That's what she said: An Empirical Investigation on the Gender Gap in Inflation Expectations*

Lovisa Reiche†

March 20, 2024

Abstract

The gender gap in inflation expectations, i.e., women reporting systematically higher inflation expectations in consumer surveys, is a well-established phenomenon. The disparity has been attributed to women's greater involvement in grocery shopping and exposure to volatile food prices. I evaluate this hypothesis using a Bayesian learning framework, which suggests that signal volatility increases mean expectations only whenever the prior is flat. Such a flat prior could be caused by low financial confidence, which is more prevalent in women. Using data from the "Bundesbank Online Panel – Households", I find that grocery shopping increases expectations only for a low confidence sample and including a control for financial confidence closes the gender gap fully. This observation has significant macroeconomic implications, including potential gender-based disparities in retirement investment and monetary policy targeting.

Keywords Consumer Inflation Expectations, Gender, Financial Literacy, Experience

JEL Codes E31, E52, E71

Repository location GitHub page EmpiricalGenderGap

*I am grateful to Michael McMahon, Alexandre Kohlhas and Ryan Rholes for their supervision and advice. Further, I thank Martin Weidner for his feedback on the econometric modelling in this paper. This work has also benefited from the feedback of conference and seminar participants at the University of Copenhagen, the Royal Economic Society, Bocconi University, the University of Bamberg, the National Bank of Belgium, Norges Bank and the CEPR Symposium. Declarations of interest: none.

†University of Oxford, Department of Economics, Manor Road Building, OX1 3UQ, Oxford, UK
Email: lovisa.reiche@economics.ox.ac.uk

1 Introduction

The gender gap in inflation expectations is an established phenomenon. Using a 1977 survey of Swedish households, Jonung (1981) first found that women had significantly higher inflation perceptions and explained his findings through the role of "lived experiences". Women traditionally engaged more in food shopping, and food prices rose faster in the 70s. Since then, it has been shown that the gap persists over time and spans a wide range of geographies, survey designs and experimental settings.[1] Recent estimates reach up to 2.44 percentage points in raw data (Armantier et al., 2016) and 1.45 percentage points when controlling for standard demographics (Brischetto & de Brouwer, 1999).[2] The gap is thus of substantial size given inflation targets around 2%. The dominant interpretation is still that traditional gender roles may cause women to engage more often in grocery shopping (D'Acunto et al., 2021). This exposes women to more volatile price signals than men. As volatility can lead to overestimation due to a disproportional focus on price increases (Dräger et al., 2014), women could on average have higher expectations.

In this paper, I show that this experience hypothesis is insufficient to explain the gender gap and provide an alternative explanation which focuses on financial confidence. To formalise my hypothesis, I employ a Bayesian framework with log-normally distributed signals and a conjugate log-normal prior. As the noise of signals increases (i.e. as the individual observes more volatile prices, for instance due to grocery price exposure), expected value as well as variance of the posterior distribution increase, suggesting higher average expectations for women in traditional gender roles. However, signal volatility only increases mean expectations whenever the prior is flat. This is an important condition for the experience channel to work. One potential cause for a flat prior may be financial confidence, which is linked to higher uncertainty around point forecasts. Hence, my alternative financial confidence hypothesis is that the gender gap is a composite of traditional gender norms amplified by a gender gap in financial confidence (Bucher-Koenen et al., 2014, 2021).

This hypothesis is also suited to explain other observations in the literature, namely that the gender gap is largest in times when inflation perceptions are close to actual inflation as

[1] Brischetto and de Brouwer (1999) for Australia; Bryan and Venkatu (2001) for a survey on Ohioan consumers; Palmqvist and Strömberg (2004) for Sweden, Pfajfar and Santoro (2008) for the US Michigan Survey; Blanchflower and Maccoille (2009) for the UK; Leung (2009) for New Zealand; Del Giovane et al. (2008) and Corduas (2022) for Italy; Bruine De Bruin et al. (2010) and Armantier et al. (2016) for the RAND American life panel; Arioli et al. (2017) and Lindén (2015) and Reiche and Meyler (2022) for a range of EU countries; D'Acunto et al. (2021) for the Chicago Booth Expectations and Attitudes Survey; Dräger and Nghiem (2020) for Germany and Abildgren and Kuchler (2021) for Denmark are a non-exhaustive list of authors mentioning this empirical finding. I include a summary of the literature in Appendix Table 8.

[2] Appendix F.1 shows that while higher age, income, and education reduce the gender gap (see interaction terms in Table 10, even in combination they only close it for very high values (see Figure F.1).

men have a lower heterogeneity in expectations then (Corduas, 2022), and that in information provision experiments gender gaps diminish after any numerical anchor has been provided (Armantier et al., 2016). My results suggest that while experiences matter, heterogeneity along the gender dimension may be driven more by differential financial confidence (Bucher-Koenen et al., 2014, 2021). This is important to emphasize, as it suggests that other demographic gaps may also be due to these more fundamental gaps in understanding of financial variables.

I conduct an empirical analysis using data from the Bundesbank Online Panel – Households (BOP-HH, 2019-2022), which measures respondents household responsibilities (including grocery shopping) and includes a financial literacy test. My empirical analysis confirms that grocery shopping alone is insignificant in explaining inflation expectations and thus the gender gap in them. In contrast, my analysis shows that grocery shopping increases mean expectations for individuals with low financial confidence, while it decreases expectations for high confidence observations. Since women more often show lower levels of financial literacy and financial confidence, the effect of grocery shopping gets amplified. This main empirical result is summarized in Figure 1. If one considers both factors in isolation, the impact of financial confidence exceeds that of experience.

Further, I show evidence supporting the financial confidence channel. In the Bayesian framework adopted, lower confidence is modelled through a higher prior variance, which reduces confidence in the prior relative to the signals received. This gives rise to an increase in expected value and variance of the posterior, similarly as the adjustment of the signal volatility. Hence, the hypothesis rests on the idea that the gap in means is driven by a heavier right skew of the female distribution. I confirm that (a) women in my data show lower financial confidence and (b) low financial confidence shifts expectations to the right. Further, I demonstrate, that there is no gender gap in inflation expectations in the high confidence sample and consequentially removing outliers beyond the 80th percentile closes the gender gap fully. When possible, I complement the German data with evidence from the Survey of Consumer Expectations by the Federal Reserve Bank of New York (2013-2020) and the Michigan Survey of Consumers (1978-2022), both set in the US. This provides external validity for my results geographically and historically.

As a robustness exercise, I test two implications of the pure experience hypothesis, whose rejection support my hypothesis: First, I test whether the gender gap persists for singles, who can be assumed to engage symmetrically in grocery shopping. I reject the hypothesis that single men have the same inflation expectations as single women, controlling for a range of demographics. This is plausible under the financial confidence hypothesis, as grocery shopping only matters in interaction with financial confidence and gender gaps in the latter exist for singles as well as non-singles. Second, I test whether periods of high food price inflation are

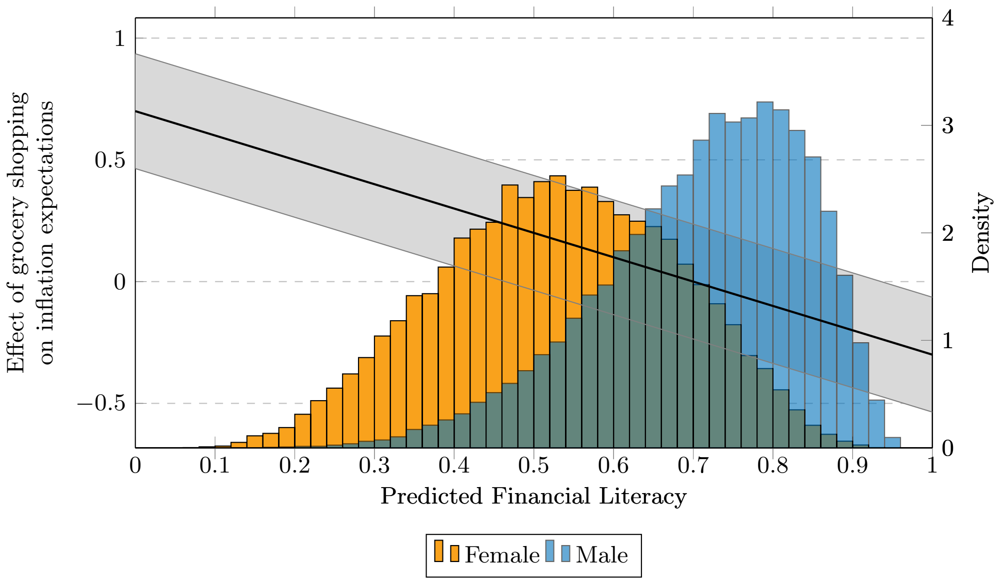


Figure 1: The effect of grocery shopping involvement on inflation expectations for different levels of financial confidence

Notes: Figure 1 plots the predicted effect of grocery shopping involvement on inflation expectations for different levels of financial confidence in the black line ($f(x) = 0.7 - x$). The full regression results are shown in Table 3, column (5). The grey area indicates 95% confidence bands (standard error: 0.1145). How grocery shopping is measured will be explained in Section 3.2.1 and financial confidence in Section 3.2.2. The histograms show the density of the male (blue) and female (orange) distribution of financial confidence scores.

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; own calculations

correlated with an increase in the gender gap. Again, I reject the hypothesis that there is a positive correlation between the food inflation gap (the difference between food and total inflation) and the gender gap. For both exercises, I use BOP-HH as well as SCE and MSC as the latter two include a much longer time-series. Thus, I provide additional evidence that the experience hypothesis cannot account for the difference between male and female inflation expectations alone.

My results speak to broadly two areas in the literature: One that explains heterogeneity in inflation expectations through heterogeneity in signals, and another one that emphasizes the role of interpretation. Initially emphasized by Jonung (1981) and later formalized by the works of Malmendier and Nagel (2016), heterogeneity in experiences (signals) may matter and explain systematic demographic differences. With regard to the gender gap, D'Acunto et al. (2021) show using intra-household data of heterosexual, married couples that the gender gap is indeed most prevalent within households when men do not partake in grocery shopping, while in households with an equal share, the gap diminishes. In addition, heterogeneity may arise due to differential ability to process signals received. Bruine De Bruin et al. (2010) suggest that individuals with lower financial literacy rely largely on personal experiences, while financially literate individuals interpret inflation as a more abstract macro concept. Indeed, there is a strong connection between cognitive abilities and inflation expectations (Burke & Manz, 2014; D'Acunto et al., 2019a, 2019b). I contribute to both literatures empirically as well as theoretically. My Bayesian framework captures the effects of the noisy information literature and shows that when the prior is log-normal, making signals noisier can increase the posterior without introducing biases. However, it links this to the interpretation literature by pointing out that for signal volatility to matter, priors need to be flat. To my knowledge, I am the first to integrate financial confidence modelled as prior precision into a Bayesian framework to show that higher uncertainty about an outcome can indeed drive average expectations upwards and is an important condition for the pure experience channel to operate. Indeed, this can be reconciled with the findings of D'Acunto et al. (2021). There are reasons to believe that women in "traditional" households differ from those who share household chores equally, for instance, in their financial confidence. Further, I run empirical tests on both channels with a focus on the gender gap in inflation expectations.

This paper is structured as follows: I first introduce the Bayesian framework (Section 2) and then data sets and measurements of experience and financial confidence employed in this study (Section 3). I present the key empirical result of this paper in Section 4. Section 5 provides evidence for the financial confidence channel, and Section 6 tests two implications of the experience hypothesis as robustness exercise. Finally, Section 7 highlights avenues of future research and Section 8 concludes.

2 Bayesian Framework

I start with a Bayesian framework which will be used to explain how the two channels hypothesized to cause the gender gap in inflation expectations, experience and confidence, can affect an individual's point forecast. I model differences in experiences, for instance through differences in grocery shopping activity, as differences in signal volatility. Those who visit grocery stores frequently observe more volatile prices as food prices are fundamentally more volatile than the core component of the consumption basket. On the other hand, I capture differences in financial confidence as differences in prior precision. Individuals with less confidence about own financial literacy will place less weight on their own forecast. The framework also highlights how these two channels can interact with each other. Intuitively, if an individuals prior is imprecise such that little weight is placed on own initial forecasts, signals matter more as they become the dominant source for information about inflation. I first present the basic framework and then explore the impact of heterogeneity in signal and prior precision.

2.1 Basic framework

There is a population of agents N of which a share $\lambda \in [0,1]$ is of type F and a share $1-\lambda$ is of type M. Let θ denote inflation 12 months ahead, an unknown random variable. Agent j of type $i = \{F, M\}$'s prior belief about future inflation is assumed to follow a log-normal distribution:

$$\theta \sim \text{LogNormal}\left(\mu_j^{i,0}, \frac{1}{\tau_j^{i,0}}\right)$$

Hence, before receiving any signals, $ln(\theta)$ is believed to be $\mu_j^{i,0}$ on average with a precision of $\tau_j^{i,0}$. The prior is heterogeneous across agents as agents receive a shock δ_j^i to their prior precision. The shock is truncated-normally distributed with a common mean δ^i for all individuals of type i.[3]

$$\tau_j^{i,0} = \tau^0 - \delta_j^i \qquad \delta_j^i \sim \psi(\bar{\delta}^i, \bar{\sigma}_\delta, a_\delta, b_\delta) \qquad a_\delta = -\infty \text{ and } b_\delta < \tau^0$$

Lower prior precision, i.e. a large δ_j^i in this framework could be caused by lower financial confidence which leads to higher uncertainty around individual point forecasts. The framework allows for prior precision to be lower on average for agents of one type relative to the other, such that gaps between both groups can arise. This will be discussed in more detail below. The prior parameter $\mu_j^{i,0}$ is defined such that the mode is constant with respect to changes in

[3]The truncated normal distribution is chosen since $\tau_j^{i,0} > 0$. The upper bound b_δ will be close to but below τ^0 and $\bar{\delta}^i$ and σ_δ small such that the distribution is approximately normal. This implies that $\delta^i \sim \bar{\delta}^i$, where the latter is the mean of the parent general normal PDF.

δ^i_j:

$$\mu^{i,0}_j = \mu^0 + \frac{\delta^i_j}{\tau^0(\tau^0 - \delta^i_j)}$$

In addition to their priors, agent j of type i receives a signal x^i_j about future inflation. Signals are noisy, reflecting heterogeneity in inflation experiences given by heterogeneous consumption baskets. I model the signals to be unbiased and follow a log-normal distribution.

$$x^i_j|\theta \sim \text{LogNormal}\left(\mu^{i,x}_j, \frac{1}{\tau^{i,x}_j}\right)$$
$$\text{where } \mu^{i,x}_j = \ln(\theta) - \frac{1}{2\tau^{i,x}_j}$$

Personal shopping experience is reflected in heterogeneous shocks to signal precision ϵ^i_j. As for shocks to prior precision, ϵ^i_j is truncated-normally distributed with a common mean ϵ^i for all individuals of type i. If an individual j typically shops for goods with volatile prices (such as groceries) she will receive signals with lower precision. As for shocks to the prior precision, I make the assumption that on average shocks to the signal precision of individuals of type F may be different to those of type M.[4]

$$\tau^{i,x}_j = \tau^x - \epsilon^i_j \qquad \epsilon^i_j \sim \psi(\bar{\epsilon^i}, \bar{\sigma}_\epsilon, a_\epsilon, b_\epsilon) \qquad a_\epsilon = -\infty \text{ and } b_\epsilon < \tau^x$$

Agents update their beliefs about θ based on the observed signal using Bayes' rule (see Appendix B.2). The agent's posterior beliefs are:

$$\theta|x^i_j \sim \text{LogNormal}\left(\hat{\mu}^i_j, \frac{1}{\hat{\tau}^i_j}\right)$$

where $\hat{\mu}^i$ represents the mean of the logged posterior inflation expectations of the agent and $\hat{\tau}^i$ the corresponding precision given by:

$$\hat{\mu}^i_j = \frac{\mu^{i,0}_j\tau^{i,0}_j + \ln(x^i_j)\tau^{i,x}_j + \frac{1}{2}}{\tau^{i,x}_j + \tau^{i,0}_j} \tag{1}$$

$$\hat{\tau}^i_j = \tau^{i,0}_j + \tau^{i,x}_j \tag{2}$$

The beliefs depend on the realisation of x^i_j. On average, $\ln(x^i_j) = \mu^{i,x}_j$. Once the posterior distribution is defined, it is easy to compute moments of the distribution. The expected value

[4]As above, the truncated normal distribution is chosen since $\tau^{i,x}_j > 0$. The upper bound b_ϵ will be close to but below τ^x and $\bar{\epsilon^i}$ and σ_ϵ small such that the distribution is approximately normal. This implies that $\epsilon^i \sim \bar{\epsilon^i}$, where the latter is the mean of the parent general normal PDF.

of θ under the posterior distribution is simply the mean of the posterior distribution, which can be computed as:

$$\mathbb{E}^i(\theta|x_j^i) = e^{\hat{\mu}^i + \frac{1}{2\hat{\tau}^i}} = e^{\frac{\tau_j^{i,0}\mu_j^{i,0} + \tau_j^{i,x}\mu_j^{i,x} + 1}{\tau_j^{i,0} + \tau_j^{i,x}}} \tag{3}$$

The variance of θ under the posterior distribution is equivalently given by:

$$\mathrm{Var}^i(\theta|x_j^i) = e^{2\hat{\mu}^i + \frac{1}{\hat{\tau}^i}} \left(e^{\frac{1}{\hat{\tau}^i}} - 1\right) = e^{2\times\frac{\tau_j^{i,0}\mu_j^{i,0} + \tau_j^{i,x}\mu_j^{i,x} + 1}{\tau_j^{i,0} + \tau_j^{i,x}}} (e^{\frac{1}{\tau_j^{i,x} + \tau_j^{i,0}}} - 1) \tag{4}$$

2.2 Features of the framework

In this section I will highlight some features of the framework before analyzing comparative statics.

2.2.1 Log-normal prior and signal

The log-normal prior is chosen because it is bounded at zero and features a heavy tail. This matches observations in the data: (1) there appears to be a zero lower bound in inflation expectations (Gorodnichenko & Sergeyev, 2021);[5] (2) a majority of agents have expectations in line with central bank targets, but there are possible tail events to the right; (3) the observed cross-section of inflation expectations is approximately log-normal (see Figure B.1 in the Appendix). Similarly, the log-normal signal is motivated by (1) behavioral biases that imply that consumers pay more attention to price increases rather than price decreases (Dräger et al., 2014), requiring a heavy tailed asymmetric distribution, (2) allowing for analytical tractability due to the conjugate prior. The asymmetrz of the log-normal distribution is a key element of the results discussed below.

2.2.2 Unbiasedness of signals

The parameter $\mu_j^{i,x}$ of the signal distribution is specified to be increasing in $\tau_j^{i,x}$ to ensure that the signals are unbiased. This is to show that signal volatility alone can affect mean expectations, i.e. purely by observing more volatile grocery prices an individual's inflation expectation can increase. Relaxing this assumption increases the effects discussed in the next section.

[5] In the surveys used in the empirical section of this paper reaches from 0.002% in the BOP-HH to 3.24% in the MSC and 7.04% in the SCE. This is in line with estimates from the authors for a range of EU countries, the US and Japan. This feature is a feature of the chosen functional form, relaxing the zero lower bound would not impact the results as long as the asymmetry is retained.

2.2.3 Mode correction

The prior parameter $\mu_j^{i,0}$ is defined as above to ensure that despite shocks to the prior precision, the mode of the prior distribution remains constant. I thus allow for anchoring, in this case for instance at inflation targets. The adjustment is required due to the asymmetry of the chosen prior, and has no qualitative implications for the results of this analysis. As for unbiasedness, if removed, the effects of the higher variance on expected inflation would be even larger.

2.3 Channels of the gender gap

My framework can be used to explain heterogeneity in observed point forecasts, in particular those between men and women. Differences in the average noisiness of signals received as well as priors between the two types of agents in the framework give rise to gaps in mean expectations.

Begin with the role of shopping experience. I assume that the composition of j's shopping basket may affect the signal precision parameter $\tau_j^{i,x}$ through the individual shock component ϵ_j^i. Shopping for baskets with volatile-priced goods, such as groceries, increases the average shock to the precision as volatile pricing makes perceiving the true level of inflation difficult for grocery shoppers. In this application, the signal remains unbiased, such that groceries are assumed to have the same level of inflation as other goods. Below I show how the moments of the posterior distribution vary with ϵ_j^i. The computations can be found in Appendix B.3.

$$\frac{d\text{Var}(\theta|x_j^i)}{d\epsilon_j^i} > 0 \tag{5}$$

$$\frac{d\mathbb{E}(\theta|x_j^i)}{d\epsilon_j^i} > 0 \iff \mu_j^{i,0} + \frac{1}{\tau_j^{i,0}} > \ln(x_j^i) \tag{6}$$

Decreasing signal precision will always reduce certainty in posterior expectations and may under certain conditions also increase posterior expectations. Notice that the condition in 6 relaxes when the prior is flat, i.e., prior precision $\tau_j^{i,0}$ is small, such that the agent relies more on the signals received.

