DOI: 10.1111/jmcb.13003

DAVID A. COMERFORD

Response Bias in Survey Measures of Expectations: Evidence from the Survey of Consumer Expectations' Inflation Module

The Survey of Consumer Expectations (SCE) infers respondents' inflation expectations from density forecasts. Using numeracy data and tests for coherence among 117,000 respondents to the SCE, I find that density forecasts suffer non-negligible reporting bias and selective nonresponse. A simple verbal question collected by the SCE suffers neither of these deficiencies and so has better properties to deliver an accurate snapshot of the population's inflation expectations than the headline measures of inflation expectations published by the SCE. I demonstrate how the verbal measure can be harnessed to improve the signal-to-noise ratio in density forecasts.

JEL codes: E31, D84
Keywords: inflation expectations, forecasting, consumer surveys, density forecast

Survey data on consumer expectations are influential. One concrete measure of their influence is that investors paid for early access to the Michigan Survey of Consumers and trading volumes increased on the release of those data (Hiban 2013). These data also speak to a core question in economics–how quickly and completely do people update their inflationary expectations? (e.g., Coibion, Gorodnichenko, and Kamdar 2018, Forder 2010). Much of what we know about expectations formation and updating comes from survey measures of expectations (e.g., Armantier et al. 2016, Binder 2017, Capistran and Timmermann 2009; Mankiw, Reis, and Wolfers 2003, Dräger and Lamla 2012, Malmendier and Nagel 2016, Clements 2018, Rudd 2022). Substantial resources are invested in collecting survey data on expectations–in the United States alone, the Survey of Consumer Expectations (SCE), the Conference Board, and the Michigan Survey of Consumers each recruit three

Received January 12, 2021; and accepted in revised form June 9, 2022.

Journal of Money, Credit and Banking, Vol. 0, No. 0 (December 2022)

nationally representative samples every month. It is important to test whether these data capture what we infer them to be capturing.

In these analyses, I investigate the state-of-the-art measure of inflation expectations. In 2013, the Federal Reserve Bank of New York initiated a Survey of Consumer Expectations (henceforth SCE). Its inflation expectations module was informed by a series of experiments devised by a team of economists and psychologists. The literature suggests that the inflation expectations measures it delivers are more meaningful than data returned by another widely used data source, the Michigan Survey of Consumers (for a summary, see de Bruin and Fischhoff 2017). Today, the Federal Reserve Bank of New York publishes as its headline metric of inflation expectations a measure inferred from the SCE's density forecasts.

Here, I test the inflation expectations returned by the SCE's density forecast procedure against a basic criterion: do they accurately capture respondents' beliefs regarding the direction of inflation? This criterion was suggested by the data; as well as eliciting density forecasts, the SCE asks respondents by a verbal procedure whether there will be inflation or deflation. Although the literature focuses on *how large* inflation is expected to be, there are advantages to using this directional criterion. First, it is impossible to accurately forecast inflation without also accurately forecasting the direction of inflation. So, if a procedure fails to capture a respondent's directional beliefs, then it could not possibly capture more precise measures of that respondent's beliefs. Second, deflationary expectations are especially important because they have a self-fulfilling character; if consumers expect prices to fall, then they will tend to defer consumption to avail of lower future prices, thereby reducing aggregate demand (Krugman 1998). In short, the question of whether a procedure accurately captures a respondent's directional beliefs is fundamental.

The SCE data also offer an unambiguous test of which of the density forecast and the verbal procedure more accurately captures respondents' beliefs. Each respondent to the SCE is quizzed on their numerical ability. Clearly, the density forecast procedure requires numerical ability–in order to answer it meaningfully, a respondent must translate their inflationary beliefs into percentages. The verbal procedure does not require anything approaching the same level of numerical ability–it simply asks will there be inflation or deflation. My hypothesis is that respondents who are identified by the SCE data as low in numeracy will be more likely to deliver density forecasts that contradict their verbal reports than respondents who are higher in numeracy. I find strong support for this hypothesis. The only circumstance in which these results could be reconciled with the density forecast procedure delivering a more accurate measure of inflationary beliefs than the verbal procedure is if the following two conditions hold. First, low numeracy respondents would have to perform better on a complex numerical task than on a simpler verbal task; they would have to do a better job of answering 10 questions that sought the percentage chance that inflation/deflation would fall within some numerical range than they did at answering the verbal question "will there be inflation or deflation?" Second, lower numeracy respondents would have to be more likely than higher numeracy respondents to report percentages that were more meaningful than their verbal responses. Since it is implausible that these conditions could hold, we must conclude that the density forecast procedure captures beliefs less accurately than the verbal procedure.

The current research speaks to any research that uses survey measures of inflation expectations. It clarifies that the current state-of-the-art measure of inflation expectations systematically misrepresents the public's beliefs. Moreover, it documents that this misrepresentation is heterogeneous. Groups who are low in numeracy (e.g., the less educated) are especially unlikely to have their beliefs represented. These results inform a fundamental question, that is, whether inflation expectations matter (Rudd 2022). To the extent that current survey measures systematically misrepresent the population's inflation expectations, we would not expect those measures to strongly predict consumption and savings behavior. I conclude with suggestions of how we might obtain measures that more accurately capture the public's beliefs.

1. BACKGROUND

An influential paper by Manski (2004) highlighted the value of collecting survey measures of expectations. Although expectations data had been collected for decades through the Michigan Survey of Consumers, that survey does not elicit expectations as precisely as it could. For instance, in its inflation expectations module, the Michigan Survey asks respondents for a point estimate of price changes. Manski (2004) and Dominitz and Manski (2004) proposed collecting data on macroeconomic expectations by asking respondents for the probability that outcomes would fall in some given range. This proposal gave rise to the density forecast procedure. In 2011, Manski and coauthors reported the properties of the data returned by a density forecast procedure, which elicited inflation expectations as subjective probabilities (De Bruin et al. 2011). Unlike the Michigan procedure, the density forecast procedure gives measures of a respondent's uncertainty regarding inflation. Recent research has highlighted that this contribution is especially important because, in the absence of any explicit question about their uncertainty, respondents to the Michigan survey signal their uncertainty by reporting round numbers (Binder 2017). Additionally, the Michigan data lack face validity; about half of responses to its point estimate question indicate inflation expectations in excess of 5%, which has been questioned as implausibly high (de Bruin and Fischhoff 2017).

In 2013, the SCE was launched and it employed the density forecast procedure as its headline measure of inflation expectations. The SCE's inflation expectations density forecast procedure is depicted in Figure 1. A series of questions ask respondents to assign probabilities to various ranges of inflation/deflation.

