$. Hence perturbation affects all histograms in exactly the same way, leading to rank correlations of one which demonstrate the robustness of the ERPS.

6 Empirical comparisons

We next provide a brief empirical comparison between the ERPS and the EMW method. We focus on the EMW-SD variant; the results based on EMW-IQR are qualitatively identical and are hence omitted for brevity.

In order to summarize the similarity of the ERPS and the EMW method for the SCE data, we compute the rank correlation of ERPS and EMW-SD, separately for each survey wave and for each of the four probability variables (inflation at two horizons, house prices, and personal earnings). The resulting correlation coefficients range from 0.871 to 0.951, indicating fairly close overall correspondence between the two uncertainty measures. Hence, the ERPS captures uncertainty in a similar way as the EMW method.

We next analyze whether respondents who express high uncertainty about inflation also express high uncertainty about house prices and their personal earnings. To this end, Figure 2 plots the rank correlation coefficient of uncertainty across variables, based on either ERPS or EMW-SD. At each date, we consider the average rank correlation over six pairs of probability variables. As shown by the figure, the ERPS produces higher correlations in each sample period. The same pattern holds for each of the six pairwise coefficients (see Figure 1 in the Online Appendix). There is hence clear

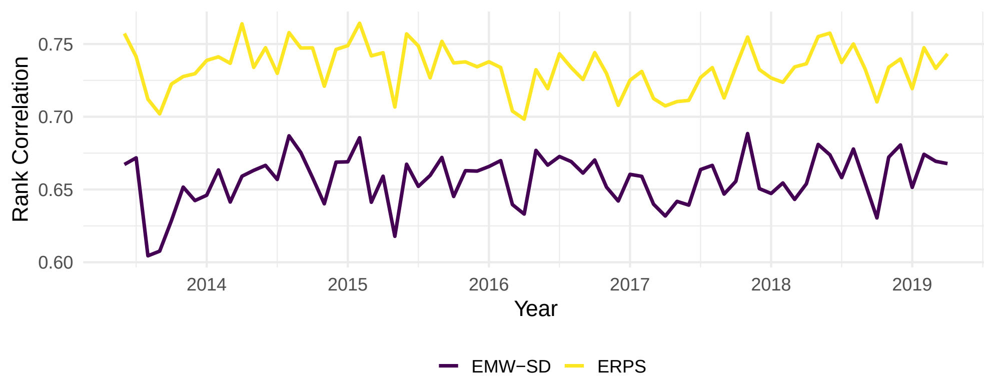


Figure 2: Average pairwise rank correlation of uncertainty across four variables (inflation at two different horizons, house prices, personal earnings).

evidence that the ERPS is more consistent across variables than the EMW method. In the absence of a 'ground truth' measure of uncertainty, we cannot tell whether this feature of the ERPS is desirable. However, these findings do point to an interesting and robust difference between both measures.

We further compare the persistence of uncertainty as measured by EMW versus ERPS. We measure persistence by the rank correlation of uncertainty in two subsequent SCE waves, for the subset of participants who are present in both waves. In principle, small rank correlation may indicate a genuine shift in relative uncertainty from one month to the next (e.g., Anne is more uncertain than Bob in January, whereas Bob is more uncertain than Anne in February). Alternatively, small rank correlation may simply reflect noise in the uncertainty measure.

Figure 3 presents results on the persistence of uncertainty, as measured by ERPS versus EMW-SD. For personal earnings, the persistence of ERPS and EMW-SD is similar. Genuine shifts in relative uncertainty seem particularly plausible for this variable since it is individual-specific and hence prone to idiosyncratic information updates (such as Anne signing a new labor contract in February). For house prices and inflation, the rank correlation for ERPS (yellow line) exceeds the rank correlation for EMW-SD (purple line) in the clear majority of time periods (100 % of time periods for house prices, 91.4% and 95.7% for inflation). Similar to the findings across variables, and reflecting a broader theme in this paper, this indicates that the ERPS is less sensitive to small changes in the raw probabilities $\underline{p}$.

7 Discussion

This paper introduces the ERPS, a new measure of uncertainty in probabilistic survey expectations. The ERPS is based on an ordinal interpretation of the survey outcome categories which obviates any parametric assumptions and explains its simplicity and robustness. The Engelberg et al. (2009, EMW) method, which is the current standard for quantifying uncertainty in economic surveys, instead uses a numerical interpreta-