In the framework, there are two types of agents. While signals differ on an individual level, the framework captures the possibility that one group of agents on average receives less precise signals than the other. For instance, if one group of individuals, say group F, systematically shops for goods with more volatile inflation (i.e., food) the groups will experience less precise signals: $\epsilon^M < \epsilon^F$. Decreasing signal-precision of group F relative to M even when the signal remains unbiased will result in an increase in the variance of the expectations of group F (see above) and may increase F's average point forecasts ($\hat{\theta}^F = \frac{1}{\lambda N}\sum_{j,i=F}\mathbb{E}(\theta|x_j^i)$) relative to those of group M ($\hat{\theta}^M = \frac{1}{(1-\lambda)N}\sum_{j,i=M}\mathbb{E}(\theta|x_j^i)$) if condition 6 holds on average. However, note that the condition depends on prior precision.

Proposition 2.1 *If $\bar{\epsilon^M} < \bar{\epsilon^F}$ it will be that $\hat{\theta}^F > \hat{\theta}^M$ whenever $\mu^0 + \frac{\delta^i+\tau^0}{\tau^0(\tau^0-\delta^i)} > \ln(\theta) - \frac{1}{2(\tau^x-\epsilon^i)}$ for both types.*

In summary, under the assumption of a log-normal signal and its conjugate prior, increases in the noise of the environment alone can indeed increase the expected value of the posterior distribution for an individual respondent. Further, if there are systematic differences in signal precision between two types of agents, the one with the lower precision is predicted to have higher expected values on average. This captures and formalizes the argument of the experience hypothesis that women observing higher volatility through higher observed food prices have increased inflation expectations. However, this is only true if the prior precision is small. In contrast, a tight prior may cause mean expectations to decrease when the environment is noisier. Thus, it is important to analyze the consequences of prior heterogeneity, which may be caused by different levels of financial confidence.

Financial confidence may impact the parameters of the prior distribution $\mu_j^{i,0}$ and $\tau_j^{i,0}$ through shocks to δ_j^i, such that agents with lower confidence have a flatter prior. It has been shown that women have lower confidence about their own financial literacy (Bucher-Koenen et al., 2021).[6] Low confidence may imply higher uncertainty over point forecasts, such as expected inflation. This can be modelled as lower precision. Intuitively, this reflects that individuals with lower financial confidence may have a less formalized idea of price changes when observing prices. Subsequently, I will discuss the case when $\delta^F > \delta^M$, i.e. negative shocks to precision are on average larger for agents of type F.

As before, I first show how the moments of the posterior distribution vary with individual shocks to precision δ_j^i. The computations can be found in Appendix B.4.

$$\frac{d\mathrm{Var}(\theta|x_j^i)}{d\delta_j^i} > 0 \tag{7}$$

$$\frac{d\mathbb{E}(\theta|x_j^i)}{d\delta_j^i} > 0 \iff \ln(x_j^i) + \frac{2}{\tau_j^{i,x}} + \frac{1}{\tau^0} > \mu_0 \tag{8}$$

Hence, the introduction heterogeneity in priors modelled as shocks to prior precision δ_j^i can also give rise to heterogeneous expectations when signals received are identical. Just as

[6] The authors show that women perform equally well in financial literacy tests when no "don't know" answer is provided, but worse when such option is not available.

the condition in 6 hinges on the prior being flat, the condition in 8 requires signals to be imprecise.

As before, if one group of individuals, say F, systematically has a lower prior precision such that $\delta^M < \delta^F$, it will be that this group has more dispersed expectations and average expectations are larger if condition 8 holds. This is an alternative channel to explain the gender gap.

Proposition 2.2 *If $\delta^M < \delta^F$ it will be that $\hat{\theta}^F > \hat{\theta}^M$ whenever $\ln(\theta) - \frac{1}{2(\tau^x - \epsilon^i)} + \frac{2}{\tau^x - \epsilon^i} + \frac{1}{\tau^0} > \mu_0$*

In the Bayesian framework with log-normal priors and signals, noisy environments caused by grocery shopping and low financial confidence can individually be a cause for higher inflation expectations. Moreover, they interact. The framework shows that noisy signals increase expectations when priors are flat. Simultaneously, low financial confidence (modelled as flat priors) increases expectations when signals are imprecise. This shows that the effects are complementing each other. Taking both conditions in Equations 6 and 8 together, I arrive at a joint condition:

$$\frac{\delta^i + \tau^0}{\tau^0(\tau^0 - \delta^i)} > -\frac{2}{\tau_x - \epsilon_i} - \frac{1}{\tau_0} \tag{9}$$

Notice that the left-hand side is increasing in δ^i_j and the right-hand side is decreasing in ϵ^i_j, a complementary effect. Further, the right hand side is always negative while the right hand side is always positive. This implies that at any point it must be that either the effect of signal volatility or the effect of prior imprecision is present.

Proposition 2.3 *The effect of shocks to prior and signal precision, δ^i_j and ϵ^i_j, on inflation expectations is complementary. It will always be that if $\frac{d\mathbb{E}(\theta|x^i_j)}{d\delta^i_j} \leq 0$ then $\frac{d\mathbb{E}(\theta|x^i_j)}{d\epsilon^i_j} > 0$ and if $\frac{d\mathbb{E}(\theta|x^i_j)}{d\epsilon^i_j} \leq 0$ then $\frac{d\mathbb{E}(\theta|x^i_j)}{d\delta^i_j} > 0$*

The framework is thus well suited to explain the interaction of the two channels hypothesized to explain the gender gap in inflation expectations. It shows that if women on average observe more volatile price signals through greater involvement in grocery shopping, they may have higher expectations than men. Similarly, if women on average have lower confidence in their own forecasts they could also have higher expectations. The framework shows that both channels complement each other: observing volatile prices will increase expectations when the individual is less confident. This makes sense intuitively if those confident about their own financial literacy rely more on aggregate news while those with lower confidence rely on their day-to-day experiences.

The remainder will show this complementarity empirically. I show that there is an interaction effect between grocery shopping and financial confidence when predicting inflation expectations. Further, I show that there is strong evidence for the financial confidence channel while grocery shopping experience, i.e. observing more volatile signals, in isolation may not be able to explain the gap between male and female expectations.

3 Data

3.1 Surveys as data source

My primary data source is the Bundesbank Online Panel of German consumers from April 2019 until September 2022.[7] This survey is particularly suited to analyze the gender gap in inflation expectations because it contains individual level data on financial literacy and household responsibilities including grocery shopping, thus allowing me to test both hypotheses on the same individuals. Data for the BOP-HH has been collected regularly since April 2020 to investigate the economic expectations of German consumers. I use data until September 2022. In addition, there are three months of pilot phase from April-June 2019. Approximately two thousand participants are initially drawn randomly from a larger pool of candidates recruited via telephone. The participants complete an online survey with various questions ranging from macroeconomic assessments and expectations to political issues. Demographic characteristics are recorded by self-assessment; therefore, the terms "female" and "women" in my analysis refer to self-identified gender. The survey includes the option to not choose a binary gender and few candidates do so, these responses are excluded from this analysis. I also exclude all participants who do not give an inflation point forecast. In the subsequent waves, new participants join while some former participants drop out, resulting in a significant turnover. Inflation expectations in the survey are measured quantitatively. At first, individuals are presented with a short definition of inflation[8] and are asked if they expect inflation or deflation in the coming 12 months. Subsequently, they indicate their anticipated inflation or deflation rate numerically. Answers are limited to a range of 0 to 100. Additionally, the survey elicits uncertainty around the point forecast through a probabilistic question.

I complement this survey using two established consumer surveys, namely the Michigan Survey of Consumers in the US from June 1978 until January 2023 (MSC);[9] and the Federal Reserve Bank of New York Survey of Consumer Expectations in the US from June 2013 until November 2020 (SCE).[10] Adding these surveys allows me to explore a longer time horizon than

[7]DOI: https://doi.org/10.12757/Bbk.BOPHH.202204.01. Disclaimer: The results published and the related observations and analysis may not correspond to results or analysis of the data producers.

[8]Inflation is the percentage increase in the general price level. It is mostly measured using the consumer price index. A decrease in the price level is generally described as "deflation".

[9]Source: University of Michigan, Survey Research Center, Surveys of Consumers, available at https://data.sca.isr.umich.edu/.

[10]Disclaimer: © 2013-2020 Federal Reserve Bank of New York (FRBNY). The SCE questions are available

Survey	Time/Place	Participants	Wording
BOP-HH	Apr.2020-Sep.2022, DE	2000/month	*inflation/deflation* + (definition) from 0-100 + Financial literacy test + Household responsibilities
SCE	Jun.2013-Nov.2020, US	1200/month	*inflation/deflation* from 0-100 + Financial literacy test
MSC	Jan.1978-Dec.2022, US	500/month	*prices in general* from 0-95, probing $> 5\%$

Table 1: Features of the three surveys

the short period of the BOP-HH, which was also heavily influenced by the Covid-19 pandemic, and provides external validity by bench-marking results to the US. Further, including the SCE adds robustness to my computation of financial confidence and addresses internal validity concerns. BOP, SCE and MSC are of panel structure and provide microdata. Details of the other surveys can be found in the Appendix C. All surveys are summarized in Table 1.

3.2 Experience and financial confidence in data

While the BOP-HH is the only dataset that contains information on the respondents' grocery shopping experience, both SCE and BOP-HH contain financial literacy questions that help me compute a measure of confidence.

3.2.1 Measuring experience

Inference of differentiated experience is possible in the BOP-HH due to a question regarding household responsibilities introduced in April 2021, namely everyday purchases (*shop_groceries*), major purchases (*shop_major*), meal preparation (*prep_meals*) and financial decisions (*decide_finance*). Respondents indicate if they are not involved in the task (1), engage jointly with other household members (2) or are solely responsible for all the work (3).[11] Since the question is only asked for the first time an individual participates in the survey, I assume that household chores remain constant over time in the panel. Further, the variable is only asked for non-singles. I make the assumption that singles will be responsible in their households for all four activities, and are thus given a score of 3 in each task. The analysis will show both full sample and non-single sub-samples separately and where household experiences are included focus on the households with more than one member.

without charge at http://www.newyorkfed.org/microeconomics/sce and may be used subject to license terms posted there. FRBNY did not participate in or endorse this work, and FRBNY disclaims any responsibility or legal liability for the administration of the survey and the analysis and interpretation of data collected.

[11] Assigned numbers are different from the original questionnaire.

Using this measure, I can confirm gender gaps in household responsibilities: Table 2 compares experience measures across male and female samples. As anticipated, the female respondents appear significantly more involved in grocery shopping and meal preparation but less involved in financial planning in households that involve more than one member. Major purchases are balanced between the samples such that no clear gender roles emerge. Men in the data are less likely to live alone.

Table 2: Confirming the presence of traditional gender roles in the BOP-HH

	shop_groceries	shop_major	prep_meals	decide_finance	single
Male	1.95	2.19	1.70	2.39	0.42
Female	2.49	2.11	2.53	2.21	0.50

Non-single sample: N=26595

Full sample: N=48146

Notes: Table 2 shows the average scores for each of the household roles and experience variables for men and women who are not living alone. All variables are ranked from 1 (partner does this alone) to 3 (respondent does this alone). The last columns shows the share of single households for men and women from the full sample.

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; own calculations

3.2.2 Measuring financial confidence

I exploit data from the BOP-HH and SCE that contain information on micro-level financial literacy, namely a standardised financial literacy test (see examples in Lusardi and Mitchell, 2008 and Bucher-Koenen et al., 2014). The value of the test score is computed from three questions. They cover one question regarding compound interest, one question on inflation and one question on risk. For each correct answer, one point is added, don't know responses are treated as false responses and don't yield a point. Thus, the score ranges from 0-3. The wording of the questions can be found in Appendices D.1 for the BOP-HH and D.2 for the SCE. The SCE allows me to compute a tested financial literacy score for each observation (assuming their score to remain constant over time) while in the BOP-HH the financial literacy test is only asked in January 2022. To measure confidence and recover a measure for the full sample, I use a combination of variables related to confidence to predict financial literacy out-of-sample. This allows me to focus of the confidence components in financial literacy. I compute robustness exercises for the BOP-HH measure using the SCE full sample. Below I introduce the variables used in the out-of-sample prediction.

The first measure is derived from the argument in Binder (2017) and Reiche and Meyler (2022): rounding. According to the "Round Numbers, Round Interpretations" hypothesis in linguistics (Krifka, 2009), individuals uncertain about the precise value use the nearest round number. Thus, I compute a dummy for rounding (*round*). In the context of inflation, this implies that those uncertain about future inflation may be inclined to predict inflation of around 0, 5 or 10 rather than more precise digits.[12] The majority of participants respond with a non-rounded number (79.43% in the BOP-HH and 66.70% in the SCE).[13]

As a second measure of financial confidence, I use self-reported interest (and difficulty in the BOP-HH) of the survey provided through feedback questions at the end of the survey. These are about all the topics of the survey, inflation being one core element. Assuming that individuals with more interest answering the questions are more confident, one can estimate confidence by responses to the feedback question (*qeasy* in the BOP-HH and *qinterest* in both surveys). Note, however, that feedback is not given purely on inflation. The surveys include many other political topics and economic indicators, so a respondent could be confident about inflation but have difficulty evaluating other topics.

Finally, I control whether someone has participated in an earlier survey wave (*refresher*) as the learning-through-survey effect can be substantial in household panels (Kim & Binder, 2020) suggesting that refreshers may be more confident about the topics of the survey.

I run an ordered logit regressions using a computed, standardized test score as dependent variable and the measures described above as explanatory variables. The out-of-sample prediction allows me to create a larger sample for the BOP-HH while verifying results with the direct measure in the SCE. Table 9 in the Appendix shows the relative importance of each of the measures described above on the tested financial literacy in the full SCE sample and the BOP-HH subsample. Most variables behave as expected, choosing *round* reduces financial literacy scores in both surveys, while *qinterest* has a positive coefficient in both surveys and *qeasy* in the BOP-HH. The coefficient for being a *refresher* is negative in both surveys and significant in the SCE which is surprising. Finally, note that being female reduces the ex-

[12]Reiche and Meyler (2022) show in detail how the distribution of the responses indicates that there are indeed three groups of respondents, those responding in digits ("certain"), those responding in fives ("uncertain") and those responding in multiples of ten ("highly uncertain"). These can also be found in a histogram of the quantitative inflation expectations in both surveys (see Appendix Figure F.3). Most prominently, it displays a heavier tail for females.

[13]This differs from the findings of Reiche and Meyler (2022), who estimate a share of precise respondents of around 25% in 2019 and after. One possible explanation for the difference could be that the authors use data from the European Commission Consumer Survey which records an inflation expectation of zero for all respondents, qualitatively indicating that inflation "will stay about the same". In contrast, the BOP-HH and SCE do not directly link qualitative and quantitative questions. This explains the lower share of "zero" respondents in both surveys, which are classified as "rounders" in the analysis.

pected financial literacy score in both surveys at a similar (and insignificantly different) rate. I construct financial confidence as the predicted probability of obtaining 3 correct answers in the financial literacy test out-of-sample in all survey waves using the ordered logistic regression .

Notice, that the framework (Section 2) models low confidence as smaller prior precision. To verify that indeed, low confidence is associated with higher uncertainty around point forecasts I show that the confidence score is inversely correlated with the interquartile range of the parametrically fitted implied probability distribution (*intqr*), an indicator of individual-level uncertainty. Probabilistic questions have been used in Engelberg et al. (2009) in the context of the Survey of Professional Forecasters and for consumers in Armantier et al. (2013). I follow the procedure applied in both papers to exploit the probabilistic question, details can be found in Appendix D.4. The Pearson correlation coefficient between the predicted financial confidence score and the fitted interquartile range is -0.1235 in the BOP-HH and -0.1312 in the SCE. Both are negative and significantly different from zero with $p < 0.001$, demonstrating that uncertainty around point forecasts measured by the interquartile range increased with higher financial confidence. Figure D.1 in the appendix shows using a binscatter plot how the interquartile range declines as the predicted financial confidence score increases in both surveys.

4 The Effects of Financial Confidence and Shopping Experience

Heterogeneity in experiences and heterogeneity in financial confidence are not mutually exclusive hypotheses to explain the gender gap in inflation expectations. In the simple Bayesian framework presented in Section 2 I have shown that both parameters are complementary – when financial confidence is low, experiences matter more and vice versa, the effect of a flat prior is stronger when signals are noisy.

This can be tested using the BOP-HH, which contains data on grocery shopping and financial confidence, as discussed above. I set up the following panel regression model:

$$\begin{aligned}\pi^E_{i,t} = \beta_0 + \beta_1 female_i + \beta_2 P(\hat{test} = 3)_{i,t} + H_i\gamma_1 + P(\hat{test} = 3)_{i,t} \times H_i\gamma_2 \\ + X_{i,t}\gamma_3 + D_t\gamma_4 + R_i\gamma_5 + \bar{X}_i\theta + \upsilon_i + \rho_{i,t}\end{aligned} \tag{10}$$

Where $\pi^E_{i,t}$ denotes individual i's inflation expectation (point forecast, 12 months ahead) at time t, $female_i$ is a dummy for self-identifying as female, $P(\hat{test} = 3)_{i,t}$ is a measure of financial confidence (see Section 3.2.2), and H_i is a vector of individual level involvement in household activities such as grocery shopping, meal preparation, purchase of major items and financial decision-making (see Section 3.2.1). Demographic controls[14] are summarized in $X_{i,t}$, R_i denotes the matrix of regional dummies, D_t denotes the matrix of time dummies and $\bar{X}_i$ denotes the time averages of individual i of age and income.[15] I initially focus on non-single household respondents but extend to the full sample for comparison.

While the results in Table 3 do not show a significant role for grocery shopping and other household chores in column (3), financial confidence has a big impact on the level of inflation expectations and including confidence as additional variable reduces the size of the gender gap substantially in column (2) relative to a baseline specification without experience and financial confidence in column (1). Including both, experience controls and financial confidence shows no qualitative difference to those observations (column (4)). However, once interaction terms for experiences with financial confidence are included, I show that experience does matter. Columns (5) and (6) introduce a range of experience variables and their interaction terms with financial confidence. I find that the coefficient for female remains largely unaffected. The coefficient for financial confidence reduces slightly but the standard error increases such that the effect becomes insignificant. In contrast, grocery shopping, initially insignificant, becomes positively significant with a negative and significant interaction term. In the BOP-HH, grocery shopping has a significantly positive effect on inflation expectations for the bottom 16.75% in terms of financial confidence (share of women: 82.25%).[16] Figure 1 visualizes this result. The reverse is true for purchasing major items for the household, traditionally a more male dominated task. Different levels of financial confidence amplify the effects of experience on inflation expectations. Finally, I show that there are no effects of living in a single household beyond the involvement in the chores.

Result 4.1 *Grocery shopping and meal preparations increase inflation expectations only for the individuals in the lowest quintile of financial confidence distribution (the lowest 18.43%),*

[14] $age_{i,t}$ records the individual i's age at time t, $educ_i$ is an ordered categorical variable of i's education, $inc_{i,t}$ is an ordered categorical variable of the household income of observation i at time t

[15] All three surveys with microdata are of panel structure. Due to my interest in time invariant variables such as *female*, I cannot use fixed effects estimation. I employ an alternative estimator to estimate time invariant variables, while maintaining robustness to endogeneity caused by time-invariant observables: I incorporate between effects of time varying variables in the existing model and apply pooled OLS to the transformed model. The estimation technique is covered in depth in Appendix E. Between effects are computed as follows:

$$\bar{x}_i = \frac{1}{T}\sum_{t=1}^{T} x_{i,t} \tag{11}$$

[16] Using the coefficients in column (5) of Table 3 and the standard error of the sum of the coefficient on shop_groceries and $P(\hat{test} = 3)\times$shop_groceries (0.1203), I compute the lower bound of the 95% confidence around the predicted effect of grocery shopping for every value of financial literacy. This yields that the effect of of grocery shopping is only statistically significantly positive at the 5% significance level for $P(\hat{test} = 3) \leq$ 0.4636.

Table 3: The role of financial confidence and experience

	Inflation expectation (12 months ahead, point estimate)					
	Non-singles					Full sample
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	4.69***	5.43***	4.65***	5.31***	5.16***	5.13***
	(0.22)	(0.25)	(0.28)	(0.30)	(0.71)	(0.74)
female	0.80***	0.53***	0.71***	0.45***	0.42***	0.43***
	(0.06)	(0.08)	(0.07)	(0.08)	(0.08)	(0.06)
$P(\widehat{test} = 3)$		−1.68***		−1.67***	−1.41	−1.43
		(0.30)		(0.30)	(1.05)	(1.04)
shop_groceries			0.03	0.03	0.70***	0.69***
			(0.05)	(0.05)	(0.21)	(0.21)
shop_major			0.02	0.03	−0.90***	−0.92***
			(0.07)	(0.07)	(0.29)	(0.29)
prep_meals			0.07	0.08	0.32	0.32
			(0.05)	(0.05)	(0.20)	(0.20)
decide_finance			−0.09*	−0.07	−0.01	−0.02
			(0.05)	(0.05)	(0.23)	(0.23)
single						−0.07
						(0.30)
$P(\widehat{test} = 3)\times$shop_groceries					−1.00***	−0.99***
					(0.31)	(0.31)
$P(\widehat{test} = 3)\times$shop_major					1.39***	1.42***
					(0.43)	(0.43)
$P(\widehat{test} = 3)\times$prep_meals					−0.38	−0.37
					(0.29)	(0.29)
$P(\widehat{test} = 3)\times$decide_finance					−0.10	−0.09
					(0.34)	(0.34)
$P(\widehat{test} = 3)\times$single						−0.17
						(0.46)
age	−0.01***	−0.01***	−0.01***	−0.01***	−0.01***	−0.01***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.001)
educ	−0.08***	−0.06***	−0.08***	−0.06***	−0.06***	−0.05***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
hhinc	−0.09***	−0.04***	−0.08***	−0.04***	−0.05***	−0.06***
	(0.01)	(0.02)	(0.01)	(0.02)	(0.02)	(0.01)
Observations	26,595	26,595	26,595	26,595	26,595	48,146
R^2	0.14	0.14	0.14	0.15	0.15	0.16
Adjusted R^2	0.14	0.14	0.14	0.14	0.14	0.16
F Statistic	185.86***	179.94***	159.61***	155.38***	137.67***	183.34***

*p<0.1; **p<0.05; ***p<0.01

Notes: Table 3 shows the regression coefficients of a pooled OLS estimation of individual inflation expectations (12 months ahead, point estimate) in the BOP-HH on the dummy variable $female_i$, predicted financial confidence and a range of experience variables defined in detail in Section 3.2.1. The full model can be found in Equation 10. All models include regional controls, between effects and time fixed effects. For models (1)-(5) the sample is restricted to those who who do not live in single households, model (6) includes singles. Standard errors in brackets below.