The specifics of the density forecast procedure employed in the SCE mean that it gives more precise data than the Michigan measure. For instance, the Michigan question asks about "prices in general," whereas the density forecast asks about "inflation." In an experiment, De Bruin et al. (2012) compared expectations of "inflation" against expectations of "prices in general" and found greater dispersal of forecasts by the "prices" question. de Bruin et al. (2017) tested whether inflation forecasts

Now we would like you to think about the different things that may happen to inflation **over the next 12 months.** We realize that this question may take a little more effort.

In your view, what would you say is the percent chance that, **over the next 12 months...**
Instruction H4.

the rate of inflation will be 12% or higher (bin 1)	_____ percent chance
the rate of inflation will be between 8% and 12% (bin 2)	_____ percent chance
the rate of inflation will be between 4% and 8% (bin 3)	_____ percent chance
the rate of inflation will be between 2% and 4% (bin 4)	_____ percent chance
the rate of inflation will be between 0% and 2% (bin 5)	_____ percent chance
the rate of deflation (opposite of inflation) will be between 0% and 2% (bin 6)	_____ percent chance
the rate of deflation (opposite of inflation) will be between 2% and 4% (bin 7)	_____ percent chance
the rate of deflation (opposite of inflation) will be between 4% and 8% (bin 8)	_____ percent chance
the rate of deflation (opposite of inflation) will be between 8% and 12% (bin 9)	_____ percent chance
the rate of deflation (opposite of inflation) will be 12% or higher (bin 10)	_____ percent chance
TOTAL	**100**

Fig 1. A Screenshot of the Density Forecast Questions from the SCE.

differ when collected by phone (as the Michigan survey does) versus when collected by online survey (as the SCE does). That work found greater dispersal of inflation forecasts collected by phone. In short, the density forecast procedure employed by the SCE appears to represent the most precise data regarding inflation expectations currently collected, at least in the United States.

Yet, both the Michigan question and the density forecast procedure seem vulnerable to mismeasuring beliefs. Both rely on respondents translating their likelihood beliefs onto a percentage scale, which many people are incapable of doing. When a representative sample of the U.S. population was asked to convert the natural frequency 20 out of 100 to a percentage, 28% could not (Kahan et al. 2012).

The economics literature is cognizant of reporting errors because they show up when data are aggregated, for example, because clusters emerge at round numbers (e.g., Hurd 2009, Manski and Molinari 2010). But dealing with reporting errors at the level of individual observations is fraught with assumptions and trade-offs. For instance, some respondents report 50% to indicate “I have no idea” (Fischhoff and de Bruin, 1999; Hurd 2009). Since it is usually impossible to distinguish uninformative round numbers from meaningful responses, researchers are often faced with admitting pure noise into their analysis or dropping signal from their data.

A handful of recent papers have delivered insight on respondents’ beliefs by moving beyond a face-value interpretation of survey reports. For instance, in the Michigan data, a face-value interpretation of a response of 5% is that this respondent's best estimate of inflation is 5%. Binder (2017) observes that some respondents report 5% (and other round numbers) as a means to convey that they lack confidence in their forecast.

In the domain of survival beliefs, knowledge of a reporting bias was used to extract a greater signal from respondents' stated beliefs. Comerford (2021) shows that respondents to the English Longitudinal Study of Ageing (ELSA) systematically misreport their survival forecasts such that their reports are closer to 50% than their beliefs. This reporting bias explains a number of previous results that had been wrongly interpreted as biases of beliefs. Moreover, model specifications that included parameters to account for the reporting bias (e.g., a quadratic term and a rank term) improved the predictive power of respondents' forecasts relative to a face-value specification. These results exemplify a general insight: identification of reporting bias in existing survey data can improve inference and forecasting.

There is reason to suspect reporting bias in density forecasts of inflation. Comerford and Robinson (2017) used a density forecast to elicit survival beliefs and found that 12% of respondents reported higher probabilities of living to an older age than to some younger age. These errors are transparently illogical in the context of survival beliefs because it is impossible to live to an older age without also living to some younger age. In the context of inflation forecasts, however, there is no way of knowing which data manifest mistakes of this type. As a result, meaningless density forecasts will contaminate the data.

Additionally, there is reason to suspect that meaningless responses will give rise to bias. The first likelihood belief that the density forecast procedure asks respondents to report is that inflation will be 12% or more (Figure 1). If respondents who mean to indicate "I have no idea" respond using 50%, then they imply a 50% chance of extremely high inflation and so would be expected to add positive bias to the aggregate measure of inflation expectations.

A separate concern is that a respondent who finds the density forecast procedure too taxing may simply skip it. If such respondents have beliefs about inflation that differ from those of the rest of the population, then the aggregate data returned by the density forecast procedure will no longer provide a representative snapshot of the population's inflation expectations.

In sum, there are two sources of bias that might result from the response burden imposed on respondents by the density forecast procedure. The first is selective nonresponse. The second is biased response. These biases would cause the SCE's inflation expectations measure to give a distorted impression of the inflation expectations that the U.S. public brings to bear on their spending and savings decisions.

Additionally, the forecast density procedure imposes a burden of responsibility on survey administrators. Assumptions must be made and choices taken when estimating a density forecast from respondents' reported probabilities and when summarizing respondents' forecast densities in a headline metric. For instance, the summary measure of inflation expectations reported on the SCE's website is the median of the mean of respondents' density forecasts (Armantier et al. 2017) but plausible alternatives exist, for example, the mean of the means; the mode of the medians; and various nonparametric measures of central tendency.

Note that the verbal procedure asked in the SCE is less vulnerable than the density forecast procedure on each of these fronts. In order to meaningfully answer the verbal question, the respondent needs to understand the terms “inflation,” “deflation,” and to have some expectation regarding inflation for the coming year. In order to meaningfully answer the density forecast procedure, all of these same conditions must be met as well as the condition that the respondent can meaningfully translate their likelihood beliefs onto a percentage scale. Additionally, the verbal question has the advantage that it elicits from respondents a single, unambiguous response, whereas the density forecast procedure requires survey administrators to interpret respondents’ forecasts.

1.1 Empirical Strategy

If two procedures are accurately measuring the same belief, then they should correlate perfectly. The key outcome in this paper is discrepancies across procedures. The first hypothesis tested in these data is that discrepancies are predicted by the SCE’s measure of respondent numeracy.

A separate question concerns how representative of the population’s beliefs are the measures returned by the density function? A second hypothesis is that numeracy will predict nonresponse to the density forecast procedure.

Then, follow a set of exploratory analysis to assess whether meaningless response and nonresponse distort the aggregate forecasts reported by the SCE.