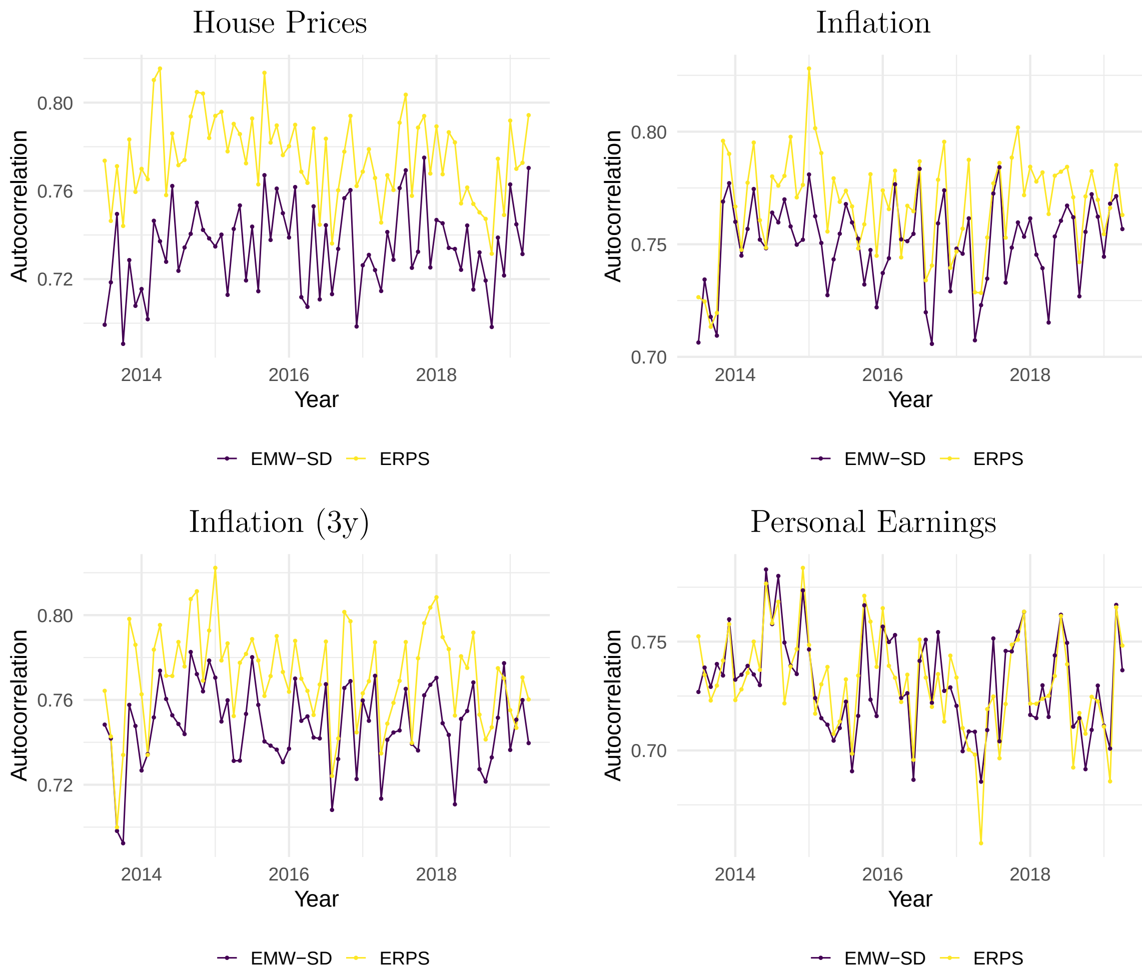


Figure 3: Rank correlation of subjective uncertainty over two subsequent survey months, based on participants who are present in both months.

tion of outcome categories. The numerical interpretation is informationally more demanding and requires the researcher to make parametric assumptions about unknown aspects of the histogram. In return, it provides a full picture of subjective uncertainty.

We think that a user's choice between the ERPS and the EMW method should depend on the signal-to-noise ratio in the subjective probability data. If this ratio is high, then the EMW method – which is more sensitive to small changes in the probabilities – seems more appropriate. Examples of this situation include average histograms across time or across socio-demographic groups (which may be based on hundreds of individual responses), and perhaps probability assessments by individual expert forecasters. By contrast, the ERPS seems preferable in the context of individual-level probabilities by consumers, such as the ones covered by the SCE. This type of data is an innovative source for monitoring the general public's inflation expectations, and similar surveys have recently been started by the Bank of Canada and the Bundesbank. In Online Appendix D, we illustrate how the ERPS can be related to the socio-demographic information available in the SCE.

While we have focused on measuring subjective uncertainty by itself, an interesting related question is whether subjective uncertainty lines up with measures of realized uncertainty based on expectation errors. This comparison is of economic relevance since over– or underestimating objective uncertainty has possibly severe implications for decision making (see e.g. Ben-David et al., 2013). In Online Appendix E, we demonstrate that the ERPS can usefully be applied in this context as well.

References

Armantier, O., G. Topa, W. van der Klaauw, and B. Zafar (2017): "An overview of the Survey of Consumer Expectations," *Federal Reserve Bank of New York Economic Policy Review*, 23, 51–72.

Baker, S. R., N. Bloom, and S. J. Davis (2016): "Measuring economic policy uncertainty," *The Quarterly Journal of Economics*, 131, 1593–1636.

Ben-David, I., E. Fermand, C. M. Kuhnen, and G. Li (2019): "Expectations uncertainty and household economic behavior," Working Paper, available at `http://public.kenan-flagler.unc.edu/faculty/kuhnenc/RESEARCH/bfkl_v7.pdf` (last accessed: August 8, 2019).

Ben-David, I., J. R. Graham, and C. R. Harvey (2013): "Managerial miscalibration," *The Quarterly Journal of Economics*, 128, 1547–1584.

Binder, C. C. (2017): "Measuring uncertainty based on rounding: New method and application to inflation expectations," *Journal of Monetary Economics*, 90, 1–12.

Bloom, N. (2009): "The impact of uncertainty shocks," *Econometrica*, 77, 623–685.

