RATIONALITY & COMPETITION
CRC TRR 190

Economic Uncertainty and Subjective Inflation Expectations

Tobias Rossmann (LMU Munich)

Discussion Paper No. 160

June 25, 2019

Collaborative Research Center Transregio 190 | www.rationality-and-competition.de
Ludwig-Maximilians-Universität München | Humboldt-Universität zu Berlin
Spokesperson: Prof. Dr. Klaus M. Schmidt, University of Munich, 80539 Munich, Germany
+49 (89) 2180 3405 | info@rationality-and-competition.de

Economic uncertainty and subjective inflation expectations*

Tobias Rossmann
University of Munich

June 2019

Abstract: Measuring economic uncertainty is crucial for understanding investment decisions by individuals and firms. Macroeconomists increasingly rely on survey data on subjective expectations. An innovative approach to measure aggregate uncertainty exploits the rounding patterns in individuals' responses to survey questions on inflation expectations (Binder, 2017). This paper uses the panel dimension of household surveys to study individual-level heterogeneity in this measure of individual uncertainty. The results provide evidence for the existence of considerable heterogeneity in individuals' response behavior and inflation expectations.

JEL codes: C10, D80, D83, D84, E31

Keywords: Uncertainty, inflation, expectations, mixture models

* I would like to thank Davide Cantoni, Helmut Farbmacher, Peter Haan, Florian Heiß and Joachim Winter as well as seminar participants at the University of Munich and the University of Southern California for helpful comments. Financial support by the Deutsche Forschungsgemeinschaft through CRC TRR 190 and through the Elite Network of Bavaria is gratefully acknowledged.

1 Introduction

Inflation expectations of individuals are crucial for understanding the economy and economic policies. Individuals' expectations are directly linked to their decision-making regarding investments, savings, retirement planning and wage negotiations. Since these decisions again directly translate into real economy transactions, modern monetary policy relies to a large extent on individuals' inflation expectations (Sims, 2009; Galí, 2015). In fact, expected or perceived inflation is often thought to be more important for monetary policy than the actual, measured inflation rate (Bernanke, 2007; Blanchard et al., 2010).

When analyzing individual data on inflation expectations, macroeconomic studies usually focus on their predictive power for actual inflation and on interpersonal heterogeneity (see, amongst others, Souleles, 2004; Blanchflower and MacCoille, 2009; Hobijn et al., 2009). These studies usually take survey answers at face value, neglecting that responses may suffer from several reporting issues, such as rounding, measurement error and non-response. However, as shown in Kleinjans and van Soest (2014), these reporting issues may not only reduce data quality, but also lead to biases in the estimates induced by selection effects. Microeconomic studies have, in contrast, a longer tradition of rigorously modeling these reporting issues. For example, when analyzing (probabilistic) stock market expectations of private households, it is common to explicitly model measurement error, rounding behavior or both.[1] In a recent contribution, Manski (2018) gives an overview of how macroeconomics can benefit from microeconomic insights, when working with subjective expectations, and encourages interactions between both fields.

In this paper, I follow the call by Manski (2018) and propose a microeconometric panel data model for inflation point (rather than probabilistic) expectations of individuals, explicitly accounting for item nonresponse and rounding behavior.[2] Specifically, I generalize

[1] See, for example, Hudomiet et al. (2011), Ameriks et al. (2018) or Heiss et al. (2019).
[2] In this paper, I abstract from measurement error other than rounding.

a model by Binder (2017) who suggests that the population can be described by a mixture of two different response types. When asked about the year-ahead inflation expectations, type NR (non-rounder) reports her true expectation, while Type RD (rounder) rounds her answer to a multiple of five percent. Binder (2017) estimates monthly RD type shares in the US between 1978 and 2014 and shows that they can serve as measure of economic uncertainty. This paper builds on her model and extends it in several dimensions. First, I introduce a third response type DK for respondents, who choose a "don't know" option, when asked about their inflation expectations. Second, I add a panel dimension to the econometric model and estimate the uncertainty index by month-year fixed effects in the model for the type probabilities, rather than by hundreds of separate estimations. Third, I allow the type probabilities to depend on both observed and unobserved heterogeneity, rather than treating them as constant scalars. I therefore contribute to the literature by providing a rich, but tractable panel data model for inflation expectations, which – in contrast to previous studies, in particular Binder (2017) – allows for an additional panel dimension, individual-specific heterogeneity and item nonresponse.

I apply the model to monthly data from the Michigan Survey of Consumers (MSC) between 1978 and 2017. Assuming type RD rounds to the next multiple of five percent, the estimated population shares of types (NR,RD,DK) are (0.65,0.28,0.07). This implies that most respondents report their true inflation expectation, while only few choose a "don't know" response. The model also identifies considerable heterogeneity in individuals' type probabilities. For example, males and respondents with at least a college degree are significantly less likely to round or to choose a "don't know" option than females and respondents without a college degree. I also find evidence for the importance in accounting for unobserved factors. The unobserved, individual-specific (random) effects for types RD and DK are positively correlated, implying that respondents who are more likely to round are, in general, also more likely to choose a "don't know" option. This also suggests that discarding non-respondents – as often done in the literature and also in Binder (2017) – is invalid, because it would only be allowed if the individual effects were uncorrelated. In addition, my model identifies considerable heterogeneity across individuals' inflation expectations, confirming previous findings from the literature.

At the individual level, I find evidence for the persistence of response types over time. If respondents are interviewed twice, the probability of being a specific response type in the first interview is positively correlated with the probability of being the same type in the second interview and negatively correlated with the probability of being another type. Furthermore, model-implied posterior type probabilities, i.e. type probabilities conditional on the reported inflation expectation, suggest that roughly every second respondent who reports an inflation expectation of zero or five percent is rounding. Almost all respondents who report more extreme multiples of five, such as 25 or minus ten percent, are predicted to round.

I then follow the insight in Binder (2017) and construct a macroeconomic uncertainty index, which is given by the monthly share of rounders (RD) and respondents choosing the "don't know" option (DK). The resulting uncertainty index spikes during periods of arguably high uncertainty, such as the financial crisis, 9/11 or the Gulf War. However, the index is almost identical to the uncertainty index by Binder (2017). Even though it is more strongly correlated with alternative state-of-the-art uncertainty measures, the advantages of the generalized model therefore vanish – at least in terms of measuring macroeconomic uncertainty.

This paper is related to three different strands of the literature. First, several studies focus on heterogeneity of inflation expectations across individuals. Most prominently, females are found to systematically report higher inflation expectations than males. This is often explained by an argument of Jonung (1981), suggesting that females are on average more exposed to food prices than males and therefore more able to predict price changes. However, this view is challenged by Bryan and Venkatu (2001a,b), showing that gender differences can also be found between single females and single males as well as during periods where food prices actually increased less than prices for other goods. More generally, systematic differences in inflation expectations between different socio-economic subgroups of the population are often related to different consumption patterns, even though this is known not to be enough to explain all the variation (see for example, Ranyard et al., 2008; Hobijn et al., 2009; Georganas et al., 2014). Malmendier and Nagel (2016) show that experienced inflation rates during a respondent's lifetime are also strong predictors for inflation expectations. Indeed, research has shown that personal inflation experiences of members of the Federal Open Market Committee (FOMC) can be used to predict their voting behavior and consequently the federal funds target rate (Malmendier et al., 2017).

Second, the paper is related to several microeconometric papers focusing on measurement and modeling of probabilistic (rather than point) expectations. Comprehensive overviews are given by Manski (2004) and Hurd (2009). Kleinjans and van Soest (2014) show that expectations in various domains in the Health and Retirement Study (HRS) are subject to rounding, nonresponse and focal values and discuss potential implications. Heiss et al. (2019) elicit individual distributions of stock market expectations, analyze how individuals differ in using past stock market returns, when forming their expectations, and explicitly model rounding behavior. Drerup et al. (2017) argue that subjective stock market expectations might only be meaningful if they are precise. Expectations with low precision may indicate that individuals base their decisions not on expectations, but rather on heuristics or rules of thumb.

A third strand of the literature is concentrated on measuring general economic uncertainty. Traditional measures are given by the realized (or implied) volatility of stock market returns, the ex-ante cross-sectional dispersion of subjective forecasts by households or professional forecasters – often referred to as "disagreement" – and the ex-post cross-sectional dispersion of stock returns, productivity and forecast errors (see, amongst others, Bloom, 2009; Bachmann et al., 2013; Rossi and Sekhposyan, 2015; Rossi et al., 2017). In a recent contribution, Baker et al. (2016) show that using newspaper coverage frequencies of specific combinations of terms, such as “uncertainty”, “economic” and “deficit”, can also be used to construct a measure of economic uncertainty. Jurado et al. (2015) propose another measure which is based on whether the economy has become more or less predictable by focusing on the volatility of expected forecast errors. As mentioned earlier, Binder (2017) introduces an uncertainty measure which is based on rounding patterns in inflation expectations of US households.

The remainder of this paper is organized as follows. I first describe the data and present basic descriptive statistics in Section 2. The econometric model is introduced in Section 3. Section 4 applies the model to data from the Michigan Survey of Consumers and presents the results, while several robustness analyses are discussed in Section 5. Section 6 concludes.

2 Data

For information on subjective inflation expectations and socio-economic characteristics, I draw on data from the Michigan Survey of Consumers (MSC).[3] Starting in 1978, this nationally representative, monthly survey asks roughly 500 respondents on a variety of topics, including personal finances, unemployment, confidence in government and economic policies, personal attitudes and expectations.[4] Most prominently, answers to some of these questions are used to construct the University of Michigan Consumer Sentiment Index, one of the leading US indicators for consumer confidence.

In every month, respondents can be divided into three different groups. One third are

[3] After registration, the data is freely available at: https://data.sca.isr.umich.edu/ [accessed August 10, 2018].

[4] American households from Alaska and Hawaii are not included in the sample. Note also that some questionnaire items from the MSC date back to the late 1940s, when surveys were conducted on a yearly or quarterly basis. The systematic rotating panel design was incorporated in January 1978, which is also the earliest date available at the University of Michigan Survey Research Center. For more details on the survey and its design see Curtin (1982).

new respondents who will be interviewed again in six months, while another third are new respondents who will not be contacted again. The last third consists of re-interviews of respondents who were already interviewed six months before. A substantial share of respondents is therefore interviewed twice, adding a panel dimension to the data, which will later be exploited by the econometric model. Focusing on the entire universe of interviews between January 1978 and December 2017, the data set contains 97,159 individuals who are interviewed twice and 77,630 individuals who are interviewed once, making a total of 271,948 observations. To reduce the computational burden, the main analysis concentrates on respondents who are interviewed twice, but the results are shown to be robust to including respondents with only one interview.[5]

As the focus of this paper lies on subjective inflation expectations, the following question from the MSC is of particular interest:

> **Q1:** *"During the next 12 months, do you think that prices in general will go up, or go down, or stay where they are now?"*

Respondents are asked about "prices in general" rather than "inflation" directly, the main reason being that researchers are afraid that "ordinary persons may not understand the professional economic use of the term [inflation]" (Manski, 2018, p.441). However, as discussed in Armantier et al. (2013), asking about "prices in general" might be problematic too, because respondents could interpret the term heterogeneously. Indeed, the authors find that some respondents focus on prices which they recently paid themselves rather than on actual inflation. While I assume that respondents think about actual inflation, this distinction becomes less important to the extent that the individual-specific (random) effects in the panel data model capture these interpersonal differences.[6]

[5] For details, see Section 5 and Appendices B and F.

[6] An in-depth analysis of the effect of the exact question wording on inflation expectations can also be found in Bruine de Bruin et al. (2010).

If respondents' answer to question **Q1** is "stay where they are", their answer is coded as zero. If respondents choose "go up" or "go down", they are asked another question:

> **Q2:** *By about what percent do you expect prices to go (up/down) on the average, during the next 12 months?*

Respondents are allowed to report any integer response. The answers to both questions are combined into an integer variable **px1**, measuring the subjective expected inflation rate in the year ahead. Note that both questions allow respondents to choose a "don't know" (DK) option, i.e. respondents are not forced to answer the questions if they cannot or do not want to.[7] Figure 1 shows the response distribution of individuals' inflation expectations (px1) in the year 2009. Overall, responses vary between −25 and 25 percent, with most respondents expecting a positive inflation rate. One quarter of respondents report an expected inflation rate of zero percent, i.e. no change in prices, and more than ten percent of respondents do not answer the questions at all (DK). Clearly, there is evidence for substantial heaping at multiples of five and ten percent, and focal values of zero, two or three percent. While these response patterns are often taken into account by microeconometric studies, they are usually neglected in macroeconomic studies that take responses at face value. However, as shown in Binder (2017), rounding is systematically related to economic uncertainty, indicating that temporal variation in these response patterns contains additional information by itself.

In addition, the questionnaire also includes questions about the average, yearly inflation rate over the next five years (px5). The elicitation and construction of these medium-run inflation expectations is almost identical to the procedure for short-run expectations and is presented in Appendix A. However, questions for px5 have not been asked in all months and years, leading to several month-year combinations where data is missing. The main analysis therefore focuses on short-run inflation expectations (px1), while the results for medium-run inflation expectations (px5) are reported as robustness check in Section 5.

[7] There are also additional questions after a "don't know" response or an extraordinarily high inflation rate to ensure respondents' understanding of the question. The exact procedure is given by the interviewer instructions summarized in Appendix A.

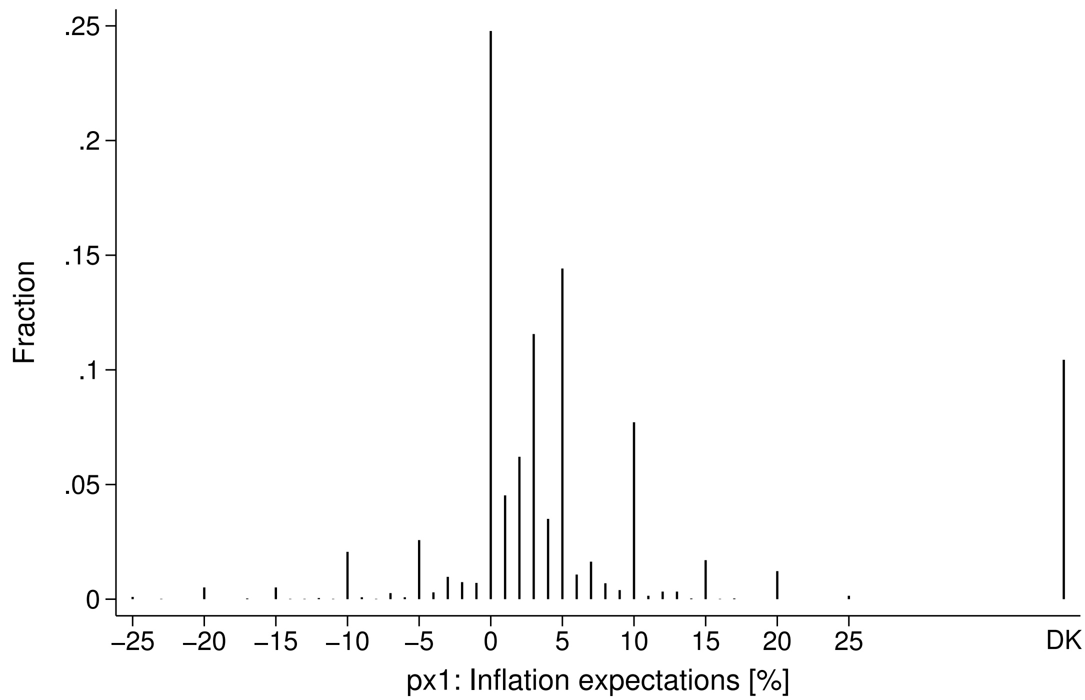


Figure 1: Response distribution of one-year inflation expectations (px1) in 2009

Table 1 reports standard summary statistics for the main sample, which consists of all respondents, who are interviewed twice between January 1978 and December 2017 (always with a six-month interval in between).[8] Panel A focuses on respondents' inflation expectations. On average, respondents expect an inflation rate of 3.92 percent for the year ahead and a slightly higher, yearly inflation rate of 3.97 percent over the next five years. The standard deviations are with 5.51 and 4.92 relatively large, hinting at substantial disagreement between respondents.