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; own calculations

which is dominated by women.

After showing that experience matters only jointly with financial confidence, I shed light on this channel. The next section verifies that indeed in the BOP-HH and the SCE, women have lower financial confidence scores and that such scores are associated with right-skewed inflation expectations. Further, I demonstrate that there is no inflation expectations gender gap in high financial confidence samples. Even if grocery shopping is unequally distributed amongst men and women in this sample, the priors are tight enough to remove the mean-increasing effect of observing volatile signals.

5 Financial Confidence Channel

To provide further evidence for the financial confidence channel, I first verify the assumptions by testing whether, women are less confident about their own literacy and whether this causes higher inflation expectations through a heavy right tail as in the Bayesian framework. I confirm both for the data at hand. Further, I show that when controlling for confidence the gap disappears and test an implication of my hypothesis, namely that if the gender gap is driven by women in the tail, the gap should close fully when the outliers are removed.

5.1 Verifying the assumptions

The financial confidence hypothesis relies on two assumptions: (1) Women have lower financial confidence, and (2) low financial confidence increases expectations through a heavy right tail. I test these assumptions and show that while we observe women to be over-represented in the tail, this can be explained predominantly through financial confidence differentials.

5.1.1 The gender gap of financial confidence

There is evidence that women perform worse in standardized financial literacy tests (Bucher-Koenen et al., 2014). Such a test has been included in the BOP-HH in January 2022 and in all waves of the SCE and can be used to test the gender gap of financial literacy in the sample at hand.

Table 4 summarizes financial confidence scores (as well as literacy test scores) for the male and female samples, as well as averages of the predictors discussed in Section 3.2.2. Women have significantly lower financial literacy in test scores than their male counterparts. They also receive lower confidence scores based on my method. This is in line with evidence from standardized tests (Bucher-Koenen et al., 2014, 2021). In my sample, women show much higher uncertainty and less interest in the topics around inflation. The only variable

Table 4: Financial literacy and confidence of men and women

	BOP-HH *N=91501 (19765)*					FRBNY *N=113165*				
	Mean		Tests			Mean		Tests		
	M	F	t	wilcox	ks	M	F	t	wilcox	ks
fin_lit_test	(2.67	2.42	***	***	***)	1.48	1.19	***	***	***
$P(\hat{test} = 3)$	0.71	0.53	***	***	***	0.27	0.17	***	***	***
refresher	0.65	0.63	***	***	***	0.87	0.86	**	**	
round	0.18	0.25	***	***	***	0.23	0.43	***	***	***
qinterest	3.78	3.59	***	***	***	3.78	3.68	***	***	***
qeasy	3.28	3.01	***	***	***					

*p<0.1; **p<0.05; ***p<0.01

Notes: Table 4 compares the mean values of the variables used to measure financial confidence between the male and female sub samples in two surveys. All variables are introduced in detail in Section 3.2.2. $P(\hat{test} = 3)$ denotes the predicted probability that the test score is equal to 3 estimated from the ordinal logit regression. *refresher* is a dummy for previous survey participation, *round* is a dummy for a rounded point forecast and *qinterest* and *qeasy* are ordered categorical feedback responses on interest in survey (1: very interesting - 4: not interesting at all) and how difficult it was (1: very difficult - 4: very easy). Welch t-test and Wilcoxon-test (both two-sided) are computed to assess the null of the same mean/median and Kolmogorov-Smirnoff test is computed to assess the null of a common distribution. I find a significant difference in financial confidence between men and women.

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; Federal Reserve Bank of New York (FRBNY), SCE, June 2013 - November 2020; own calculations

that shows no difference compared to men is *refresher* in the SCE which may cause accurate forecasts (Kim & Binder, 2020) but shouldn't have a gender gap. However, in the BOP-HH, even refresher is higher for men, indicating that they are more likely to stay in the survey.

Observation 5.1 *Women in the BOP-HH and SCE perform worse on standardized financial literacy tests. They are also shown to be less confident, have higher uncertainty of their expectations, round more often and find the questions less easy and interesting.*

5.1.2 Financial confidence and inflation expectations

The second assumption of the hypothesis is that financial confidence impacts inflation expectations through higher uncertainty, i.e., a flat prior. Financial literacy is known to be an important upward driver of inflation expectations (Burke & Manz, 2014; D'Acunto et al., 2019a, 2019b). Confidence in own literacy is a related concept, partially because one should be less confident if actual literacy is lower and partially because performance in tests decreases when confidence is low (Bucher-Koenen et al., 2021). One channel which may cause this result is that lower financial confidence causes higher uncertainty, which skews expectations to the right. The direction is determined by the survey design. Consumers are typically asked for their point forecasts in the range 0-100. Thus, the survey elicitation is asymmetric around the inflation target of 2%. If consumers with a flatter prior choose some value towards the middle of the range, it will lead to an upward bias. This has been shown in Section 2 where a flatter log-normal prior gives rise to increased average expectations despite unchanged signals.

Using the confidence measures derived in Section 3.2.2 I compute kernel densities of the inflation expectations of high and low financial confidence respondents, where I define low confidence as being strictly below the median of the full sample. Figures 2a to 2d show that high point forecasts beyond 10% are more often chosen by individuals classified as low financial confidence (predicted) or literacy (tested). Large peaks at rounded numbers (i.e., 5% or 10%) are driven by low confidence/literacy individuals, in line with the rounding theory in Binder (2017) and Reiche and Meyler (2022).

Observation 5.2 *There is a confidence gap in inflation expectations, driven by the heavy tail of the low confidence sample.*

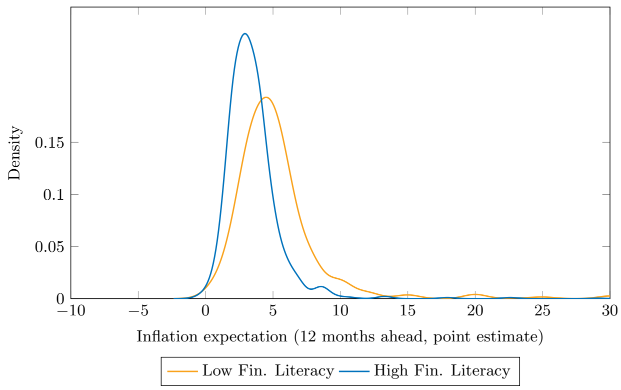
5.2 Financial confidence controls reduce the gender gap

After having confirmed that indeed, women in the BOP-HH and SCE are characterized by lower financial confidence and financial confidence decreases inflation expectations, it remains to show that when controlling for financial confidence the gender gap diminishes. To test this, again, I apply the controlled version of a pooled OLS estimation with point forecasts as dependent variable and a range of demographic explanatory variables. In contrast to previous model, I include my estimate of financial confidence and interact it with female.

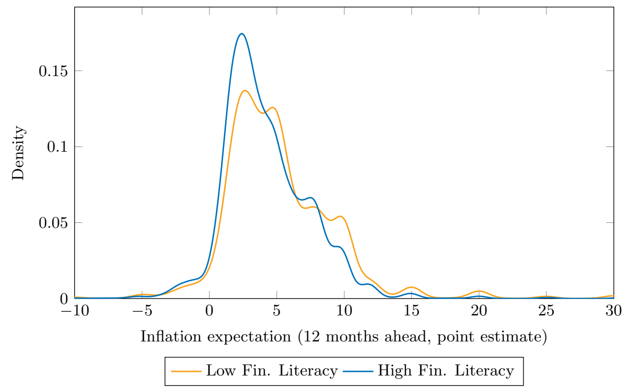
$$\begin{aligned}\pi^E_{i,t} = \alpha + \beta_1 female_i + \beta_2 P(\hat{test} = 3)_{i,t} + \beta_3 female_i \times P(\hat{test} = 3)_{i,t} + X_{i,t}\gamma_1 + D_t\gamma_2 \\ + R_i\gamma_3 + \bar{X}_i\theta + v_i + \rho_{i,t}\end{aligned} \tag{12}$$

Where all variables are defined as before. I test the null hypothesis that the gender gap reduces when financial confidence increases: $H_0 : \beta_3 \leq 0$.

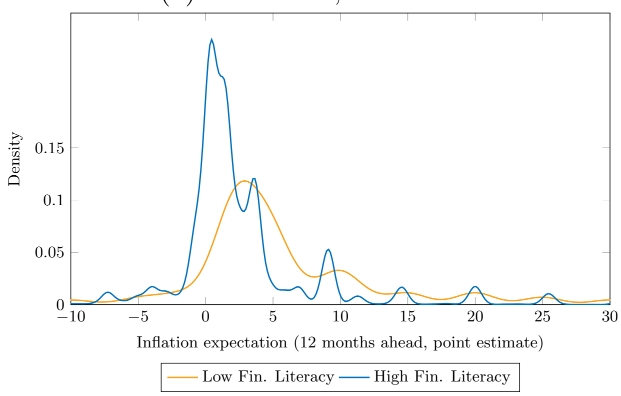
The results are summarized in Tables 5 for the BOP-HH. In the Appendix, I include Table 12 for the SCE. For both surveys, I find that the interaction term in model (3) and (5) indicates that while women with low financial confidence/literacy have much higher expectations than their male counterparts holding all other demographics constant, as confidence increases, the



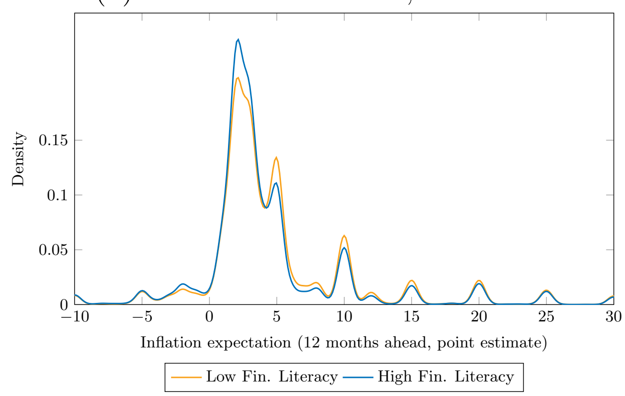
(a) Test score, BOP-HH



(b) Predicted confidence, BOP-HH



(c) Test score, SCE



(d) Predicted confidence, SCE

Figure 2: Kernel density of inflation expectations split by financial confidence and literacy

Notes: Figures 2a to 2d compare the high financial literacy/confidence (in blue) and low financial literacy/confidence (in orange) sample distribution of inflation expectations (measured as point forecasts over 12 months) pooled across all time periods. There is one plot for literacy using the test scores (left hand side) and one plot for confidence (right hand side) for both surveys. For all, low confidence is defined as a value less than the median and both estimates stem from the ordered logit regression described in Section 3.2.2. The Kernel density is plotted using two times Silverman's rule-of-thumb bandwidth (Silverman, 1986).[17]

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; Federal Reserve Bank of New York (FRBNY), SCE, June 2013 - November 2020; own calculations

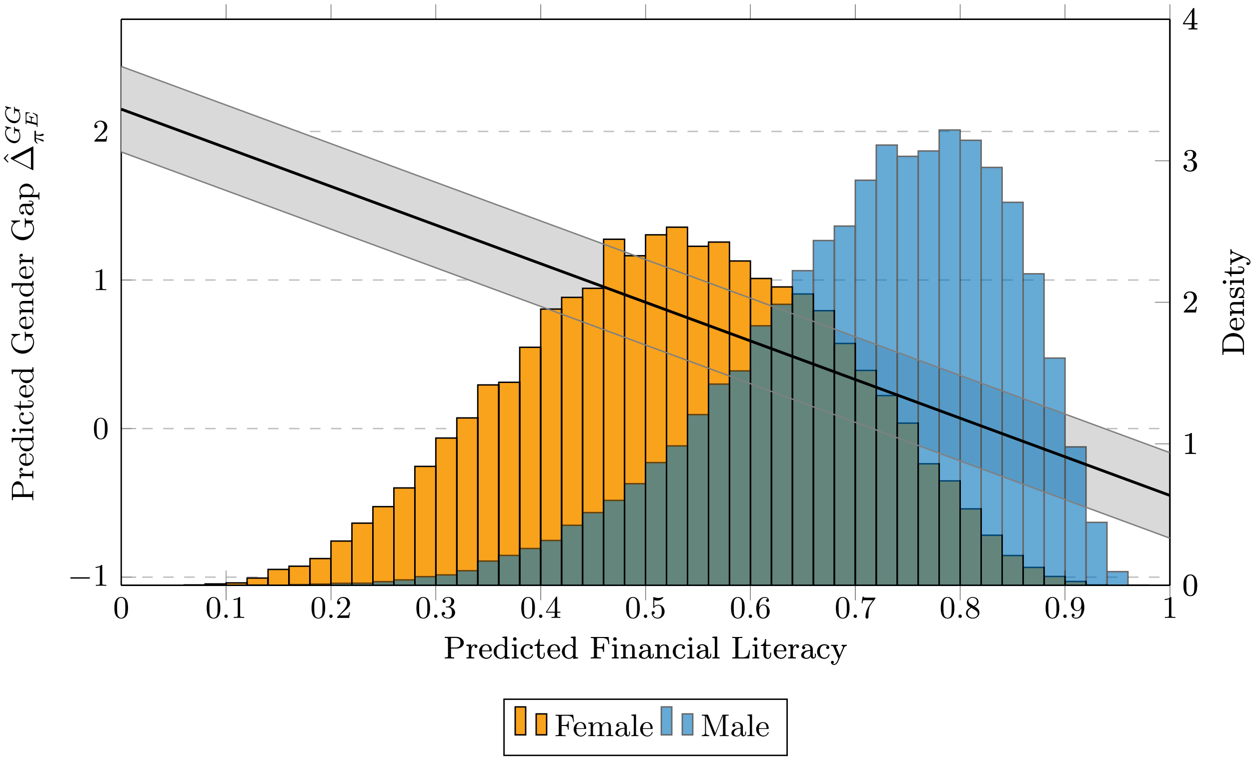


Figure 4: The gender gap for different levels of financial confidence (BOP-HH)

Notes: The black line in Figure 4 plots the predicted gender gap along all possible values of the financial confidence score in the black line ($\hat{\Delta}^{GG}_{\pi^E}(x) = 2.15 - 2.60x$). The full regression results are shown in Table 5, column (3). The grey area indicates 95% confidence bands (standard error: 0.1468). The histograms show the density of the male (blue) and female (orange) distribution of financial confidence scores.

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; own calculations

gap diminishes. This effect is stronger for the confidence score than for the pure literacy tests core. Figure 4 plots the predicted gender gap along different financial confidence scores computed as $\hat{\Delta}^{GG}_{\pi^E} = \beta_1 + \beta_3 \times \mathit{finlit}$. The same plot for the SCE can be found in the Appendix (Figure F.5).

Observation 5.3 *The gender gap in inflation expectations diminishes as financial confidence increases and is zero (or negative) for high financial confidence individuals.*

5.3 The gender gap driven by the tails

One implication of the financial confidence hypothesis is that the gender gap should disappear when the sample is trimmed as outliers in the right-skewed distribution are removed. I test

Table 5: The impact of financial confidence on the gender gap (BOP-HH)

Survey: BOP-HH	Inflation expectation (12 month ahead, point forecast)				
	(1)	(2)	(3)	(4)	(5)
Constant	4.69***	5.39***	4.56***	8.03***	7.58***
	(0.17)	(0.18)	(0.19)	(0.56)	(0.62)
female	0.87***	0.54***	2.15***	0.43***	1.30**
	(0.03)	(0.04)	(0.15)	(0.16)	(0.58)
$P(\hat{test} = 3)$		−2.03***	−0.73***		
		(0.16)	(0.20)		
$P(\hat{test} = 3)\times$female			−2.60***		
			(0.24)		
fin_lit_test				−0.70***	−0.53***
				(0.11)	(0.15)
fin_lit_test × female					−0.35
					(0.22)
age	−0.01***	−0.01***	−0.01***	−0.01**	−0.01**
	(0.001)	(0.001)	(0.001)	(0.01)	(0.01)
single	−0.34***	−0.21***	−0.23***	−0.17	−0.16
	(0.04)	(0.04)	(0.04)	(0.20)	(0.20)
educ	−0.07***	−0.05***	−0.05***	−0.03	−0.03
	(0.005)	(0.01)	(0.01)	(0.02)	(0.02)
hhinc	−0.14***	−0.08***	−0.09***	−0.04	−0.05
	(0.01)	(0.01)	(0.01)	(0.04)	(0.04)
Time dummies	Yes	Yes	Yes	No	No
Region dummies	Yes	Yes	Yes	Yes	Yes
Between effects	Yes	Yes	Yes	No	No
Observations	91,501	91,501	91,501	2,916	2,916
R^2	0.16	0.16	0.16	0.03	0.03
Adjusted R^2	0.15	0.16	0.16	0.02	0.02
F Statistic	430.80***	424.64***	417.80***	8.43***	7.84***

Note: *p<0.1; **p<0.05; ***p<0.01

Notes: Table 5 shows the regression coefficients of a pooled OLS estimation of individual inflation expectations (12 months ahead, point estimate) in the BOP-HH on the dummy variable $female_i$, other demographics and confidence and literacy scores.The full model can be found in Equation 12. Standard errors in brackets below.

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; own calculations

this hypothesis by computing the gender gap along deciles in the expectations' distribution, controlling for demographics and time periods in three surveys (BOP-HH, SCE and MSC). Kernel densities of the expectations' distribution and higher moments can be found in Appendix F.5. I run a regression on each decile in the expectations' distribution. The setup resembles the previous regression models. The dependent variable is denoted by $I_{\pi^E_{i,t} \leq \pi^E_q} \times \pi^E_{i,t}$ where $I_{\pi^E_{i,t} \leq \pi^E_q}$ is an indicator variable for whether an observation has inflation point forecasts below percentile q of the full sample.

$$\begin{aligned} I_{\pi^E_{i,t} \leq \pi^E_q} \times \pi^E_{i,t} = I_{\pi^E_{i,t} \leq \pi^E_q} \times (&\alpha + \beta_1 female_i + X_{i,t}\gamma_1 + D_t\gamma_2 \\ &+ R_i\gamma_3 + \bar{X}_i\theta + v_i + \rho_{i,t}) \end{aligned} \tag{13}$$

I test the following null hypothesis: The gender gap closes when outliers are removed: $H_0 : \beta_1^{(q_1)} \leq \beta_1^{(q_2)} \Rightarrow q_1 \leq q_2$. Figure 5 plots the coefficient of *female* for different percentiles. It increases steadily, as predicted by the financial confidence hypothesis. In samples restricted to inflation expectations strictly below the 50th percentile in a multivariate setting, no positive gender gap can be identified. In contrast, for low percentiles, the gender gap is negative. The regression table can be found in the Appendix, Table 14.

Observation 5.4 *The gender gap in means is driven by the heavy tail in the female distribution. When the sample is restricted to the lowest 50% of inflation expectations, there is no positive gap and at lower percentiles a significantly negative gender gap emerges.*

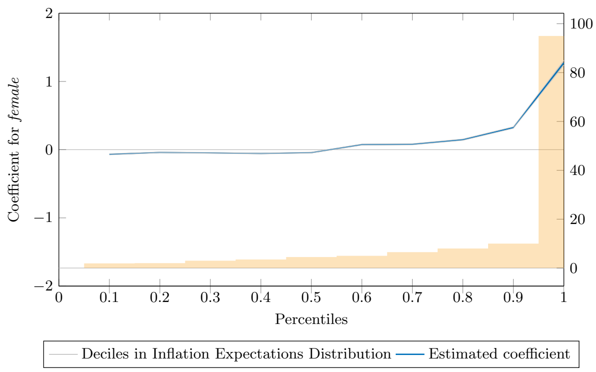
5.4 Summary of the financial confidence channel

This section provides evidence for the financial confidence channel of my hypothesis. I establish that, as predicted in the literature, female respondents in the BOP-HH and SCE have lower financial literacy and confidence therein and that lower financial confidence is associated with higher inflation expectations. Further, I use interaction terms to show that the gender gap is positive only for low financial confidence individuals and diminishes as confidence increases.