Boero, G., J. Smith, and K. F. Wallis (2011): "Scoring rules and survey density forecasts," *International Journal of Forecasting*, 27, 379–393.

Brier, G. W. (1950): "Verification of forecasts expressed in terms of probability," *Monthly Weather Review*, 78, 1–3.

Carriero, A., T. E. Clark, and M. Marcellino (2018): "Measuring uncertainty and its impact on the economy," *The Review of Economics and Statistics*, 100, 799–815.

Coibion, O., Y. Gorodnichenko, and S. Kumar (2018): "How do firms form their expectations? New survey evidence," *American Economic Review*, 108, 2671–2713.

Croushore, D. D. (1993): "Introducing: The Survey of Professional Forecasters," *Federal Reserve Bank of Philadelphia Business Review*, 6, 3–15.

Dominitz, J. and C. F. Manski (1997): "Using expectations data to study subjective income expectations," *Journal of the American Statistical Association*, 92, 855–867.

ECB (2019): "Inflation expectations and the conduct of monetary policy," Speech by Benoît Cœuré, Member of the Executive Board of the ECB, at an event organised by the SAFE Policy Center, Frankfurt am Main, 11 July 2019. Available at `https://www.ecb.europa.eu/press/key/date/2019/html/ecb.sp190711~6dcaf97c01.en.html` (last accessed: August 8, 2019).

Engelberg, J., C. F. Manski, and J. Williams (2009): "Comparing the point predictions and subjective probability distributions of professional forecasters," *Journal of Business & Economic Statistics*, 27, 30–41.

Epstein, E. S. (1969): "A scoring system for probability forecasts of ranked categories," *Journal of Applied Meteorology*, 8, 985–987.

Federal Reserve Bank of New York (2020): "Survey of Consumer Expectations," `https://www.newyorkfed.org/microeconomics/sce` (last accessed: March 9, 2020).

Federal Reserve Bank of Philadelphia (2020): "Survey of Professional Forecasters," `https://www.philadelphiafed.org/research-and-data/real-time-center/survey-of-professional-forecasters/` (last accessed: March 9, 2020).

Garcia, J. A. (2003): "An introduction to the ECB's Survey of Professional Forecasters," ECB Occasional Paper no. 8.

Glas, A. (2019): "Five dimensions of the uncertainty–disagreement linkage," *International Journal of Forecasting*, forthcoming.

Gneiting, T. and M. Katzfuss (2014): "Probabilistic forecasting," *Annual Review of Statistics and Its Application*, 1, 125–151.

Gneiting, T. and A. E. Raftery (2007): "Strictly proper scoring rules, prediction, and estimation," *Journal of the American Statistical Association*, 102, 359–378.

Good, I. (1952): "Rational decisions," *Journal of the Royal Statistical Society. Series B (Methodological)*, 14, 107–114.

Gosselin, M.-A. and M. Khan (2015): "A Survey of Consumer Expectations for Canada," *Bank of Canada Review*, 2015, 14–23.

Grishchenko, O., S. Mouabbi, and J.-p. Renne (2019): "Measuring inflation anchoring and uncertainty: A US and Euro area comparison," *Journal of Money, Credit and Banking*, 51, 1053–1096.

López-Menéndez, A. J. and R. Pérez-Suárez (2019): "Acknowledging uncertainty in economic forecasting. Some insight from confidence and industrial trend surveys," *Entropy*, 21, 413.

Mackowiak, B. and M. Wiederholt (2009): "Optimal sticky prices under rational inattention," *American Economic Review*, 99, 769–803.

Manski, C. F. (2004): "Measuring expectations," *Econometrica*, 72, 1329–1376.

——— (2018): "Survey measurement of probabilistic macroeconomic expectations: progress and promise," *NBER Macroeconomics Annual*, 32, 411–471.

Rich, R. and J. Tracy (2010): "The relationships among expected inflation, disagreement, and uncertainty: evidence from matched point and density forecasts," *Review of Economics and Statistics*, 92, 200–207.

Schotter, A. and I. Trevino (2014): "Belief elicitation in the laboratory," *Annual Review of Economics*, 6, 103–128.

Shannon, C. E. (1948): "A mathematical theory of communication," *Bell System Technical Journal*, 27, 379–423.

Sims, C. A. (2003): "Implications of rational inattention," *Journal of Monetary Economics*, 50, 665–690.

Quantifying Subjective Uncertainty in Survey Expectations: Online Appendix*

Fabian Krüger
Karlsruhe Institute of Technology

Lora Pavlova
Karlsruhe Institute of Technology

March 30, 2020

This online appendix is structured as follows: Section A formally describes the Engelberg et al. (2009, EMW) method. Section B proves a claim made in Section 4.3 of the paper. Sections C and D provide additional empirical results. Section E describes a framework for comparing subjective and objective uncertainty based on the ERPS.

A Details on the EMW method

Here we provide details related to the informal discussion in Section 3.1 of the paper.

Case A: Forecaster uses one or two bins

Following Engelberg et al. (2009, EMW), we construct isosceles triangles that are completely characterized by their support which we denote by $[a, b]$. The mode of the distribution is located at $c = (a + b)/2$.

