How Much Can I Make? Insights on Belief Updating in the Labor Market*

Sebastian Brown†, Kenneth Chan‡

November 14, 2024

Abstract

We use a nationally representative survey, the labor supplement of the Survey of Consumer Expectations, to study how people update their wage expectations over time. Using a recently developed excess belief movement test for the martingale property (Augenblick and Rabin, 2021), we find strong evidence of non-Bayesian learning at the aggregate level. Among survey respondents who responded at least twice to the survey, we find an average movement in beliefs that is roughly 517% of the reduction in their beliefs' uncertainty, 417% more than the Bayesian benchmark. This result is consistent with base rate neglect and overreaction to signals, and we found suggestive evidence that people exhibit base rate neglect. Our simulations show that this result is unlikely to be explained solely by measurement error or bias in prior beliefs. We also found patterns of asymmetric updating, where individuals update their beliefs more when they receive good wage offers relative to bad wage offers.

JEL: D83, D84, D90, J64

Keywords: Wage Expectations, Belief Updating, Job Search, Base Rate Neglect, Excess Belief Movement

*We are grateful to Gary Charness, Erik Eyster, Lorenz Goette, Peter Kuhn, Ryan Oprea, Jessica Pan, Peter Rupert, Dick Startz, Sevgi Yuksel, and Basit Zafar for their helpful comments and suggestions. All remaining errors are our own.

†Department of Economics, UC Santa Barbara. sebastian_brown@ucsb.edu.

‡Department of Economics, UC Santa Barbara. kenneth_chan@ucsb.edu.

1 Introduction

Canonical job search models typically assume that the agent knows the underlying wage distribution (Mortensen, 1970; McCall, 1970). This class of model has an optimal decision rule of setting a reservation value and accepting the first offer that is better than the reservation value (Weitzman, 1979). While these models give us valuable insights into job search behavior, it is often unrealistic to assume that people know the underlying wage distribution.

Another class of search models assumes an unknown underlying distribution (Rothschild, 1978; Rosenfield and Shapiro, 1981; Talmain, 1992; Li and Yu, 2018), where the agent has to learn about the underlying distribution over time by observing the offers they receive.[1] The actions that the agent takes are a function of the agent's beliefs.

While these models are more realistic, the empirical application of these models has been limited due to insufficient high-quality data on job searchers' beliefs. Most studies in this area assume Bayesian updating in the calibrated model (e.g. Potter (2021)). Consequently, there has been a lack of research on how people learn and update their beliefs in light of new information. Our paper hopes to fill this gap by providing some insights into how job searchers incorporate new information into existing beliefs.

Over the last two decades, nudges (Thaler and Sunstein, 2008) have become a popular policy tool for policymakers to influence people's behavior. One of the most popular forms of nudges is information provision, which is a potent policy tool for changing people's beliefs and behavior. Examining the process of belief updating can provide valuable insights into the effectiveness of information provision policy in job search.

In a series of recent field experiments, providing information has been shown to help students make more informed decisions about their college majors (Wiswall and Zafar, 2015), improve people's knowledge about COVID-19 (Sadish et al., 2021), eliminate statistical

[1] These models are search models where an agent is looking for a good with the lowest price or a product with the highest quality. These models can be extended easily to a job search context.

racial discrimination in a patient's choice of medical professional (Chan, 2022), and reduce disagreement about the extent of racial discrimination (Haaland and Roth, 2023).

In a job search context, Arni (2016) found that a coaching program was successful in increasing job-finding rates among the treated job searchers by 9 percent. The author argues that coaching helped workers to have more realistic expectations and search more effectively. In another experiment, Gee (2018) randomly displayed information to LinkedIn users about the number of workers applying for a specific job. She showed that this additional labor market information increased the probability of a worker completing a job application by 2.5 percent. Finally, Jiang and Zen (2023) conducted a field experiment at UC Berkeley, showing that information provision can correct incorrect wage beliefs, and affect people's reservation wage. Overall, these studies highlighted have shown the effectiveness of information treatments in affecting people's behavior in various settings.