Panel B displays summary statistics regarding several binary socio-demographic characteristics. Overall, the sample contains slightly more females than males. One in five respondents is 65 or older; one in three respondents is younger than 40. 61 percent of respondents report to be living with a partner, and 40 percent to hold at least a college

[8] Table B1 in Appendix B reports summary statistics for the full sample, adding respondents who are interviewed only once. The results are very similar.

degree. Starting in October 1979, respondents are also asked about their total income (all sources including job) from the previous year. In every given month-year combination, this information is used to classify respondents into income quartiles, which are also presented in Panel B. Panel C reports coarse information on the region of residence at the time of the interview.[9]

Table 1: Summary statistics for the main sample

	Mean	SD	p5	p95	Min	Max	Observations
A: Inflation expectations [%]							
Short-run (px1)	3.92	5.51	0	11	-50	50	179,483
Medium-run (px5)	3.97	4.92	0	10	-50	50	139,897
B: Sociodemographics [0/1]							
Male	0.46	0.50	0	1	0	1	194,065
Partner	0.61	0.49	0	1	0	1	193,224
Age > 64	0.21	0.41	0	1	0	1	193,379
Age < 40	0.36	0.48	0	1	0	1	193,379
College	0.40	0.49	0	1	0	1	193,174
1st income quartile	0.20	0.40	0	1	0	1	183,104
2nd income quartile	0.21	0.41	0	1	0	1	183,104
3rd income quartile	0.28	0.45	0	1	0	1	183,104
4th income quartile	0.30	0.46	0	1	0	1	183,104
C: Regional information [0/1]							
West	0.20	0.40	0	1	0	1	194,269
Northcentral	0.27	0.45	0	1	0	1	194,269
Northeast	0.19	0.39	0	1	0	1	194,269
South	0.33	0.47	0	1	0	1	194,269

Notes: This table is based on all 97,159 respondents from the MSC who are interviewed twice within a six-month interval between January 1978 and December 2017, making a total of 194,318 observations. Number of observations differs due to item nonresponse. Panel B and C report dummy variables. Information on income (1st-4th quartile) not available before October 1979. For details see text.

[9] US states are classified into the four statistical regions "West", "Northcentral", "Northeast" and "South", as defined by the United States Census Bureau.

Lastly, I draw on data on the official US inflation rate from the Organisation for Economic Co-operation and Development (OECD).[10] More specifically, I use monthly inflation rates between January 1956 and June 2018, measured by the annual growth rate of the US Consumer Price Index (CPI). Its time series is presented in Appendix C.

3 Econometric model

In the following paragraphs, I introduce the econometric panel data model used in this paper and discuss the main differences to the model by Binder (2017).

3.1 Panel data model and likelihood function

Assume that the population can be described by three distinct response types, who differ in how they report their true inflation point expectation. Type NR (non-rounder) always reports her true inflation expectation. In contrast, type RD (rounder) always rounds to the next multiple of m, say for example five percent. Type DK chooses the "don't know" option, when asked about her one-year inflation expectation. Note that individuals' responses partially identify individuals' response types. For example, both reporting non-multiples of m and not answering at all uniquely classifies respondents as type NR and type DK, respectively. However, every individual reporting a multiple of m is always consistent with both type NR and type RD (but not with type DK).

I assume that the true inflation expectations of all individuals approximately follow a normal distribution, whose two parameters (mean μ and variance σ^2) are allowed to differ across types NR and RD:[11]

$$y_{it}^* \sim N(\mu^R, (\sigma^R)^2), \qquad R \in \{NR, RD\} \tag{1}$$

[10] The data is freely available at https://data.oecd.org/price/inflation-cpi.html [accessed August 10, 2018].

[11] In a robustness analysis, I later relax the normality assumption and report estimates under the assumption that inflation expectations follow alternative distributions (cf. Table F1).

where y^*_{it} is the true (and partially) unobserved inflation expectation of individual i in period t. The main difference between type NR and type RD is given by the mapping from the reported inflation expectation y_{it} to the true inflation expectation y^*_{it}. Abstracting from measurement error other than rounding, type NR reports *by definition* her true expectation, i.e. $y^*_{it} = y_{it}$. In contrast, type RD always reports a rounded value (to the next multiple of m), implying that we can only identify a symmetric interval for the true inflation expectation, i.e. $y^*_{it} \in [y_{it} - \frac{m}{2}; y_{it} + \frac{m}{2}]$. Using this insight as well as the fact that the reported values of y_{it} partially identify individuals' response types, the probabilities of observing y_{it} conditional on response type T_{it} are then given by:

$$
\begin{aligned}
P(y_{it}|T_{it}) &= \begin{cases} f_{NR}(y_{it}) & \text{if } T_{it} = NR \\ f_{RD}(y_{it}) & \text{if } T_{it} = RD \\ 0 & \text{if } T_{it} = DK \end{cases} && \text{\& } y_{it} \text{ is a multiple of } m \\
P(y_{it}|T_{it}) &= \begin{cases} f_{NR}(y_{it}) & \text{if } T_{it} = NR \\ 0 & \text{if } T_{it} = RD \\ 0 & \text{if } T_{it} = DK \end{cases} && \text{\& } y_{it} \text{ is not a multiple of } m \qquad (2) \\
P(y_{it}|T_{it}) &= \begin{cases} 0 & \text{if } T_{it} = NR \\ 0 & \text{if } T_{it} = RD \\ f_{DK}(y_{it}) & \text{if } T_{it} = DK \end{cases} && \text{\& } y_{it} \text{ is missing}
\end{aligned}
$$

with

$$
\begin{aligned}
f_{NR}(y_{it}) &= \phi(y_{it}; \mu^{NR}; \sigma^{NR}) \\
f_{RD}(y_{it}) &= \Phi\left(\frac{y_{it} + \frac{m}{2} - \mu^{RD}}{\sigma^{RD}}\right) - \Phi\left(\frac{y_{it} - \frac{m}{2} - \mu^{RD}}{\sigma^{RD}}\right) \qquad (3) \\
f_{DK}(y_{it}) &= 1
\end{aligned}
$$

where $\phi(\cdot)$ denotes the probability density function (p.d.f.) of the standard normal distribution and $\Phi(\cdot)$ denotes the standard normal cumulative distribution function (c.d.f.). Equation 2 illustrates the partial identification of reporting types. If an individual reports a missing value or an inflation expectation which is not a multiple of m, her type is uniquely identified as type DK or NR, respectively. Reporting a multiple of m, however, is consistent with two types, namely NR and RD. As shown in Equation 3, the p.d.f. for type NR, $f_{NR}(\cdot)$, is given by a Gaussian density function with mean μ^{NR} and standard deviation σ^{NR}, whereas the p.d.f. for type RD, $f_{RD}(\cdot)$, is given by the difference between two normal c.d.f.s with mean μ^{RD} and standard deviation σ^{RD} evaluated at $y_{it} \pm \frac{m}{2}$, respectively. For completeness, the p.d.f. for type DK, $f_{DK}(\cdot)$, is equal to one and therefore independent of any parameters.

In addition, the model allows the (type-specific) mean of the inflation expectation distribution to vary across individuals and time by using the following linear parameterization:

$$\mu_{it}^{R} = \boldsymbol{w}_{it}\boldsymbol{\beta}^{R}, \qquad R \in \{NR, RD\} \tag{4}$$

where $\boldsymbol{w}_{it}$ is a vector of potentially time-varying covariates of respondent i in period t. This formulation allows to capture systematic differences in inflation expectations between individuals, as often found in the literature.

I model response type probabilities in a standard random effects multinomial logit model with three outcomes:

$$\begin{aligned} u_{it}^{j} &= \boldsymbol{x}_{it}\boldsymbol{\beta}^{j} + \alpha_i^{j} + \varepsilon_{it}^{j}, && j = 1,2,3 \\ T_{it} &= j \text{ if } u_{it}^{j} \geq u_{it}^{k}, && k = 1,2,3 \\ P(\varepsilon_{it}^{j} \leq z) &= \exp(-\exp(-z)) && \text{(standard Gumbel)} \end{aligned} \tag{5}$$

where $\boldsymbol{x}_{it}$ is a vector of covariates of respondent i in period t, potentially including period fixed effects. α_i^{j} is an unobserved respondent-specific effect for type j and ε_{it}^{j} denotes an i.i.d. standard Gumbel error term. Without loss of generality, type NR is taken as benchmark outcome $T_{it} = 1$, leading to the standard normalizations $\boldsymbol{\beta}^1 = 0$ and $\alpha_i^1 = 0$. The other outcomes are type RD ($T_{it} = 2$) and type DK ($T_{it} = 3$). The response type probabilities conditional on the observed covariates $\boldsymbol{x}_{it}$ and the unobserved effects α_i^2 and α_i^3 can then be derived from the distributional assumptions on the error term ε_{it}^j and are given by:

$$P(T_{it} = j|\boldsymbol{x}_{it}, \alpha_i^2, \alpha_i^3) = \frac{\exp(\alpha_i^j + \boldsymbol{x}_{it}\boldsymbol{\beta}^j)}{\sum\limits_{k=1}^{3} \exp(\alpha_i^k + \boldsymbol{x}_{it}\boldsymbol{\beta}^k)}; \qquad j = 1, 2, 3 \tag{6}$$

In addition to the previous assumptions, I impose the following assumption on the vector of unobserved heterogeneity $\boldsymbol{\alpha}$:

$$\boldsymbol{\alpha} = \begin{pmatrix} \alpha_i^2 & \alpha_i^3 \end{pmatrix} = \begin{pmatrix} \alpha_i^{RD} & \alpha_i^{DK} \end{pmatrix} \sim N(\mathbf{0}, \boldsymbol{\Sigma}) \tag{7}$$

Equation 7 implies that the individual (random) effects are i.i.d. jointly normal with mean zero and arbitrary variance-covariance matrix $\boldsymbol{\Sigma}$ and independent of x_{it} and ε_{is}^j for $j = 1, 2, 3$ and $s = 1, ..., T$. Note that both rounding and not answering at all can be seen as indicators for individual uncertainty. I therefore expect a positive correlation between α_i^{RD} and α_i^{DK}, indicating that individuals who do not answer at all are also more likely to round, unlike in a standard multinomial logit model.[12]

Under these assumptions the likelihood function conditional on the unobserved individual effects α_i^{RD} and α_i^{DK} can be written as:

$$L^c(\alpha_i^{RD}, \alpha_i^{DK}) = \prod_{i=1}^{N} \prod_{t=1}^{T} L_{it}^c(\alpha_i^{RD}, \alpha_i^{DK}) \tag{8}$$

[12] As shown by Revelt and Train (1998), adding unobserved heterogeneity to a multinomial logit model breaks the independence of irrelevant alternatives (IAA) assumption. See also Kleinjans and van Soest (2014) for a similar application.

with

$$
\begin{aligned}
L^c_{it}(\alpha^{RD}_i, \alpha^{DK}_i) &= P\left(T_{it} = NR|\boldsymbol{x}_{it}, \alpha^{RD}_i, \alpha^{DK}_i\right) f_{NR}(y_{it}|\boldsymbol{w}_{it}) + \\
&+ P\left(T_{it} = RD|\boldsymbol{x}_{it}, \alpha^{RD}_i, \alpha^{DK}_i\right) f_{RD}(y_{it}|\boldsymbol{w}_{it}) \qquad \text{if } y_{it} \text{ is a multiple of } m \\
L^c_{it}(\alpha^{RD}_i, \alpha^{DK}_i) &= P\left(T_{it} = NR|\boldsymbol{x}_{it}, \alpha^{RD}_i, \alpha^{DK}_i\right) f_{NR}(y_{it}|\boldsymbol{w}_{it}) \qquad \text{if } y_{it} \text{ is not a multiple of } m \\
L^c_{it}(\alpha^{RD}_i, \alpha^{DK}_i) &= P\left(T_{it} = DK|\boldsymbol{x}_{it}, \alpha^{RD}_i, \alpha^{DK}_i\right) f_{DK}(y_{it}|\boldsymbol{w}_{it}) \qquad \text{if } y_{it} \text{ is missing}
\end{aligned}
$$

where $f_{NR}(\cdot)$, $f_{RD}(\cdot)$ and $f_{DK}(\cdot)$ are given by Equation 3 and the conditional type probabilities $P(T_{it}|\cdot)$ by Equation 6. Again this conditional likelihood function illustrates the partial identification of response types, as already discussed before.

The unconditional likelihood function can be derived by integrating out the individual effects:[13]

$$
L = \prod_{i=1}^{N} \int_{\mathbb{R}^2} \prod_{t=1}^{T} L^c_{it}(\alpha^{RD}_i, \alpha^{DK}_i) f(\boldsymbol{\alpha}) d\boldsymbol{\alpha}. \tag{9}
$$

To avoid numerical integration in multiple dimensions, I use Maximum Simulated Likelihood (MSL) and replace the integral by a simulated mean. The simulated sample likelihood (SL) is then given by

$$
SL = \prod_{i=1}^{N} \frac{1}{Q} \sum_{q=1}^{Q} \prod_{t=1}^{T} L^c_{it}(\alpha^{RD}_{iq}, \alpha^{DK}_{iq}) \tag{10}
$$

where $\alpha^{RD}_{iq}, \alpha^{DK}_{iq}$ are simulated random effects for a given draw q. Applying a Cholesky decomposition of the variance-covariance matrix $\mathbf{\Sigma}$, yields a positive semi-definite lower diagonal matrix $\mathbf{L}$ such that $\mathbf{\Sigma} = \mathbf{LL}'$, with the elements of $\mathbf{L}$ to be estimated. For a given draw q, the unobserved heterogeneity is then calculated by $\boldsymbol{\alpha} = \mathbf{L}\boldsymbol{\tau}$, where $\boldsymbol{\tau}$ con-

[13] Appendix D discusses the derivation of the likelihood function in greater detail. Note also that the likelihood function is written for a respondent who participates in every period. If a respondent did not participate in one particular period, her likelihood contribution for this period (L^c_{it}) can be replaced by one.

tains simulated vectors of the independent standard normal distribution. As suggested by Train (2003), I use draws from Halton sequences to obtain the independent standard normal variables $\boldsymbol{\tau}$ to reduce the variance induced by the simulation.