In case a forecaster uses only one bin, we use a triangular distribution with support equal to the support of the bin used. This approach, which is recommended in EMW's Section 4.1.1, differs from the SCE, which assumes a uniform distribution over the support of the bin (Armantier et al., 2017, Footnote 28).

To discuss the two-bin case, suppose that a forecaster uses two adjacent bins, $[l, m)$ and $[m, r)$, with $l < m < r$, probability mass α in the left bin $[l, m)$, and probability mass $(1-\alpha)$ in the right bin $[m, r)$.

If $\alpha < 1/2$, we set $b = r$ and

$$a = m - \frac{(r-m)\ (\alpha + \sqrt{2\alpha})}{2-\alpha}.$$

Hence the triangular distribution satisfies $\mathcal{T}([a, m)) = \alpha$ and $\mathcal{T}([m, r)) = 1 - \alpha$, where the notation $\mathcal{T}(I)$ indicates the probability mass assigned to the interval I. In that sense,

*Financial support from the German Research Foundation (DFG) via grant KR 5214/1-1 is gratefully acknowledged.

the triangular distribution matches the bin $[m, r)$ containing more probability mass. As $\alpha \to 1/2, a \to m - (r - m)$, such that the fitted triangle is symmetric around m, and the triangle's base length is $2\ (r - m)$.

If $\alpha \geq 1/2$, we set $a = l$ and

$$b = m + \frac{(m - l)\ \left(1 - \alpha + \sqrt{2\ (1 - \alpha)}\right)}{1 + \alpha}.$$

Hence it holds that $\mathcal{T}([l, m)) = \alpha$ and $\mathcal{T}([m, b)) = 1 - \alpha$, i.e. the triangular distribution matches the bin $[l, m)$ containing more probability mass. For $\alpha = 1/2$, the fitted triangle is symmetric around m, with base length equal to $2\ (m - l)$.

There are two scenarios that are not covered by the preceding description:
- The forecaster uses two non-adjacent bins such as $[0, 2)$ and $[4, 8)$.
- The forecaster uses one or two bins, including one of the outer bins (i.e., $p_1 > 0$ or $p_K > 0$).

The EMW method does not prescribe a solution for the former scenario. In the latter scenario, any solution would seem to depend on an arbitrary choice of support limit. In our simulation analysis (Section 5 of the paper), we hence drop observations from either of the two scenarios in order not to distort our findings on the EMW method.

Case B: Forecaster uses three or more bins

If the forecaster uses three or more bins, EMW propose to fit a generalized Beta distribution given by

$$\begin{aligned}
F_{\text{gBeta}}(x; a, b, l, r) &= \begin{cases} 0 & x \leq l, \\ \frac{1}{B(a,b)} \int_l^x \frac{(u-l)^{a-1}(r-u)^{b-1}}{(r-l)^{a+b-1}} du & l < x \leq r, \\ 1 & x > r, \end{cases} \qquad (1) \\
B(a, b) &= \frac{\Gamma(a)\Gamma(b)}{\Gamma(a + b)}, \\
\Gamma(a) &= \int_0^\infty u^{a-1} \exp(-u)\ du.
\end{aligned}$$

Instead of the limits 0 and 1 of the regular Beta distribution, F_{gBeta} entails flexible left and right limits $l, r \in \mathbb{R}$ with $l < r$. The two shape parameters $a, b \in \mathbb{R}_+$ play the same role as in regular Beta distributions. EMW impose the constraint that $a > 1$ and $b > 1$ in order to obtain a unimodal shape, which seems plausible in the present context.

In order to fit the distribution at (1) to a vector of histogram probabilities $\underline{p}$, EMW propose to fix the limits l and r at the endpoints of the bins that are being used. If one or both of the two outer bins are being used, the authors propose to treat the limits l and/or r as free parameters to be estimated. That is, l is a free parameter if $p_1 > 0$, and r is a free parameter if $p_K > 0$, where $K = 10$ in the case of the SCE. Following Armantier et al. (2017, Appendix C), we impose the constraint that $l > -38$ and that $r < 38$ when estimating l and/or r. We further impose that $l < -12$ and $r > 12$, as is logically required by the SCE's bin design. The shape parameters a and b are estimated in either case. In the most general case where l and r are both estimated, the fitting problem is thus given by

$$\max_{\substack{a>1, b>1, \\ -38<l<-12, \\ 12<r<38}} \sum_{k=1}^{K} \left[F_{\text{gBeta}}(x_k; a, b, l, r) - P_k\right]^2,$$

where x_k is the right endpoint of the kth histogram bin, and $P_k = \sum_{j=1}^{k} p_j$ is the cumulative probability of the first k bins.

B Maximal ERPS

Here we prove a claim made in Section 4.3 of the paper.