Unlike most of the above papers which use a treatment intervention to manipulate people's beliefs, in this paper, we utilize a government dataset (the Survey of Consumer Expectations) to study an individual's capacity to learn about the wage distribution without any experimental manipulation. As our data is representative and free of the cognitive influences of an experimental setting, it is also a more externally valid measure of how individuals process information in the labor market.

Our main analysis employs an excess belief movement test Augenblick and Rabin (2021) that tests for the martingale property in belief updating. The martingale property requires that before seeing a signal, the expected updated beliefs over the signal realizations should be equal to the prior. The intuition of this property is the agent should not expect his beliefs to change before seeing a signal. There are two statistics required for this test. Belief movement is defined as the squared difference of changes in belief, and uncertainty reduction is defined as the amount of reduction in the belief's variance from updating. If the agent's updating rule satisfies the martingale property, the average movement in beliefs will be equal to the average uncertainty reduction. Using this test, we find an average movement in beliefs that is over five times the average reduction in uncertainty. We can reject that people are updating their beliefs in a Bayesian manner, and our results suggest that on average people over-update relative to the Bayesian benchmark. This is consistent with biases such as overreaction to signals and base rate neglect. This also provides a possible explanation of why information provision policy has been effective in these settings (Jiang and Zen, 2023).

Knowing that an information provision policy is likely to be effective, the remaining of our analysis looks at whether we need a policy intervention. One of the justifications for a policy intervention is whether there is base rate neglect. Base rate neglect can prevent individuals from learning about their actual wage distribution even in an information-rich environment. The second justification for policy intervention is the presence of asymmetric updating where individuals respond differently to good and bad news. Asymmetric updating can also hinder the learning process and in extreme cases, the individual's belief may converge to the wrong distribution.

The paper that is closest to ours is Conlon, Pilossoph, Wiswall and Zafar (2018) who study non-Bayesian updating. Despite having a similar research question and using the same dataset, our papers differ in four ways. Firstly, at the general level, our paper is primarily interested in using the data to test various updating rules and identifying models that can best describe people's updating behavior, while their paper primarily focuses on how information friction affects job search behavior. In our paper, we further examine and identify different belief updating patterns instead of just comparing them against the Bayesian benchmark. We found asymmetric belief updating patterns and suggestive evidence of base rate neglect.

Secondly, we adopt a different methodology in answering the research question. Their analysis assumes that the agent is updating in a Gaussian framework where the priors and signals (with a correctly perceived variance) are normally distributed. The wage offers the respondents received between the surveys are assumed to be the signals. This allows them to compute the Bayesian benchmark and compare the survey respondent updating behavior against this Bayesian benchmark, where they found over-updating relative to the Bayesian benchmark. Our approach uses the excess belief movement test (Augenblick and Rabin, 2021) which circumvents the need to compute the Bayesian benchmark to compare if people are Bayesian, which allows us to make fewer and more conservative assumptions.

Thirdly, our approach does not restrict our sample to respondents who received a wage offer. The excess belief movement test only requires the initial beliefs and updated beliefs to test for the martingale property. This allows us to examine belief updating without receiving a job offer as the offers are not the only signals that reveal information about one's wage distribution. For example, not receiving a wage offer may also provide some information about the respondent's wage offer distribution (Milgrom, 1981; Jin, Luca and Martin, 2021) or there may be social learning (Mobius and Rosenblat, 2014) where individuals can learn from their peers. This allows us to use a larger sample from the SCE data, which makes our result more representative of the population.

Lastly, our test allows us to reject all updating rules without the martingale property, which is a stronger test than their paper. Bayesian updating is one of the updating rules that has the martingale property. Rejecting the martingale property will allow us to reject a larger class of updating rules, as well as identify mistakes like having incorrect priors.[2]

The remainder of the paper proceeds as follows. Section 2 reviews relevant literature. Section 3 describes the data and our process of fitting distributions to the data. Section 4 outlines a descriptive theoretical model of labor market updating that relates to the SCE survey questionnaire. Section 5 explains our empirical strategy and the excess belief movement test. Section 6 then lists our main results. We conclude in section 7 with a broader discussion of our approach, suggesting future directions for research.