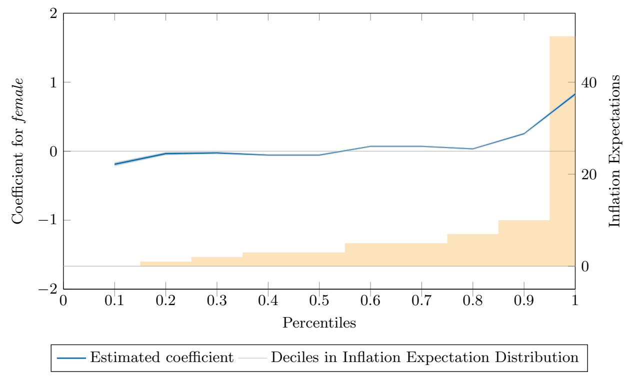
Result 5.5 *Observations 5.1-5.4 suggests that the gender gap in means of inflation expectations is driven at least partially by low financial confidence of women, which skews the distribution to the right.*

6 Robustness Exercises

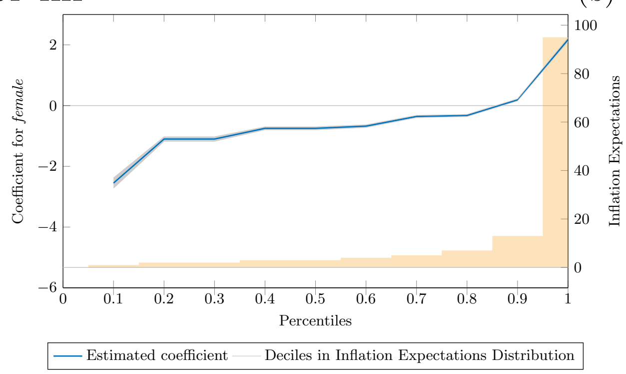
I compute two robustness checks in support of the argument made above, namely that experience alone cannot account for the gender gap in inflation expectations: First, I test whether



(a) BOP-HH



(b) MSC



(c) SCE

Figure 5: The gender gap along deciles in the inflation expectations distribution

Notes: Figures 5a, 5b and 5c show the estimated regression coefficient for the dummy variable $female_i$ in regression for each decile in the inflation expectations distribution from 0.1 up to 0.9 in three surveys, BOP-HH, MSC and SCE as the blue line. The regression model can be found in Equation 13. 95% confidence bands are shaded in grey. The orange bars indicate the percentiles in the inflation expectations distribution.

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; Federal Reserve Bank of New York (FRBNY), SCE, June 2013 - November 2020; University of Michigan, Survey Research Center, MSC, January 1978 - January 2023; own calculations

there is a gender gap amongst singles. There should be no gender gap amongst singles under the experience hypothesis, as both men and women have to engage in symmetrically in grocery shopping if they live alone. However, there may still be a financial confidence gender gap for singles and hence the symmetric shopping can still cause asymmetric expectations. Second, I test whether the gender gap increases in periods of high food price inflation. Under the experience hypothesis, the gap should increase whenever food prices are particularly high or volatile. This need not be true if financial confidence is the key channel through which the gender gap occurs. I show that both tests are rejected in three surveys, providing evidence in favor of my hypothesis.

6.1 The gender gap amongst singles

The first implication of the pure experience hypothesis is that there should not be a gender gap in inflation expectations for single men and women. This is because singles can be assumed to engage in grocery shopping irrespective of gender, and should thus have no systematic differences in the inflation levels and volatility observed.

I test whether there is a gender gap for singles. To do so, I run a simple panel regression of point forecasts of inflation expectations on a female dummy controlling for age, income, and education. The model is written as:

$$\pi^E_{i,t} = \alpha + \beta_1 female_i + X_{i,t}\gamma_1 + D_t\gamma_2 + R_i\gamma_3 + \bar{X}_i\theta + +v_i + \rho_{i,t} \tag{14}$$

Where all variables are as defined before. The model is run for two subsamples of each survey, that of singles (S) and that of non-singles, i.e., individuals who record that their household size exceeds 1 (N). The null hypothesis under the experience hypothesis is that the magnitude of the gender gap is higher in the non-single sample: $H_0 : \beta_1^{(N)} - \beta_1^{(S)} > 0$

Table 6 shows that for all surveys (a) there is a persistent and significant gender gap in both, single and non-single samples and (b) it is not statistically smaller for singles. In fact, in the SCE singles show a larger gender gap than those living in households with more than one person. This rejects H_0. D'Acunto et al. (2021) show no evidence for non-married and single individuals, and Jonung (1981) shows no treatment of disaggregated data.

Observation 6.1 *The gender gap is significant and no different between singles and non-singles.*

6.2 The gender gap correlated with historical food prices

Under the experience hypothesis, the gender gap in inflation expectations should be increasing in periods where food price inflation (volatility) exceed CPI core inflation (volatility) and

Table 6: Comparing the gender gap in inflation expectations for singles and non-singles

	Inflation expectation (12 months ahead, point estimate)					
	BOP-HH		SCE		MSC	
	Non-Single	Single	Non-Single	Single	Non-Single	Single
	(1)	(2)	(3)	(4)	(5)	(6)
constant	6.58***	5.38***	8.66***	8.39***	7.27***	7.37***
	(0.22)	(0.38)	(0.41)	(0.80)	(0.29)	(0.60)
female	1.45***	1.37***	1.39***	2.01***	0.85***	0.89***
	(0.04)	(0.08)	(0.08)	(0.13)	(0.03)	(0.05)
age	−0.02***	−0.01***	0.01**	0.02***	−0.01***	−0.02***
	(0.001)	(0.002)	(0.003)	(0.004)	(0.001)	(0.001)
hhinc	−0.24***	−0.23***	−0.37***	−0.41***	−0.0000***	−0.0000***
	(0.01)	(0.02)	(0.02)	(0.03)	(0.0000)	(0.0000)
educ	−0.12***	−0.09***	−0.46***	−0.58***	−0.24***	−0.26***
	(0.01)	(0.01)	(0.03)	(0.05)	(0.01)	(0.02)
Time dummies	Yes	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes	Yes
Between effects	Yes	Yes	No	No	Yes	Yes
$\beta_1^{(N)} - \beta_1^{(S)}$	*0.09*		*-0.61****		*-0.04*	
	(0.09)		*(0.15)*		*(0.06)*	
Observations	83,704	27,381	74,385	41,106	195,107	66,268
R^2	0.13	0.11	0.03	0.03	0.13	0.10
Adjusted R^2	0.13	0.11	0.03	0.03	0.12	0.09
F Statistic	334.21***	93.17***	21.91***	12.37***	52.04***	13.62***

*p<0.1; **p<0.05; ***p<0.01

Notes: Table 6 compares the regression coefficients of a pooled OLS estimation of individual inflation expectations (12 months ahead, point estimate), on demographics, including the female dummy. Standard errors in brackets below. The full model can be found in Equation 14. Since the SCE's panel structure cannot be used due to crucial demographic questions being asked only once, there are no between effects for this survey. Models (1) and (2) compare coefficients for the non-single and single sub samples in the BOP-HH, (3) and (4) for the SCE and (5) and (6) for the MSC. The italics below compute the implied gap between the coefficient on female in non-single and single samples and the standard error.

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; Federal Reserve Bank of New York (FRBNY), SCE, June 2013 - November 2020; University of Michigan, Survey Research Center, MSC, January 1978 - January 2023; own calculations

shrink in periods where the opposite occurs. This is because in those periods, household members with grocery shopping exposure observe particularly high/volatile prices changes to their non-shopping counterparts, which increases the level/noisiness of their signals in the Bayesian learning framework.

I analyze the correlation of high food inflation and the gender gap on the micro level. Using the same regression set-up as before but replacing time dummies with a measure for the difference in food price and total inflation as well as the moving coefficients of variations of those variables.

$$\begin{aligned}\pi^E_{i,t} &= \beta_0 + \beta_1 female_i + \beta_2 female_i \times (CPI^{food}_t - CPI^{total}_t) + \beta_3(CPI^{food}_t - CPI^{total}_t) \\ &+\beta_4 female_i \times (\rho^{food}_{t,6} - \rho^{total}_{t,6}) + \beta_5(\rho^{food}_{t,6} - \rho^{total}_{t,6}) + X_{i,t}\gamma_1 + R_i\gamma_2 + \bar{X}_i\theta + v_i + \rho_{i,t}\end{aligned} \tag{15}$$

Here, $(CPI^{food}_t - CPI^{total}_t)$ measures the gap between a given period's food inflation and total inflation and $(\rho^{food}_{t,6} - \rho^{total}_{t,6})$ the gap between the coefficient of variation of food and total inflation in a 6 month moving window.[18] The results of this regression can be found in Table 7. Under the experience hypothesis, I expect that $H_{0,1} : \beta_2 > 0$ and $H_{0,2} : \beta_4 > 0$. In contrast, the data reveals that in both BOP-HH and MSC the coefficient for the interaction term with the absolute inflation gap is significantly negative, suggesting that the gender gap reduces in size whenever food prices are very high. For the gap in inflation variability the interaction with female is insignificant. β_3 and β_5 are estimated significantly positive in the BOP but negative in the SCE, hence there is also no clear indication whether these factors alone increase or decrease inflation expectations.

Additional evidence on the time series correlation of the gender gap in inflation expectations and the food inflation gap using aggregate data can be found in Appendix F.7.

Observation 6.2 *The magnitude of the gender gap is unresponsive to the size of food price inflation relative to total inflation.*

6.3 Summary of the robustness exercises

This section has provided evidence that rejects the pure experience hypothesis as it has been presented in D'Acunto et al. (2021). This hypothesis would imply that the gap should also disappear for singles, who can be assumed to participate in grocery shopping to equal shares. I find no evidence for this in an analysis of three large consumer surveys in DE and the US. In contrast, singles show significant gender gaps that are no different to those of non-singles. Further, the authors only provide snapshots from 2015 and 2016. Using the entire time series available to me, I show that the gap between male and female inflation expectations is

[18]Details of the computation can be found in Appendix F.7.

Table 7: Microlevel effects of high food prices

	Inflation expectation (12 months ahead, point estimate)		
	BOP-HH	SCE	MSC
constant	7.96***	7.77***	7.13***
	(0.12)	(0.20)	(0.06)
female	1.51***	1.61***	0.75***
	(0.04)	(0.07)	(0.02)
female × $(CPI_t^{food} - CPI_t^{total})$	−0.04***	−0.03	−0.09***
	(0.01)	(0.03)	(0.01)
female × $(\rho_{t,6}^{food} - \rho_{t,6}^{total})$	−0.005	0.01	−0.01
	(0.01)	(0.02)	(0.01)
$CPI_t^{food} - CPI_t^{total}$	0.46***	0.03	−0.06***
	(0.01)	(0.02)	(0.01)
$\rho_{t,6}^{food} - \rho_{t,6}^{total}$	0.08***	−0.01	−0.03***
	(0.005)	(0.01)	(0.01)
age	0.67***	0.01***	−0.04***
	(0.05)	(0.002)	(0.02)
hhinc	0.04	−0.38***	−0.0000***
	(0.03)	(0.02)	(0.0000)
single	−0.53***	−0.09	−0.16***
	(0.05)	(0.08)	(0.03)
educ	−0.10***	−0.50***	−0.34***
	(0.01)	(0.02)	(0.01)
Year dummies	No	No	No
Between effects	Yes	No	Yes
Regional dummies	Yes	Yes	Yes
Observations	111,085	115,491	259,755
R^2	0.09	0.03	0.03
Adjusted R^2	0.09	0.03	0.03
F Statistic	789.19***	257.20***	619.13***

*p<0.1; **p<0.05; ***p<0.01

Notes: Table 7 compares the regression coefficients of a pooled OLS estimation of individual inflation expectations (12 months ahead, point estimate) on demographics, the per period difference between food and total inflation and their moving coefficient of variation. Interaction terms with female are included. Standard errors in brackets below. The full model can be found in Equation 15. No time dummies are included due to multicollinearity with the time varying variables. Since the SCE's panel structure cannot be used due to crucial demographic questions being asked only once, there are no between effects for this survey.

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; Federal Reserve Bank of New York (FRBNY), SCE, June 2013 - November 2020; University of Michigan, Survey Research Center, MSC, January 1978 - January 2023; OECD, Prices: Consumer prices, Main Economic Indicators (database), January 1978 - January 2023; own calculations significantly smaller in periods when food inflation rises above total CPI inflation. Further, in an aggregate time series analysis, I find no evidence for the gender gap to increase when food prices rise faster (Appendix F.7). As a robustness exercise I compute a basic OLS regression with expectations on prices of specific items (gas, food, education, medical care, rent, and gold) in Appendix F.3 using data from the SCE. I find that the gender gap for food is positive and insignificantly smaller than for the full basket. The only item with a significantly larger gender gap is education. This is an additional indicator that the gender gap isn't driven solely driven by traditional gender norms enforcing a differential shopping experience.

I have shown that experience matters a lot, jointly with financial confidence. However, it is important to consider both factors together. The gender gap in inflation expectations appears to be driven of a composite of both factors, and looking at experience alone ignores the importance of financial confidence.

Result 6.3 *Observations 6.1 and 6.2 suggest that gendered experiences cannot fully account for the gender gap in inflation expectations.*

7 Discussion

There are potentially other drivers of the gender gap in inflation expectations that are not explored in this paper. One hypothesis suggested by the literature may be gender gaps in pessimism. Future inflation is an unknown variable, and individuals with a more pessimistic sentiment could skew their expectations towards the unfavorable outcome – high inflation. Hey (1984) defines economic pessimism as revising "down the probabilities of favorable events; and [...] up the probabilities of unfavorable events." Thus, sentiment is a likely driver of inflation expectations. A recent example is the response of household inflation expectations to the Covid-19 pandemic (Binder, 2020). However, while there is evidence that women may be more pessimistic about some uncertain outcomes (see Jacobsen et al. (2014) on consumer sentiment and stock market performance, Chaney et al. (1998) on election outcomes, Garbarino and Strahilevitz (2004) on online shopping, Lin and Raghubir (2005) on marriage, Lyons et al. (2009) on health and Gwartney-Gibbs and Lach (2016) on war), there is no evidence that women may be more pessimistic about inflation. A simple t-test on general mood in the BOP-HH shows that women's pessimism is no different from that of men in the survey (means 2.12 for men, 2.14 for women, scale 1: very optimistic to 4: very pessimistic, p-value: 0.0544). Thus, this hypothesis is not further explored in this paper. Note that the question addresses general mood and is not targeted at inflation in particular. in a recent analysis Garriga, 2023 shows that women are less satisfied with the work of the Bank of England which could be linked to a more negetaive perception of inflation in general. Further research could also investigate the possibility of “pinkflation”. While my analysis shows that grocery shopping is not the main driver of the gender gap, it might be that products bought dominantly by women experience higher inflation than those bought by men. I am unaware of research in this area. Scanner data could be helpful to answer such questions.

The empirical section of this paper focuses on financial confidence measured by filtering out the confidence component of financial literacy. However, the theoretical framework models this as flat prior, a phenomenon that would be more accurately describes as uncertainty. It is difficult to disentangle financial confidence, uncertainty about the concept of inflation and well-informed uncertainty about the future from the data at hand. I show in Section 3.2.2 that low financial confidence is associated with a larger interquartile range in the probabilistic question – a measure that could capture both types of uncertainty. However, it is also associated with rounding, which is more likely to be associated with uncertainty about inflation as a concept (Reiche & Meyler, 2022, see). Hence, both measures appear highly correlated such that a measure of financial confidence captures uncertainty to a sufficient extent. Disentangling the effects of those three drivers is an avenue for future research and would enable researchers to design questionnaires more accurately.

Another avenue for future research could be the behavioral consequences of the gender gap. The data at hand doesn’t allow for an in-depth analysis of the validity of the Euler equation for men and women, as the surveys only capture indent to spend but no information on realized spending. The literature appears split on whether consumers behave according to an Euler equation. While Dräger and Nghiem (2020) find supporting evidence using a new survey of German consumers, Bachmann et al. (2015) show that spending intent (as measured in the MSC) is unresponsive to changes in inflation expectations. It is possible that those mixed results are the result of different survey designs, such as measuring spending intent versus actual spending and geographical differences. However, an additional factor, as shown in this paper, may be heterogeneity in financial confidence. It appears plausible that spending responds to inflation expectations as predicted by an Euler equation only for those individuals that are (a) not liquidity constrained and (b) have sufficient financial literacy to understand the consequences of inflation for their personal savings and (c) trust their own literacy such that they will base financial decisions on their beliefs. It would be interesting to see future research on this.

8 Conclusion

The contribution of this paper to the literature is theoretical as well as empirical: I first show that heterogeneity in observed point forecasts can be driven by parameter heterogeneity in priors as well as signals received. In fact, when assuming a log-normal prior distribution, it is enough to adjust the volatility of noise to increase average inflation expectations. This is in line with the experience hypothesis that has become standard in the literature (D'Acunto et al., 2021; Jonung, 1981). However, the model also allows for an alternative interpretation. Flatter priors due to uncertainty can also increase expected value and variance of the posterior distribution. In fact, both forces interact. Signal volatility will increase average expectations whenever the prior is relatively flat and vice versa. My empirical findings support this hypothesis. Using data from the BOP-HH, I show that grocery shopping alone has no explanatory power in explaining inflation expectations. However, interactions with financial confidence measures indicate that grocery shopping increases expectations for those observations with low confidence and decreases them for the high confidence individuals. Hence, my results shed new light on the gender gap: It appears to be the composite of traditional gender norms and a reflection of lower financial confidence of women. This low confidence increases uncertainty and thus reliance on prices observed in the daily shopping experience.

I proceed by focusing in more detail on the financial confidence channel. I verify that there is a gender gap in financial confidence in my data and that the distribution of the low confidence sample is skewed to the right. Further, I find that the gender gap in expectations only exists for the low confidence individuals. Consequentially, the gender gap in expectations is a phenomenon of the upper tail. When restricting expectations to the lowest 50th percentile, the gap diminishes. According to my understanding, I am the first to give an in-depth treatment of the higher moments in the distribution. My findings are important to interpret survey data on inflation expectations beyond the gender gap: if an upward bias may be driven by uncertainty when inflation is low through rounding, the bias may in fact be downwards in high inflation periods and thus could make survey expectations appear more anchored than they are.

To support my argument, I examine two implications from the standard experience hypothesis about the gender gap in inflation expectations from a comparison of three surveys. Namely, that the gender gap should vanish for singles and that the gender gap should increase in periods when food price inflation is high. I reject both implications, suggesting that experience cannot be the sole driver of the gender gap.

The evidence suggests that the dominating experience hypothesis in the literature (D'Acunto et al., 2021) is not enough to explain the gender gap in inflation expectations. Financial confidence is an important channel, as grocery shopping increases expectations only in the low confidence sample, as predicted by the framework. My alternative hypothesis for the gender gap in inflation expectations can reconcile the stylized facts above with the evidence presented by D'Acunto et al. (2021) if in fact married women in traditional gender roles have lower financial confidence.

The fact that the gender gap appears to be driven largely by financial confidence has policy implications. While more than half of all women in the sample (those with expectations below the 50th percentile) have similar expectations to those of men and appear equally financially literate, there exists a large upper tail of women with lower confidence. This translates into rounded and less precise estimates. This matters firstly for female investment and saving behavior. Lusardi and Mitchell (2008) show that women often under-save for retirement, which is worsened by the fact that many reach an older age than male spouses. Expecting higher levels of inflation due to lower confidence rationalizes this result. Further, lower confidence may lead to lower perceptibly to policy communicated in expert language. If women pay less attention or are less likely to draw the correct conclusions from policy messages due to low levels of financial literacy in the tails and trust in own abilities, they will not adjust behavior as expected. This suggests that policy should focus on (a) improving financial literacy of women and (b) communicate monetary policy in simpler language to address individuals with lower literacy.

References

Abildgren, K., & Kuchler, A. (2021). Revisiting the inflation perception conundrum. *Journal of Macroeconomics*, *67*. https://doi.org/10.1016/j.jmacro.2020.103264

Arioli, R., Lim, Y. C., Sek, S. K., Easaw, J., Golinelli, R., Malgarini, M., Morvan, C., Jenkins, B., Carroll, C. D., Axtell, R., Branch, W., Christ, C., Dickens, W., Dotsey, M., Epstein, J., Goodfriend, M., Manning, J., Mccallum, B., Mehra, Y., . . . Schulhofer-Wohl, S. (2017). EU Consumers' Quantitative Inflation Perceptions and Expectations: an Evaluation. *Occasional Paper Series, European Central Bank*, *91*(186). https://www.ecb.europa.eu/pub/pdf/scpops/ecb.op186.en.pdf?c58df51446c7605dd5f939bcc7cd35b3

Armantier, O., Bruine De Bruin, W., Potter, S., Topa, G., Van Der Klaauw, W., & Zafar, B. (2013). Measuring inflation expectations. https://doi.org/10.1146/annurev-economics-081512-141510

Armantier, O., Nelson, S., Topa, G., van der Klaauw, W., & Zafar, B. (2016). The price is right: Updating inflation expectations in arandomized price information experiment. *Review of Economics and Statistics*, *98*(3), 503–523. https://doi.org/10.1162/REST{_}a{_}00499

Armantier, O., Topa, G., van der Klaauw, W., & Zafar, B. (2017). An Overview of the Survey of Consumer Expectations. *FRBNY Economic Policy Review*, *23*(2), 51–72.