The first question concerns selection on beliefs, that is, do the beliefs of nonrespondents to the density forecast procedure differ from the beliefs of respondents? I can conduct this test because the verbal question gives a measure of beliefs from both groups. For robustness, I supplement this test with data from the Michigan Survey of Consumers, which also asks respondents for a directional forecast but does so without employing the technical terms “inflation” and “deflation.” It simply asks the respondent whether they expect prices to go up or down.

I then test for variation in response rates over time. We might expect nonresponse to be greater when uncertainty is greater. The SCE data return a respondent-specific uncertainty measure–the interquartile range (IQR) from respondents’ density forecasts. In periods where respondents feel confident that they can predict inflation with a high degree of precision, they will return density forecasts with the mass of probabilities assigned to a narrow range of inflationary outcomes, resulting in a small IQR. In periods of high uncertainty, respondents will distribute their probabilities across a wider range of outcomes and the IQR will be larger. If biases are especially prevalent at times of high uncertainty, we would expect a positive correlation between measures of bias (e.g., contradictions; nonresponse) and the mean IQR returned by respondents.

I close by considering the bias induced by meaningless responses to the density forecast procedure. My approach in this analysis is to test the effect on aggregate forecasts of dropping from the sample those responses where there were unambiguous contradictions across a respondent's density forecast and their verbal report. I define an unambiguous contradiction as one where there is no parametric measure of central tendency that matches the directional forecast delivered by the verbal response, that is, a respondent is categorized as delivering a meaningless density function if the mean, median, and mode of their forecast density are of a different sign than their verbal report.

A typology of the three possible permutations of contradictions is useful for thinking through the consequences of dropping these unambiguous contradictions. The first type of contradiction is caused by respondents who gave meaningless answers to both formats, for example, because they do not understand the terms "inflation" and "deflation." By definition, these respondents do not know what they are talking about and so are simply adding noise. Dropping them from the data must increase the signal-to-noise ratio. The second type of contradiction results from respondents who gave a meaningful response to the verbal question but not to the density forecast procedure. By definition, their density forecasts imply the opposite direction of forecast from their inflation beliefs. Again, dropping these respondents will only improve the signal-to-noise ratio. The third type concerns individuals whose density forecasts are meaningful but whose verbal responses are meaningless. By definition, any individual who made this type of contradiction invested effort in answering the density forecast procedure and so demonstrated themselves to be engaged respondent. That they contradicted their density forecast implies that–despite their engagement–they made some random error when answering the simple verbal question, for example, they happened to click the wrong button by mistake. Given that these errors happened at random, dropping such respondents from the sample is equivalent to dropping observations at random throughout the distribution of meaningful responses. It will dilute the proportion of meaningful density forecasts in the data and so will reduce the signal-to-noise ratio in any given wave. There are two reasons to infer that this type of contradiction is rare. First, a prerequisite for it occurring is that a respondent meaningfully answered each of the 10 cognitively taxing questions that form the density forecast procedure but failed to meaningfully answer a far more straightforward verbal question. That particular pattern of response would be expected to occur only very rarely. Second, this third type of contradiction is inconsistent with the result that contradictions are negatively predicted by respondent numeracy. As such, this third type constitutes a small minority of observed contradictions. In expectation then, dropping any given contradictory response from the sample would have a positive impact on the signal-to-noise ratio in any given wave of data collection. Also, because there is no reason to expect a respondent to be more likely to click on the wrong button in one wave of data collection than in any other, mistaken clicks in response to the verbal question could not account for the time variation in bias. In what follows, therefore, I refer to the data in which contradictions are dropped as signal-rich data.

1.2 *Empirics*

1.2.1 Data. The data come from the SCE between June 2013 and November 2020.[1] The SCE elicits responses from a representative sample of approximately 1,300 U.S. residents in an online survey each month.

The SCE helps respondents construct their density forecast. It reminds respondents that percentages "can range from 0 to 100, where 0 means there is absolutely no chance, and 100 means that it is absolutely certain" and offers some additional verbal descriptions of specific numerical percentages and ranges: for example, "83 percent or so may mean a 'very good chance'." Also, an automated process embedded within the web survey ensures that respondents deliver responses that sum to 100 exactly. From the reported probabilities, the SCE parametrically estimates the underlying forecast density function (for details, see section 5.2 of Armantier et al. 2017).

As well as the density forecast, I use the verbal question that opens the SCE's module on inflation expectations. It reads "Over the next 12 months, do you think that there will be inflation or deflation? (Note: deflation is the opposite of inflation)" and is answered with either "inflation" or "deflation."[2] The final measure I use is a binary variable in the SCE data that categorizes respondents as either low/not low in numeracy. This categorization was made by the SCE data administrators on the basis of respondents' answers to a series of math questions (sample item QNUM5: If the chance of getting a disease is 10%, how many people out of 1,000 would be expected to get the disease?). Cumulatively, my data comprise 117,832 individual-level observations across 90 waves of data collection.

2. RESULTS

This analysis focuses on forecasts of inflation for the coming 12 months; I consider longer-term expectations in a robustness test. Each respondent provides two data points regarding their belief as to whether there will be inflation or deflation over the coming 12 months. The first is a verbal measure which takes a value of 1 if the respondent answered "inflation" and takes a value of zero otherwise. The second is the density forecast summary measure recorded by the SCE–the mean of their density function. My outcome of interest is discrepancies across these two measures, for example, a respondent answers the verbal question with "deflation" but the SCE records them as having indicated "inflation" because the mean of their density forecast was positive.

1. Source: Survey of Consumer Expectations, © 2013–2020 Federal Reserve Bank of New York (FRBNY). The SCE data are available without charge at https://www.newyorkfed.org/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.

2. In response to a follow-up question, respondents could indicate 0% inflation.

In practice, there is one uninteresting mechanism that would give rise to discrepancies and it is specific to respondents who believe that inflation will be 0%. In answer to the verbal question, these respondents cannot reply "no change," they must reply either "deflation" or "inflation." If they also return a density forecast that is even slightly noisy, then there is a high chance that they would contradict themselves even though the SCE is meaningfully (though imperfectly) capturing their inflationary beliefs. A specific example is a respondent who answers "inflation" and delivers a density forecast with a mean forecast of –0.02%. The SCE data allow me to identify respondents who believe inflation will be 0% because immediately after eliciting their verbal forecast it asks respondents to report a point estimate of inflation.[3] I drop from the analysis respondents who made point estimates of 0% inflation and so are at risk of contradicting a potentially meaningful density forecast. These responses are rare. In total, they come to 662 out of over 117,000. At their most prevalent, in April of 2020, there are 18 forecasts of 0% out of a total sample of 1,300.