Note that after solving the maximization problem the estimated parameter vector can be used to predict individual-specific (i) unconditional type probabilities as well as (ii) posterior type probabilities, i.e. response type probabilities conditional on the reported value of y_{it}. The calculation of (i) is based on Equation 6 with the true parameter vectors $\boldsymbol{\beta}^j$ being replaced by their respective estimates $\hat{\boldsymbol{\beta}}^j$ and the individual effects being integrated out by simulation or quadrature methods. Specifically, I use 151 draws from Halton sequences and simulate the normal individual effects with mean zero and a variance-covariance matrix which is given by the estimates of $\hat{\boldsymbol{\Sigma}}$. (ii) can be calculated using Bayes' theorem. Not surprisingly, the posterior probability of being type DK is one if y_{it} is missing. Similarly, if the respondent does not report a multiple of m, her posterior probability of being type NR is one. In contrast, if she reports a multiple of m, the probability of being type NR and RD, respectively, is strictly between zero and one and can be calculated applying Bayes' theorem. See Appendix D for derivations, formulas and further details.

3.2 Comparison to Binder (2017)

This econometric panel data model is essentially a generalization of the model by Binder (2017) and nests it as a special case. First, Binder (2017) models the population as a mixture of two response types only (rounders and non-rounders) and drops respondents with missing information on inflation expectations.[14] This is equivalent to restricting the unconditional probability for type DK to zero in my model. Second, she does not allow for either observed or unobserved interpersonal heterogeneity in the type probabilities, which corresponds to restricting all coefficients other than the constants in the random effects

[14]She does so for the estimation of her empirical model. For the construction of her uncertainty index, the DK share is added ad-hoc after the estimation.

multinomial logit model (Equation 6) to zero. Third and most importantly, she ignores the panel dimension of the data and rather estimates cross-sectional models, separately for every month between January 1978 and July 2014. This is equivalent to restricting the variances and covariances of the individual effects to zero, i.e. $\mathbf{\Sigma} = \mathbf{0}$, and applying the restricted model to each month separately. An important difference between both models is therefore that temporal variation in the unconditional rounding probabilities – which will later be used to construct the macroeconomic uncertainty index – comes from hundreds of separate estimations in Binder (2017) and from the month-year fixed effects of the joint model in this paper.

My approach offers several advantages over the model by Binder (2017). First, one can expect considerable gains in efficiency, mainly stemming from two sources. On the one hand, the model additionally uses information of respondents with missing information on inflation expectations; on the other hand, the entire model is estimated jointly for all months in the estimation sample. Second, my approach allows for the identification of interpersonal heterogeneity in the response type probabilities. Third, leveraging the existence of the panel dimension in the MSC data also allows to model unobserved heterogeneity via the inclusion of individual-specific (random) effects. By allowing for arbitrary correlations between the individual effects, the model can actually test whether or not dropping respondents with missing responses – as often done in the literature – is valid. This will only be allowed if the individual-specific effects are uncorrelated.

4 Results

4.1 Interpersonal heterogeneity

I apply the econometric model to monthly data from the Michigan Survey of Consumers between 1978 and 2017.[15] Assuming that type RD rounds her true inflation expectation to the next multiple of five percent, i.e. $m = 5$, Table 2 reports one model specification excluding and one specification including month-year fixed effects in the random effects multinomial logit model (Equation 6), respectively.[16]

Columns 1a, 1b and 2a, 2b of Table 2 report coefficients of the random effects multinomial logit model for the type probabilities (Equation 6). Recall that the baseline category is type NR (non-rounder). Interestingly, males are found to be significantly less likely to round or report a “don’t know” response than females. This finding could be driven by the fact that men are on average more financially literate than women and therefore more certain of and confident about their inflation predictions (see, for example, van Rooij et al., 2011). It could also correspond to general overconfidence of men, as often found in behavioral studies (cf. Niederle and Vesterlund, 2007). Unfortunately, the MSC does neither include a measure of financial literacy nor (over-)confidence to further analyze these patterns. Education is also significantly associated with type probabilities. Respondents holding at least a college degree are less likely type RD (rounder) or type DK (don’t know), compared to respondents without a college degree. This seems intuitive, because more educated people are arguably more likely to know the concept of inflation.

[15] The estimation sample is based on all respondents, who are interviewed twice and who have full information on all socio-economic characteristics and the exact month and year of the interview. Note that respondents with missing information on inflation expectations via choosing a “don’t know” option are explicitly allowed in the model and thus not excluded from the analysis. To make the results comparable to Binder (2017), I exclude extreme inflation expectations that are smaller than minus ten and larger than 25 percent. Including these outliers, however, yields almost identical results. These data requirements result in a total of 172,548 observations.

[16] As a robustness check, I also estimate the model for $m = 10$, i.e. type RD rounds her true inflation expectation to the next multiple of ten percent, as well as a model which includes both rounding types at the same time. Results are discussed in Section 5.

Wealthy individuals are more likely to be type NR, i.e. these individuals tend to provide exact answers, compared to less affluent respondents. Comparing specifications 1 and 2, the coefficients of the socio-economic covariates are remarkably similar. Therefore, including month-year fixed effects in the random effects multinomial logit model leaves the effects of the covariates on the type probabilities almost unchanged.[17] In summary, there is strong evidence for the fact that socio-economic characteristics predict individual type probabilities. Recall that Binder (2017) models these type probabilities as constant scalars, which would require all coefficients in columns 1a, 1b, 2a and 2b in Panel A of Table 2 other than the constants to be statistically indistinguishable from zero.

Columns 1c, 1d and 2c, 2d report estimates for the parameterized subjective mean of inflation expectations for type NR (non-rounder) and RD (rounder), respectively (Equation 4). Even though the magnitude of the coefficients slightly varies between both types, the effect of the covariates is qualitatively the same. Men report significantly lower inflation expectations than women, while less educated and less affluent respondents tend to report higher inflation expectations, independent of response type. Overall, these findings confirm findings from the previous literature (cf. Section 1).

Panel B focuses on the estimated standard deviation of the type-specific normal distribution of inflation expectations. Interestingly, rounders seem to have a more dispersed distribution of inflation expectations than non-rounders. The estimated standard deviations for both types differ, in fact, by a factor of two. This is in line with arguing that rounders perceive a higher level of uncertainty than non-rounders. It is, however, important to distinguish this estimated standard deviation from the cross-sectional standard deviation of individual beliefs, which is also often used in the literature as measure of uncertainty (see, for example, Bachmann et al., 2013).

[17] The results are also shown to be robust to including month-year fixed effects in the model of the parameterized mean of the normal inflation expectations (Equation 4). For further details, see Section 5.

Table 2: Model estimates

	Excluding month-year FE				Including month-year FE			
	P(T=RD) (1a)	P(T=DK) (1b)	Mean NR (1c)	Mean RD (1d)	P(T=RD) (2a)	P(T=DK) (2b)	Mean NR (2c)	Mean RD (2d)
Panel A								
Male	-1.00***	-1.40***	-0.12***	-1.10***	-0.93***	-1.40***	-0.12***	-1.12***
	[0.03]	[0.03]	[0.02]	[0.06]	[0.02]	[0.03]	[0.02]	[0.06]
Partner	-0.06**	-0.12***	-0.02	0.32***	-0.05*	-0.13***	-0.01	0.29***
	[0.03]	[0.03]	[0.02]	[0.06]	[0.02]	[0.03]	[0.02]	[0.06]
College	-0.66***	-0.58***	-0.12***	-0.37***	-0.55***	-0.67***	-0.13***	-0.39***
	[0.03]	[0.03]	[0.02]	[0.06]	[0.03]	[0.03]	[0.02]	[0.06]
1st income quartile	0.69***	1.76***	0.22***	1.37***	0.71***	1.71***	0.23***	1.37***
	[0.04]	[0.04]	[0.03]	[0.09]	[0.04]	[0.05]	[0.03]	[0.09]
2nd income quartile	0.26***	0.87***	0.01	1.14***	0.37***	0.81***	0.02	1.12***
	[0.03]	[0.04]	[0.03]	[0.09]	[0.03]	[0.04]	[0.03]	[0.09]
3rd income quartile	0.15***	0.33***	0.00	0.66***	0.18***	0.31***	0.00	0.67***
	[0.03]	[0.04]	[0.02]	[0.08]	[0.03]	[0.04]	[0.02]	[0.08]
West	-0.14***	-0.10***	0.06***	-0.01	-0.15***	-0.12***	0.06**	0.01
	[0.03]	[0.04]	[0.02]	[0.08]	[0.03]	[0.04]	[0.02]	[0.08]
Northcentral	-0.02	-0.21***	-0.01	-0.17**	-0.04	-0.20***	-0.01	-0.17**
	[0.03]	[0.04]	[0.02]	[0.07]	[0.03]	[0.04]	[0.02]	[0.07]
Northeast	0.16***	0.10***	0.03	-0.06	0.13***	0.10***	0.03	-0.05
	[0.03]	[0.04]	[0.02]	[0.08]	[0.03]	[0.04]	[0.02]	[0.08]
Constant	-0.70***	-2.92***	3.16***	4.84***	1.19***	-0.51*	3.17***	4.83***
	[0.04]	[0.05]	[0.03]	[0.09]	[0.24]	[0.31]	[0.03]	[0.09]
Panel B								
σ^{NR}	2.82***				2.81***			
	[0.00]				[0.00]			
σ^{RD}	5.87***				5.89***			
	[0.02]				[0.02]			
Panel C								
$Var(\alpha^{RD})$	3.46***				2.88***			
	[0.11]				[0.10]			
$Var(\alpha^{DK})$	4.24***				4.04***			
	[0.12]				[0.12]			
$Corr(\alpha^{RD}, \alpha^{DK})$	0.71***				0.71***			
	[0.01]				[0.01]			
Panel D								
Implied share NR	0.6485				0.6491			
Implied share RD	0.2834				0.2829			
Implied share DK	0.0681				0.0680			
Month-year FE	no				yes			
Observations	172,548				172,548			

Notes: This table reports model estimates for the dependent variable on short-run inflation expectations (px1). Response types are non-rounders (NR), rounders (RD) and respondents who choose a "don't know" answer (DK). Specification 1 (2) excludes (includes) month-year fixed effects in the random effects multinomial logit model for type probabilities (Equation 6). Panel A reports estimates for interpersonal heterogeneity. Columns a and b focus on the random effects multinomial logit model for type probabilities. Omitted category is type NR. Columns c and d report estimates for the parameterized mean of inflation expectations for type NR and RD (Equation 4), respectively. Panel B displays type-specific estimates for the standard deviation of the normal distribution of inflation expectations. Panel C reports the estimated variances of the individual specific random effects and its correlations. Panel D reports averages of model-implied unconditional type probabilities. For details see text. Standard errors in brackets. *** p <0.01, ** p <0.05, * p <0.1.

Panel C reports the estimated variances and correlation of the two random individual effects, which are derived from the entries of the estimated Cholesky matrix $\hat{\mathbf{L}}$. As shown, the variances of the individual effects are both significantly different from zero, confirming the importance in accounting for unobserved heterogeneity in the model. The individual effects are – as suspected in Section 3 – positively correlated ($\rho = 0.71$ in both specifications), implying that individuals who are more likely to round are also more likely not to respond at all. It is important to note that this correlation could not have been identified if type probabilities were modeled in a standard multinomial logit model. The positive correlation between the individual effects thus reinforces the necessity of joint estimation of the model. In fact, separate estimation – as often done in the literature by discarding item nonrespondents – would only be valid if the individual effects were uncorrelated.

Panel D displays the model-implied response type distribution in the sample, which is given by the unconditional type probabilities, averaged over time and individuals. With almost no differences between the two specifications, the average probability for type DK is given by 6.8 percent. This is almost identical to the crude DK share in the data set, which is given by 6.7 percent (11,490 out of 172,548 respondents choose the "don't know" option), strengthening the validity of the model. The average share of non-rounders is given by roughly 65 percent, implying that almost two in three respondents report an exact inflation expectation. The remaining 28 percent are consistent with type RD, implying that roughly one in four respondents rounds her inflation expectations to the next multiple of five percent. In comparison, the crude share of responses which are multiples of five percent is given by roughly 43 percent (74,161 out of 172,548 respondents) and clearly overestimates the true rounding share in the population, as identified by the model.

4.2 Type transitions and posterior probabilities

Recall that the methodology in this paper does not uniquely classify respondents into the three response types NR (non-rounder), RD (rounder) and DK (don't know), but rather assigns individual-specific probabilities to each of the three types. The panel dimension of the data allows me to analyze how these type probabilities change between the six-monthsapart interviews. Figure 2 plots individual, unconditional type probabilities for a specific response type in the first interview against the type probability of the same type in the second interview, based on the results of specification 2 of Table 2 (including month-year fixed effects). Clearly, there is evidence for a strong, positive correlation. For all three response types, the Pearson correlation coefficient is between 0.85 and 0.89. This indicates that, for example, individuals with a high probability of being type NR in the first wave, also have a high probability of being type NR in the second interview. The same applies to DK and RD type probabilities, even though the levels are considerably smaller. Overall, the strong, positive correlation across time can be explained by the fact that several covariates, such as gender and education, are time-constant for most respondents in the sample. Therefore, temporal variation in the unconditional type probabilities mainly stems from time-varying covariates and from the month-year fixed effects. Note that unobserved heterogeneity – modeled via the individual-specific random effects – only contributes to variation in the type probabilities across respondents, but not over time.

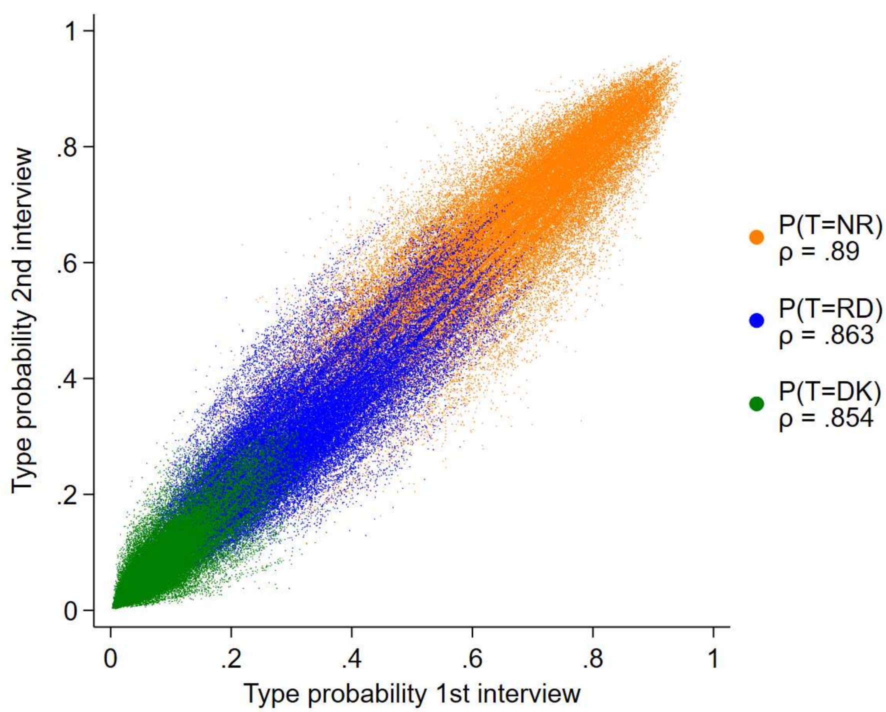


Figure 2: Type probability correlation between 1st and 2nd interview

Moreover, I am interested in how the probability of a specific type in the first interview is related to the probability of being another type in the second interview. Figure 3 therefore plots the RD probability in the first interview (horizontal axis) against the two other type probabilities (NR and DK) in the second interview (vertical axis). Mirroring the findings from the previous figure, there is a strong, negative correlation between types NR and RD. The lower the RD probability in the first interview, the higher the NR probability in the second interview ($\rho = -0.810$). The correlation with the type DK probability in the second interview is much weaker. Despite being slightly correlated ($\rho = 0.295$), a higher RD probability in the first interview seems to be rather unrelated to the DK probability in the second interview.