Bachmann, R., Berg, T. O., & Sims, E. R. (2015). Inflation Expectations and Readiness to Spend: Cross-Sectional Evidence. *American Economic Journal: Economic Policy*, *7*(1), 1–35. https://doi.org/10.1257/pol.20130292

Binder, C. (2017). Measuring uncertainty based on rounding: New method and application to inflation expectations. *Journal of Monetary Economics*, *90*. https://doi.org/10.1016/j.jmoneco.2017.06.001

Binder, C. (2020). Coronavirus Fears and Macroeconomic Expectations. *The Review of Economics and Statistics*, *102*(4), 721–730. https://doi.org/10.1162/REST{_}A{_}00931

Blanchflower, D. G., & Maccoille, C. (2009). *The formation of inflation expectations: an empirical analysis for the UK*, NBER.

Breusch, T. S., & Pagan, A. R. (1979). A Simple Test for Heteroscedasticity and Random Coefficient Variation. *Econometrica*, *47*(5), 1287. https://doi.org/10.2307/1911963

Brischetto, A., & de Brouwer, G. (1999, March). *Householders' Inflation Expectations*, Reserve Bank of Australia. https://catalogue.nla.gov.au/Record/1434103/Details

Bruine De Bruin, W., Van der Klaauw, W., Downs, J. S., Fischhoff, B., Topa, G., Armantier, O., Chien, S., Colvin, T., Forman, D., Fielding, P., Greenwald, D., Gutsche, T., Holbrook, M., & Weerman, B. (2010). Expectations of Inflation: The Role of Demographic Variables, Expectation Formation, and Financial Literacy. *Journal of Consumer Affairs*, *44*(2), 381–402. https://doi.org/10.1111/J.1745-6606.2010.01174.X

Bruine de Bruin, W., van der Klaauw, W., van Rooij, M., Teppa, F., & de Vos, K. (2017). Measuring expectations of inflation: Effects of survey mode, wording, and opportunities to revise. *Journal of Economic Psychology*, *59*, 45–58. https://doi.org/10.1016/J.JOEP.2017.01.011

Bryan, M., & Venkatu, G. (2001). The Curiously Different Inflation Perspectives of Men and Women. *Economic Commentary*, (November 2001). https://www.clevelandfed.org/newsroom-and-events/publications/economic-commentary/economic-commentary-archives/2001-economic-commentaries/ec-20011101-the-curiously-different-inflation-perspectives-of-men-and-women

Bucher-Koenen, T., Alessie, R. J., Lusardi, A., & Rooij, M. v. (2021, May). *Fearless Woman: Financial Literacy and Stock Market Participation*, National Bureau of Economic Research. https://doi.org/10.3386/W28723

Bucher-Koenen, T., Lusardi, A., Alessie, R., & Rooij, M. v. (2014, December). *How financially literate are women? An overview and new insights*, National Bureau of Economic Research. https://doi.org/10.3386/W20793

Burke, M. A., & Manz, M. (2014). Economic Literacy and Inflation Expectations: Evidence from a Laboratory Experiment. *Source: Journal of Money, Credit and Banking*, *46*(7), 1421–1456. https://doi.org/10.1111/jmcb.12144

Chaney, C. K., Alvarez, R. M., & Nagler, J. (1998). Explaining the gender gap in U.S. presidential elections, 1980-1992. *Political Research Quarterly*, *51*(2), 311–339. https://doi.org/10.1177/106591299805100202

Corduas, M. (2022). Gender differences in the perception of inflation. *Journal of Economic Psychology*, *90*, 102522. https://doi.org/10.1016/J.JOEP.2022.102522

D'Acunto, F., Hoang, D., Paloviita, M., & Weber, M. (2019a, January). *IQ, Expectations, and Choice*, National Bureau of Economic Research. https://doi.org/10.3386/W25496

D'Acunto, F., Hoang, D., Paloviita, M., & Weber, M. (2019b). Cognitive Abilities and Inflation Expectations. *AEA Papers and Proceedings*, *109*, 562–66. https://doi.org/10.1257/PANDP.20191050

D'Acunto, F., Malmendier, U., & Weber, M. (2021). Gender Roles produce divergent economic expectations. *Proceedings of the National Academy of Science*, *118*(21). https://doi.org/10.1073/pnas.2008534118

Del Giovane, P., Fabiani, S., & Sabbatini, R. (2008). What's Behind 'Inflation Perceptions'? A Survey-Based Analysis of Italian Consumers. In P. Del Giovane & R. Sabbatini (Eds.), *The euro, inflation and consumer's perceptions* (pp. 157–197). Springer. https://doi.org/https://doi.org/10.1007/978-3-540-78370-1{_}7

Dräger, L., Menz, J. O., & Fritsche, U. (2014). Perceived inflation under loss aversion. *Applied Economics*, *46*(3), 282–293. https://doi.org/10.1080/00036846.2013.844328

Dräger, L., & Nghiem, G. (2020). Are Consumers' Spending Decisions in Line with an Euler Equation? *The Review of Economics and Statistics*, *103*(3), 580–596. https://doi.org/10.1162/rest{_}a{_}00909

Engelberg, J., Manski, C. F., & Williams, J. (2009). Comparing the Point Predictions and Subjective Probability Distributions of Professional Forecasters. *Journal of Business & Economic Statistics*, *27*(1), 30–41. https://doi.org/10.1198/jbes.2009.0003

Garbarino, E., & Strahilevitz, M. (2004). Gender differences in the perceived risk of buying online and the effects of receiving a site recommendation. *Journal of Business Research*, *57*(7), 768–775. https://doi.org/10.1016/S0148-2963(02)00363-6

Garriga, A. C. (2023). Box A: Public Confidence in the Bank of England. *National Institute UK Economic Outlook*.

Gillitzer, C., Prasad, N., & Robinson, T. (2021). Political Attitudes and Inflation Expectations: Evidence and Implications. *Journal of Money, Credit and Banking*, *53*(4), 605–634. https://doi.org/10.1111/JMCB.12797

Gorodnichenko, Y., & Sergeyev, D. (2021). *Zero Lower Bound on Inflation Expectations*, NBER.

Gwartney-Gibbs, P. A., & Lach, D. H. (2016). Sex Differences in Attitudes toward Nuclear War: *Journal of Peace Research*, *28*(2), 161–174. https://doi.org/10.1177/0022343391028002003

Hausman, J. A. (1978). Specification tests in econometrics. *Econometrica*, *46*(6), 1251–1271. https://doi.org/10.2307/1913827

Hausman, J. A., & Taylor, W. E. (1981). Panel Data and Unobservable Individual Effects. *Econometrica*, *49*(6), 1377–1398. https://doi.org/10.2307/1911406

Hey, J. D. (1984). The Economics of Optimism and Pessimism. *Kyklos*, *37*(2), 181–205. https://doi.org/10.1111/J.1467-6435.1984.TB00748.X

Jacobsen, B., Lee, J. B., Marquering, W., & Zhang, C. Y. (2014). Gender differences in optimism and asset allocation. *Journal of Economic Behavior & Organization*, *107*, 630–651. https://doi.org/10.1016/J.JEBO.2014.03.007

Jonung, L. (1981). Perceived and Expected Rates of Inflation in Sweden. https://www.jstor.org/stable/1803477

Kim, G., & Binder, C. (2020). Learning-through-Survey in Inflation Expectations. *SSRN Electronic Journal*. https://doi.org/10.2139/SSRN.3790834

Krifka, M. (2009). *Approximate Interpretations of Number Words: A Case for Strategic Communication*, Humboldt Universität, Zentrum für Allgemeine Sprachwissenschaft.

Lamla, M. J., & Lein, S. M. (2008). The Role of Media for Consumers' Inflation Expectation Formation. *Journal of Economic Behavior & Organization*, *106*, 62–77. https://doi.org/10.1016/j.jebo.2014.05.004

Leung, C. (2009). The demographics of household inflation perceptions and expectations. *Reserve Bank of New Zealand Bulletin*, *72*(2), 34–42.

Lin, Y. C., & Raghubir, P. (2005). Gender differences in unrealistic optimism about marriage and divorce: Are men more optimistic and women more realistic? *Personality and Social Psychology Bulletin*, *31*(2), 198–207. https://doi.org/10.1177/0146167204271325

Lindén, S. (2015). Quantified Perceived and Expected Inflation in the Euro Area - How Incentives Improve Consumers Inflation Forecasts. *Joint European Commission - OECD Workshop on International Development of Business and Consumer Tendency Surveys*.

Lusardi, A., & Mitchell, O. S. (2008). Planning and Financial Literacy: How Do Women Fare? *American Economic Review*, *98*(2), 413–417. https://doi.org/10.1257/aer.98.2.413

Lyons, K. S., Stewart, B. J., Archbold, P. G., & Carter, J. H. (2009). Optimism, Pessimism, Mutuality, and Gender: Predicting 10-Year Role Strain in Parkinson's Disease Spouses. *The Gerontologist*, *49*(3), 378–387. https://doi.org/10.1093/GERONT/GNP046

Malmendier, U., & Nagel, S. (2016). Learning from Inflation Experiences. *The Quarterly Journal of Economics*, *131*(1), 53–87. https://doi.org/10.1093/QJE/QJV037

Palmqvist, S., & Strömberg, L. (2004). Households' inflation opinions - a tale of two surveys. *Economic Review*, (4), 23–42. https://www.researchgate.net/publication/237826248_Households'_inflation_opinions_-_a_tale_of_two_surveys

Pfajfar, D., & Santoro, E. (2008). *Asymmetries in Inflation Expectation Formation Across Demographic Groups*.

Reiche, L., & Meyler, A. (2022, February). *Making Sense of Consumer Inflation Expectations*, European Central Bank. https://www.ecb.europa.eu/pub/pdf/scpwps/ecb.wp2642~f96823e5de.en.pdf?a6b9965facced0d1ecc25de9a5a4f94d

Silverman, B. W. (1986). Density Estimation for Statistics and Data Analysis. In Chapman & Hall (Eds.), *Monographs on statistics and applied probability*.

Wallace, T. D., & Hussain, A. (1969). The Use of Error Components Models in Combining Cross Section with Time Series Data. *Econometrica*, *37*(1), 55. https://doi.org/10.2307/1909205

A The Gender Gap in the Literature

Table 8: The gender gap in the literature

Year	Author	Period	Country	Survey	N	GG raw data	GG multivariate	Controls
1981	Jonung	1978	Sweden		6019	0.8	-0.2	perceived inflation, age, income
1999	Brischetto & de Brouwer	1995-1998	Australia	Melbourne Institute	28077	-	1.45***	age, occupation, education, political view, income, location
2001	Bryan & Venkatu	1998-2001	US	FRBC/OSU Inflation Psychology Survey	~20000	2.3	2.1[1]	age, race, education, income
2004	Palmqvist & Strömberg	2001-2004	Sweden	GfK survey	~46500	0.98	0.91***	marital status, children, education, income, age, work status, housing status
2005	Linden	2003-2005	Euro area	Harmonised Business and Consumer Survey	-	0.7	-	-
2008	Pfajfar & Santoro	1978-2008?	US	Michigan Survey	-	1.09	-	-
2009	Blanchflower & MacCoille	2009	England	BoE Inflation Attitudes Survey	5243	-	0.19	age, worker, education, housing status
2009	Leung	1998-2008	New Zealand	Marketscope survey	16044	-	0.79***	age, employment, ethnicity, main grocery shopper, income, occupational skill, region, season dummies
2010	Bruine de Bruin et al.	2007-2008	US	RAND American Life Panel	299	0.39	0.00	thoughts about finances, financial literacy, race,marital status, education, income, age
2016	Armantier et al.	2011	US	RAND American Life Panel	244	2.44	-	-
2016	Ballantyne et al.	1995-2014	Australia	Melbourne Institute	262878	-	0.80***	age, education, income
2017	Arioli et al.	2004-2015	European Union	Harmonised Business and Consumer Survey	-	1.9	-	-
2020	D'Acunto et al.	2015-2016	US	Chicago Booth Expectations and Attitudes Survey	20866	0.4	0.16[2]	grocery shopper, age, employment, income, home ownership, marital status, education, race, risk tolerance, confidence in forecast accuracy
2020	Dräger	2015-2016	Germany	University of Hamburg Survey	421	0.65	-	-
2021	Reiche & Meyler	2004-2020	European Union	Harmonised Business and Consumer Survey	2539617	-	1.02***	age, education, income, economic sentiment, macroeconomic environment, country fixed effects
2021	Abildgren & Kuchler	2007-2016	Denmark	Danish Consumer Expectations Survey	52957	-	1.44***[3]	income, age, education, pessimism, overpessimism

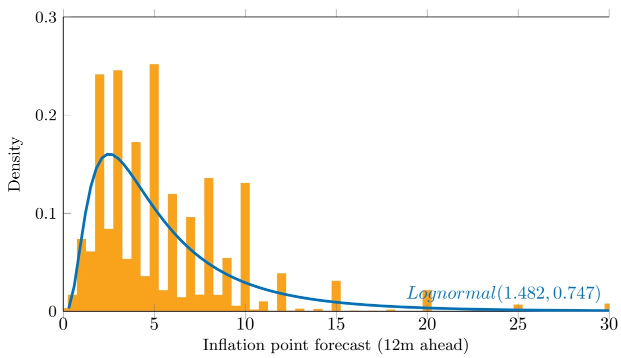
For multivariate gender gaps: *p<0.1; **p<0.05; ***p<0.01

[1] No significance results reported, [2] Focus on intra-household gender gaps, [3] Use perceptions instread of expectations

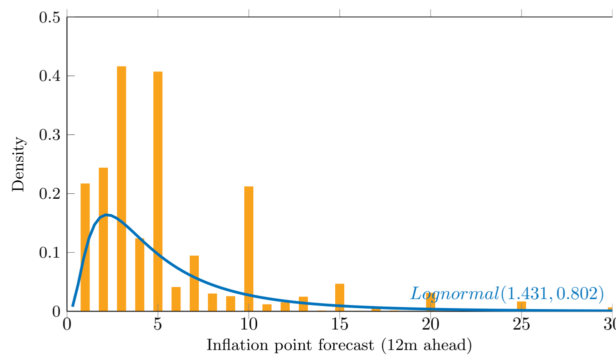
B Additional Material for Bayesian Framework

B.1 Motivating Log-Normality

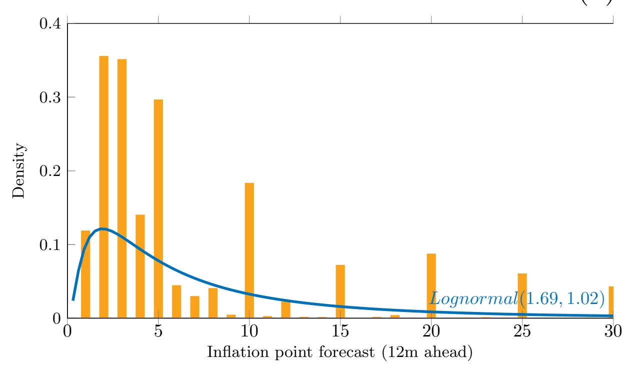
The cross-sectional distribution of point forecasts is in line with a log-normal posterior. Figure B.1 shows the histograms of the pooled cross-sections of three surveys, the BOP-HH, the MSC and the SCE. The fitted log-normal parameters are shown for each survey.



(a) BOP-HH



(b) MSC



(c) SCE

Figure B.1: Histogram and fitted distribution of inflation expectation point forecasts

Notes: Figure B.1 compares the distribution of inflation expectations (measured as point forecasts over 12 months) pooled across all time periods in three surveys (BOP-HH, MSC and SCE). There is one plot per survey. The log-normal distribution is fitted to the data. The figure shows that in all surveys the distribution is right skewed. Rounded numbers (i.e. multiples of 5 or 10) are chosen more frequently than estimated. Reiche and Meyler (2022) show how the fit can be improved by acknowledging the existence of two groups, rounders and non-rounders. This method is beyond the scope of this simple exercise. The authors chose a log-logistic distribution instead of a log-normal distribution but report similar features.

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; Federal Reserve Bank of New York (FRBNY), SCE, June 2013 - November 2020; University of Michigan, Survey Research Center, MSC, January 1978 - January 2023; own calculations

B.2 Algebraic manipulations to derive the log-normal posterior

The prior is defined as

$$\theta \sim \text{LogNormal}\left(\mu_j^{i,0}, \frac{1}{\tau_j^{i,0}}\right)$$

$$p(\theta) = \frac{\sqrt{\tau_j^{i,0}}}{\theta\sqrt{2\pi}} \exp\left(-\frac{\tau_j^{i,0}(\ln\theta - \mu_0)^2}{2}\right)$$

I define the signal to be unbiased. This requires that:

$$\mathbb{E}(x|\theta) = \theta$$

$$e^{\mu_j^{i,x} + \frac{1}{2\tau_j^{i,0}}} = \theta$$

$$\mu_j^{i,x} = \ln(\theta) - \frac{1}{2\tau_j^{i,x}}$$

Hence, the likelihood:

$$p(x|\theta) = \frac{\sqrt{\tau_j^{i,x}}}{x\sqrt{2\pi}} \exp\left(-\frac{\tau_j^{i,x}(\ln(x) - \ln\theta + \frac{1}{2\tau_j^{i,x}})^2}{2}\right)$$

Then, I can compute the posterior following Bayesian updating:

$$\begin{aligned}
&p(\theta|x) \propto p(\theta)p(x|\theta) \\
&= \frac{\sqrt{\tau_j^{i,0}}}{\theta\sqrt{2\pi}} \exp\left(-\frac{\tau_j^{i,0}(\ln\theta - \mu_j^{i,0})^2}{2}\right) \frac{\sqrt{\tau_j^{i,x}}}{x\sqrt{2\pi}} \exp\left(-\frac{\tau_j^{i,x}(\ln(s) - \ln\theta + \frac{1}{2\tau_j^{i,x}})^2}{2}\right) \\
&= \frac{\sqrt{\tau_j^{i,0}}\sqrt{\tau_j^{i,x}}}{\theta x 2\pi} \exp\left(-\frac{1}{2}\left[\tau_j^{i,0}\left(\ln(\theta)^2 - 2\mu_j^{i,0}\ln(\theta) - (\mu_j^{i,0})^2\right)\right.\right. \\
&\left.\left. + \tau_j^{i,x}\left(\ln(x)^2 + \ln(\theta)^2 + (\frac{1}{2\tau_j^{i,x}})^2 + 2\left(\ln(x)\frac{1}{2\tau_j^{i,x}} - \ln(s)\ln(\theta) - \ln(\theta)\frac{1}{2\tau_j^{i,x}}\right)\right)\right) \\
&\propto \frac{1}{\theta} \exp\left(-\frac{1}{2}\left[\tau_j^{i,0}\left(\ln(\theta)^2 - 2\ln(\theta)\mu_j^{i,0}\right) + \tau_j^{i,x}\left(\ln(\theta)^2 - 2\ln(sx\ln(\theta) - \ln(\theta)\right)\right]\right) \\
&\propto \frac{1}{\theta} \exp\left(-\frac{1}{2}\left[(\tau_j^{i,0} + \tau_j^{i,x})\ln(\theta)^2 - 2(\tau_j^{i,0}\mu_j^{i,0} + \tau_j^{i,x}\ln(x) + \frac{1}{2})\ln(\theta)\right]\right)
\end{aligned}$$

Note that this is proportional to a log-normal distribution:

$$p(\theta|x) \propto \frac{1}{\theta} \exp\left(-\frac{\hat{\tau}^i(\ln(\theta) - \hat{\mu}^i)^2}{2}\right)$$

where

$$
\begin{aligned}
\hat{\mu}^i_j &= \frac{\mu^{i,0}_j \tau^{i,0}_j + \ln(x^i_j)\tau^{i,x}_j + \frac{1}{2}}{\tau^{i,x}_j + \tau^{i,0}_j} \\
\hat{\tau}^i &= \tau^{i,0}_j + \tau^{i,x}_j
\end{aligned}
$$

B.3 The effect of ϵ_i on parameters and moments

Parameters

$$
\begin{aligned}
\hat{\mu}^i(\epsilon_i) &= \frac{\mu^i_0\tau^i_0 + \ln(x)\tau^i_x(\epsilon_i) + \frac{1}{2}}{\tau^i_x(\epsilon_i) + \tau^i_0} \\
&= \frac{\mu^i_0\tau^i_0 + \ln(x)(\tau_x - \epsilon_i) + \frac{1}{2}}{(\tau_x - \epsilon_i) + \tau^i_0} \\
\frac{d\hat{\mu}^i(\epsilon_i)}{d\epsilon_i} &= \frac{\mu^i_0\tau^i_0 - \ln(x) + \frac{1}{2}}{(\tau_x + \tau^i_0 - \epsilon_i)^2} > 0 \text{ whenever } \mu^i_0 + \frac{1}{2\tau^i_0} > \ln(s) \\
\hat{\tau}^i(\epsilon_i) &= \tau^i_x(\epsilon_i) + \tau^i_0 \\
&= \tau_x + \tau^i_0 - \epsilon_i \\
\frac{d\hat{\tau}^i(\epsilon_i)}{d\epsilon_i} &= -1 < 0
\end{aligned}
$$