The ERPS of a distribution $\underline{p}$ is given by

$$\text{ERPS}(\underline{p}) = \sum_{k=1}^{K} P_k(1 - P_k)$$

In matrix notation, let $\underline{p}$ be the $K \times 1$ vector with probabilities p_k, and $\underline{P}$ be the corresponding vector of cumulative probabilities P_k. We have that $\underline{P} = C'\underline{p}$, where C is a $K \times K$ upper triangular matrix with all elements above the main diagonal equal to one, and all diagonal elements equal to one. We can write

$$\text{ERPS}(\underline{p}) = \underline{P}'(\tau - \underline{P}) = \underline{p}'C\tau - \underline{p}'CC'\underline{p},$$

where τ is a $K \times 1$ vector of ones. To find the maximand of the ERPS, we solve the following problem:

$$\arg\max_{\underline{p}} \text{ERPS}(\underline{p}) \text{ such that } \underline{p}'\tau = 1;$$

note that the constraint that probabilities be nonnegative need not be enforced explicitly. Setting up the Langrangian and solving the resulting quadratic problem then shows that the maximand is given by

$$\underline{p}^* = \begin{pmatrix} 1/2, & 0, & \dots, & 0, & 1/2 \end{pmatrix}';$$

note that the second-order condition for a maximum is satisfied since CC' is strictly positive definite.

C Details on empirical comparisons

Here we provide additional details on Section 6 of the paper. In particular, Figure 1 presents the rank correlation coefficient of uncertainty for six pairs of variables, separately for ERPS

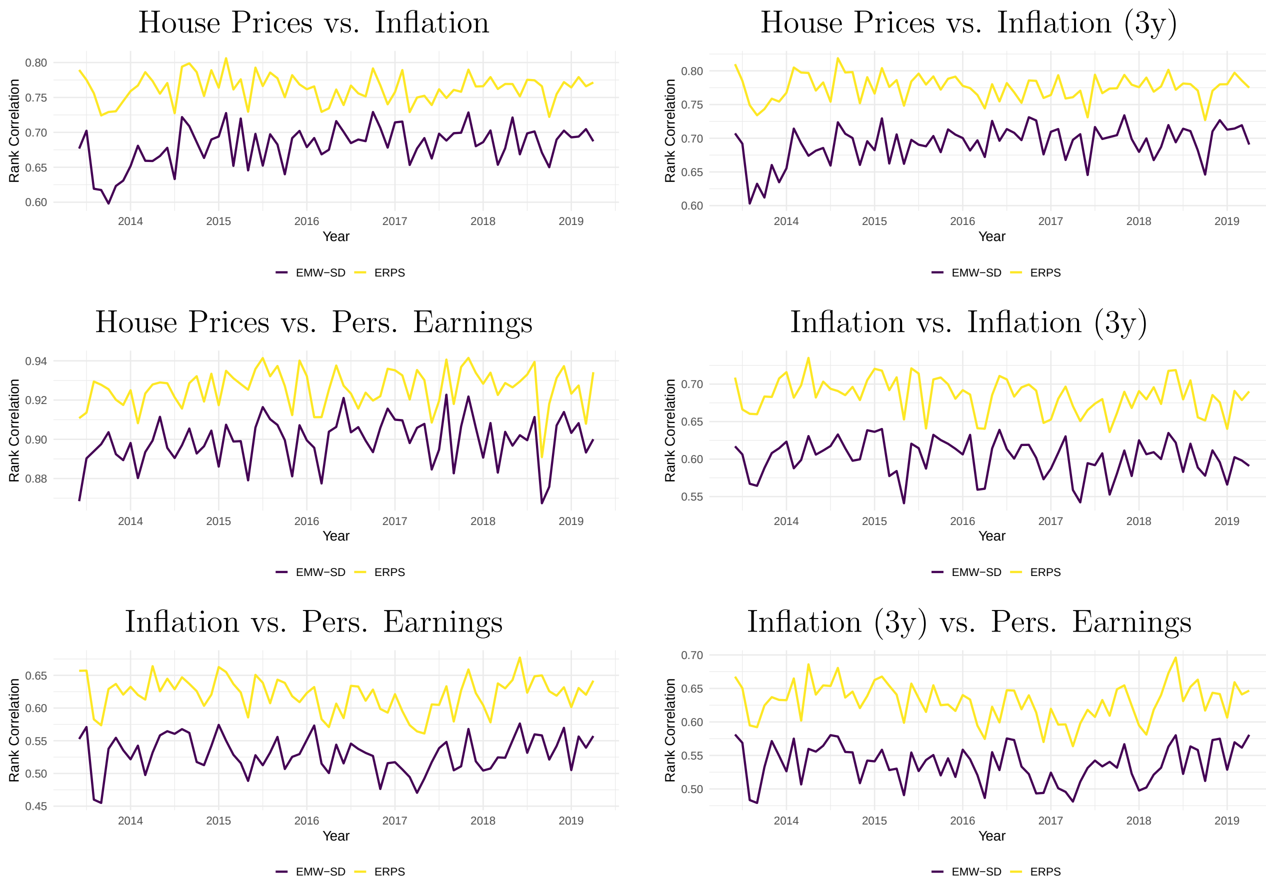


Figure 1: Rank correlation of subjective uncertainty for six pairs of variables.

versus EMW-SD. The figure shows that the correlation is higher for ERPS than for EMW-SD, for each of the variable pairs and in each time period.