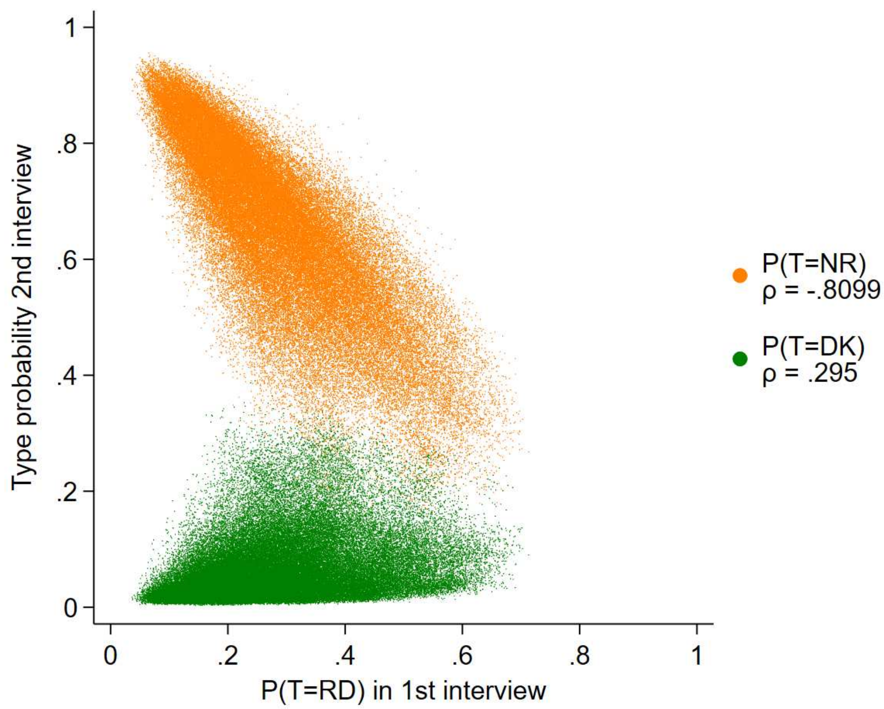


Figure 3: Type transition probabilities between 1st and 2nd interview

As highlighted in Figures 2 and 3, the levels of the unconditional type probabilities differ considerably across respondents and types. For example, the highest probability of being type DK is predicted to be "only" 0.390, the average being 0.0681 (cf. Table 2). In contrast, the highest probability of being type RD is given by 0.722, with an average of 0.283 (cf. Table 2). However, the probabilities discussed so far are (individual-specific) unconditional type probabilities. Posterior probabilities, i.e. type probabilities conditional on the reported inflation expectations (px1), may in contrast be very different.[18]

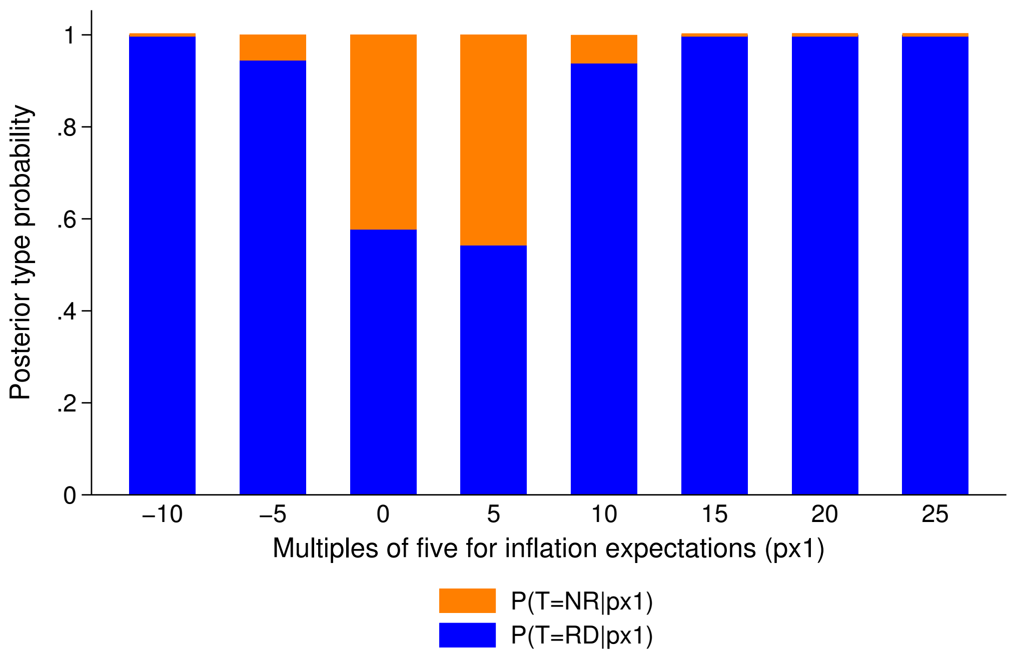


Figure 4: Posterior type probabilities conditional on reporting multiples of five

In fact, conditional on reporting a missing value, the posterior probability of type DK is one and, consequently, the posterior probabilities of types RD or NR are zero. For non-missing inflation expectations, there are two cases. First, if the respondent reports a non-multiple of five, the posterior probability of type NR is one; the posterior probabilities of type NR or DK are then zero. Second, if the reported value is a multiple of five, the posterior probability of being type DK is (exactly) zero, while the posterior probabilities of types RD and NR are strictly positive and can be calculated using Bayes' theorem. Figure 4 shows these probabilities for several multiples of five. To respondents reporting

[18]The exact calculation of the posterior probabilities is described in Section 3 and Appendix D.

extreme expectations, such as a future inflation rate of 25 percent or minus ten percent (deflation), the model assigns a posterior RD probability of almost one, implying that these respondents are rounders with almost certainty. This pattern changes when looking at more moderate inflation expectations. For example, conditional on reporting a predicted inflation rate of zero (five) percent, the posterior probability of being type RD is 0.57 (0.54). This implies that every second respondent reporting an inflation prediction of zero (five) percent does not have an exact inflation expectation of zero (five) percent in mind, but rather some different value and reports a rounded value.

4.3 Uncertainty index

Next, I use the insight of Binder (2017) and argue that rounding patterns in individuals' inflation expectations can serve as measure of economic uncertainty. I follow her analysis and calculate an uncertainty index as average, unconditional rounding (RD) probability, augmented by the average unconditional probability for type DK. Thus, the index is essentially an estimate for the population shares of types RD and DK. While Binder (2017) estimates this index separately for each month, the temporal variation in my index comes from the month-year fixed effects in the random effects multinomial logit model.

Figure 5 plots the uncertainty index as well as the average DK share for every month between 1980 and 2017, based on the results from specification 2 of Table 2. The share of rounders (RD share) is implicitly given by the difference between both lines. Clearly, there is evidence for meaningful variation over time. Respondents' reported inflation expectations display more rounding in times of higher economic uncertainty, compared to times of lower uncertainty. For example, the uncertainty index increases shortly after the terrorist attacks in September 2001 or Hurricane Katrina in August 2005. Moreover, the variation in the index is almost exclusively driven by temporal variation in the RD share rather than variation in the DK share. In fact, the latter is relatively constant across time and on average around seven percent. Therefore, respondents seem to systematically use rounding rather than the "don't know" option to express uncertainty. These results confirm the findings in Binder (2017).

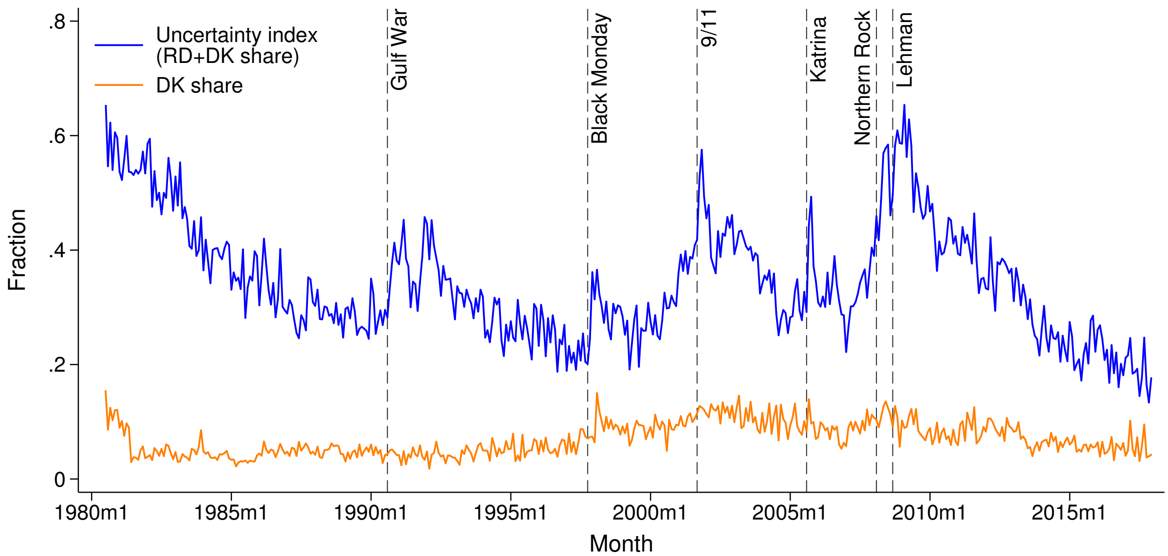


Figure 5: Model-implied uncertainty index over time

Figure 6 compares the uncertainty index to several other measures.[19] Most importantly, Panel A shows that the index is highly correlated with the original Binder (2017) index (Pearson's $\rho = 0.964$). Even though some minor differences exist, both indices display almost identical variation over time. My index is also shown to be correlated with two other measures of economic uncertainty (Panel B and C).[20] The first index by Baker et al. (2016) is based on newspaper coverage frequencies of specific combinations of terms, such as "uncertainty", "economic" and "deficit". In contrast to my uncertainty measure, this index does not spike after Hurricane Katrina, but does spike during the European sovereign debt crisis in 2012. The overall correlation between both indices is 0.527. The second index by Jurado et al. (2015) measures uncertainty by the volatility of expected forecast errors over a one-year horizon. The correlation with my measure of economic uncertainty

[19]Figure 6 is inspired by Figure 3 in Binder (2017, p.8).

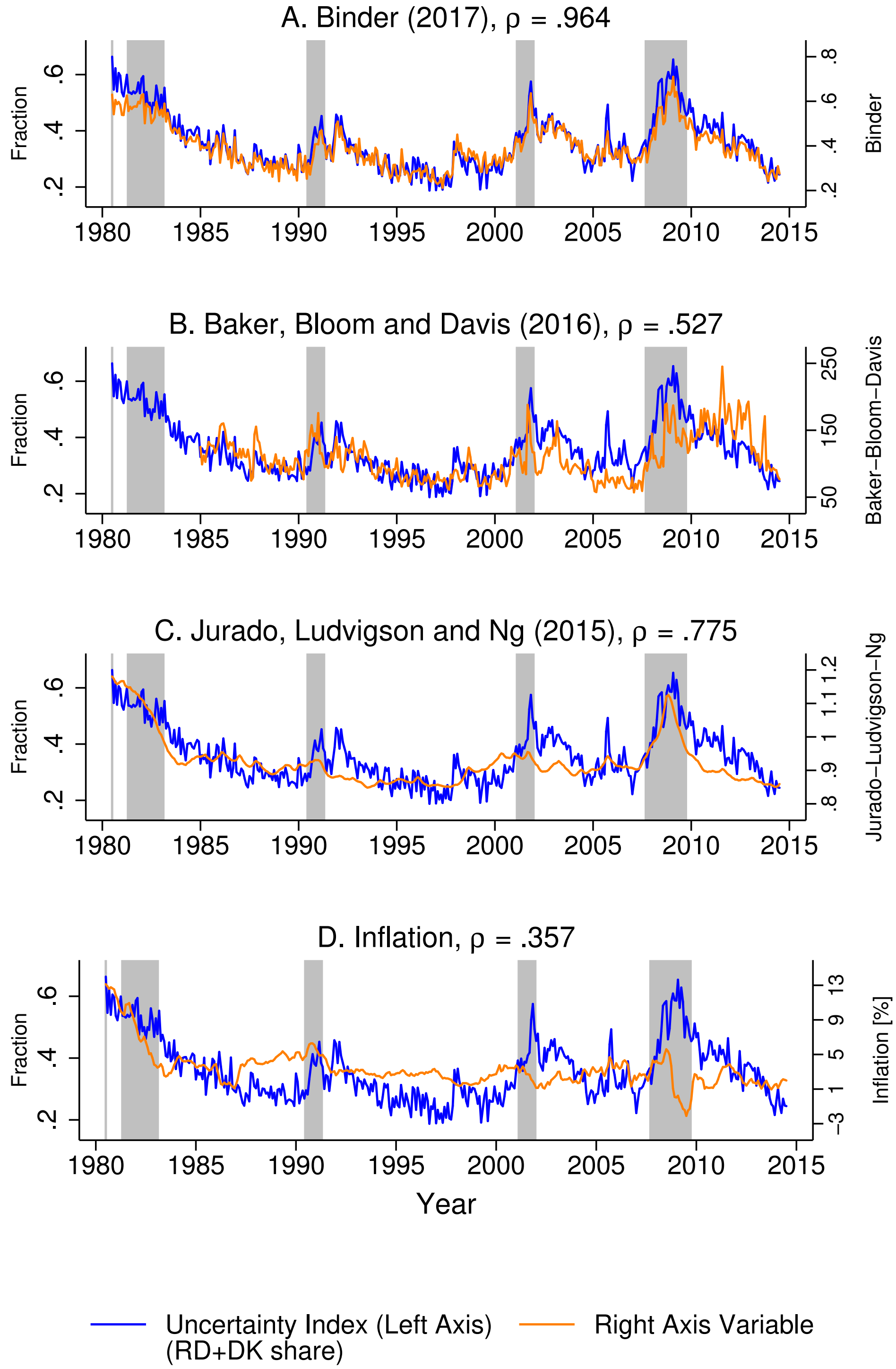
[20]The data for both indices is freely available at the authors' websites: http://www.policyuncertainty.com and https://www.sydneyludvigson.com [accessed October 8, 2018].

is a lot higher with a correlation coefficient of 0.775. While the increase after 9/11 and Hurricane Katrina is less pronounced, the increase during the financial crisis in 2008 is similar. Overall, my uncertainty index shows a slightly higher correlation with the two alternative uncertainty measures than the original Binder (2017) index.[21]

Last, Panel D shows that there is only a small, positive correlation of the uncertainty index with the actual US inflation rate across time ($\rho = 0.357$). In particular, it is reassuring that the variation in the index is not driven by the level of the current inflation. The uncertainty index spikes during both times of high inflation, such as the Gulf War in late 1990, and times of low inflation, such as the financial crisis in 2008.