Moments

$$
\begin{aligned}
\mathrm{E}(\theta|x,\epsilon_i) &= e^{\frac{\mu_0\tau^i_0 + \ln(x)(\tau_x - \epsilon_i) + 1}{\tau_x + \tau^i_0 - \epsilon_i}} \\
\frac{d\mathrm{E}(\theta|x,\epsilon_i)}{d\epsilon_i} &= \mathrm{E}(\theta|x,\epsilon_i) \times \left(\frac{\mu^i_0\tau^i_0 - \ln(x)\tau^i_0 + 1}{(\tau_x + \tau^i_0 - \epsilon_i)^2}\right) > 0 \text{ whenever } \mu^i_0 + \frac{1}{\tau^i_0} > \ln(x) \\
\frac{d\mathrm{Var}(\theta|x,\epsilon_i)}{d\epsilon_i} &= e^{\frac{2\mu^i_0\tau^i_0 - 2\ln(x)(\tau_x - \epsilon_i) + 3}{\tau_x + \tau^i_0 - \epsilon_i}} - e^{\frac{2\mu^i_0\tau^i_0 - 2\ln(sx(\tau_x - \epsilon_i) + 2}{\tau_x + \tau^i_0 - \epsilon_i}} \\
\frac{d\mathrm{Var}(\theta|s,\epsilon_i)}{d\epsilon_i} &= e^{\frac{2\mu^i_0\tau^i_0 - 2\ln(s)(\tau_x - \epsilon_i) + 3}{\tau_x + \tau^i_0 - \epsilon_i}} \times \left(\frac{2\mu^i_0\tau^i_0 - 2\ln(s)\tau^i_0 + 3}{(\tau_x + \tau^i_0 - \epsilon_i)^2}\right) \\
&\quad - e^{\frac{2\mu^i_0\tau^i_0 - 2\ln(x)(\tau_x - \epsilon_i) + 2}{\tau_x + \tau^i_0 - \epsilon_i}} \times \left(\frac{2\mu^i_0\tau^i_0 - 2\ln(x)\tau^i_0 + 2}{(\tau_x + \tau^i_0 - \epsilon_i)^2}\right) > 0
\end{aligned}
$$

B.4 The effect of δ_i on parameters and moments

Parameters

$$
\begin{aligned}
\hat{\mu}^i(\delta_i) &= \frac{\mu_0^i(\delta_i)\tau_0^i(\delta_i) + \ln(x)\tau_x^i + \frac{1}{2}}{\tau_x^i + \tau_0^i(\delta_i)} \\
&= \frac{(\mu_0 + \frac{\delta_i}{\tau_0(\tau_0-\delta_i)})(\tau_0 - \delta_i) + \ln(x)\tau_x^i + \frac{1}{2}}{\tau_x^i + (\tau_0 - \delta_i)} \\
\frac{d\hat{\mu}^i(\delta_i)}{d\delta_i} &= \frac{\tau_x^i\left(-\mu_0 + \frac{1}{\tau_0} + \frac{3}{2\tau_x^i} + \ln(x)\right)}{(\tau_x^i + \tau_0 - \delta_i)^2} > 0 \text{ whenever } \ln(x) + \frac{3}{2\tau_x^i} + \frac{1}{\tau_0} > \mu_0 \\
\hat{\tau}^i(\delta_i) &= \tau_x^i + \tau_0^i(\delta_i) \\
&= \tau_x^i + \tau_0 - \delta_i \\
\frac{d\hat{\tau}^i(\delta_i)}{d\delta_i} &= -1 < 0
\end{aligned}
$$

Moments

$$
\begin{aligned}
\mathrm{E}(\theta|x,\delta_i) &= e^{\frac{\mu_0(\tau_0-\delta_i) + \frac{\delta_i}{\tau_0} + \ln(x)\tau_x^i + 1}{\tau_x^i + \tau_0 - \delta_i}} \\
\frac{d\mathrm{E}(\theta|x,\delta_i)}{d\delta_i} &= \mathrm{E}(\theta|x,\delta_i) \times \left(\frac{\tau_x^i\left[-\mu_0 + \frac{1}{\tau_0} + \frac{2}{\tau_x^i} + \ln(x)\right]}{(\tau_x^i + \tau_0 - \delta_i)^2}\right) \\
\mathrm{Var}(\theta|x,\delta_i) &= e^{\frac{2\mu_0(\tau_0-\delta_i) + \frac{2\delta_i}{\tau_0} + 2\ln(x)\tau_x^i + 3}{\tau_x^i + \tau_0 - \delta_i}} - e^{\frac{2\mu_0(\tau_0-\delta_i) + \frac{2\delta_i}{\tau_0} + 2\ln(x)\tau_x^i + 2}{\tau_x^i + \tau_0 - \delta_i}} \\
\frac{d\mathrm{Var}(\theta|x,\delta_i)}{d\delta_i} &= e^{\frac{2\mu_0(\tau_0-\delta_i) + \frac{2\delta_i}{\tau_0} + 2\ln(x)\tau_x^i + 3}{\tau_x^i + \tau_0 - \delta_i}} \times \left(\frac{\tau_x^i\left[-2\mu_0 + \frac{2}{\tau_0} + \frac{5}{\tau_x^i} + 2\ln(x)\right]}{(\tau_x^i + \tau_0 - \delta_i)^2}\right) \\
&\quad - e^{\frac{2\mu_0(\tau_0-\delta_i) + \frac{2\delta_i}{\tau_0} + 2\ln(x)\tau_x^i + 2}{\tau_x^i + \tau_0 - \delta_i}} \times \left(\frac{\tau_x^i\left[-2\mu_0 + \frac{2}{\tau_0} + \frac{4}{\tau_x^i} + 2\ln(x)\right]}{(\tau_x^i + \tau_0 - \delta_i)^2}\right) > 0
\end{aligned}
$$

C Details of SCE and MSC

C.1 Federal Reserve Bank of New York Survey of Consumer Expectations (SCE)

The SCE was launched as a monthly panel from June 2013 until November 2020 and is publicly available. Approximately 1200 household heads participate up to 12 times, resulting in a similar rotational structure as the BOP-HH. The SCE also includes a combination of core questions and ad-hoc questions. Gender is identified by self-assessment, but only a binary option is given. Some respondents refuse to indicate such that an error message is caused. Again, those respondents are excluded from the survey. Inflation is measured in a two-step procedure, as in the BOP-HH. Respondents indicate if they expect inflation or deflation over the next 12 months. They are given no definition of inflation but told that deflation is the opposite of inflation. Then they are asked to give a best guess for the rate of inflation/deflation over the next 12 months. Finally, as in the BOP-HH, respondents indicate probabilities attached to inflation bins. In addition, the core survey includes numeracy questions that capture the same financial literacy test as featured in the 25th survey wave in the BOP-HH. Note that the SCE is used in the analysis of D'Acunto et al. (2021). Since household roles are not elicited in this survey, the authors proxy them through age and region. The survey features an additional survey question on price expectations for specific items, namely gas, food, college education, medical care, rent, and gold (used as robustness checks in Appendix F.3).

C.2 Michigan Survey of Consumers (MSC)

The MSC is one of the oldest running household panels, with monthly microdata available publicly since 1978. I use data until January 2023. Each month, approximately 500 consumers across the United States respond to the 50 core questions. This smaller size of respondents allows for a less in-depth analysis of demographic factors and has lower spacial representation in each month, however, the longer time series is helpful to show that effects persist over time. Recruitment is via random telephone sampling. As in the BOP-HH participants rotate, in the MSC each month about 60% of respondents have participated before. Further, inflation expectations are elicited similarly. Respondents are first asked to indicate if they expect prices in general to go up, stay the same or go down in the next 12 months. Then they are asked by how much in the next step. There are two key differences to the BOP-HH and SCE: First, the survey focuses on *prices in general* rather than *inflation*. Bruine De Bruin et al. (2010) and Bruine de Bruin et al. (2017) have shown that this biases consumers to think more about their personal situation and increases expectations on average. Secondly, the MSC allows for the option to indicate that prices stay about the same. These respondents will be probed to specify if they mean inflation to remain constant or prices to remain constant. If prices are chosen, no numerical answer is given and inflation expectations are marked as zero. This leads to a large fraction of consumers responding zero (16.45% compared to less than 0.27% in the BOP-HH where consumers are initially given the binary choice of choosing "inflation" or "deflation"). This difference has been pointed out by Armantier et al. (2017) who argue that the design as it is featured in the BOP-HH and the SCE is more precise. Further, the survey probes consumers with expectations in excess of 5% if they are really sure about their estimate. Probing reduces average expectations and thus biases averages downwards (Bruine de Bruin et al., 2017).

D Financial Literacy

D.1 Questions in BOP-HH survey wave 25, January 2022

W25: In the following section, we would like to ask you a few more questions on general economic topics.

Question: Let us assume you have a balance of €100 in your savings account. This balance bears interest at an annual rate of 2%, and you leave it there for five years. How high do you think your balance will be after five years?

1. Higher than €102
2. Exactly €102
3. Lower than €102

Don't know
No answer

Question: Let us assume that the interest paid on your savings account is 1% per year and the inflation rate is 2% per year. After one year, do you think you will be able to buy just as much, more, or less than you could today with the balance in your savings account?

1. More than today
2. Just as much as today
3. Less than today

Don't know
No answer

Question: Do you agree with the following statement? "Investing in shares of a single company is less risky than investing in a fund containing shares of similar companies."

1. Agree
2. Disagree

Don't know
No answer

D.2 Questions in the SCE, asked only new respondents

QnumIntro. Next, we would like to ask you five questions to see how people use numbers in everyday life. Please answer the following questions by filling in the blank.

QNUM2. Let's say you have $200 in a savings account. The account earns ten percent interest per year. Interest accrues at each anniversary of the account. If you never withdraw money or interest payments, how much will you have in the account at the end of two years?
$ _

No answer

QNUM8. Imagine that the interest rate on your savings account was 1% per year and inflation was 2% per year. After one year, how much would you be able to buy with the money in this account?

1. More than today
2. Just as much as today
3. Less than today

No answer

QNUM9. Please tell me whether this statement is true or false: Buying a single company's stock usually provides a safer return than a stock mutual fund.

1. True
2. False

No answer

D.3 Predicting financial confidence scores

Table 9: Explaining financial confidence through financial confidence variables

	Correct answers in financial literacy test (0-3)	
	BOP-HH	SCE
	(1)	(2)
age	−0.002	0.01***
	(0.003)	(0.0004)
female	−0.59***	−0.69***
	(0.08)	(0.01)
single	0.28**	0.06***
	(0.11)	(0.01)
hhinc	0.12***	0.10***
	(0.02)	(0.003)
educ	0.07***	0.26***
	(0.01)	(0.004)
round	−0.16	−0.61***
	(0.10)	(0.01)
refresher	−0.04	−0.41***
	(0.26)	(0.02)
qeasy	0.44***	
	(0.06)	
qinterest	0.25***	0.06***
	(0.05)	(0.01)
Time dummies	No	Yes
Between effects	No	No
Region dummies	Yes	Yes
Observations	2,916	113,165

*p<0.1; **p<0.05; ***p<0.01

Notes: Table 9 shows the regression coefficients of an ordered logistic regression of demographic variables (age, female, single, hhinc and educ) and financial confidence predictors discussed in section 3.2.2 on the number of correct responses in the financial literacy test.

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; own calculations

D.4 Probabilistic survey responses

I follow the method of Engelberg et al. (2009). In the simple case of the respondent using only one or two bins, I fit an isosceles triangle using the bin edges as the limits to the support. In contrast, I estimate an unimodal generalized beta distribution with two parameters for three or more bins. In the cumulative distribution shown in Equation 16, l and r indicate the limits to the support taken directly from the bin edges.

$$Beta(t,a,b,l,r) = \left\{ \begin{array}{cc} 0 & \text{if} \quad t \leq l \\ \frac{1}{B(a,b)} \int_l^t \frac{(x-1)^{a-1}(r-x)^{b-1}}{(a-l)^{a-b-1}} dx & \text{if} \quad l < t \leq r \\ 1 & \text{if} \quad t \geq r \end{array} \right\} \tag{16}$$
$$\text{where} \quad B(a,b) = \frac{\Gamma(a)\Gamma(b)}{\Gamma(a+b)} \quad \text{and} \quad \Gamma(a) = \int_0^\infty x^{a-1} e^{-1} dx$$

I then estimate the shape parameters a and b, minimizing the sum of squared differences of the implied beta distribution and the probability mass allocated by the respondent, where the sum is taken over the right-hand edges of each bin (Equation 17).

$$\min_{a>1,b>1} \sum_{i=1}^{9} (Beta(t_i,a,b,l,r) - F(t_i))^2 \tag{17}$$

Since there are ten bins in total, but the last one is unbounded to the right, the sum is taken over nine values. Once the distribution is computed for each observation, the interquartile range can be inferred. It is a good indicator of uncertainty and shows how spread out the probability mass is. A problem with this approach is that I lose many observations, as those with probability mass in non-contiguous bins cannot be rationalized. Further, I cannot fit a probability distribution when the highest mass is in an infinite bin ('less / higher than 12%'). Some authors set open-ended intervals twice the width of the nearest closed intervals (here, this would mean 12-20%). However, given that point estimates reach up to 50% and higher, this assumption appears not justified in consumer data.

D.5 Financial Confidence and Uncertainty

In the framework low financial confidence is modelled as a flat prior, i.e. as high uncertainty around one's point forecast. I show that my financial confidence score is indeed negatively correlated with uncertainty around point forecasts. For this is use the interquartile range computed as in the section above (Section D.4) and plot a binscatter in Figure D.1.

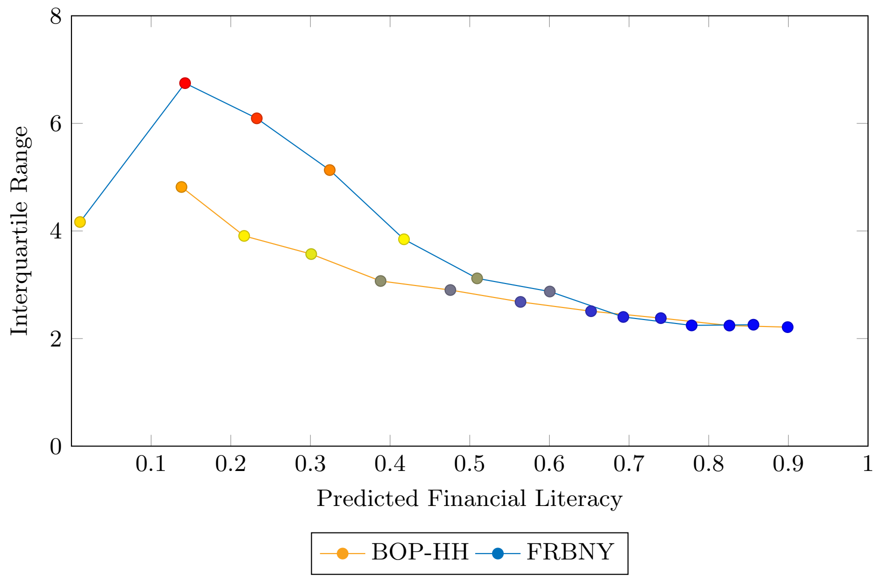


Figure D.1: Binscatter of interquartile range and predicted financial confidence

Notes: Figure D.1 plots the average interquartile range per predicted financial confidence bin, where the full sample is split into 10 bins of equal size. The BOP-HH is shown in the dots connected by the orange line and the SCE is connected by the blue line. The data is pooled across all time periods. No controls are included.

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; Federal Reserve Bank of New York (FRBNY), SCE, June 2013 - November 2020; own calculations

E Estimation Technique

The basic model for a panel regression can be written as

$$y_i t = \beta \times female_i + X_{it} \times \theta + W_i \times \gamma + D_t \times \delta + (v_i + \rho_{it}) \tag{18}$$

In the baseline regression, on the left-hand side, $y_i t$ captures the reported quantitative inflation expectations of individual i at time t. X_{it} contains time-varying individual-specific explanatory variables, W_i time-invariant explanatory variables, and D_t time fixed-effects. The error term has two components. In the error term, v_i contains the unobserved time-invariant component, potentially including topic-specific education and occupation (i.e., an Economics degree and a job at a bank) and overall financial confidence. Given that these may be correlated with both inflation expectations and being female, the dummy variable 'female' might be endogenous, leading to an omitted variable bias. Lastly, ρ_{it} is the time-variant error term. Possible time-variant factors could be current news consumption (Lamla & Lein, 2008), the political environment (Gillitzer et al., 2021), or the recent consumption of goods with exceptionally high or low inflation.

Assumptions about the error term have implications for the efficiency and consistency of the estimator chosen. Candidate estimators for a panel model are standard Ordinary Least Squares (OLS), Pooled OLS (PO) - a least squared dummy variable estimation with time fixed effects, Random Effects (RE) – a feasible GLS estimation applying Wallace and Hussain (1969) method of identifying the variance-covariance matrix of the combined error terms $\Omega = E(uu')$ through $\hat{u}_{OLS}$, and Fixed Effects/Within Groups (FE) estimator. In addition, I develop an alternative estimator that combines elements of the between groups and pooled OLS (LSDV_control). Consistency of the OLS estimator requires linearity, exogeneity, homoscedasticity and no multicollinearity. Homoscedasticity is too restrictive in this context as the variance of the error term increases for larger values of inflation expectations, and the Breusch-Pagan test for heteroscedasticity rejects the null of homoskedasticity (Breusch & Pagan, 1979). Controlling for heteroscedasticity, the Pooled OLS estimator is consistent when maintaining the assumptions of exogeneity with respect to time-invariant ($E(X_{it}\eta_i) = 0$) and predeterminedness with respect to time-varying error term components ($E(X_{it}v_{is}) = 0$ for $s \geq t$). I find that OLS and Pooled OLS estimates are similar. If the assumption of predeterminedness is strengthened to strict exogeneity ($E(X_{it}v_{is}) = 0$ for all s, t), the Random Effects GLS estimator has higher efficiency. Coefficients are found to be almost identical, suggesting no concerns about the strict exogeneity assumption. This is reasonable given the short time frame in which the concern is greater for time invariant unobservables.

Time invariant variables are unidentified in the fixed effects estimator but applying the

Hausman test (Hausman, 1978) to the time varying subset of explanatory variables rejects the null of random effects. Therefore, I develop an alternative estimator to recover estimates of time invariant variables. I incorporate between effects in the existing model and apply the least squared dummy variable estimator to the transformed model (see equation 19). Thus, the estimator is a controlled version of the LSDV estimator. As long as fixed effects can be explained by the between effects β and γ can be identified. I argue, that this estimator is preferred because of its robustness to fixed effects while maintaining identification of time invariant coefficients. The assumption of strict exogeneity with respect to time varying unobservables is acceptable given the short time frame. This estimator is not unique with these characteristics. An alternative method to estimate time invariant variables in a panel with fixed effects has been proposed by Hausman and Taylor (1981).[19] A clear disadvantage of their estimator is that the researcher must decide arbitrarily which are the endogenous variables in this model. Further, it is not clear whether there are any endogenous time invariant variables in the present model. Thus, the estimator introduced above is chosen in this analysis.

$$y_it = \beta \times female_i + X_{it} \times \theta_1 + \bar{X}_i \times \theta_2 + W_i \times \gamma + D_t \times \delta + (v_i + \rho_{it}) \tag{19}$$

F Additional Empirical Material

F.1 The role of demographics

I verify that the gender gap in inflation expectations cannot be explained by standard demographic variables such as age, income and education, which may be distributed differently for men and women. To do so, Table 10 shows their interaction effects with female and Figure F.1 computes the implied gender gap for different demographics on this basis.

[19] They develop an estimator in two stages which is asymptotically efficient. In the first stage, the coefficients on time-varying variables are estimated using Within Groups. Then a two-stage least squares regression is performed where the endogenous time-invariant variables are estimated using Pooled OLS. Finally, the fitted values are used in a regression of the residuals from the first stage on time-invariant exogenous and endogenous variables.