Figure 2 depicts the proportion of low numeracy respondents (27.7% of the sample, top panel) and other respondents (the remaining 72.3%, bottom panel), indicating inflation by each procedure. For respondents low in numeracy, the correlation across procedures is far from perfect ($r = 0.47$). For the rest of the sample, the correlation is significantly higher than it was for the low numeracy group ($z = 35.84$, $p < 0.001$), but it is still less than perfect ($r = 0.63$).

The chief source of discrepancy is that individual respondents contradict themselves. Of those who answered both procedures, 6.1% indicated inflation by one and deflation by the other. Three thousand four hundred and ninety-two respondents, or 33.4% of those who answered "deflation," returned density forecasts that had a positive mean. Five thousand five hundred and forty-four respondents, or 5.2% of those who answered "inflation" implied deflation by their density forecasts. Removing self-contradicting respondents from the sample raises the correlation coefficient from $r = 0.61$ to $r = 1$.

The remaining source of discrepancy across procedures is differential rates of response. Whereas fewer than two in a thousand respondents failed to answer the verbal question, more than 10 times as many (2.2% of the sample) failed to respond to the density forecast procedure over these 90 waves of data. Restricting the sample to those who answered by both procedures raises the correlation coefficient from $r = 0.57$ to $r = 0.61$.

Both sources of discrepancy are predicted by numeracy. Among those identified as low in numeracy, 4.5% failed to answer the density forecast questions compared with 1.3% of the rest of the sample ($z = 22.12$, $p < 0.001$). Low numeracy respondents were over twice as likely as other respondents to deliver density forecasts that contradicted their verbal responses (low numeracy = 11.0% vs. rest of sample = 4.3%).

3. The question is worded as: "What do you expect the rate of [inflation (if Q8v2=inflation)/deflation (if Q8v2=deflation)] to be **over the next 12 months**? Please give your best guess."

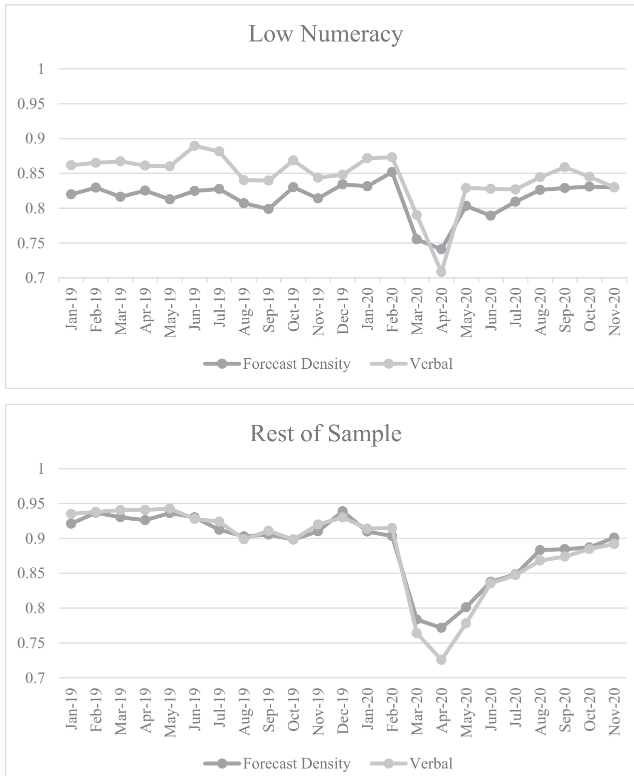

Fig 2. Proportion Low Numeracy (top) and Remaining Respondents (bottom) Forecasting Inflation by the Verbal Procedure and by the Density Forecast Procedure.

Being low in numeracy is a highly significant predictor of contradicting oneself ($z = 40.11$, $p < 0.001$).

2.1 Robustness Checks and Exploratory Analyses

I opened these analyses by observing that respondents who predict precisely zero inflation might deliver contradictions that are nondistortionary. A related point is that some respondents might deliver contradictions that are relatively harmless because they imply only low levels of inflation or deflation. Of those who contradicted themselves, over half delivered forecast densities that implied inflation or deflation of greater than 1%. Over 2% of the total sample answered that they expected price changes to move in one direction but returned forecast densities that implied a larger than 2% price change in the opposite direction. The next section further investigates the scale of the distortions induced by nonresponse and reporting bias but for now, it suffices to say that these sources of bias present a nontrivial threat to data quality.

The above analysis focused on the mean of the density forecast because that is the headline measure reported by the SCE. In their daily lives, some people might make decisions with reference to the mean of their distribution, others might draw on the median and others might draw on something else. Whereas the SCE has to make an assumption about which measure of central tendency respondents use, the verbal question vests this judgment with the respondent. Consider a respondent who delivers a forecast density with a mean that implies inflation and a median that implies deflation. In any such case, the SCE assumes the mean to be the valid measure (Armantier et al. 2017). But the verbal question resolves any such conflict by simply asking the respondent what they meant: inflation or deflation? I conclude that, when the mean and median conflict, we should refer to the direction indicated by the verbal question. Some of those 6.1% of responses that appeared contradictory might be reconciled by this mechanism. To quantify this, I consider two additional measures of central tendency–the parametric median and the mode.[4] A respondent is coded as making an unambiguous contradiction if the mean, median, and mode of their density forecast are all contrary to their answer to the verbal question. Unambiguous contradictions were made in 4.0% of the ca. 117,000 responses to the SCE. I use the term ambiguous contradictions to describe the set of remaining contradictions, 2.1% of the total sample. This set comprises two discrete subsets of contradictions made by two distinct types of respondent. One subset contains meaningless density forecasts and these will predominantly be made by low numeracy respondents. The other contains technical contradictions that derive from the SCE's reliance on the mean as opposed to other measures of central tendency that respondents use. The type of respondent who manifests a technical contradiction is someone who can formulate a meaningful density forecast. Since numerical ability is a prerequisite for returning a meaningful density forecast, these technical contradictions will predominantly be made by high numeracy respondents. The testable implications are that (1) being low in numeracy is a positive predictor of making unambiguous contradictions and (2) being low in numeracy is an increasingly *negative* predictor of making ambiguous contradictions as the proportion of technical contradictions in the set of ambiguous contradictions approaches 1, that is, if all ambiguous contradictions were explained by the fact that respondents return meaningful density forecasts from which the SCE wrongly infers the mean as the crucial measure of central tendency, then we would see low numeracy negatively predicting ambiguous contradictions. Unambiguous contradictions are positively predicted by low numeracy (odds ratio: 3.06, 95% CI: 2.89–3.26). Ambiguous contradictions are also positively predicted by low numeracy, but significantly less so than unambiguous contradictions (odds ratio: 1.89, 95% CI: 1.75–2.06). Though they are less predicted by numeracy than unambiguous contradictions, the fact remains that ambiguous contradictions are about twice as likely to be committed by those low in numeracy than those higher

4. There are also many nonparametric measures of central tendency that could conceivably be what respondents draw on when making their inflation-relevant decisions.

in numeracy. This result implies that a substantial proportion of ambiguous contradictions are a symptom of low numeracy respondents returning meaningless density forecasts.