D Further empirical illustration

This section presents an empirical illustration of the ERPS as a measure of individual uncertainty, in the spirit of Ben-David et al. (2019) who analyze uncertainty using the EMW method. We consider respondents' probabilistic forecasts about the inflation rate, change in the average home price, and change in personal earnings, for the June 2013 to April 2019 waves of the SCE. All of these expectations are one year ahead and are taken from the core module of the survey. Figure 2 plots the average ERPS across all respondents within each wave. For comparability, it is based on respondents who provide expectations for all variables. This implies in particular that it covers only participants who are employed since earnings uncertainty is only available for these participants. Interestingly, respondents are consistently more uncertain about the future development of the two macroeconomic outcomes, as compared to uncertainty about personal earnings. The figure further shows that average uncertainty about the three variables is fairly constant over time. This result can perhaps be rationalized by the relatively long one-year horizon to be forecasted (such that conditioning information may be of limited subjective relevance), and by the short sample period of the SCE we cover.

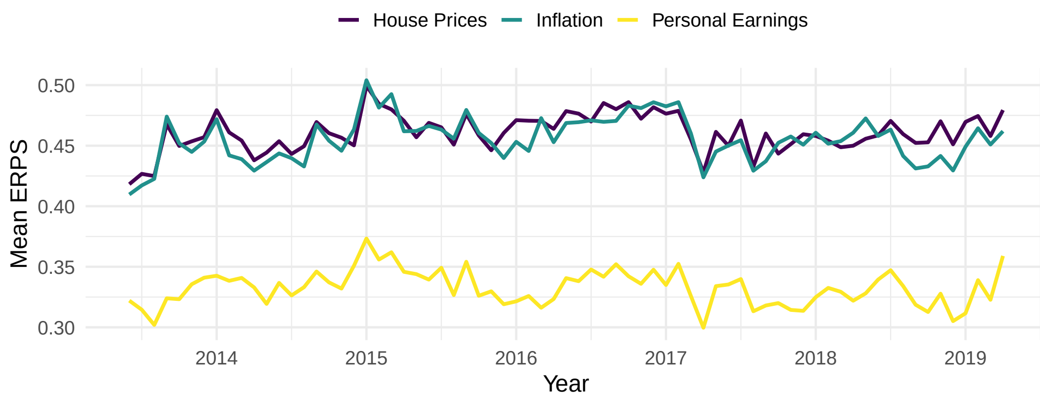


Figure 2: Mean ERPS for each month and variable.

We next consider heterogeneity in subjective uncertainty. Following Federal Reserve Bank of New York (2020), we distinguish respondents according to a number of demographic and socioeconomic characteristics: Age, educational attainment, household income, as well as financial literacy and numeracy skills. (The official SCE website by Federal Reserve Bank of New York 2020 publishes graphical summaries of uncertainty for these demographic groups. Regarding their age, survey participants are classified into three groups: 'Under 40', '40 to 60' and 'Over 60' years. In terms of educational attainment, the SCE allows to distinguish between respondents with no college education, some college education and a fully accomplished college degree. Household income is reported in three categories: 'Under 50k', '50k to 100k' and 'Over 100k'. Finally, a measure of the respondents' numeracy and financial literacy is introduced, such that one can distinguish between respondents with high and low literacy. Following a widely used approach, five questions in the survey aim to evaluate respondents' knowledge of concepts used in financial decision making such as interest compounding, understanding of inflation and risk diversification. Respondents who give a correct answer to four out of the five questions are categorized as having high numeracy and financial literacy skills.) Figure 3 plots patterns in uncertainty by demographics, focusing on inflation uncertainty for brevity. The figure indicates that younger, poorer, less educated and less financially literate survey participants experience higher inflation uncertainty. This is consistent with previous findings in the inflation expectations literature, which suggest that the respective groups have shorter financial planning horizons, are more exposed to transitory price shocks (Bruine de Bruin et al., 2010), and thus experience higher uncetrainty. The differences across demographic and socioeconomic groups just reported are broadly similar for inflation at a three-year horizon as well as for house prices and personal earnings at a one-year horizon (not reported for brevity). Furthermore, our findings in Figures 2 and 3 are qualitatively very similar to the ones reported by Federal Reserve Bank of New York (2020), which are based on the EMW method.

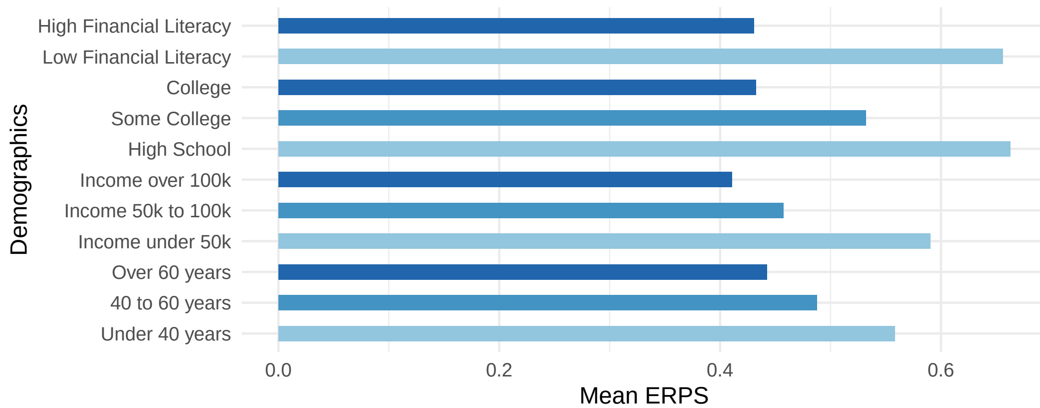


Figure 3: Mean ERPS across sociodemographic groups (inflation, one year ahead).