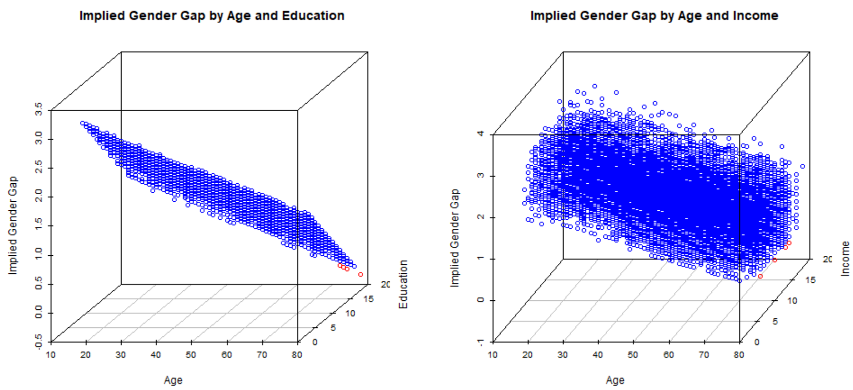
Table 10: The gender gap and demographic controls

	Inflation expectation (12 months ahead, point estimate)					
	BOP-HH		SCE		MSC	
	(1)	(2)	(3)	(4)	(5)	(6)
Constant	6.40***	5.06***	8.50***	7.66***	7.30***	6.20***
	(0.19)	(0.20)	(0.38)	(0.41)	(0.26)	(0.26)
female	1.45***	5.02***	1.64***	3.46***	0.86***	3.17***
	(0.04)	(0.20)	(0.07)	(0.35)	(0.02)	(0.10)
age	−0.02***	−0.01***	0.01***	0.01***	−0.02***	−0.01***
	(0.001)	(0.002)	(0.002)	(0.003)	(0.001)	(0.001)
hhinc	−0.24***	−0.17***	−0.38***	−0.34***	−0.0000***	−0.0000***
	(0.01)	(0.01)	(0.02)	(0.02)	(0.0000)	(0.0000)
educ	−0.11***	−0.07***	−0.50***	−0.33***	−0.24***	−0.10***
	(0.01)	(0.01)	(0.02)	(0.03)	(0.01)	(0.01)
single	−0.56***	−0.54***	−0.07	−0.09	−0.004	0.002
	(0.05)	(0.05)	(0.08)	(0.08)	(0.03)	(0.03)
female x age		−0.02***		0.01*		−0.02***
		(0.002)		(0.005)		(0.001)
female x hhinc		−0.16***		−0.10***		−0.0000***
		(0.01)		(0.03)		(0.0000)
female x educ		−0.12***		−0.36***		−0.30***
		(0.01)		(0.05)		(0.02)
Year dummies	Yes	Yes	Yes	Yes	Yes	Yes
Between effects	Yes	Yes	No	No	No	No
Observations	111,085	111,085	115,491	115,491	261,375	261,375
R^2	0.13	0.13	0.03	0.03	0.12	0.12
Adjusted R^2	0.13	0.13	0.03	0.03	0.12	0.12
F Statistic	415.61***	396.22***	35.66***	35.75***	64.00***	65.02***

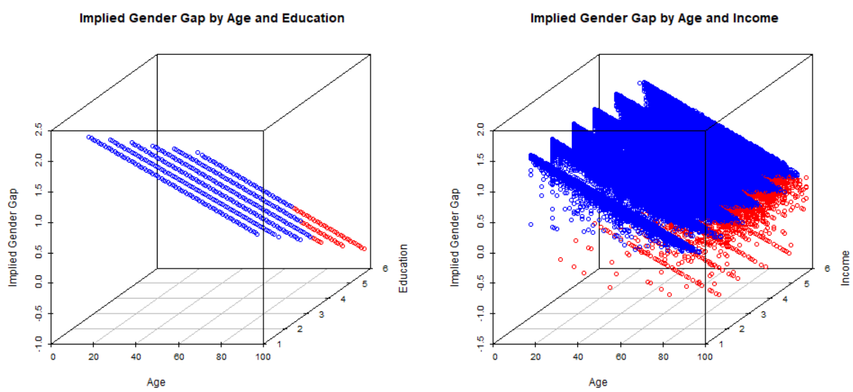
*p<0.1; **p<0.05; ***p<0.01

Notes: Table 10 compares the regression coefficients of a pooled OLS estimation of individual inflation expectations (12 months ahead, point estimate), on the dummy variable female, a continuous variable age, and the ordered categorical variables education and household income. Standard errors in brackets below. The estimation technique is discussed in Section E. Since the SCE's panel structure cannot be used due to crucial demographic questions being asked only once, there are no between effects for this survey. Models (1) and (2) compare coefficients including and exclusing interactions of female with age, education and income in the BOP-HH, (3) and (4) for the SCE and (5) and (6) for the MSC.

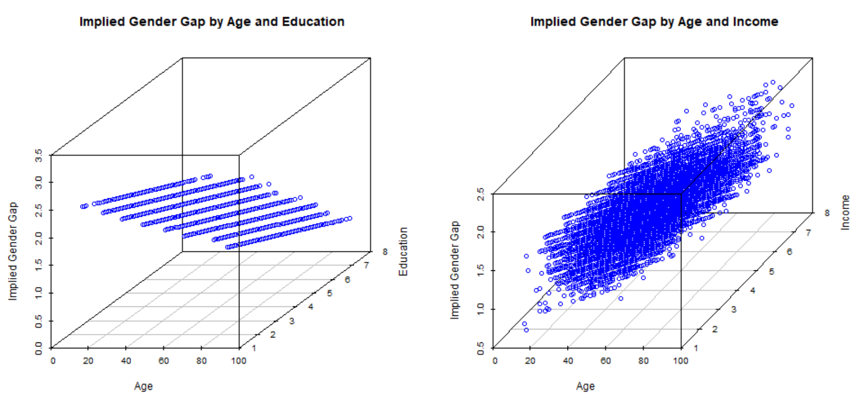
Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; Federal Reserve Bank of New York (FRBNY), SCE, June 2013 - November 2020; University of Michigan, Survey Research Center, MSC, January 1978 - January 2023; own calculations



(a) BOP-HH



(b) MSC



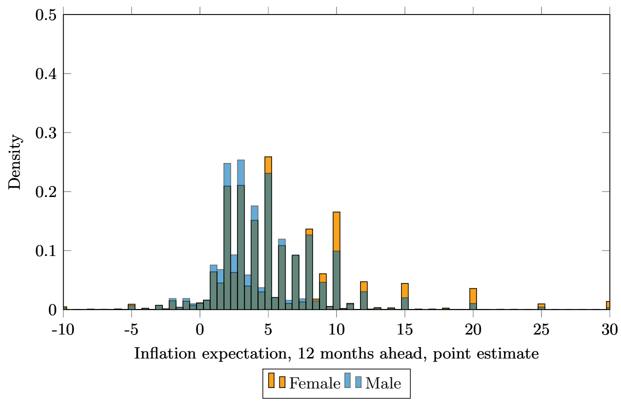
(c) SCE

Figure F.1: Implied gender gap by demographics

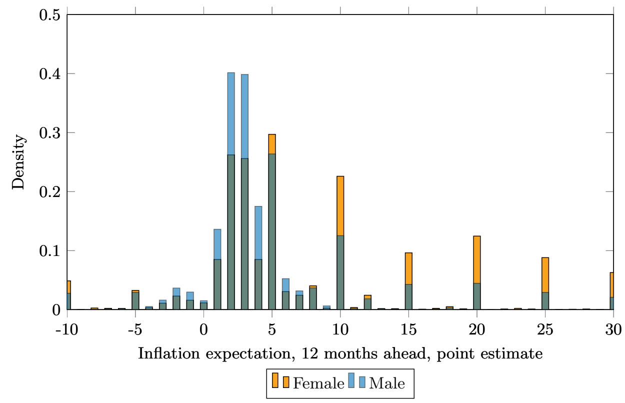
Notes: Figure F.1 plots the implied gender gap for combinations of age and education (holding income at median value, left hand panel) and age and income (holding education at median level, right hand panel) using the regression results from Table 10. Red dots mark a negative implied gender gap.

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; Federal Reserve Bank of New York (FRBNY), SCE, June 2013 - November 2020; University of Michigan, Survey Research Center, MSC, January 1978 - January 2023; own calculations

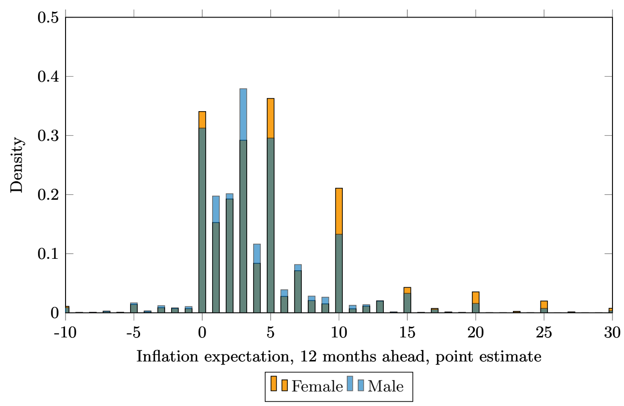
F.2 Histogram



(a) BOP-HH



(b) SCE



(c) MSC

Figure F.3: Histogram of inflation expectation point forecasts of men and women

Notes: Figure F.3 compares the distribution of male and female inflation expectations (measured as point forecasts over 12 months) pooled across all time periods. There is one plot per survey. The figures show that the distribution is more right skewed for women and rounded numbers (i.e. multiples of 5 or 10) are chosen more frequently.

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; Federal Reserve Bank of New York (FRBNY), SCE, June 2013 - November 2020; University of Michigan, Survey Research Center, MSC, January 1978 - January 2023; own calculations

F.3 Specific expectations

As a robustness exercise, I check if the gender gap is driven by a specific item in the consumption basket. I do so by comparing gender gaps in price expectations for different items from a standard basket, namely gas, food, college education, medical care and gold in the US. Under the experience hypothesis, the gender gap should be particularly large for food, as this has the most differential shopping experience in traditional households. However, the analysis in Table 11 shows that the gender gap is smaller for food than for the full basket. The only category for which it significantly increases is education. The categories with the smallest gender gaps are gas and gold. The gap is insignificant for both.

Table 11: Expectations about specific prices

	Inflation expectation (12 months ahead, point estimate)						
	Full basket	Gas	Food	Education	Medical	Rent	Gold
Constant	11.41***	10.96***	7.24***	13.76***	15.74***	8.56***	8.76***
	(1.80)	(1.85)	(1.32)	(1.92)	(2.77)	(1.47)	(2.78)
female	1.35***	0.56	0.90**	3.11***	1.45*	1.63***	0.25
	(0.51)	(0.52)	(0.37)	(0.54)	(0.78)	(0.41)	(0.78)
age	−0.02	−0.01	0.04***	0.02	0.04	0.04***	0.01
	(0.02)	(0.02)	(0.01)	(0.02)	(0.03)	(0.01)	(0.03)
single	0.52	−0.96	−1.52***	−1.85**	−2.40**	−0.96	0.79
	(0.74)	(0.76)	(0.54)	(0.79)	(1.14)	(0.61)	(1.15)
educ	−0.53***	−0.21	−0.30**	−0.61***	−0.62**	−0.64***	−0.47*
	(0.17)	(0.17)	(0.12)	(0.18)	(0.26)	(0.14)	(0.26)
hhinc	−0.33***	−0.43***	−0.38***	−0.48***	−0.51***	−0.31***	−0.54***
	(0.12)	(0.12)	(0.09)	(0.13)	(0.18)	(0.10)	(0.18)
decide_finance	−0.32	0.20	0.39*	0.12	0.50	0.21	−0.61
	(0.32)	(0.33)	(0.24)	(0.34)	(0.49)	(0.26)	(0.50)
fin_lit_test	−0.59	0.16	0.58	−0.06	0.88	0.24	−0.20
	(0.75)	(0.77)	(0.55)	(0.80)	(1.15)	(0.61)	(1.16)
$\beta^{\text{Full}}_{\text{female}} - \beta^{\text{Cat}}_{\text{female}}$		0.79	0.45	-1.76**	-0.10	-0.28	1.11
		(0.73)	(0.63)	(0.74)	(0.93)	(0.65)	(0.93)
Observations	1,184	1,184	1,184	1,184	1,184	1,184	1,184
R^2	0.05	0.03	0.06	0.08	0.05	0.07	0.03
Adjusted R^2	0.04	0.02	0.05	0.07	0.04	0.06	0.02
F Statistic	5.10***	3.38***	6.39***	8.20***	5.20***	7.16***	2.58***

*p<0.1; **p<0.05; ***p<0.01

Notes: Table 11 compares the regression coefficients of a pooled OLS estimation of individual inflation expectations (12 months ahead, point estimate) for the overall index to that of several sub-components, namely gas, food, costs of college education, medical costs, rent and gold on the dummy variable female, a continuous variable age, and the ordered categorical variables education and household income in the SCE. Standard errors in brackets below. Since the SCE's panel structure cannot be used due to crucial demographic questions being asked only once, there are no between effects for this survey. All models include regional controls and time fixed effects. The bottom line computes the difference between the female coefficient for the full basket and for the six sub-categories.

Sources: Federal Reserve Bank of New York (FRBNY), SCE, June 2013 - November 2020; own calculations

F.4 The role of financial confidence in the SCE

I replicate my results in section 5.2 for the BOP-HH with similar data from the SCE which provide external validity by extending to a different time period and geography and internal validity by providing robustness of the financial confidence measure. Table 12 shows the regression output and Figure F.5 visualizes them. Note, there are no between effects included in the SCE analysis, as demographic variables are only elicited once per respondent and are thus by design never time varying.

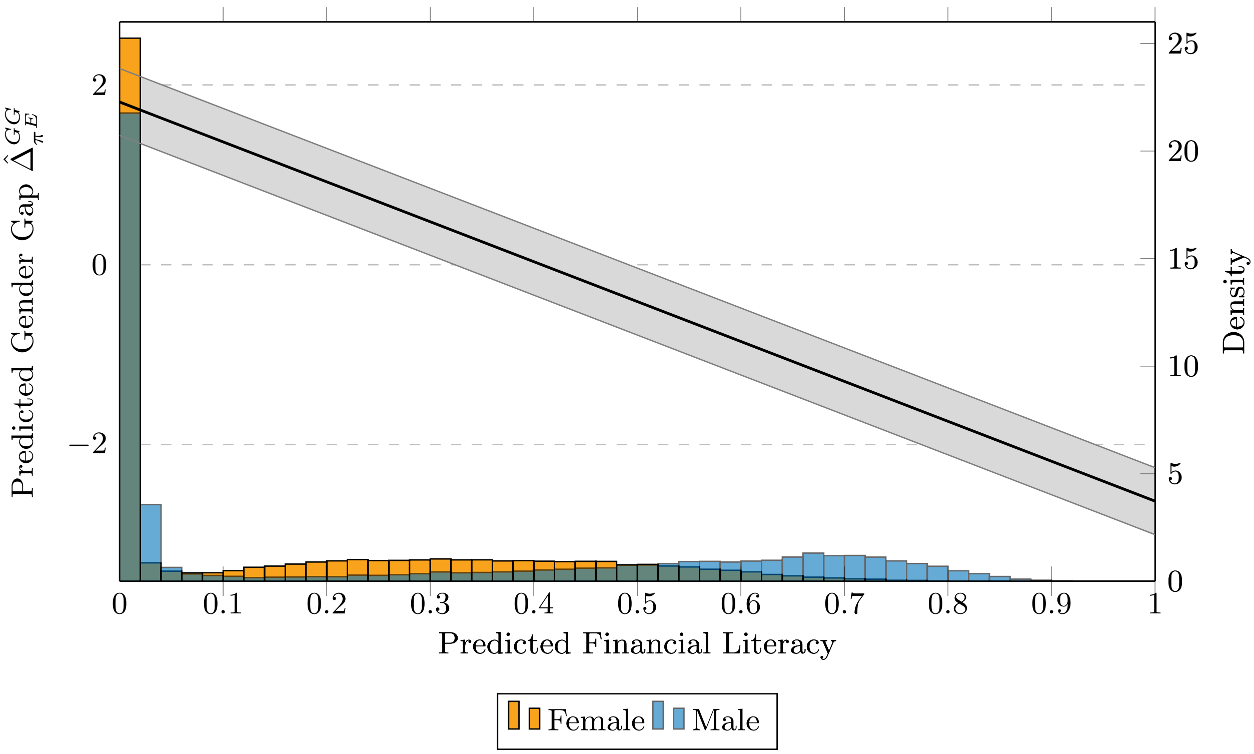


Figure F.5: The gender gap for different levels of financial confidence (SCE)

Notes: The black line in Figure 4 plots the predicted gender gap along all possible values of the financial confidence score in the black line ($\hat{\Delta}^{GG}_{\pi^E}(x) = 1.81 - 4.44x$). The full regression results are shown in Table 12, column (3). The grey area indicates 95% confidence bands (standard error: 0.3725). The histograms show the density of the male (blue) and female (orange) distribution of financial confidence scores.

Sources: Federal Reserve Bank of New York (FRBNY), SCE, June 2013 - November 2020; own calculations

F.5 Kernel densities and moments of the male and female distributions

Figure F.6 plots kernel densities for the pooled data from three surveys. It shows a higher fraction of men centered around 0-5% inflation while there is a visible mass of women with

Table 12: The impact of financial confidencebon the gender gap (SCE)

Survey: SCE	Inflation expectation (12 month ahead, point forecast)				
	(1)	(2)	(3)	(4)	(5)
Constant	8.40***	7.13***	6.15***	7.86***	7.68***
	(0.37)	(0.38)	(0.38)	(0.37)	(0.37)
female	1.63***	1.10***	1.81***	1.37***	1.81***
	(0.07)	(0.07)	(0.09)	(0.07)	(0.11)
$P(\hat{test} = 3)$		−6.18***	−7.13***		
		(0.35)	(0.36)		
$P(\hat{test} = 3)$×female			−4.44***		
			(0.30)		
fin_lit_test				−1.02***	−0.86***
				(0.05)	(0.06)
fin_lit_test × female					−0.33***
					(0.06)
age	0.01***	0.02***	0.02***	0.02***	0.02***
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
single	−0.05	−0.001	−0.01	−0.02	−0.03
	(0.08)	(0.08)	(0.08)	(0.08)	(0.08)
educ	−0.51***	−0.33***	−0.24***	−0.41***	−0.41***
	(0.02)	(0.03)	(0.03)	(0.02)	(0.02)
hhinc	−0.37***	−0.29***	−0.25***	−0.33***	−0.33***
	(0.01)	(0.02)	(0.02)	(0.01)	(0.01)
Time dummies	Yes	Yes	Yes	Yes	Yes
Region dummies	Yes	Yes	Yes	Yes	Yes
Between effects	No	No	No	No	No
Observations	113,165	113,165	113,165	113,165	113,165
R^2	0.03	0.03	0.03	0.03	0.03
Adjusted R^2	0.03	0.03	0.03	0.03	0.03
F Statistic	35.67***	38.54***	40.38***	39.70***	39.61***
Note:					*p<0.1; **p<0.05; ***p<0.01

Notes: Table 12 shows the regression coefficients of a pooled OLS estimation of individual inflation expectations (12 months ahead, point estimate) in the SCE on the dummy variable $female_i$, predicted confidence and actual financial literacy test scores. Since the SCE's panel structure cannot be used due to crucial demographic questions being asked only once, there are no between effects for this survey. The full model can be found in Equation 12. Standard errors in brackets below.

Sources: Federal Reserve Bank of New York (FRBNY), SCE, June 2013 - November 2020; own calculations expectations at 5, 10, 15 and 20%. The densities resemble those of the low and high confidence sample closely, providing the first suggestive evidence that the gender gap is indeed driven by low confidence outliers. These results are also reflected in Table 13. The higher moments of the male and female distribution of expectations strongly differ. Women have a higher standard deviation across all surveys. Further, while there is no gap between the lower percentiles of the male and female inflation expectations distribution, there appears to be a significant gap at the 75th and 90th percentile. The negative gap at the lowest percentile in the SCE can be explained by the female distribution being wider (thus also more negative values). This is a further indication of greater uncertainty rather than greater price experiences. While the survey design of BOP-HH and MSC discourage negative values, they are much more typical in the SCE (see discussion in Section 3). Further, not only do the Welch t-test and Wilcoxon rank test reject a common mean and median, the Kolmogorov-Smirnoff test also rejects the null of a common distribution. In a next step, I restrict the sample to observations with expectations below the 90th percentile of each survey (pooled over all time periods and for men and women). In this sample, I fail to reject the null of a common mean in two surveys (MSC and SCE), and that of a common median in one survey (MSC). Nonetheless, the null of a common distribution is rejected.

F.6 Regression in the deciles

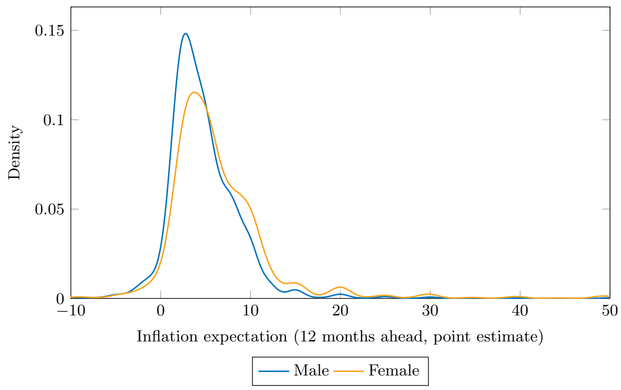
Table 14 provides the full regression output for Figure 5 in the main body. The full model can be found in Equation 13.

F.7 Time series evidence

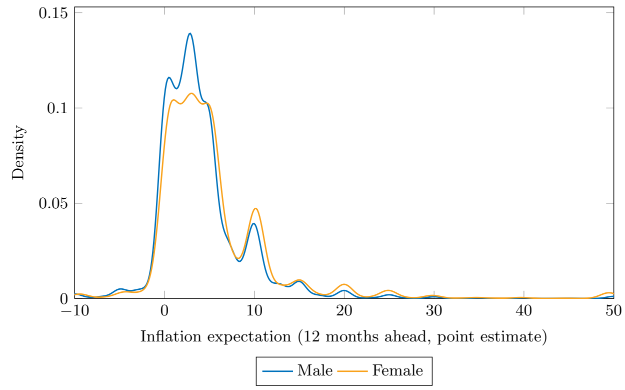
I further analyze the correlation between the gender gap in inflation expectations and the food inflation gap using time series aggregates. The survey with the longest time periods covered is the MSC (1978-2023). Indicative evidence from plotting the gender gap in mean inflation expectations against the difference in CPI food and CPI total (called henceforth the *food inflation gap*) in Figure F.8 for the US shows that there is no strong correlation between the gender gap and the food CPI gap (volatility). While in the US there appears to be some co-movement during the Great Financial Crisis and the onset of the Covid-19 pandemic and the high inflation period thereafter, in the euro area the gender gap appears to mirror the food inflation gap.