In summary, at least for the construction of the uncertainty index, the advantages of the generalized model introduced in this paper over the original Binder (2017) model become small. Both models yield almost identical uncertainty indices, with only minor advantages for my uncertainty index in terms of correlation with alternative uncertainty measures.

[21]The correlations of the original Binder index with the Baker et al. (2016) index and the Jurado et al. (2015) index are given by 0.470 and 0.755 (not reported), respectively.



Note: Gray bars denote NBER recessions.

Figure 6: Comparison of the uncertainty index to other measures

5 Robustness

This section provides several robustness checks to variations in methodology and sample size. All tables and figures are presented in Appendix F. To reduce the computational burden, some specifications are estimated under the restriction that the variances of the individual effects are zero, as indicated in the tables.[22]

Logistic inflation expectations. To check whether the results are sensitive to the assumption that inflation expectations follow a normal distribution, I repeat the analysis under the assumption of logistic inflation expectations (Table F1). The implied type distribution as well as the effects of the covariates on individual type probabilities and inflation expectations are almost unchanged.

Medium-run inflation expectations. I also use data on medium-run inflation expectations, which are based on the expected yearly inflation rate over the next five years (px5). Note, however, that I lose several month-year combinations, since this question is not asked throughout all waves. As shown in Table F2, the implied NR share increases to roughly 77 percent, implying that respondents round on average less when asked about medium-run inflation expectations, compared to short-run expectations (px1). In addition, the DK share increases to almost nine percent. Both effects are shown not to be driven by differences in sample size and periods, but rather by the difference between short-run and medium-run inflation expectations (not reported).

Rounding to multiples of ten. The main analysis assumes that type RD respondents round to the next multiple of five percent. I repeat the analysis for rounding to the next multiple of ten percent in Table F3. Since multiples of ten are by definition also multiples of five, the new share of rounders should decrease, as it does. In fact, the type shares are similar

[22]By restricting the variances of the individual effects to zero, these models ignore the panel structure of the data and essentially become pooled ordinary least squares models.

in magnitude to the ones from the analysis on medium-run inflation expectations. Again, the effects of the covariates on individual type probabilities and mean inflation expectations remain the same.

Two rounding types. I also estimate a variant of the model with two rounding types: type RD5 rounds to the next multiple of five percent, while type RD10 rounds to the next multiple of ten percent. Together with type NR and DK, this model is then a mixture of four different response types. Results are presented in Table F4, suggesting that the aggregated NR and DK share are 65% and 6.7%, respectively, and thus literally identical to the shares from the main model. The remaining 28% of rounders are split between 19% of respondents who round to the multiple of five, and 9% who round to the next multiple of ten percent. The effect of the covariates on the type probabilities is very similar for types RD5 and RD10 and qualitatively close to the main findings. Further analysis shows that the increase in the uncertainty index after 9/11 and the Lehman collapse are mainly driven by an increase in the RD10 share, while the increase after hurricane Katrina and Northern Rock is driven by an increase in the RD5 share (cf. Figure F1).

Level of current inflation. Rounding may also depend on the current level of the inflation rate. For example, one might to be more willing to round to five percent if the current inflation rate is given by 4.8 percent rather than 4.0 percent. Table F5 therefore includes month-year fixed effects not only in the random effects multinomial logit model, but also in the equation for the mean of the inflation expectation distribution (Equation 4). The time effects capture all variables affecting respondents similarly across time, such as the current inflation rate. The implied type distribution is literally unaffected by this specification and the associations with the covariates get even stronger.

Full sample. I also estimate the model for the full data set, i.e. I add data from the 77,630 respondents, who are interviewed only once (Table F6). The average NR share slightly decreases by two percentage points, while the RD and DK share increase by one percentage point each. The resulting type distribution is thus almost identical to the one from the main section. Due to the increase in sample size, the standard errors of the estimates get – as expected – even smaller. The other results are unchanged.

6 Conclusion

This paper introduces a microeconometric panel data model for inflation point expectations of US households. In contrast to previous studies, in particular Binder (2017), I explicitly model a panel dimension and allow for individual heterogeneity and item non-response. The population is described as a finite mixture of three distinct response types, who differ in how they report their inflation expectations: rounders (RD), non-rounders (NR) and respondents who choose a “don’t know” response (DK). Type probabilities are allowed to depend on both observed and unobserved heterogeneity.

The estimated average population shares of types (NR,RD,DK) are given by (0.65,0.28,0.07), implying that most respondents actually report their true inflation expectation rather than some rounded value. However, the results suggest that more than a quarter of respondents round their inflation expectations to the next multiple of five, with meaningful variation over time. Rounding is more prevalent in times of higher economic uncertainty compared to times of lower economic uncertainty. Moreover, I find that response type probabilities can be predicted by both observed and unobserved heterogeneity. For example, males and respondents with at least a college degree are significantly less likely to round and to choose a “don’t know” option than females and respondents without a college degree. Respondents who are more likely to round are also more likely to choose “don’t know”, questioning the standard procedure of dropping missing answers. I also find evidence for type stability across interviews.

This generalized model of Binder (2017) allows to increase efficiency of the estimates, to identify meaningful heterogeneity in the type probabilities and to explicitly take item nonresponse into account. However, in terms of the construction of the uncertainty index, there seems to be little difference across both models. In fact, the resulting uncertainty indices are almost identical.

This paper has several implications. For example, the insight that rounding behavior systematically varies with socio-economic characteristics may guide future survey design and improve data quality. Furthermore, since rounding patterns in inflation expectations are systematically linked with economic uncertainty over time, this information may be used to determine or at least improve existing estimates for the current level of uncertainty in the economy. Future research should, in addition, analyze if this also applies to other domains, i.e. if economic uncertainty is also related to rounding behavior in expectations questions in other domains or survey questions unrelated to expectations.

More generally, this paper demonstrates the usefulness of survey data, that goes beyond the face value of individuals' responses. Researchers have recently started to extensively rely on so-called paradata. This includes, for example, respondent-level information on the amount of time spent on a specific survey question, the number of adjustments, the number of mouse clicks as well as the exact mouse movement pattern. The latter, for example, has already been used for PC user verification (Pusara and Brodley, 2004; Zheng et al., 2011). Clearly, these novel approaches have the potential to improve not only data quality, but also the understanding of the decision-making process of individuals itself.

Appendix

A Questionnaire for price expectations

Figures A1 and A2 describe the exact procedure for the elicitation of inflation expectations in the short-run (px1) and the medium-run (px5), respectively. The entire questionnaire and interviewer instructions are available at the University of Michigan Survey Research Center and are described in Curtin (1996).

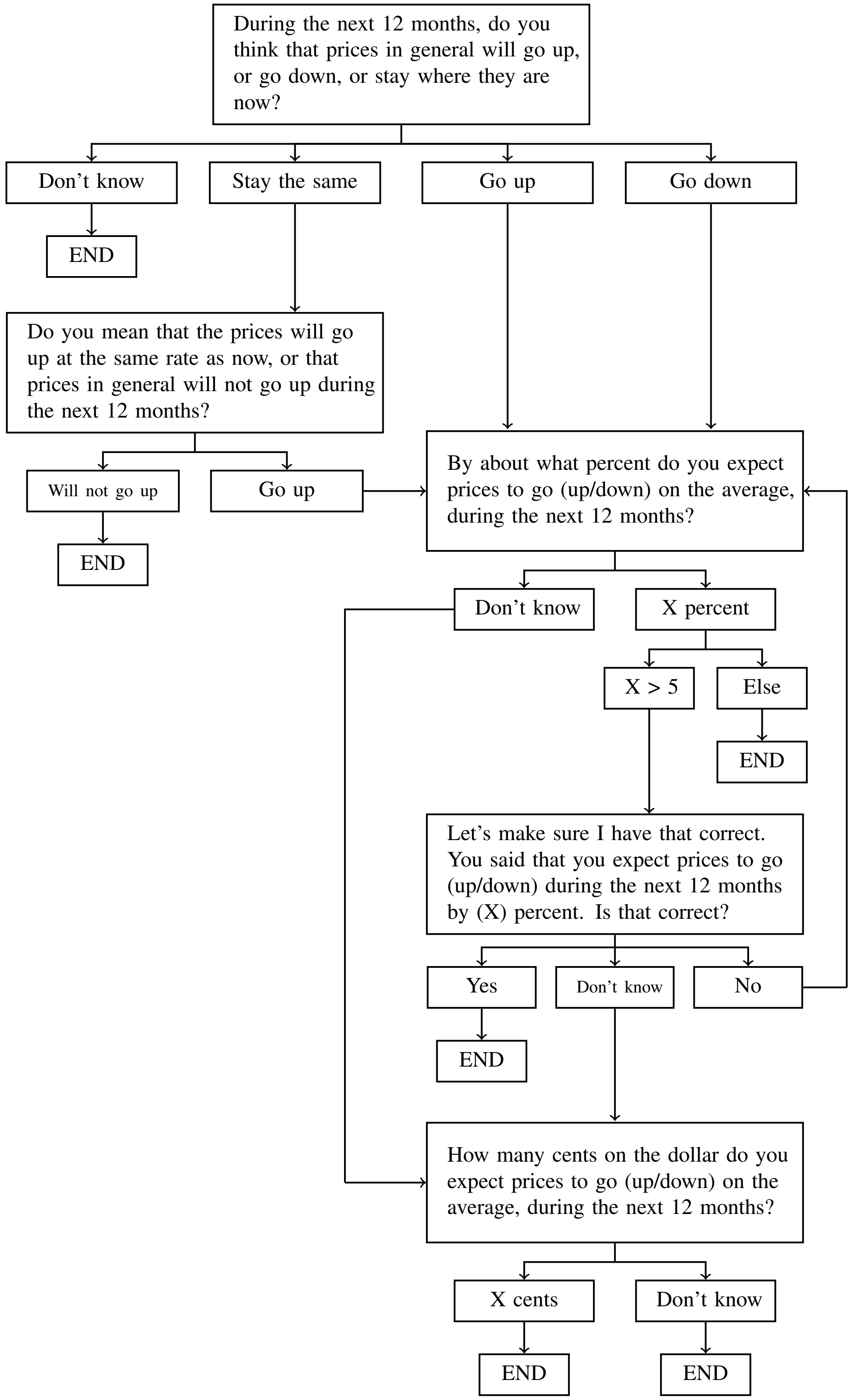


Figure A1: Questionnaire for short-run inflation expectations (px1)

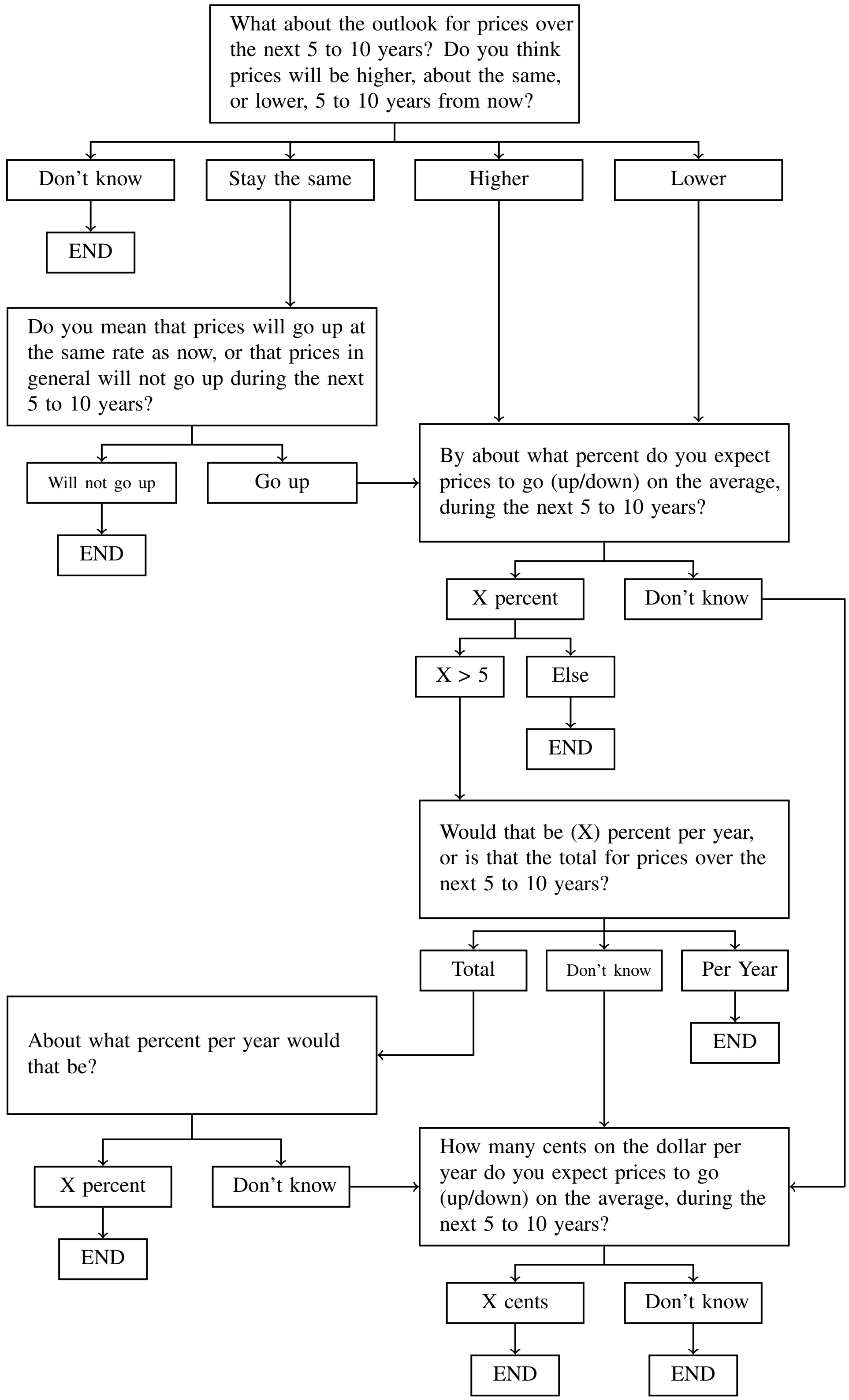


Figure A2: Questionnaire for medium-run inflation expectations (px5)