E Comparing subjective and objective uncertainty

It is often relevant to ask whether a person's subjective uncertainty is in line with an objective measure of uncertainty. In particular, miscalibrated probabilistic expectations (with subjective uncertainty exceeding objective uncertainty or vice versa) may lead to suboptimal decisions in a wide range of situations (see e.g. Ben-David et al., 2013, and the references therein). In the macroeconomic literature, subjective and objective uncertainty are often called 'ex ante uncertainty' and 'ex post uncertainty' (see e.g. Clements, 2014). This terminology reflects the fact that subjective uncertainty is typically based on forecasts, whereas objective uncertainty is based on subsequent realizations.

Following recent proposals by Clements (2014) and Galvao and Mitchell (2019), comparing a forecaster's ERPS to their RPS (on average across several time periods) yields a simple and theoretically appealing comparison of ex ante and ex post uncertainty. (As detailed in their Section 3.4, Galvao and Mitchell's notion of uncertainty corresponds to the difference between ex ante and ex post uncertainty. By contrast, we follow Jurado et al. 2015 and others in measuring uncertainty via ex ante uncertainty.) We next provide a formal treatment tailored to our setup. To this end, we consider a so-called prediction space setup (Gneiting and Ranjan, 2013) that models the joint distribution of expectations and realizations. We treat the K histogram probabilities $\underline{\mathbf{p}}$ as a random vector, and denote the bin containing the realization by the discrete random variable $\mathbf{k}^* \in \{1, \ldots, K\}$. The sample space of interest, Ω, consists of forecast-observation pairs $(\underline{\mathbf{p}}, \mathbf{k}^*)$. We omit time indexes for simplicity; to obtain an intuition, subsequent realizations of $(\underline{\mathbf{p}}, \mathbf{k}^*)$ can be thought of as independent (whereas one would expect contemporaneous dependence between $\underline{\mathbf{p}}$ and $\mathbf{k}^*$, of course). (It can be shown that the methodology of comparing ERPS to RPS remains valid under serial dependence in the forecast-observation tuples, as long as their joint process is strictly stationary. See Strähl and Ziegel 2017 for a technical treatment of a prediction space under serial dependence.) As in Ehm et al. (2016, Section 3.1), let $\mathbb{Q}$ be a probability measure on $(\mathcal{A}, \Omega)$, where $\mathcal{A}$ is a σ-field on Ω. The following result then provides a formal condition under which ex ante

uncertainty and ex post uncertainty coincide in expectation.
Assumption 1. *Assume that there is some information set $\mathcal{F} \subseteq \mathcal{A}$ such that*

$$\mathbb{Q}(\mathbf{k}^* = k|\mathcal{F}) = \mathbf{p}_k$$

holds almost surely for $k = 1, \ldots, K$, where $\mathbb{Q}(\mathbf{k}^* = k|\mathcal{F})$ is the true conditional probability that $\mathbf{k}^* = k$ (conditional on the information set $\mathcal{F}$), and $\mathbf{p}_k$ is the kth element of $\underline{\mathbf{p}}$.

Proposition 1. *Under Assumption 1, it holds that $\mathbb{E}(RPS(\underline{\mathbf{p}}, k^*)) = \mathbb{E}(ERPS(\underline{\mathbf{p}}))$.*

Proof. We have that

$$\begin{aligned}
\mathbb{E}(\mathrm{RPS}(\underline{\mathbf{p}}, \mathbf{k}^*)) &= \mathbb{E}(\mathbb{E}(\mathrm{RPS}(\underline{\mathbf{p}}, \mathbf{k}^*)|\mathcal{F})) \\
&= \mathbb{E}(\sum_{k=1}^{K} \mathbf{p}_k \ \mathrm{RPS}(\underline{\mathbf{p}}, k)) \\
&= \mathbb{E}(\mathrm{ERPS}(\underline{\mathbf{p}})),
\end{aligned}$$

where the first equality follows from the law of iterated expectations, the second equality follows from Assumption 1, and the final equality follows from the definition of ERPS. □

Assumption 1 requires that the probability forecast $\underline{\mathbf{p}}$ is correctly specified, in the sense that there is *some* information set relative to which the forecast is optimal. Under this assumption, Proposition 1 states that the RPS and ERPS of $\underline{\mathbf{p}}$ coincide in expectation. As a simple example (loosely following Gneiting et al., 2007, Table 1), let $Y = X + \varepsilon$, where both variables on the right are independently standard normal. Suppose for simplicity that there are only two outcome bins, $r_1 = (-\infty, 0]$ and $r_2 = (0, \infty)$. Consider forecaster A with $\mathbf{p}_1^{\mathrm{A}} = \Phi(-X), \mathbf{p}_2^{\mathrm{A}} = 1 - \Phi(-X) = \Phi(X)$. For forecaster A, Assumption 1 is satisfied with $\mathcal{F} = \sigma(X)$, the sigma algebra generated by X. In line with Proposition 1, it can be shown that the expected RPS and expected ERPS of forecaster A are both equal to 1/6. (In the notation of Proposition 1, it holds that $\mathbb{E}(\mathrm{RPS}(\underline{\mathbf{p}}^{\mathrm{A}}, k^*)) = \mathbb{E}(\mathrm{ERPS}(\underline{\mathbf{p}}^{\mathrm{A}})) = 1/6$.) For a second forecaster B with $\mathbf{p}_1^{\mathrm{B}} = \mathbf{p}_2^{\mathrm{B}} = 0.5$, Assumption 1 is satisfied with $\mathcal{F} = \emptyset$, the empty information set. The expected ERPS and expected RPS of forecaster B are both equal to 1/4, confirming the intuition that B's forecast is less informative than A's forecast.