[2]We are grateful to Sevgi Yuksel for pointing out that incorrect priors can cause the martingale property to fail even when agents are Bayesian.

2 Literature Review

Our work is related to three strains of literature: search theoretic models with unknown distributions, non-Bayesian updating, and learning in job search. In these models, the underlying distribution from which offers are drawn, while fixed, is also unknown to the agent.[3] A key characteristic of this class of model is that the agent has to update his beliefs about the underlying distribution as he searches.

Rothschild (1978) wrote the first search model with unknown distribution, and he showed that the static reservation value no longer applies in most scenarios as the agent will continuously update his expectations about the future. He also pointed out that having the ability to recall a previously rejected offer will affect the optimal search strategy when the distribution is unknown. When recall is not possible, the optimal strategy can be characterized by reservation value that is a function of the agent's beliefs (Rothschild, 1978; Rosenfield and Shapiro, 1981). When recall is possible, the strategy is to set a decreasing sequence of reservation values (Talmain, 1992).

For tractability of the model, the agent is typically assumed to have a Dirichlet prior on the wage distribution (Rothschild, 1978; Rosenfield and Shapiro, 1981; Talmain, 1992) or a Gamma distribution on the arrival rate of offers Potter (2021). This is because the Dirichlet and Gamma distribution has the conjugacy property[4] which provides a closed-form expression for the posterior. Recently, Li and Yu (2018) deviated from the conjugate prior assumption and they assume the agent believes that there is a set of possible distributions and all of the distributions in this set satisfy the monotone likelihood ratio property (MLRP). The MLRP assumption allows them to perform monotone comparative statics (Milgrom and Shannon, 1994; Athey, 2002) on the reservation value.

While these search models typically assume people are updating their beliefs in a

[3] An alternative way to model this is to let the arrival rate of offers be unknown, see Potter (2021). However, assuming the underlying distribution is unknown is the main way of introducing uncertainty into the model.

[4] Under Bayesian updating, the posterior belief and the prior beliefs are from the same parametric family of distribution.

Bayesian manner, behavioral economists have shown that people are not Bayesian. Behavioral theorists have come up with belief updating models to accommodate these non-Bayesian behavior (Grether, 1980; Epstein, Noor and Sandroni, 2010; Hagmann and Loewenstein, 2017). Some of the more well-known biases include base rate neglect (Kahneman and Tversky, 1972; Esponda, Vespa and Yuksel, 2024), a phenomenon where people underweight their priors, and conservatism bias (Phillips and Edwards, 1966), a situation where people are insensitive to new information. While most of the empirical evidence is from lab experiments (see Benjamin (2019) for a survey of the experimental literature on belief updating), some recent papers have used field data to show that people update their beliefs in a non-Bayesian manner (Conlon, Pilossoph, Wiswall and Zafar, 2018; Bordalo, Gennaioli, Porta and Shleifer, 2019; Bordalo, Gennaioli, Ma and Shleifer, 2020; Augenblick and Rabin, 2021).

There is a burgeoning literature that empirically studies learning in job search and how people's behavior deviates from the Bayesian benchmark. Firstly, Kudlyak, Lkhagvasuren and Sysuyev (2014) finds that job seekers first apply to jobs that match their education levels. But with prolonged unemployment, they apply to jobs that require a lower education level. They argue that this is evidence that searching workers learn to adjust their expectations downward over the unemployment spell. Similarly, Mueller, Spinnewijn and Topa (2021) found that unemployed workers adjust their beliefs downwards but not sufficiently. This results in the long-term unemployed displaying an optimistic bias in job finding.

Contrary to the findings of Mueller, Spinnewijn and Topa (2021) about job finding rates, Conlon, Pilossoph, Wiswall and Zafar (2018) use the SCE data and find that people overupdate their wage beliefs relative to the Bayesian benchmark. They estimated that on average, wage expectations increase by \$0.47 for every one-dollar increase in observed wage offer, while the Bayesian benchmark is estimated to be \$0.16.