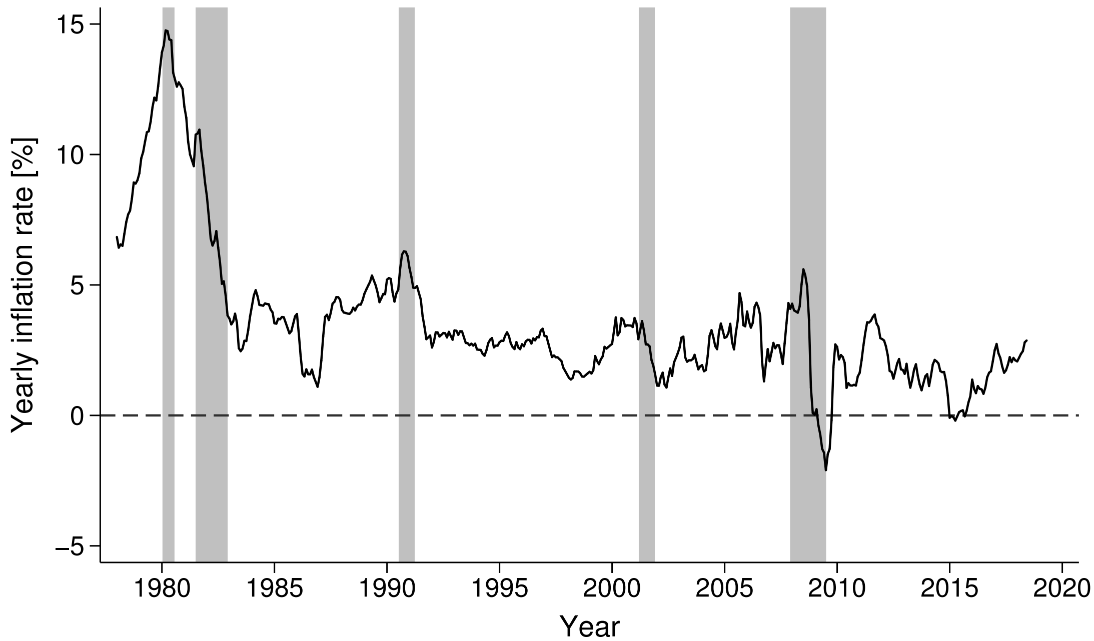
B Full sample summary statistics

Table B1: Summary statistics for the full sample

	Mean	SD	p5	p95	Min	Max	Observations
A: Inflation expectations [%]							
Short-run (px1)	4.55	6.30	0	15	-50	50	246,683
Medium-run (px5)	4.06	5.17	0	10	-50	50	176,177
B: Sociodemographics [0/1]							
Male	0.46	0.50	0	1	0	1	271,277
Partner	0.60	0.49	0	1	0	1	268,594
Age > 64	0.20	0.40	0	1	0	1	269,899
Age < 40	0.39	0.49	0	1	0	1	269,899
College	0.37	0.48	0	1	0	1	268,579
1st income quartile	0.21	0.41	0	1	0	1	234,095
2nd income quartile	0.21	0.41	0	1	0	1	234,095
3rd income quartile	0.28	0.45	0	1	0	1	234,095
4th income quartile	0.30	0.46	0	1	0	1	234,095
C: Regional information [0/1]							
West	0.20	0.40	0	1	0	1	271,853
Northcentral	0.27	0.44	0	1	0	1	271,853
Northeast	0.19	0.39	0	1	0	1	271,853
South	0.33	0.47	0	1	0	1	271,853

Notes: This Table is based on all 77,630 respondents who are interviewed once and all 97,159 respondents from the MSC who are interviewed twice between January 1978 to December 2017, making a total of 271,948 observations. Number of observations differ due to item nonresponse. Panel B and C report dummy variables if not indicated differently. Information on income (1st-4th quartile) not available before October 1979. For details see text.

C US inflation between 1978 and 2018



Note: Gray bars denote NBER recessions. Inflation rates are measured as CPI growth rates.
Data source: OECD

Figure C1: Yearly inflation rates in the US between 1978 and 2018

D Derivation of the likelihood function

Recall that the probabilities of observing y_{it} conditional on type T_{it} are given by:

$$P(y_{it}|T_{it}) = \begin{cases} f_{NR}(y_{it}) & \text{if } T_{it} = NR \\ f_{RD}(y_{it}) & \text{if } T_{it} = RD \\ 0 & \text{if } T_{it} = DK \end{cases} \quad \& \ y_{it} \text{ is a multiple of } m$$

$$P(y_{it}|T_{it}) = \begin{cases} f_{NR}(y_{it}) & \text{if } T_{it} = NR \\ 0 & \text{if } T_{it} = RD \\ 0 & \text{if } T_{it} = DK \end{cases} \quad \& \ y_{it} \text{ is not a multiple of } m$$

$$P(y_{it}|T_{it}) = \begin{cases} 0 & \text{if } T_{it} = NR \\ 0 & \text{if } T_{it} = RD \\ f_{DK}(y_{it}) & \text{if } T_{it} = DK \end{cases} \quad \& \ y_{it} \text{ is missing}$$

with

$$\begin{aligned} f_{NR}(y_{it}) &= \phi(y_{it}; \mu^{NR}; \sigma^{NR}) \\ f_{RD}(y_{it}) &= \Phi\left(\frac{y_{it} + \frac{m}{2} - \mu^{RD}}{\sigma^{RD}}\right) - \Phi\left(\frac{y_{it} - \frac{m}{2} - \mu^{RD}}{\sigma^{RD}}\right) \\ f_{DK}(y_{it}) &= 1 \end{aligned}$$

By definition, the unconditional probability of observing y_{it} is given by:

$$\begin{aligned} P(y_{it}) = \ & P(y_{it}|T_{it} = NR) \cdot P(T_{it} = NR) + \\ & P(y_{it}|T_{it} = RD) \cdot P(T_{it} = RD) + \\ & P(y_{it}|T_{it} = DK) \cdot P(T_{it} = DK) \end{aligned}$$

which can then be simplified to:

$$P(y_{it}) = \begin{cases} P(T_{it} = DK) & \text{if } y_{it} \text{ is missing} \\ f_{NR} \cdot P(T_{it} = NR) & \text{if } y_{it} \text{ is not a multiple of } m \\ f_{NR} \cdot P(T_{it} = NR) + f_{RD} \cdot P(T_{it} = RD) & \text{if } y_{it} \text{ is a multiple of } m \end{cases}$$

Taking the product over individuals and time and parameterizing the type probabilities $P(T_{it} = j)$ results in the likelihood function presented in the main section. Note also that – after maximization of the likelihood function – the estimated (unconditional) individual type probabilities can be used to calculate posterior type probabilities conditional on the reported values of y_{it}. More specifically, those are given by Bayes' theorem:

$$P(T_{it} = j|y_{it}) = P(y_{it}|T_{it} = j)\frac{P(T_{it} = j)}{P(y_{it})}$$

Using the definitions introduced earlier, it is straightforward to show that

$$P(NR|y_{it}) = \begin{cases} 0 & \text{if } y_{it} \text{ is missing} \\ 1 & \text{if } y_{it} \text{ is not a multiple of } m \\ f_{NR} \cdot \frac{P(NR)}{f_{NR} \cdot P(NR) + f_{RD} \cdot P(RD)} & \text{if } y_{it} \text{ is a multiple of } m \end{cases}$$

$$P(RD|y_{it}) = \begin{cases} 0 & \text{if } y_{it} \text{ is missing} \\ 0 & \text{if } y_{it} \text{ is not a multiple of } m \\ f_{RD} \cdot \frac{P(RD)}{f_{NR} \cdot P(NR) + f_{RD} \cdot P(RD)} & \text{if } y_{it} \text{ is a multiple of } m \end{cases}$$

$$P(DK|y_{it}) = \begin{cases} 1 & \text{if } y_{it} \text{ is missing} \\ 0 & \text{if } y_{it} \text{ is not a multiple of } m \\ 0 & \text{if } y_{it} \text{ is a multiple of } m. \end{cases}$$

E Computational issues for the Hessian matrix

The default optimization method in Stata®15 is given by a (modified) Newton-Raphson algorithm, which is based on the calculation of the gradient and the Hessian matrix. While this algorithm is known to work fine for many applications, it becomes computationally very costly as the number of parameters increases. In fact, calculating the Hessian matrix for a k-dimensional parameter vector requires $O(k^2)$ evaluations of the log-likelihood function (Jeliazkov and Lloro, 2011). In the application of my model, I use monthly data over a 40-year period, which implies that adding month-year fixed effects increases the dimension of the parameter vector by almost 500 per response type. In combination with the Maximum Simulated Likelihood approach, which requires a repeated calculation of the likelihood function at every iteration, calculating the Hessian matrix and thus using the Newton-Raphson algorithm becomes computationally too costly and in fact infeasible.

I therefore rely on Quasi-Newton, gradient-based optimization methods, which replace the Hessian matrix by some other – computationally less costly – measure. For example, the Berndt-Hall-Hall-Hausmann (BHHH) algorithm replaces the negative Hessian by the outer product of the gradients. Similarly, the Broyden-Fletcher-Goldfarb-Shanno (BFGS) algorithm replaces the Hessian by a function of the gradient, which aims for an ever-improving estimate of the Hessian at every iteration.[23] One fundamental advantage of these algorithms is that they only require $O(k)$ evaluations of the likelihood function (Jeliazkov and Lloro, 2011). My specific optimization routine switches between the BHHH algorithm (5 iterations) and the BFGS algorithm (10 iterations) and focuses on BFGS only, when BHHH is not applicable.

By default, Stata declares convergence if the following two conditions are met: First, the scaled gradient is sufficiently small, i.e. $\boldsymbol{g}\boldsymbol{H}^{-1}\boldsymbol{g}' < 10^{-5}$, where $\boldsymbol{g}$ is the gradient (row) vector and $\boldsymbol{H}$ is the Hessian matrix of the parameter vector $\hat{\boldsymbol{\theta}}$. Second, either the relative

[23]See Gould et al. (2006) for more details on both algorithms.

change in the parameter vector $\hat{\boldsymbol{\theta}}$ or the relative change in the value of the log-likelihood function $L(\hat{\boldsymbol{\theta}})$ from one iteration to the next is sufficiently small. As the first criterion requires again the calculation of the Hessian matrix, I use Stata's `qtolerance()` option, which causes Stata to use the modified (gradient-based) version of the Hessian matrix as final check for convergence rather than the actual Hessian. Note that this procedure has been the default option in Stata until version 12. The second criterion remains unchanged.

Similarly, I estimate the variance-covariance matrix of my parameter vector and therefore the standard errors of my estimates by the outer product of the gradients (Gould et al., 2006). Again, Stata's default estimator would require the calculation of the Hessian matrix.

F Additional Figures and Tables

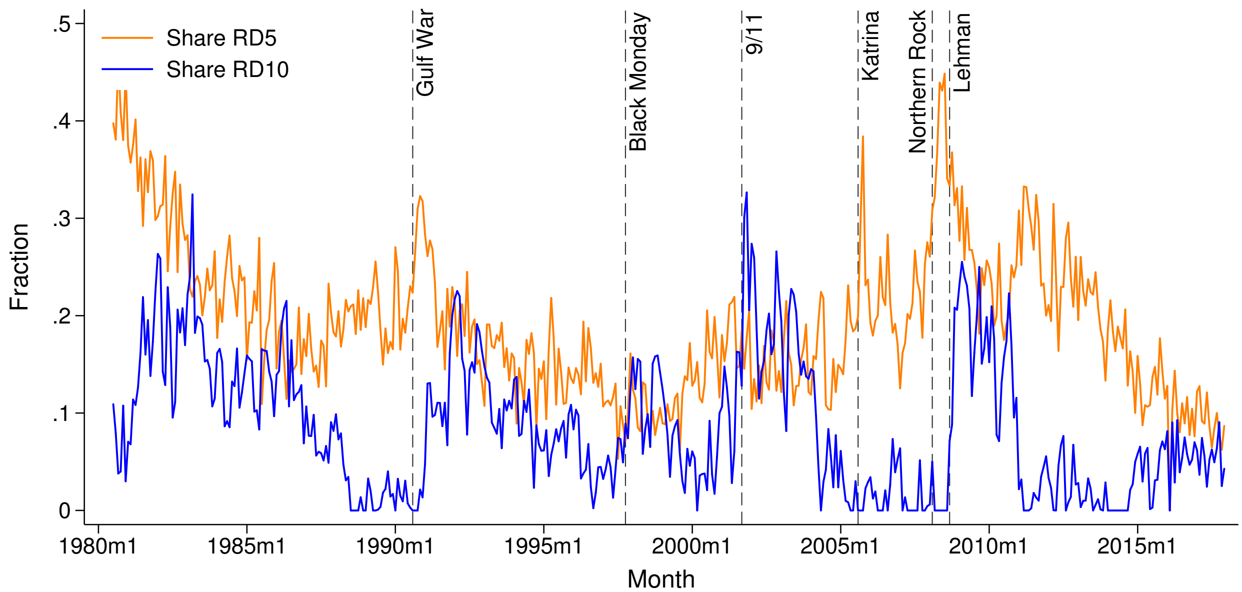


Figure F1: Rounding shares with two distinct rounding types

Table F1: Model estimates for logistic inflation expectations

	Logistic distribution				Normal distribution			
	Mean NR (1a)	Mean RD (1b)	P(T=RD) (1c)	P(T=DK) (1d)	Mean NR (2a)	Mean RD (2b)	P(T=RD) (2c)	P(T=DK) (2d)
Panel A								
Male	-0.01	-1.05***	-0.59***	-0.99***	-0.13***	-1.12***	-0.64***	-0.99***
	[0.02]	[0.06]	[0.01]	[0.02]	[0.02]	[0.06]	[0.02]	[0.02]
Partner	0.01	0.31***	-0.02	-0.06***	0.02	0.34***	-0.02	-0.06***
	[0.02]	[0.06]	[0.02]	[0.02]	[0.02]	[0.07]	[0.02]	[0.02]
Age	-0.01***	-0.02***	-0.00	0.02***	-0.01***	-0.02***	-0.00*	0.02***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
College	-0.01	-0.22***	-0.35***	-0.43***	-0.11***	-0.37***	-0.38***	-0.43***
	[0.02]	[0.06]	[0.02]	[0.02]	[0.02]	[0.07]	[0.02]	[0.02]
1st income quartile	0.15***	1.61***	0.44***	1.20***	0.30***	1.59***	0.49***	1.21***
	[0.03]	[0.09]	[0.02]	[0.03]	[0.03]	[0.10]	[0.03]	[0.03]
2nd income quartile	-0.01	1.27***	0.23***	0.58***	0.05*	1.22***	0.26***	0.59***
	[0.02]	[0.08]	[0.02]	[0.03]	[0.03]	[0.09]	[0.02]	[0.03]
3rd income quartile	-0.03	0.67***	0.10***	0.23***	0.01	0.67***	0.12***	0.24***
	[0.02]	[0.07]	[0.02]	[0.03]	[0.02]	[0.08]	[0.02]	[0.03]
West	0.06***	0.08	-0.10***	-0.06**	0.06**	-0.00	-0.10***	-0.06**
	[0.02]	[0.08]	[0.02]	[0.03]	[0.03]	[0.09]	[0.02]	[0.03]
Northcentral	0.03	-0.17**	-0.02	-0.16***	-0.02	-0.18**	-0.02	-0.16***
	[0.02]	[0.07]	[0.02]	[0.03]	[0.02]	[0.08]	[0.02]	[0.03]
Northeast	0.04	-0.02	0.08***	0.06**	0.02	-0.04	0.08***	0.06**
	[0.02]	[0.08]	[0.02]	[0.03]	[0.03]	[0.08]	[0.02]	[0.03]
Constant	3.09***	5.05***	0.68***	-1.35***	3.50***	5.78***	0.67***	-1.36***
	[0.03]	[0.11]	[0.17]	[0.23]	[0.04]	[0.12]	[0.18]	[0.23]
Panel B								
σ^{NR}	1.39***				2.81***			
	[0.00]				[0.01]			
σ^{RD}	3.10***				5.87***			
	[0.01]				[0.02]			
Panel C								
	Random effects are restricted to zero							
Panel D								
Implied share NR	0.629				0.652			
Implied share RD	0.305				0.281			
Implied share DK	0.067				0.067			
Month-year FE	yes				yes			
Observations	172,548				172,548			

Notes: This table repeats the main analysis under the assumption of logistic inflation expectations (specification 1). Dependent variable is short-run inflation expectations (px1). Response types are non-rounders (NR), rounders (RD) and respondents who choose a "don't know" answer (DK). All columns include month-year fixed effects in the random effects multinomial logit model for type probabilities (Equation 6). Panel A reports estimates for interpersonal heterogeneity. Columns a and b report estimates for the parameterized mean of inflation expectations for type NR and RD (Equation 4), respectively. Columns c and d focus on the random effects multinomial logit model for type probabilities. Omitted category is type NR. Panel B displays type-specific estimates for the standard deviation of the logistic or normal distribution, respectively. The individual effects are normalized to zero (Panel C). Panel D reports averages of model-implied unconditional type probabilities. For details see text. Standard errors in brackets. *** p <0.01, ** p <0.05, * p <0.1.