We think that the ideas just sketched are a useful first step toward comparing subjective and objective uncertainty based on the (E)RPS. That said, further questions need to be addressed before applying the comparison to the SCE data, notably relating to the panel structure of the data (with many cross-sectional units and relatively few time periods). We leave these questions for future research.

References

Armantier, O., G. Topa, W. van der Klaauw, and B. Zafar (2017): "An overview of the Survey of Consumer Expectations," *Federal Reserve Bank of New York Economic Policy Review*, 23, 51–72.

Ben-David, I., E. Fermand, C. M. Kuhnen, and G. Li (2019): "Expectations uncertainty and household economic behavior," Working Paper, available at http://public.kenan-flagler.unc.edu/faculty/kuhnenc/RESEARCH/bfkl_v7.pdf (last accessed: August 8, 2019).

Ben-David, I., J. R. Graham, and C. R. Harvey (2013): "Managerial miscalibration," *The Quarterly Journal of Economics*, 128, 1547–1584.

Bruine de Bruin, W., W. Van der Klaauw, J. S. Downs, B. Fischhoff, G. Topa, and O. Armantier (2010): "Expectations of inflation: The role of demographic variables, expectation formation, and financial literacy," *Journal of Consumer Affairs*, 44, 381–402.

Clements, M. P. (2014): "Forecast uncertainty-ex ante and ex post: US inflation and output growth," *Journal of Business & Economic Statistics*, 32, 206–216.

Ehm, W., T. Gneiting, A. Jordan, and F. Krüger (2016): "Of quantiles and expectiles: consistent scoring functions, Choquet representations and forecast rankings," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 78, 505–562.

Engelberg, J., C. F. Manski, and J. Williams (2009): "Comparing the point predictions and subjective probability distributions of professional forecasters," *Journal of Business & Economic Statistics*, 27, 30–41.

Federal Reserve Bank of New York (2020): "Survey of Consumer Expectations," https://www.newyorkfed.org/microeconomics/sce (last accessed: March 9, 2020).

Galvao, A. B. and J. Mitchell (2019): "Measuring data uncertainty: An application using the Bank of England's "fan charts" for historical GDP growth," Working Paper, available at https://docs.google.com/viewer?a=v&pid=sites&srcid=ZGVmYXVsdGRvbWFpbnxhbmFiZ2FsdmFvfGd4OjY0MDBjMzA1YTc0NDE4OWU (last accessed: August 8, 2019).

Gneiting, T., F. Balabdaoui, and A. E. Raftery (2007): "Probabilistic forecasts, calibration and sharpness," *Journal of the Royal Statistical Society: Series B (Statistical Methodology)*, 69, 243–268.

Gneiting, T. and R. Ranjan (2013): "Combining predictive distributions," *Electronic Journal of Statistics*, 7, 1747–1782.

Jurado, K., S. C. Ludvigson, and S. Ng (2015): "Measuring uncertainty," *American Economic Review*, 105, 1177–1216.

Strähl, C. and Ziegel (2017): "Cross-calibration of probabilistic forecasts," *Electronic Journal of Statistics*, 11, 608–639.

Working Paper Series in Economics

recent issues

No. 139 *Fabian Krüger and Lora Pavlova:* Quantifying subjective uncertainty in survey expectations, March 2020

No. 138 *Michael Müller and Clemens Puppe:* Strategy-proofness and responsiveness imply minimal participation, January 2020

No. 137 *Andranik S. Tangian:* Tackling the Bundestag growth by introducing fraction-valued votes, October 2019

No. 136 *Susanne Fuchs-Seliger:* Structures of rational behavior in economics, September 2019

No. 135 *Cornelia Gremm, David Bälz, Chris Corbo and Kay Mitusch:* Intermodal competition between intercity buses and trains – A theoretical model, September 2019

No. 134 *Marta Serra-Garcia and Nora Szech:* The (in)elasticity of moral ignorance, August 2019

No. 133 *Andranik S. Tangian:* Composite indicators for computer-aided collective bargaining, July 2019

No. 132 *Andranik S. Tangian:* Combining the third vote with traditional elections, May 2019

No. 131 *Klaus Nehring and Clemens Puppe:* Resource allocation by frugal majority rule, April 2019

No. 130 *Benedikt Notheisen and Christof Weinhard:* The blockchain, plums, and lemons: Information asymmetries & transparency in decentralized markets, February 2019

No. 129 *Benedikt Notheisen, Vincenzo Marino, Daniel Englert and Christof Weinhard:* Trading stocks on blocks: The quality of decentralized markets, February 2019