In addition, I run a simple time series model:

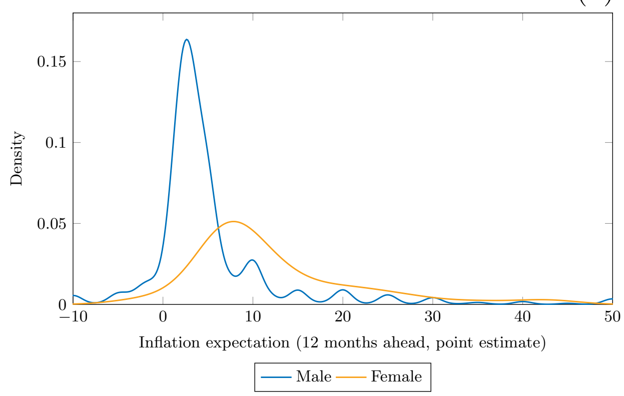
$$\Delta_t^{GG} = \alpha + \beta_1 B_t + \beta_2 CPI_t^{food} + \beta_3 CPI_t^{total} + \beta_4 CPI_{t-1}^{food} + \beta_5 CPI_{t-1}^{total} + \beta_6 \rho_{t,6}^{food} + \beta_7 \rho_{t,6}^{total} + \epsilon_t \quad (20)$$



(a) BOP-HH



(b) MSC



(c) SCE

Figure F.6: Kernel density of male and female inflation expectations

Notes: Figures F.6a, F.6b and F.6c compare the male (in blue) and female (in orange) distribution of inflation expectations (measured as point forecasts over 12 months) pooled across all time periods in the three surveys (BOP-HH, MSC and SCE). There is one plot per survey. The Kernel density is plotted using three times Silverman's rule-of-thumb bandwidth (Silverman, 1986).[20] The figure shows that in all surveys the distributions of both men and women are right skewed. However, very high, rounded numbers $> 10\%$ are more often chosen by women, visible as the orange line exceeds the blue line. In contrast, values around the inflation target of 2% are chosen more frequently by men.

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; Federal Reserve Bank of New York (FRBNY), SCE, June 2013 - November 2020; University of Michigan, Survey Research Center, MSC, January 1978 - January 2023; own calculations

Table 13: Higher moments of the gender gap in inflation expectations

	BOP-HH		SCE		MSC	
	Male	Female	Male	Female	Male	Female
	(1)	(2)	(3)	(4)	(5)	(6)
mean	5.13	6.77	4.05	6.12	4.04	4.99
median	4.0	5.0	3.0	3.2	3.0	3.0
std	4.88	7.79	8.29	14.08	5.23	6.98
p10	1.7	2.0	1.0	−1.0	0.0	0.0
p25	2.5	3.0	2.0	2.0	1.0	1.0
p75	7	8	5	10	5	6
p90	10	12	10	20	10	10
t-test	***		***		***	
wilcoxon-test	***		***		***	
ks-test	***		***		***	
Sub sample: $\pi^E < \pi^E_{90\%}$						
mean	4.11	4.37	2.69	2.04	2.61	2.54
median	4	4	3	3	3	3
std	2.23	2.30	5.04	8.69	2.93	3.15
p10	1.5	1.8	1.0	−2.0	0.0	0.0
p25	2.0	2.5	2.0	2.0	1.0	1.0
p75	5.5	6.0	4.0	5.0	5.0	5.0
p90	8	8	6	10	6	5
t-test	***					
wilcox-test	***		***			
ks-test	***		***		***	

*p<0.1; **p<0.05; ***p<0.01

Notes: Table 13 shows moments and percentiles of the distribution of quantitative inflation expectations (12 months ahead, point forecast) across the three surveys. I compare those of the male distribution (Columns (1),(3),(5)) to those of the female sub sample (Columns (2),(4),(6)). I compute Welch t-test and Wilcoxon-test (null: women and men have the same expectations versus alternative: women have higher expectations) to investigate if mean and median are the same across the sub samples and reject this. Further, I compute the Kolmogorov-Smirnoff test for common distribution which is also rejected. Finally, I compute the tests for a sub sample with inflation expectations below the 90th percentile for each survey. For this sub sample the male and female mean and median are not consistently different but the distributions differ.

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; Federal Reserve Bank of New York (FRBNY), SCE, June 2013 - November 2020; University of Michigan, Survey Research Center, MSC, January 1978 - January 2023; own calculations

Table 14: Quantile regression

	Inflation expectations (12 months ahead, point estimate)				
	Bottom 20%	Bottom 40%	Bottom 60%	Bottom 80%	Full Sample
Survey: BOP-HH					
female	-0.04^{***}	-0.06^{***}	0.07^{***}	0.15^{***}	1.27^{***}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.04)
single	-0.01	0.01	-0.03^{***}	-0.05^{***}	-0.50^{***}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.05)
age	0.002^{***}	0.004^{***}	0.002^{***}	0.004^{***}	-0.02^{***}
	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.001)
educ	-0.005^{***}	-0.01^{***}	-0.01^{***}	-0.02^{***}	-0.10^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.01)
hhinc	0.0004	0.002	-0.01^{***}	-0.02^{***}	-0.22^{***}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.01)
Observations	23,473	44,788	69,819	91,682	111,085
R^2	0.08	0.21	0.31	0.51	0.15
Adjusted R^2	0.07	0.21	0.31	0.51	0.15
F Statistic	46.96^{***}	291.15^{***}	751.44^{***}	$2{,}309.63^{***}$	463.38^{***}
Survey: MSC					
female	-0.03^{**}	-0.06^{***}	0.07^{***}	0.03^{***}	0.83^{***}
	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)
single	-0.04^{*}	-0.02^{**}	-0.02^{**}	-0.02	0.10^{***}
	(0.02)	(0.01)	(0.01)	(0.01)	(0.02)
age	0.001^{**}	-0.0000	-0.002^{***}	-0.002^{***}	-0.01^{***}
	(0.001)	(0.0003)	(0.0003)	(0.0003)	(0.001)
educ	0.02^{***}	0.04^{***}	0.03^{***}	0.03^{***}	-0.30^{***}
	(0.01)	(0.004)	(0.004)	(0.004)	(0.01)
hhinc	0.0000^{***}	0.0000^{***}	-0.0000^{***}	-0.0000^{***}	-0.0000^{***}
	(0.0000)	(0.0000)	(0.0000)	(0.0000)	(0.0000)
Observations	73,834	143,278	199,150	213,444	261,374
R^2	0.38	0.54	0.56	0.56	0.32
Adjusted R^2	0.37	0.54	0.56	0.56	0.32
F Statistic	80.29^{***}	303.04^{***}	469.71^{***}	487.72^{***}	225.55^{***}
Survey: SCE					
female	-1.10^{***}	-0.75^{***}	-0.68^{***}	-0.33^{***}	2.17^{***}
	(0.07)	(0.05)	(0.04)	(0.04)	(0.06)
single	0.16^{*}	0.13^{**}	0.12^{**}	0.05	-0.20^{***}
	(0.08)	(0.06)	(0.05)	(0.04)	(0.07)
age	-0.01^{***}	-0.004^{**}	-0.0003	0.003^{***}	-0.002
	(0.002)	(0.002)	(0.001)	(0.001)	(0.002)
educ	0.33^{***}	0.21^{***}	0.19^{***}	0.08^{***}	-0.71^{***}
	(0.03)	(0.02)	(0.02)	(0.01)	(0.02)
hhinc	0.21^{***}	0.14^{***}	0.12^{***}	0.06^{***}	-0.46^{***}
	(0.02)	(0.01)	(0.01)	(0.01)	(0.01)
Observations	43,196	65,389	74,175	92,712	115,491
R^2	0.37	0.41	0.42	0.43	0.24
Adjusted R^2	0.37	0.40	0.41	0.43	0.24
F Statistic	258.29^{***}	450.08^{***}	532.27^{***}	698.22^{***}	360.93^{***}
Time dummies	Yes	Yes	Yes	Yes	Yes
Regional dummies	Yes	Yes	Yes	Yes	Yes
Between effects	Yes	Yes	Yes	Yes	Yes
Note:					*p<0.1; **p<0.05; ***p<0.01

Notes: Table 14 shows the regression coefficients of the panel model in Equation 13 for the bottom 20% (1), 40% (2), 60% (3) and 80% (4) in the inflation expectations distribution. The table is separated by survey, showing first results from the BOP-HH, then the MSC and finally, SCE. Standard errors are shown in brackets below.

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; Federal Reserve Bank of New York (FRBNY), SCE, June 2013 - November 2020; University of Michigan, Survey Research Center, MSC, January 1978 - January 2023; own calculations

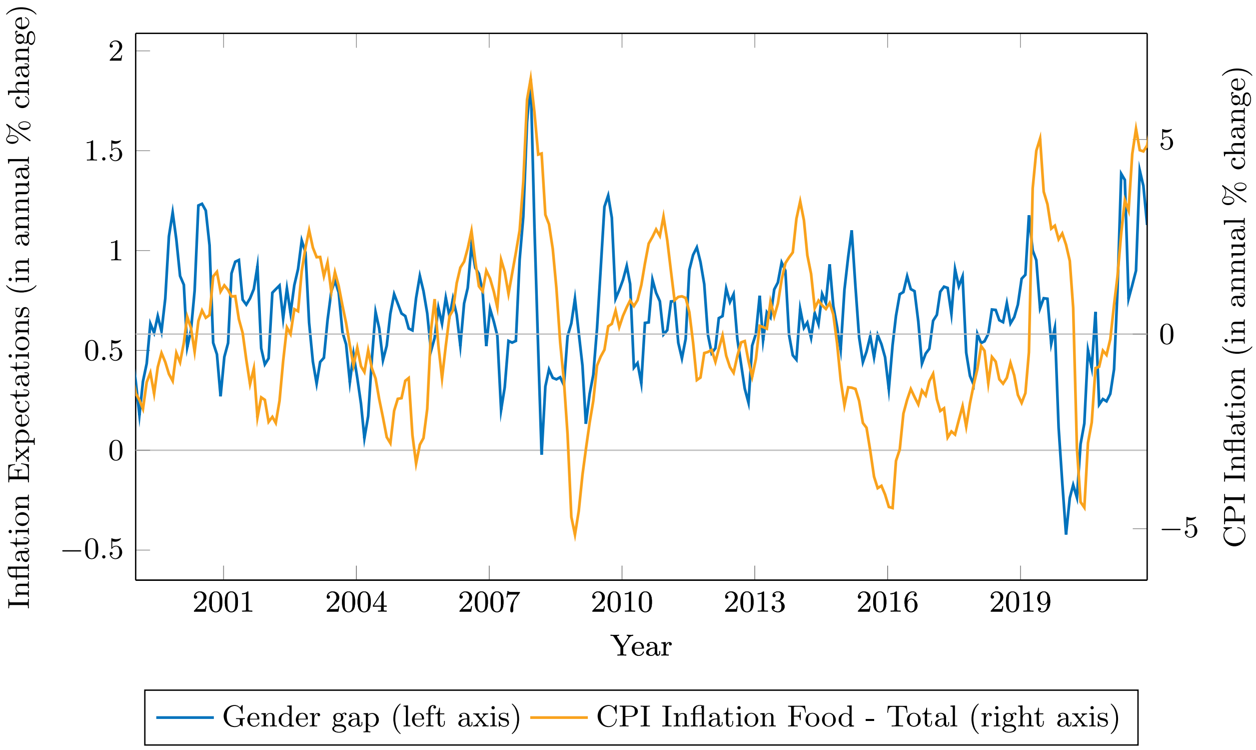


Figure F.8: Time series of the gender gap (MSC) and the food inflation gap (US)

Notes: Figure F.8 plots the time-series of the gender gap in inflation expectations per month (MSC) in blue and the distance between food inflation and total inflation, i.e. the *food inflation gap*, in orange. The gender gap in mean inflation expectations is defined as the difference between the mean of the female and male sub samples in 12 month ahead point forecast (Δ_t^{GG} as defined in Equation 21). In contrast the food inflation gap denotes the difference between the CPI food inflation and total inflation ($\%\Delta CPI_t^{food} - \%\Delta CPI_t^{total}$).

Sources: University of Michigan, Survey Research Center, MSC, January 1978 - January 2023; OECD, Prices: Consumer prices, Main Economic Indicators (database), January 1978 - January 2023; own calculations

Δ_t^{GG} denotes the gender gap in mean inflation expectations (point forecast, 12 months ahead) at time t computed as the difference between mean expectations of women and men:

$$\Delta_t^{GG} = \bar{\pi}_{F,t}^{E} - \bar{\pi}_{M,t}^{E} \tag{21}$$

where $\bar{\pi}_F^E$ is the mean of the female distribution and $\bar{\pi}_M^E$ is the mean of the male distribution. CPI_t^{food} denotes the annual %-change in CPI food at time t, CPI_t^{total} denotes the annual %-change in total CPI at time t. The coefficient of the food inflation gap on the gender gap is thus given by $\beta_2 - \beta_3$. Further, the model includes lags of food and total inflation as well as moving coefficients of variations $\rho_{t,6}^{food}$ and $\rho_{t,6}^{total}$.[21] The coefficient of the food inflation volatility gap on the gender gap is thus given by $\beta_6 - \beta_7$. I test the null hypothesis that the gender gap increases whenever food inflation exceeds core inflation: $H_{0,1} : \beta_2 - \beta_3 > 0$ and whenever food inflation volatility exceeds core inflation $H_{0,2} : \beta_6 - \beta_7 > 0$. Further, the full model includes a qualitative measure of inflation expectations, the balance statistic B_t. It measures a weighted share of respondents expecting a price increase.[22]

Table 15 shows different models for the survey with the most time-series observations, the Michigan Survey. It shows that the moving coefficient of variation is insignificant in all models, rejecting $H_{0,2}$. In contrast, current inflation does seem to impact the gender gap significantly. Table 16 shows that the implied effect of the food price inflation gap is estimated insignificant. While the insignificance result may be driven by the low power of the short time series in SCE and BOP-HH, the full period of the MSC is estimated tightly around zero. Overall, there is no evidence for the gender gap to increase whenever food price inflation exceeds core inflation,

[21] The moving coefficient of variation is defined as follows:

$$\rho_{t,n} = \frac{\sigma_{t,n}}{\bar{x_{t,n}}} \times 100$$

where t denotes the current period, n is the number of periods over which to calculate the moving average and standard deviation, $\bar{x_{t,n}}$ is the moving average and $\sigma_{t,n}$ describes the moving standard deviation computed as

$$\text{Moving SD}_t = \sqrt{\frac{1}{n-1} \sum_{i=t-n+1}^{t} (x_i - \overline{x}_t)^2}$$

where x_i is the value at time i.

[22] The BOP-HH records qualitative expectations with 5 possible options ranging from decrease significantly (MM, 1), to decrease slightly (M, 2), stay roughly the same (E, 3) to increase slightly (P, 4) and increase significantly (PP, 5). For the MSC and SCE I compute the measure from an initial question asking consumers whether they expect prices in general to go up (equivalent to PP, 5) or go down (equivalent to MM, 1), or stay where they are (equivalent to E, 3) over the next 12 months. The balance statistic is computed as follows:

$$B = (PP + \frac{1}{2} \times P) - (\frac{1}{2} \times M + MM) \tag{22}$$

hence I reject $H_{0,1}$ as well.

Table 15: Time series regression MSC

	Gender gap in inflation expectations Δ_t^{GG}							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Constant	0.77***	0.62***	0.76***	0.76***	0.63***	0.62***	0.76***	0.62***
	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)	(0.04)
B_t		4.30***			4.06***	4.28***		4.03***
		(0.44)			(0.44)	(0.44)		(0.44)
CPI_t^{food}	−0.002	−0.002	0.04	−0.002	0.01	−0.002	0.04	0.01
	(0.01)	(0.01)	(0.04)	(0.01)	(0.04)	(0.01)	(0.04)	(0.04)
CPI_t^{total}	0.04***	0.06***	−0.25***	0.04***	−0.14**	0.06***	−0.25***	−0.14**
	(0.01)	(0.01)	(0.06)	(0.01)	(0.06)	(0.01)	(0.06)	(0.06)
CPI_{t-1}^{food}			−0.05		−0.01		−0.05	−0.01
			(0.04)		(0.04)		(0.04)	(0.04)
CPI_{t-1}^{total}			0.29***		0.20***		0.29***	0.20***
			(0.06)		(0.06)		(0.06)	(0.06)
$\rho_{t,n}^{food}$				0.01		0.01	0.01	0.01
				(0.01)		(0.01)	(0.01)	(0.01)
$\rho_{t,n}^{total}$				0.02		0.01	0.02	0.01
				(0.02)		(0.02)	(0.02)	(0.02)
Observations	536	536	536	536	536	536	536	536
R^2	0.03	0.18	0.06	0.03	0.19	0.18	0.07	0.19
Adjusted R^2	0.02	0.17	0.06	0.02	0.18	0.17	0.06	0.18
Residual Std. Error	0.59	0.54	0.58	0.59	0.54	0.54	0.58	0.54
F Statistic	6.97***	37.79***	9.03***	4.08***	25.23***	22.85***	6.36***	18.13***

*p<0.1; **p<0.05; ***p<0.01

Notes: Table 15 compares the coefficients of a linear time-series regression of the per period mean gender gap in inflation point forecasts 12 months ahead (Δ_t^{GG}) in the MSC. The full model can be found in Equation 20. All models include contemporaneous inflation of CPI food (CPI_t^{food}) and CPI total (CPI_t^{total}). Their lags are included in models (3), (5), (7) and (8). Further, models (2), (5), (6) and (8) include the balance statistic of qualitative inflation expectations (computation can be found in Appendix **??**). Finally, models (4), (6) and (8) incorporates the moving coefficient of variation for food and total CPI inflation. It measures the volatility of the six month period around the observation.

Sources: University of Michigan, Survey Research Center, MSC, January 1978 - January 2023; OECD, Prices: Consumer prices, Main Economic Indicators (database), January 1978 - January 2023; own calculations

Table 16: Time series regression of the gender gap in mean inflation expectations

	Gender gap in inflation expectations Δ_t^{GG}								
	MSC 1978-2023			SCE 2013-2020			BOP-HH 2019-2022		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Constant	0.77***	0.63***	0.64***	1.94***	2.57***	2.78***	2.15***	2.16***	2.09***
	(0.04)	(0.04)	(0.04)	(0.19)	(0.22)	(0.22)	(0.14)	(0.14)	(0.16)
B_t		−4.57***	−4.34***		−17.95***	−19.83***		−0.52	−0.74
		(0.41)	(0.41)		(4.12)	(4.02)		(2.71)	(2.87)
CPI_t^{food}	−0.002	−0.001	0.01	−0.03	−0.09*	0.34**	−0.001	0.001	−0.06
	(0.01)	(0.01)	(0.04)	(0.05)	(0.04)	(0.14)	(0.06)	(0.07)	(0.10)
CPI_t^{total}	0.04***	0.06***	−0.12**	0.11	0.18*	0.84***	−0.14**	−0.13	0.04
	(0.01)	(0.01)	(0.06)	(0.10)	(0.09)	(0.23)	(0.06)	(0.09)	(0.22)
CPI_{t-1}^{food}			−0.01			−0.45***			0.09
			(0.04)			(0.15)			(0.11)
CPI_{t-1}^{total}			0.18***			−0.75***			−0.19
			(0.06)			(0.23)			(0.24)
$CPI_t^{food} - CPI_t^{total}$	-0.04**	-0.06***	0.13*	-0.14	-0.27***	-0.5*	0.14	0.13	-0.1
	(0.02)	(0.02)	(0.07)	(0.11)	(0.1)	(0.27)	(0.09)	(0.11)	(0.25)
Observations	536	536	536	89	89	89	22	22	22
R^2	0.03	0.21	0.23	0.43	0.20	0.32	0.48	0.39	0.42
Adjusted R^2	0.02	0.21	0.22	0.0004	0.17	0.28	0.32	0.29	0.23
Residual Std. Error	0.59	0.53	0.53	0.71	0.65	0.60	0.35	0.36	0.38
F Statistic	6.97***	47.58***	30.81***	1.02	7.13***	7.82***	5.98***	3.80**	2.28*

*p<0.1; **p<0.05; ***p<0.01

Notes: Table 16 compares the coefficients of a linear time-series regression of the per period mean gender gap in inflation point forecasts 12 months ahead (Δ_t^{GG}) in three surveys (MSC, SCE and BOP-HH). The full model can be found in Equation 20. All models include contemporaneous inflation of CPI food (CPI_t^{food}) and CPI total (CPI_t^{total}). Their lags are included in models (3), (6), and (9). Further, models (2) & (3), (5) & (6), and (8) & (9)include the balance statistic of qualitative inflation expectations (computation can be found in Appendix **??**). The separated line below computes the implied effect of the food inflation gap discussed above. I show the results for the full periods of each survey. Standard errors in brackets below. I find that estimates for the implied food inflation gap are small and in most surveys insignificant. There is no evidence for the gender gap increasing when food price inflation exceeds core inflation.

Sources: Research Data and Service Centre (RDSC) of the Deutsche Bundesbank, BOP-HH, April 2020 - September 2022; Federal Reserve Bank of New York (FRBNY), SCE, June 2013 - November 2020; University of Michigan, Survey Research Center, MSC, January 1978 - January 2023; OECD, Prices: Consumer prices, Main Economic Indicators (database), January 1978 - January 2023; own calculations