Table F2: Model estimates for medium-run inflation expectations

	Excluding month-year FE				Including month-year FE			
	Mean NR	Mean RD	P(T=RD)	P(T=DK)	Mean NR	Mean RD	P(T=RD)	P(T=DK)
	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Panel A								
Male	0.02	-0.30***	-0.95***	-0.84***	0.00	-0.44***	-0.89***	-0.82***
	[0.02]	[0.11]	[0.02]	[0.02]	[0.02]	[0.11]	[0.02]	[0.02]
Partner	-0.02	0.46***	-0.08***	-0.10***	-0.02	0.43***	-0.05**	-0.08***
	[0.02]	[0.11]	[0.02]	[0.02]	[0.02]	[0.10]	[0.02]	[0.02]
Age	-0.01***	-0.06***	-0.01***	0.02***	-0.01***	-0.06***	-0.00***	0.02***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
College	-0.11***	-0.00	-0.84***	-0.43***	-0.13***	-0.08	-0.58***	-0.34***
	[0.02]	[0.12]	[0.03]	[0.02]	[0.02]	[0.11]	[0.03]	[0.02]
1st income quartile	0.03	-0.12	0.93***	1.09***	0.05*	0.10	1.05***	1.14***
	[0.03]	[0.16]	[0.04]	[0.03]	[0.03]	[0.16]	[0.04]	[0.03]
2nd income quartile	-0.15***	-0.26	0.44***	0.49***	-0.15***	-0.09	0.62***	0.54***
	[0.02]	[0.16]	[0.03]	[0.03]	[0.02]	[0.15]	[0.03]	[0.03]
3rd income quartile	-0.10***	-0.20	0.29***	0.21***	-0.10***	-0.07	0.35***	0.23***
	[0.02]	[0.15]	[0.03]	[0.03]	[0.02]	[0.15]	[0.03]	[0.03]
West	0.09***	0.12	-0.11***	-0.01	0.09***	0.11	-0.11***	-0.01
	[0.02]	[0.14]	[0.03]	[0.03]	[0.02]	[0.14]	[0.03]	[0.03]
Northcentral	-0.02	-0.06	-0.11***	-0.20***	-0.02	-0.14	-0.12***	-0.20***
	[0.02]	[0.12]	[0.03]	[0.03]	[0.02]	[0.12]	[0.03]	[0.03]
Northeast	0.02	-0.17	0.01	0.06**	0.03	-0.16	-0.05	0.04
	[0.02]	[0.14]	[0.03]	[0.03]	[0.02]	[0.14]	[0.03]	[0.03]
Constant	3.88***	10.31***	-1.01***	-3.02***	3.88***	10.22***	0.54***	-1.75***
	[0.03]	[0.21]	[0.05]	[0.05]	[0.03]	[0.20]	[0.17]	[0.22]
Panel B								
σ^{NR}	2.39***				2.39***			
	[0.01]				[0.01]			
σ^{RD}	5.76***				5.77***			
	[0.03]				[0.03]			
Panel C								
	Random effects are restricted to zero							
Panel D								
Implied share NR	0.777				0.774			
Implied share RD	0.134				0.137			
Implied share DK	0.089				0.089			
Month-year FE	no				yes			
Observations	136,264				136,264			

Notes: This table repeats the main analysis for the alternative dependent variable of medium-run inflation expectations (px5). Response types are non-rounders (NR), rounders (RD) and respondents who choose a "don't know" answer (DK). Specification 1 (2) excludes (includes) month-year fixed effects in the random effects multinomial logit model for type probabilities (Equation 6). Panel A reports estimates for interpersonal heterogeneity. Columns a and b report estimates for the parameterized mean of inflation expectations for type NR and RD (Equation 4), respectively. Columns c and d focus on the random effects multinomial logit model for type probabilities. Omitted category is type NR. Panel B displays type-specific estimates for the standard deviation of the normal distribution of inflation expectations. The individual effects are normalized to zero (Panel C). Panel D reports averages of model-implied unconditional type probabilities. For details see text. Standard errors in brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table F3: Model estimates for rounding to the next multiple of ten percent

	Excluding month-year FE				Including month-year FE			
	Mean NR	Mean RD	P(T=RD)	P(T=DK)	Mean NR	Mean RD	P(T=RD)	P(T=DK)
	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Panel A								
Male	-0.51***	-1.38***	-0.52***	-0.88***	-0.51***	-1.37***	-0.48***	-0.88***
	[0.02]	[0.08]	[0.02]	[0.02]	[0.02]	[0.08]	[0.02]	[0.02]
Partner	0.06***	0.33***	-0.03	-0.05**	0.07***	0.30***	-0.03	-0.06**
	[0.02]	[0.09]	[0.02]	[0.02]	[0.02]	[0.09]	[0.02]	[0.02]
Age	-0.01***	-0.03***	0.00***	0.02***	-0.01***	-0.03***	0.00***	0.02***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
College	-0.34***	-0.37***	-0.48***	-0.32***	-0.34***	-0.41***	-0.37***	-0.38***
	[0.02]	[0.09]	[0.02]	[0.02]	[0.02]	[0.09]	[0.02]	[0.02]
1st income quartile	0.62***	2.00***	0.32***	1.12***	0.64***	1.95***	0.32***	1.11***
	[0.04]	[0.13]	[0.03]	[0.03]	[0.04]	[0.12]	[0.03]	[0.03]
2nd income quartile	0.21***	1.54***	0.02	0.57***	0.22***	1.46***	0.09***	0.52***
	[0.03]	[0.12]	[0.02]	[0.03]	[0.03]	[0.12]	[0.03]	[0.03]
3rd income quartile	0.08***	0.85***	0.02	0.22***	0.09***	0.80***	0.03	0.21***
	[0.03]	[0.11]	[0.02]	[0.03]	[0.03]	[0.11]	[0.02]	[0.03]
West	-0.02	0.09	-0.09***	-0.05	-0.02	0.09	-0.10***	-0.05
	[0.03]	[0.11]	[0.02]	[0.03]	[0.03]	[0.11]	[0.02]	[0.03]
Northcentral	-0.05*	-0.16	-0.01	-0.17***	-0.04*	-0.18*	-0.01	-0.15***
	[0.03]	[0.10]	[0.02]	[0.03]	[0.03]	[0.10]	[0.02]	[0.03]
Northeast	0.03	0.09	0.10***	0.05	0.03	0.07	0.07***	0.05*
	[0.03]	[0.11]	[0.02]	[0.03]	[0.03]	[0.11]	[0.02]	[0.03]
Constant	4.35***	4.59***	-1.20***	-3.35***	4.33***	4.69***	-0.20	-1.81***
	[0.04]	[0.16]	[0.03]	[0.05]	[0.04]	[0.16]	[0.17]	[0.22]
Panel B								
σ^{NR}	3.55***				3.55***			
	[0.01]				[0.01]			
σ^{RD}	5.50***				5.47***			
	[0.03]				[0.03]			
Panel C								
	Random effects are restricted to zero							
Panel D								
Implied share NR	0.761				0.759			
Implied share RD	0.172				0.175			
Implied share DK	0.067				0.067			
Month-year FE	no				yes			
Observations	172,548				172,548			

Notes: This table repeats the main analysis under the assumption that rounders (RD) round to the next multiple of ten rather than five percent. Other response types are non-rounders (NR) and respondents who choose a "don't know" answer (DK). Dependent variable is short-run inflation expectations (px1). Specification 1 (2) excludes (includes) month-year fixed effects in the random effects multinomial logit model for type probabilities (Equation 6). Panel A reports estimates for interpersonal heterogeneity. Columns a and b report estimates for the parameterized mean of inflation expectations for type NR and RD (Equation 4), respectively. Columns c and d focus on the random effects multinomial logit model for type probabilities. Omitted category is type NR. Panel B displays type-specific estimates for the standard deviation of the normal distribution of inflation expectations. The individual effects are normalized to zero (Panel C). Panel D reports averages of model-implied unconditional type probabilities. For details see text. Standard errors in brackets. *** p <0.01, ** p <0.05, * p <0.1.

Table F4: Model estimates for four response types

	Excluding month-year FE						Including month-year FE					
	Mean NR (1a)	Mean RD5 (1b)	Mean RD10 (1c)	P(T=RD5) (1d)	P(T=RD10) (1e)	P(T=DK) (1f)	Mean NR (2a)	Mean RD5 (2b)	Mean RD10 (2c)	P(T=RD5) (2d)	P(T=RD10) (2e)	P(T=DK) (2f)
Panel A												
Male	-0.14***	-1.24***	0.27*	-0.77***	-0.32***	-0.98***	-0.15***	-1.21***	0.21*	-0.71***	-0.32***	-0.98***
	[0.02]	[0.09]	[0.15]	[0.02]	[0.03]	[0.02]	[0.02]	[0.09]	[0.13]	[0.02]	[0.03]	[0.02]
Partner	0.00	0.36***	0.00	0.00	-0.08**	-0.06**	0.02	0.35***	0.12	-0.01	-0.05*	-0.06***
	[0.02]	[0.10]	[0.17]	[0.02]	[0.03]	[0.02]	[0.02]	[0.09]	[0.14]	[0.02]	[0.03]	[0.02]
Age	-0.01***	-0.02***	0.00	-0.00***	0.00***	0.02***	-0.01***	-0.02***	-0.00	-0.00***	0.01***	0.02***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
College	-0.17***	-0.96***	-0.07	-0.36***	-0.59***	-0.36***	-0.16***	-0.76***	-0.25*	-0.36***	-0.33***	-0.43***
	[0.02]	[0.10]	[0.17]	[0.02]	[0.03]	[0.02]	[0.02]	[0.09]	[0.13]	[0.02]	[0.03]	[0.02]
1st income quartile	0.27***	1.37***	0.47*	0.64***	0.06	1.21***	0.27***	1.30***	-0.10	0.66***	0.01	1.21***
	[0.03]	[0.15]	[0.27]	[0.03]	[0.05]	[0.03]	[0.03]	[0.13]	[0.21]	[0.03]	[0.04]	[0.03]
2nd income quartile	-0.01	0.58***	0.40	0.40***	-0.32***	0.62***	0.01	0.69***	-0.07	0.44***	-0.18***	0.58***
	[0.03]	[0.13]	[0.26]	[0.03]	[0.05]	[0.03]	[0.03]	[0.12]	[0.19]	[0.03]	[0.04]	[0.03]
3rd income quartile	-0.02	0.41***	0.06	0.22***	-0.16***	0.25***	-0.01	0.49***	-0.06	0.21***	-0.10***	0.24***
	[0.02]	[0.13]	[0.17]	[0.03]	[0.04]	[0.03]	[0.02]	[0.12]	[0.15]	[0.03]	[0.03]	[0.03]
West	0.04	-0.17	0.08	-0.07**	-0.15***	-0.06**	0.04	-0.11	-0.30*	-0.07**	-0.19***	-0.06**
	[0.03]	[0.13]	[0.22]	[0.03]	[0.04]	[0.03]	[0.03]	[0.12]	[0.17]	[0.03]	[0.04]	[0.03]
Northcentral	-0.02	-0.21*	0.12	-0.03	0.01	-0.17***	-0.01	-0.14	0.02	-0.04*	0.03	-0.16***
	[0.02]	[0.11]	[0.17]	[0.03]	[0.04]	[0.03]	[0.02]	[0.11]	[0.15]	[0.02]	[0.03]	[0.03]
Northeast	0.02	-0.04	-0.30	0.13***	0.05	0.06**	0.02	0.01	-0.33**	0.09***	0.06*	0.06**
	[0.03]	[0.12]	[0.19]	[0.03]	[0.04]	[0.03]	[0.03]	[0.12]	[0.16]	[0.03]	[0.04]	[0.03]
Constant	3.72***	7.65***	0.30	-0.81***	-1.62***	-3.10***	3.71***	7.74***	0.00	0.78***	-0.95***	-0.75***
	[0.04]	[0.19]	[0.33]	[0.05]	[0.06]	[0.05]	[0.04]	[0.17]	[0.33]	[0.20]	[0.34]	[0.22]
Panel B												
σ^{NR}			2.79***						2.80***			
			[0.01]						[0.01]			
σ^{RD5}			5.95***						5.86***			
			[0.04]						[0.03]			
σ^{RD10}			3.15***						2.69***			
			[0.06]						[0.09]			
Panel C												
	Random effects are restricted to zero											
Panel D												
Implied share NR			0.646						0.645			
Implied share RD5			0.194						0.197			
Implied share RD10			0.093						0.091			
Implied share DK			0.067						0.067			
Month-year FE			no						yes			
Observations			172,548						172,548			

Notes: This table repeats the main analysis under the assumption of four response types: respondents who round to the next multiple of five (RD5) and ten (RD10), non-rounders (NR) and respondents who choose a "don't know" answer (DK). Dependent variable is short-run inflation expectations (px1). Specification 1 (2) excludes (includes) month-year fixed effects in the random effects multinomial logit model for type probabilities. Panel A reports estimates for interpersonal heterogeneity. Columns a, b and c report estimates for the parameterized mean of inflation expectations for type NR, RD5 and RD10, respectively. Columns d, e and f focus on the random effects multinomial logit model for type probabilities. Omitted category is type NR. Panel B displays type-specific estimates for the standard deviation of the normal distribution of inflation expectations. The individual effects are normalized to zero (Panel C). Panel D reports averages of model-implied unconditional type probabilities. For details see text. Standard errors in brackets. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Table F5: Model estimates for full month-year fixed effects