The data offer additional opportunities to test if difficulties working with percentages are what is driving the observed results. For instance, we would expect to see that contradictions are similarly commonplace when respondents are asked by the density forecast procedure about other outcomes, for example, inflation over the coming 3 years, inflation in their own earnings, and house price inflation. We should also see that contradictions are again concentrated among those identified as low in numeracy. This analysis is preregistered[5] and it concerns data that I had not yet downloaded at the time of preregistration. Hence, it looks only at data collected after 2016 (and again drops respondents who forecast precisely 0% inflation). Among these responses, 7.1% contradict their forecasts of inflation over the next 36 months, 5.7% contradict forecasts of their own future earnings, and 7.6% contradict their forecasts of house price inflation.[6] As predicted, contradictions are concentrated among low numeracy respondents, 12.1% of whom contradicted their long-term inflation forecasts, 9.8% of whom contradicted their earnings forecasts, and 10.8% of whom contradicted their house price inflation forecasts.

Digging into the data gives some clues as to which response processes lead less numerate respondents to deliver biased density forecasts. Recall that respondents often express "I have no idea" by answering 50% (Fischhoff and de Bruin, 1999). When answering the first question of the density forecast procedure, 1.8% of responses were precisely 50%. Some of these responses may be sincere expressions of belief, but two results suggest that many were not. First, the question asks about extremely high inflation, 12% or more. Second, answers of 50% were over three times as likely among low numeracy respondents (odds ratio = 3.22, $z = 25.18$, $p < 0.001$).

Drilling down into the numeracy measure gives further insight into the mechanism. I ran a logistic regression of contradictions on the seven numeracy items (coded 1 if correct, 0 otherwise).[7] There are two numeracy items that ask respondents to translate natural frequencies into percentages or vice versa, QNUM3 and QNUM5. QNUM3 was answered correctly by 81%, and QNUM5 was answered correctly by 88% of the sample. The logistic regression shows that getting QNUM3 incorrect increases the likelihood of contradicting the direction of one's 12-month forecast by 32% (z

5. Preregistered hypotheses and data coding here: https://aspredicted.org/3ix86.pdf

6. We should be careful not to read much into the variation in contradictions across outcomes. The samples differ. One reason for this is because I drop respondents who forecast precisely zero inflation in each outcome. A second is because of the SCE's design: only certain respondents are asked about house price inflation and only employed respondents are asked about earnings inflation.

7. Up to March 2015, only five numeracy questions were asked of respondents. This analysis is restricted to data collected after 2017 and so considers all seven numeracy items. Respondents typically completed more than one wave of data collection but were asked the numeracy questions in only one wave of data collection. These analyses consider forecasts from the wave in which the respondent completed numeracy measures (n = 7,283).

$= 2.83$, $p = 0.005$) and that getting QNUM5 incorrect independently increases the likelihood of contradicting by 39% ($z = 2.99$, $p = 0.003$).

The strongest independent predictor of contradictions among the numeracy questions, however, is a question that measures something over and above an ability to work with percentages; QNUM2, on compound interest, appears to measure cognitive reflection. The question reads: "Let's say you have $200 in a savings account. The account earns ten per cent interest per year. Interest accrues at each anniversary of the account. If you never withdraw money or interest payments, how much will you have in the account at the end of two years?" Whereas the other questions delivered unimodal responses, answers to this question clustered at two responses: 25% of respondents answered $240 and 50% answered $242. Each of these responses demonstrates an ability to work with percentages. The correct answer of $242 additionally demonstrates something akin to cognitive reflection, which is defined as "the tendency to check and detect intuitive errors" (Sinayev and Peters 2015). Respondents who got this question incorrect were 66% more likely to contradict themselves, over and above the contribution of any other numeracy question ($z = 5.63$, $p < 0.001$). Restricting the sample to just those who answered $240 or $242, answering with the intuitive though incorrect $240 more than doubles the likelihood of contradicting one's directional forecast relative to answering $242 ($n = 5{,}578$, $z = 7.34$, $p < 0.001$). In other words, among people who were manifestly capable of calculating percentages, contradictions were predominantly made by those who failed to engage in cognitive reflection.

Cognitive reflection requires cognitive resources. To the extent that it is cognitively taxing to simulate uncertain outcomes, this cognitive reflection mechanism predicts a higher rate of contradictions when respondents are reporting their expectations regarding uncertain outcomes. We can test this prediction because we have a measure of uncertainty: the IQRs of respondents' density forecasts. This suggestion is supported both at the level of the individual respondent and at the level of the survey wave. A logistic regression finds that individuals whose IQRs are higher are more likely to contradict their 12-month inflation forecasts ($z = 69.25$, $p < 0.001$), though this result is also consistent with a tendency for meaningless density forecasts to have especially high IQRs. More persuasive support of this prediction is that there is a positive correlation between the mean IQRs from noncontradicting respondents in a given wave and the prevalence of unambiguous contradictions in that same wave, $r = 0.29$ ($p = 0.005$).[8]

8. I thank an anonymous reviewer for suggesting a normative explanation for why contradictions would be more prevalent when uncertainty is high: some respondents might foresee two paths that are similarly likely—an inflationary path and a deflationary path. These respondents could use their density forecasts to indicate their bimodal beliefs but could not express those complex beliefs by the simple verbal question. All it would take for their density forecast to contradict the answer they give when forced to call "inflation" or "deflation" is that their inflationary beliefs are inadequately captured by the mean of the density forecast procedure (e.g., Ryngaert, 2021). To reduce the influence of such cases, this analysis focuses on unambiguous contradictions, that is, cases in which all of the mean, median, and mode contradict the respondent's verbal forecast.

2.2 Error or Bias?

The goal of this section is to get an indication of the direction and scale of distortion to aggregate inflation expectations returned by density forecasts. It is helpful to keep in mind that when I speak of distortions here I am referring to distorted measures of *beliefs*. To clarify, these analyses are unconcerned with whether respondents hold accurate beliefs. The accuracy of beliefs is an independent question that is beyond the scope of these analyses. The questions of interest here are by how much and in what direction does the SCE data misrepresent the inflation expectations of the U.S. population?