	Excluding month-year FE				Including month-year FE			
	Mean NR	Mean RD	P(T=RD)	P(T=DK)	Mean NR	Mean RD	P(T=RD)	P(T=DK)
	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Panel A								
Male	-0.12***	-1.09***	-0.66***	-0.99***	-0.10***	-1.18***	-0.64***	-0.99***
	[0.02]	[0.06]	[0.02]	[0.02]	[0.02]	[0.06]	[0.02]	[0.02]
Partner	0.01	0.36***	-0.02	-0.06**	0.04*	0.31***	-0.01	-0.06**
	[0.02]	[0.07]	[0.02]	[0.02]	[0.02]	[0.06]	[0.02]	[0.02]
Age	-0.01***	-0.02***	-0.00***	0.02***	-0.00***	-0.03***	0.00	0.02***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
College	-0.10***	-0.35***	-0.45***	-0.37***	0.07***	-0.49***	-0.38***	-0.43***
	[0.02]	[0.07]	[0.02]	[0.02]	[0.02]	[0.07]	[0.02]	[0.02]
1st income quartile	0.28***	1.57***	0.48***	1.22***	0.36***	1.52***	0.50***	1.22***
	[0.03]	[0.10]	[0.02]	[0.03]	[0.03]	[0.09]	[0.02]	[0.03]
2nd income quartile	0.05*	1.21***	0.19***	0.62***	0.21***	1.13***	0.28***	0.59***
	[0.03]	[0.09]	[0.02]	[0.03]	[0.03]	[0.09]	[0.02]	[0.03]
3rd income quartile	0.00	0.68***	0.10***	0.25***	0.08***	0.65***	0.13***	0.24***
	[0.02]	[0.08]	[0.02]	[0.03]	[0.02]	[0.08]	[0.02]	[0.03]
West	0.06**	-0.00	-0.09***	-0.06**	0.05**	0.02	-0.10***	-0.06**
	[0.03]	[0.09]	[0.02]	[0.03]	[0.02]	[0.08]	[0.02]	[0.03]
Northcentral	-0.02	-0.16**	-0.01	-0.17***	-0.03	-0.24***	-0.02	-0.16***
	[0.02]	[0.08]	[0.02]	[0.03]	[0.02]	[0.07]	[0.02]	[0.03]
Northeast	0.02	-0.04	0.11***	0.06**	-0.03	-0.04	0.08***	0.06**
	[0.03]	[0.08]	[0.02]	[0.03]	[0.03]	[0.08]	[0.02]	[0.03]
Constant	3.50***	5.73***	-0.43***	-3.10***	7.13***	8.19***	0.63***	-1.37***
	[0.04]	[0.12]	[0.03]	[0.05]	[0.34]	[0.59]	[0.17]	[0.23]
Panel B								
σ^{NR}	2.82***				2.67***			
	[0.01]				[0.01]			
σ^{RD}	5.86***				5.67***			
	[0.02]				[0.02]			
Panel C								
	Random effects are restricted to zero							
Panel D								
Implied share NR	0.652				0.650			
Implied share RD	0.281				0.283			
Implied share DK	0.067				0.067			
Month-year FE	no				yes			
Observations	172,548				172,548			

Notes: This table repeats the main analysis and adds month-year fixed effects in the equation of the parameterized mean of inflation expectations for types NR and RD (Equation 4). Response types are non-rounders (NR), rounders (RD) and respondents who choose a "don't know" answer (DK). Dependent variable is short-run inflation expectations (px1). Specification 1 (2) excludes (includes) month-year fixed effects in the random effects multinomial logit model for type probabilities (Equation 6). Panel A reports estimates for interpersonal heterogeneity. Columns a and b report estimates for the parameterized mean of inflation expectations for type NR and RD (Equation 4), respectively. Columns c and d focus on the random effects multinomial logit model for type probabilities. Omitted category is type NR. Panel B displays type-specific estimates for the standard deviation of the normal distribution of inflation expectations. The individual effects are normalized to zero (Panel C). Panel D reports averages of model-implied unconditional type probabilities. For details see text. Standard errors in brackets. *** p <0.01, ** p <0.05, * p <0.1.

Table F6: Model estimates for the full sample

	Excluding month-year FE				Including month-year FE			
	Mean NR	Mean RD	P(T=RD)	P(T=DK)	Mean NR	Mean RD	P(T=RD)	P(T=DK)
	(1a)	(1b)	(1c)	(1d)	(2a)	(2b)	(2c)	(2d)
Panel A								
Male	-0.12***	-1.03***	-0.65***	-0.95***	-0.13***	-1.03***	-0.63***	-0.96***
	[0.02]	[0.06]	[0.01]	[0.02]	[0.02]	[0.06]	[0.01]	[0.02]
Partner	0.03	0.36***	-0.02	-0.06***	0.04*	0.34***	-0.02	-0.06***
	[0.02]	[0.06]	[0.01]	[0.02]	[0.02]	[0.06]	[0.01]	[0.02]
Age	-0.01***	-0.03***	-0.00***	0.02***	-0.01***	-0.03***	-0.00**	0.02***
	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]	[0.00]
College	-0.22***	-0.51***	-0.47***	-0.44***	-0.23***	-0.53***	-0.37***	-0.47***
	[0.02]	[0.06]	[0.01]	[0.02]	[0.02]	[0.06]	[0.02]	[0.02]
1st income quartile	0.23***	1.30***	0.44***	1.24***	0.25***	1.29***	0.50***	1.26***
	[0.03]	[0.08]	[0.02]	[0.03]	[0.03]	[0.09]	[0.02]	[0.03]
2nd income quartile	0.00	1.02***	0.17***	0.60***	0.01	1.01***	0.27***	0.59***
	[0.03]	[0.08]	[0.02]	[0.03]	[0.03]	[0.08]	[0.02]	[0.03]
3rd income quartile	-0.01	0.49***	0.09***	0.27***	-0.00	0.47***	0.13***	0.28***
	[0.02]	[0.07]	[0.02]	[0.03]	[0.02]	[0.07]	[0.02]	[0.03]
West	0.08***	-0.02	-0.08***	-0.08***	0.07***	-0.01	-0.09***	-0.09***
	[0.02]	[0.08]	[0.02]	[0.02]	[0.02]	[0.08]	[0.02]	[0.02]
Northcentral	-0.06**	-0.14**	-0.02	-0.19***	-0.05**	-0.16**	-0.03	-0.18***
	[0.02]	[0.07]	[0.02]	[0.02]	[0.02]	[0.07]	[0.02]	[0.02]
Northeast	0.03	-0.04	0.11***	0.02	0.03	-0.03	0.08***	0.02
	[0.03]	[0.08]	[0.02]	[0.02]	[0.03]	[0.08]	[0.02]	[0.02]
Constant	3.89***	6.61***	-0.33***	-2.81***	3.87***	6.72***	0.00	-1.66***
	[0.04]	[0.11]	[0.03]	[0.04]	[0.04]	[0.11]	[0.08]	[0.10]
Panel B								
σ^{NR}	3.06***				3.05***			
	[0.01]				[0.01]			
σ^{RD}	6.11***				6.14***			
	[0.02]				[0.02]			
Panel C								
	Random effects are restricted to zero							
Panel D								
Implied share NR	0.632				0.635			
Implied share RD	0.290				0.288			
Implied share DK	0.077				0.077			
Month-year FE	no				yes			
Observations	228,151				228,151			

Notes: This table repeats the main analysis for the full sample, thus adding respondents who are interviewed only once. Response types are non-rounders (NR), rounders (RD) and respondents who choose a "don't know" answer (DK). Dependent variable is short-run inflation expectations (px1). Specification 1 (2) excludes (includes) month-year fixed effects in the random effects multinomial logit model for type probabilities (Equation 6). Panel A reports estimates for interpersonal heterogeneity. Columns a and b report estimates for the parameterized mean of inflation expectations for type NR and RD (Equation 4), respectively. Columns c and d focus on the random effects multinomial logit model for type probabilities. Omitted category is type NR. Panel B displays type-specific estimates for the standard deviation of the normal distribution of inflation expectations. The individual effects are normalized to zero (Panel C). Panel D reports averages of model-implied unconditional type probabilities. For details see text. Standard errors in brackets. *** p <0.01, ** p <0.05, * p <0.1.

References

Ameriks, J., G. Kézdi, M. Lee, and M. D. Shapiro (2018). Heterogeneity in expectations, risk tolerance, and household stock shares: The attenuation puzzle. *NBER Working Paper, No. 25269*.

Armantier, O., W. Bruine de Bruin, S. Potter, G. Topa, W. van der Klaauw, and B. Zafar (2013). Measuring inflation expectations. *Annual Review of Economics 5*(1), 273–301.

Bachmann, R., S. Elstner, and E. R. Sims (2013). Uncertainty and economic activity: Evidence from business survey data. *American Economic Journal: Macroeconomics 5*(2), 217–49.

Baker, S. R., N. Bloom, and S. J. Davis (2016). Measuring economic policy uncertainty. *The Quarterly Journal of Economics 131*(4), 1593–1636.

Bernanke, B. S. (2007). Inflation expectations and inflation forecasting. Speech at the Monetary Economics Workshop of the National Bureau of Economic Research Summer Institute, Cambridge, Massachusetts, July 10.

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

Blanchard, O., G. Dell'Ariccia, and P. Mauro (2010). Rethinking macroeconomic policy. *Journal of Money, Credit and Banking 42*(s1), 199–215.

Blanchflower, D. G. and C. MacCoille (2009). The formation of inflation expectations: An empirical analysis for the UK. *NBER Working Paper, No. 15388*.

Bloom, N. (2009). The impact of uncertainty shocks. *Econometrica 77*(3), 623–685.

Bruine de Bruin, W., W. van der Klaauw, J. S. Downs, B. Fischhoff, G. Topa, and O. Armantier (2010). The effect of question wording on reported expectations and perceptions of inflation. *Federal Reserve Bank of New York, Staff Report No. 443*.

Bryan, M. F. and G. Venkatu (2001a). The curiously different inflation perspectives of men and women. *Federal Reserve Bank of Cleveland, Economic Commentary, November issue*.

Bryan, M. F. and G. Venkatu (2001b). The demographics of inflation opinion surveys. *Federal Reserve Bank of Cleveland, Economic Commentary, October issue*.

Curtin, R. T. (1982). Indicators of consumer behavior: The University of Michigan Surveys of Consumers. *Public Opinion Quarterly 46*(3), 340–352.

Curtin, R. T. (1996). Procedure to estimate price expectations. *University of Michigan: Survey of Consumers* (available at https://data.sca.isr.umich.edu/survey-info.php [accessed August 17, 2018]).

Drerup, T., B. Enke, and H.-M. Von Gaudecker (2017). The precision of subjective data and the explanatory power of economic models. *Journal of Econometrics 200*(2), 378–389.

Galí, J. (2015). *Monetary policy, inflation, and the business cycle: An introduction to the new Keynesian framework and its applications*. Princeton University Press.

Georganas, S., P. J. Healy, and N. Li (2014). Frequency bias in consumers' perceptions of inflation: An experimental study. *European Economic Review 67*, 144–158.

Gould, W., J. Pitblado, and W. Sribney (2006). *Maximum likelihood estimation with Stata*. Stata press.

Heiss, F., M. Hurd, M. van Rooij, T. Rossmann, and J. Winter (2019). Dynamics and heterogeneity of subjective stock market expectations. *mimeo*.

Hobijn, B., K. Mayer, C. Stennis, and G. Topa (2009). Household inflation experiences in the US: a comprehensive approach. *Federal Reserve Bank of New York, Working Paper Series No. 09-19*.

Hudomiet, P., G. Kézdi, and R. J. Willis (2011). Stock market crash and expectations of American households. *Journal of Applied Econometrics 26*(3), 393–415.

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

Jeliazkov, I. and A. Lloro (2011). Maximum simulated likelihood estimation: Techniques and applications in economics. In S. Koziel and X.-S. Yang (Eds.), *Computational Optimization, Methods and Algorithms*, Chapter 5, pp. 85–100. Springer.

Jonung, L. (1981). Perceived and expected rates of inflation in Sweden. *American Economic Review 71*(5), 961–968.

Jurado, K., S. C. Ludvigson, and S. Ng (2015). Measuring uncertainty. *American Economic Review 105*(3), 1177–1216.

Kleinjans, K. J. and A. van Soest (2014). Rounding, focal point answers and nonresponse to subjective probability questions. *Journal of Applied Econometrics 29*(4), 567–585.

Malmendier, U. and S. Nagel (2016). Learning from inflation experiences. *The Quarterly Journal of Economics 131*(1), 53–87.

Malmendier, U., S. Nagel, and Z. Yan (2017). The making of hawks and doves: Inflation experiences on the FOMC. *NBER Working Paper, No. 23228*.

Manski, C. F. (2004). Measuring expectations. *Econometrica 72*(5), 1329–1376.

Manski, C. F. (2018). Survey measurement of probabilistic macroeconomic expectations: progress and promise. *NBER Macroeconomics Annual 32*(1), 411–471.

Niederle, M. and L. Vesterlund (2007). Do women shy away from competition? Do men compete too much? *The Quarterly Journal of Economics 122*(3), 1067–1101.

Pusara, M. and C. E. Brodley (2004). User re-authentication via mouse movements. *Proceedings of the 2004 ACM workshop on visualization and data mining for computer security*, 1–8.

Ranyard, R., F. Del Missier, N. Bonini, D. Duxbury, and B. Summers (2008). Perceptions and expectations of price changes and inflation: A review and conceptual framework. *Journal of Economic Psychology 29*(4), 378–400.

Revelt, D. and K. Train (1998). Mixed logit with repeated choices: Households' choices of appliance efficiency level. *Review of Economics and Statistics 80*(4), 647–657.

Rossi, B. and T. Sekhposyan (2015). Macroeconomic uncertainty indices based on nowcast and forecast error distributions. *American Economic Review 105*(5), 650–55.

Rossi, B., T. Sekhposyany, and M. Souprez (2017). Understanding the sources of macroeconomic uncertainty. *Barcelona GSE Working Paper, No. 920*.

Sims, C. A. (2009). Inflation expectations, uncertainty and monetary policy. *Bank for International Settlements Working Paper, No. 275*.

Souleles, N. S. (2004). Expectations, heterogeneous forecast errors, and consumption: Micro evidence from the Michigan Consumer Sentiment Surveys. *Journal of Money, Credit and Banking 36*(1), 39–72.

Train, K. E. (2003). *Discrete choice methods with simulation.* Cambridge University Press.

van Rooij, M., A. Lusardi, and R. Alessie (2011). Financial literacy and stock market participation. *Journal of Financial Economics 101*(2), 449–472.

Zheng, N., A. Paloski, and H. Wang (2011). An efficient user verification system via mouse movements. *Proceedings of the 18th ACM conference on computer and communications security*, 139–150.