An additional goal is to assess whether distortions are stable over time. If they are stable over time, then they might be corrected via a straightforward calibration procedure. If the data show a bias that is varying over time, then the aggregate inflation forecasts broadcast by the SCE are confounded.

The results that follow are descriptive statistics. Where regression results are reported, they derive from the simplest model specifications, that is, they include no covariates or wave dummies.

2.3 Selection on Beliefs

The data show that respondents who failed to respond to the density forecast procedure were over twice as likely as the rest of the sample to answer the verbal question with "deflation" ($z = 22.10$, $p < 0.001$).[9] Twenty percent of them answered "deflation" compared with 8.7% of the sample who delivered density forecasts. Caution should be taken interpreting this result. There may be some respondents who answered the verbal question despite not understanding what the terms "deflation" and "inflation" mean. But a result from an independent data set supports the inference that nonrespondents are more likely than respondents to believe that prices will fall. The Michigan survey avoids the technical terms "inflation" and "deflation." Instead, it asks whether prices in general will go up over the coming 12 months, go down, or stay the same. The category of respondent who is least likely to answer the density forecast procedure in the SCE is the same category of respondent who is most likely to answer the Michigan survey by saying that prices in general will go down. That category of respondent is one whose education ended at high school. Binary logistic regressions on having no higher than a high school level of education show that over the period June 2013–December 2019 these least educated respondents were over twice as likely as other respondents to leave the density forecast procedure unanswered (SCE data: $n = 103{,}558$, odds ratio $= 2.05$, $z = 10.22$, $p < 0.001$) and were 66% more likely than other respondents to answer that prices would go "down" (Michigan data: $n = 44{,}146$, $z = 3.53$, $p < 0.001$). I conclude that the forecasts reported by the SCE underestimate the prevalence of deflationary beliefs in the population.

9. For this specific analysis, I revert to using the full sample. Since I am no longer concerned with contradictions, the analysis is untroubled by including the 662 respondents who forecast exactly 0% inflation.

2.4 Time-Variation in Nonresponse

The rates of nonresponse to the verbal question are small and constant over time. In the first wave of data collection, 14 out of the 1,253 people recruited failed to answer the verbal question. Although low, this rate of nonresponse was significantly higher than in any subsequent wave ($z = 8.25$, $p = 0.001$), which may indicate respondent learning (see also Kim and Binder, 2020) or may be explained by some amendment to the survey instrument. The highest rate of nonresponse thereafter was six respondents out of 1,330 in June of 2019. In 27 of the 90 waves, every respondent answered the verbal question.

The response rate to the density forecast procedure shows a markedly different pattern. Nonresponse ranges from 1.5% of the sample in August 2018 to 3.6% of the sample in March of 2020.

Nonresponse to the density forecast is greater when uncertainty is greater. There is a positive correlation between the frequency of nonresponse and the mean IQR from respondents' density forecasts in a given wave ($r = 0.19$, $p = 0.072$).

The degree to which nonresponse to the density forecast procedure is explained by respondent beliefs and characteristics changes over time. In most waves of data collection, answering "deflation" to the verbal question is a strong positive predictor of nonresponse, for example, October 2020 ($n = 1{,}192$, $t = 3.32$, $p = 0.001$). But in a few waves, it is nonsignificant, for example, its sign turns negative in September 2019 ($n = 1{,}319$, $t = 0.74$, $p = 0.458$). Similarly, in most waves, numeracy is a significant positive predictor of responding to the density forecast procedure but its predictive power varies a lot, and in April 2014, its sign turns negative ($n = 1{,}214$, $z = 0.15$, $p = 0.884$).

To summarize, the density forecast procedure that underpins the inflation forecasts reported by the SCE manifests considerable and time-varying nonresponse. Moreover, another inflation measure collected as part of the SCE–the verbal question–suffers none of these deficiencies because it consistently shows trivial rates of nonresponse.

2.5 Distortions Induced by Meaningless Responses

The previous analysis answered the question: are the aggregate inflation expectations delivered by the density forecast procedure representative of the beliefs of the U.S. population? The current question is whether the density forecast procedure gives a biased measure of respondents' own beliefs.

The most recent wave of the SCE in my data is from November 2020. Excluding five respondents who gave a point forecast of precisely zero inflation, it returned density forecasts from 1,202 respondents and the median rate of inflation forecast by those density forecasts was 2.83%. In that wave of data, 62 respondents made unambiguous contradictions, that is, their density forecast delivered three measures of central tendency–mean, median, and mode–that all contradicted their verbal report. The remaining 1,140 respondents were consistent in the direction of their verbal answers and at least one of the mean, median, or mode of their density forecasts. The signal-rich data from just these 1,140 respondents deliver a median rate of inflation forecast that is 5% higher, at 2.98%.

Repeating this exercise for each wave of data shows three important results. The distortion induced by unambiguous contradictions varies in magnitude over time. In 13 of the 90 waves, error is zero. In the remaining 77 waves, the error is negative. In 71 out of the 90 waves, the raw data show an expected inflation rate at least 0.1 percentage points lower than the signal-rich data.

The correlation between mean IQRs from the full sample and the scale of bias is nonsignificant, $r = 0.10$; $p = 0.348$. This implies that the bias in reported inflation beliefs is unpredictable.

The foregoing suggests that bias in density forecast is large and time-varying. Consider, for instance, that over the month August 2015, the median of the mean rate of inflation forecast decreased to 2.8% from 3.0% in July. It would be natural to interpret this reduction as implying that over this month the public's expectations of inflation decreased. But the equivalent measure from the signal-rich data implies no such decrease–they remain steady at 3.0% throughout. The implication is that the density forecast data on inflation expectations are not a reliable indicator of whether respondents' inflation expectations are increasing or decreasing over time.

The analysis reported above corrects for just one egregious symptom of bias–unambiguously contradictory forecasts made by a respondent at a moment in time. It will do nothing to treat meaningless density forecasts that are in a direction consistent with verbal reports, since I cannot identify these. The point of the analysis above is not to recover the uncontaminated truth regarding respondents' inflation beliefs. That would be impossible with the current data. Rather, it is to describe and quantify the distortion induced by one form of bias. Doing so indicates that a face-value interpretation of published data on a leading measure of inflation beliefs can be misleading.

3. DISCUSSION

This research investigated whether the inflation expectations returned by the state-of-the-art density forecast procedure capture the population's beliefs. My analyses demonstrate two sources of bias: selective nonresponse and reporting bias.

How are we to reconcile these results with the experiments that recommended using density forecasts to elicit inflation expectations (see de Bruin and Fischhoff 2017)? The introduction offered one explanation: this study tested density forecasts against a directional criterion, whereas the previous literature did not. A second explanation is that the SCE was competing against an incumbent; the Michigan Survey of Consumers has been collecting inflation expectations since 1978. Experiments conducted in developing the SCE took the Michigan procedure as a benchmark, for example, by comparing the Michigan survey's phone response against online data collection; by comparing the Michigan survey's questions about "prices in general" against questions about "inflation" (de Bruin and Fischhoff 2017). The authors of these experiments concluded that SCE's density forecast procedure is superior to the Michigan procedure. The current research does not call that conclusion into question. It simply adds that the SCE's forecast density procedure is itself delivering substantively biased data, even if that data are more precise than the data returned by the Michigan survey.

With that in mind, it is worth taking a step back to consider what it is that we seek to measure with population-wide surveys of expectations. We can think of two independent goals. One is to measure the expectations that the public bring to bear on their economic decisions. A second is to harness private information held by the population to predict future inflation. Each requires different properties. When it comes to capturing the population's expectations, the verbal question asked by the SCE has good properties that the density forecast measure lacks. The verbal question successfully returns responses from all but a trivial number of respondents, which satisfies the necessary condition of giving a population overview from a representative sample. It also avoids one major source of mismeasurement–a widespread inability to articulate beliefs on a percentage scale.

On the second goal–predicting future inflation–the current analysis suggests that neither the verbal question nor the density forecast procedure is adequate. The density forecast procedure is inadequate because a substantial proportion of respondents fail to translate their beliefs onto a percentage scale accurately. Moreover, their attempts to translate their beliefs onto a percentage scale result in response bias. The verbal question included in the SCE is also ill-suited to serve as a forecasting tool because it is insensitive to the magnitude of expected inflation. The verbal question would only capture that subset of changes in expectations that shift respondents from forecasting inflation to deflation or vice versa. For example, any respondents who forecast 1% inflation in May and 5% inflation in June would be coded as having stable expectations over the period May–June even though they have substantially updated their forecast.

There is another measure of inflation expectations collected by the SCE each month that is beyond the scope of this analysis but that may prove useful for forecasting. It is the follow-up to the verbal question and it asks respondents for a point estimate of the scale of inflation or deflation that they expect. Unlike the verbal question itself, this question would be expected to capture a month-on-month change in beliefs regarding the scale of future inflation. Like the density forecast measure, however, it asks respondents to answer with a percentage and so is vulnerable to distortion from respondents' inability to express their beliefs on a percentage scale.

The current research has not sought to investigate the forecasting properties of the various survey measures of inflation expectations. That seems an important question for future research. As well as testing the point estimate forecasts alluded to in the previous paragraph, it would be interesting to test the forecasting accuracy of the signal-rich data that derive from dropping contradictions across the verbal and forecast density measures.

Of course, when it comes to the task of predicting future inflation, there are alternatives to surveys of the general population. The Survey of Professional Forecasters might better serve that goal, although it should be noted that even these professional forecasters have been documented to deliver error-prone density forecasts (Clements 2010).

There are a number of conclusions that might be mistakenly inferred from the current research. First, these analyses do not imply that surveys of the expectations held by the general public are useless. As noted in the introduction, there is a revealed preference for survey measures of the population's expectations and people are willing to pay for the data they generate. Moreover, such data have been shown to have unique predictive power. Over and above contemporaneous economic indicators, the forecasts returned by random U.S. households for the Michigan Survey of Consumers and the Conference Board predict household spending a year hence (Ludvigson 2004). It is noteworthy that the survey data that delivers this predictive power were in every case answers to verbal measures. For instance, in the Conference Board survey, respondents were asked "Six months from now, do you think there will be more/ same / fewer jobs available in your area?" It is an open question whether a survey item that asked respondents to forecast the number of available jobs in their area or to construct a density forecast of the number of jobs in their area would be more, less, or equally predictive of household spending. In any case, the Ludvigson (2004) result suggests that SCE's verbal question on inflation expectations might also be predictive of household economic behavior. Stepping back to the big picture, the message to take from the current analysis is that existing data from surveys of the general population are useful and that there is scope to make them even more so.

Second, it might be inferred that those collecting expectations, data are faced with a choice between collecting meaningful responses to a less precise directional measure versus collecting less meaningful responses to a more precise quantitative measure. It is premature to conclude that we must sacrifice precision for meaning. Recent research suggests a survey question that appears to deliver both meaningful and precise responses. Comerford (2019) compares the subjective probabilities respondents returned when asked for a percentage chance against those returned when asked for a natural frequency (e.g., out of every 100, how many…?). Relative to the percentage chance responses, the probabilities returned by the natural frequency question showed greater validity. They were more coherent (i.e., less vulnerable to a framing effect) and they corresponded better with objective data (i.e., the survival probabilities reported better matched those returned in mortality tables of the target population). Perhaps density forecasts would be more meaningful if they were elicited by questions seeking natural frequencies instead of percentages. Also, the current research suggests a need for novel methods of eliciting precise and meaningful measures of inflation expectations.

4. CONCLUSIONS

These analyses demonstrate that a simple directional question asked in the SCE delivers data that offer a more accurate snapshot of the population's inflation expectations than the density forecast procedure. This is because the verbal procedure does not suffer nonresponse bias and because is less vulnerable to reporting bias. These results suggest scope for new and more meaningful measures of expectations. The current research reinforces de Bruin and Fischhoff's (2017) conclusion that knowledge of measurement bias offers low-hanging fruit when it comes to improving forecast accuracy.

LITERATURE CITED

Armantier, Olivier, Scott Nelson, Giorgio Topa, Wilbert Van der Klaauw, and Basit Zafar. (2016) "The Price is Right: Updating Inflation Expectations in a Randomized Price Information Experiment." *Review of Economics and Statistics*, 98, 503–23.

Armantier, Olivier, Giorgio Topa, Wilbert Van der Klaauw, and Basit Zafar. (2017) "An Overview of the Survey of Consumer Expectations." *Economic Policy Review*, 23, 51–72.

Binder, Carola C. (2017) "Measuring Uncertainty Based on Rounding: New Method and Application to Inflation Expectations." *Journal of Monetary Economics*, 90, 1–12.

Capistrán, Carlos, and Allan Timmermann. (2009) "Disagreement and Biases in Inflation Expectations." *Journal of Money, Credit, and Banking*, 41, 365–96.

Carroll, Christopher D. (2003) "Macroeconomic Expectations of Households and Professional Forecasters." *Quarterly Journal of Economics*, 118, 269–98.

Clements, Michael P. (2010) "Explanations of the Inconsistencies in Survey Respondents' Forecasts." *European Economic Review*, 54, 536–49.

Clements, Michael P. (2018) "Are Macroeconomic Density Forecasts Informative?" *International Journal of Forecasting*, 34, 181–98.

Coibion, Olivier, Yuriy Gorodnichenko, and Rupal Kamdar. (2018) "The Formation of Expectations, Inflation, and the Phillips Curve." *Journal of Economic Literature*, 56, 1447–91.

Comerford, David A. (2021) "Apparent Age and Gender Differences in Survival Optimism: To What Extent are They a Bias in the Translation of Beliefs Onto a Percentage Scale?" *Judgment & Decision Making*, 16, 1072–1096.

Comerford, David A. (2019) "Asking for Frequencies Rather than Percentages Increases the Validity of Subjective Probability Measures: Evidence from Subjective Life Expectancy." *Economics Letters*, 180, 33–35.

Comerford, David A., and Jenny Robinson. (2017) "Die-by Framing Both Lengthens and Shortens Life: Further Evidence on Constructed Beliefs in Life Expectancy." *Journal of Behavioral Decision Making*, 30, 1104–12.

Dana, Jason, Pavel Atanasov, Philip Tetlock, and Barbara Mellers. (2019) "Are Markets More Accurate than Polls? The Surprising Informational Value of "Just Asking."" *Judgment and Decision Making*, 14, 135–47.

De Bruin, Wändi Bruine, Charles F. Manski, Giorgio Topa, and Wilbert Van Der Klaauw. (2011) "Measuring Consumer Uncertainty about Future Inflation." *Journal of Applied Econometrics*, 26, 454–78.

De Bruin, Wändi Bruine, Wilbert Van der Klaauw, Giorgio Topa, Julie S. Downs, Baruch Fischhoff, and Olivier Armantier. (2012) "The Effect of Question Wording on Consumers' Reported Inflation Expectations." *Journal of Economic Psychology*, 33, 749–57.

de Bruin, Wändi Bruine, Wilbert van der Klaauw, Maarten van Rooij, Federica Teppa, and Klaas de Vos. (2017) "Measuring Expectations of Inflation: Effects of Survey Mode, Wording, and Opportunities to Revise." *Journal of Economic Psychology*, 59, 45–58.

de Bruin, Wandi Bruine, and Baruch Fischhoff. (2017) "Eliciting Probabilistic Expectations: Collaborations between Psychologists and Economists." *Proceedings of the National Academy of Sciences*, 114, 3297–304.

Dominitz, Jeff, and Charles F. Manski. (2004) "How Should We Measure Consumer Confidence?" *Journal of Economic Perspectives*, 18, 51–66.

Dräger, Lena, and Michael J. Lamla. (2012) "Updating Inflation Expectations: Evidence from Micro-Data." *Economics Letters*, 117, 807–10.

Drerup, Tilman, Benjamin Enke, and Hans-Martin Von Gaudecker. (2017) "The Precision of Subjective Data and the Explanatory Power of Economic Models." *Journal of Econometrics*, 200, 378–89.

Forder, James. (2010) "The Historical Place of the 'Friedman—Phelps' Expectations Critique." *European Journal of the History of Economic Thought*, 17, 493–511.

Hiban, Yousuf. (2013) "Thomson Reuters Ends Early Access to Key Market Data." https://money.cnn.com/2013/07/08/investing/thomson-reuters-consumer-sentiment/index.html. Accessed November 25, 2021.

Kahan, Dan M., Ellen Peters, Maggie Wittlin, Paul Slovic, Lisa Larrimore Ouellette, Donald Braman, and Gregory Mandel. (2012) "The Polarizing Impact of Science Literacy and Numeracy on Perceived Climate Change Risks." *Nature Climate Change*, 2, 732–5.

Kim, GwangMin, and Carola Binder. (2020) "*Learning-Through-Survey in Inflation Expectations*." Available at SSRN 3790834.

Krugman, Paul. (1998) "Japan's Trap." MIT.

Ludvigson, Sydney C. (2004) "Consumer Confidence and Consumer Spending." *Journal of Economic Perspectives*, 18, 29–50.

Madeira, Carlos, and Basit Zafar. (2015) "Heterogeneous Inflation Expectations and Learning." *Journal of Money, Credit and Banking*, 47, 867–96.

Malmendier, Ulrike, and Stefan Nagel. (2016) "Learning from Inflation Experiences." *Quarterly Journal of Economics*, 131, 53–87.

Mankiw, N. Gregory, Ricardo Reis, and Justin Wolfers. (2003) "Disagreement about Inflation Expectations." *NBER Macroeconomics Annual*, 18, 209–48.

Manski, Charles F. (2004) "Measuring Expectations." *Econometrica*, 72, 1329–76.

Manski, Charles F., and Francesca Molinari. (2010) "Rounding Probabilistic Expectations in Surveys." *Journal of Business & Economic Statistics*, 28, 219–31.

Ryngaert, Jane. (2021) "Balance of Risks and the Anchoring of Consumer Inflation Expectations." Accessed at 10.10.2020.

Rudd, Jeremy B. (2022) "Why Do We Think that Inflation Expectations Matter for Inflation? (And Should We?)." *Review of Keynesian Economics*, 10, 25–45.

SCE. (2021) "Consumers' Inflation Expectations Remain Elevated for the Short- and Medium-Term." Press release. https://www.newyorkfed.org/newsevents/news/research/2021/20211012. Accessed October 12, 2021.

SCE. (2020) © 2013–2020 Federal Reserve Bank of New York (FRBNY). http://www.newyorkfed.org/microeconomics/sce. Accessed 10.10.2020.

Sinayev, Aleksandr, and Ellen Peters. (2015) “Cognitive Reflection vs. Calculation in Decision Making.” *Frontiers in Psychology*, 6, 532–8.

Rudd, Jeremy B. (2022) “Why do we think that inflation expectations matter for inflation?(And should we?).” *Review of Keynesian Economics*, 10, 25–45.

Hurd, Michael D. (2009) “Subjective probabilities in household surveys.” *Annual Review Economics*, 1, 543–562.

Fischhoff, Baruch and Wändi, Bruine De Bruin. (1999) “Fifty.fifty= 50%?” *Journal of Behavioral Decision Making*, 12, 149–163.

Kim, GwangMin and Carola, Binder. (2020) “Fifty–fifty=50%?” *Journal of Behavioral Decision Making*, 12, 149–163.