Instead of focusing on wage expectation, Potter (2021) examines how the expectation of offer arrival rate changes over time using data from the Great Recession. Using a calibrated model, he showed that learning can explain the job search dynamics during the Great

Recession.

Aside from these papers, our work also relates more generally to a broader, yet still relatively new, literature on behavioral job search (e.g. DellaVigna et al. (2017); see Cooper and Kuhn (2020) for a recent review), as well as more general work studying employed job search (e.g. Faberman et al. (2022), Ahn and Shao (2021)).

3 Data

3.1 Overall Description

We are using a public dataset from the Survey of Consumer Expectations (SCE), which is a nationally representative survey administered monthly by the New York Federal Reserve. The survey is divided into two parts, a core set of questions that remain the same every month and a supplementary set of questions that rotates between several different economic topics. Our analysis focuses on the labor market supplement, which is administered every March, July, and November. Subjects can be surveyed for up to a year, and are replaced on a rolling basis. This feature allows us to observe the same individual for up to a maximum of three times for the labor force questions. In return for completing the survey, respondents are paid $15 for each survey. For a more in-depth description of the structure and administration of the SCE, see Armantier et al. (2017).

The main advantage of this dataset is that it explicitly elicits expectations from survey respondents, meaning there is no need to indirectly infer beliefs as in many previous papers (for example Potter (2021) and Spinnewijn (2015)). One of the unique questionnaires in this survey elicits a belief distribution over wages. This is the main feature that allows us to implement the excess belief movement test. The last benefit is that the survey is representative of the American population and it features mostly employed individuals, allowing us to study the wage expectations of employed individuals who is previously understudied (as highlighted in Faberman et al. (2022)). However, this comes at the cost of including fewer unemployed individuals.[5]

3.2 Survey Questions

Our main analysis centers around the questions about wage expectations which merit more detailed explanations. Figure 1 shows the survey questionnaire used to elicit the respondent's wage expectations and belief distribution. It is important to note that most of the survey respondents are not looking for a job, we interpret their responses as a hypothetical response assuming that they are searching.

OO2a2 - OO2a2 **(Added March 2015)**

Think about the job offers that you may receive within the coming four months. Roughly speaking, what do you think the annual salary for the best offer will be for the first year?

Note the best offer is the offer you would be most likely to accept.

_______________ dollars

OO2b - OO2b **(shown if OO2a2 > 0 each response is % of OO2a2 ranging from .8 to 1.2)** **(Added November 2014)**

Think again about the job offers that you may receive within the coming four months. What do you think is the percent chance that the job with the best offer will have an annual salary for the first year of...

The best offer is the offer you would be most likely to accept.

Less than [0.8* OO2a2] dollars (1)	_______________	% (1)
Between [0.8* OO2a2] dollars and [0.9* OO2a2] dollars (2)	_______________	% (2)
Between [0.9* OO2a2] dollars and [1.0* OO2a2] dollars (3)	_______________	% (3)
Between [1.0* OO2a2] dollars and [1.1* OO2a2] dollars (4)	_______________	% (4)
Between [1.1* OO2a2] dollars and [1.2* OO2a2] dollars (5)	_______________	% (5)
More than [1.2* OO2a2] dollars (6)	_______________	% (6)

OO2a - OO2a **(Added November 2014)**

Think about the job offers that you may receive within the coming four months. Roughly speaking, what do you think the average annual salary for these offers will be for the first year?

_______________ dollars

Figure 1: Survey Questions OO2a2, OO2b and OO2a from SCE

[5] While the SCE has been included in analyses of unemployed workers (e.g. Mueller, Spinnewijn and Topa (2021); Conlon, Pilossoph, Wiswall and Zafar (2018); Faberman, Mueller, Şahin and Topa (2022)), the fraction of the sample that is unemployed is very small, chiefly because the SCE is a representative sampling of the entire population, not of the unemployed. As noted in table 1 in section 3.3.2, only 3.5% of our final sample is unemployed.

Here the label "OO2a2" refers to the following question: "*Think about the job offers that you may receive within the coming four months. Roughly speaking, what do you think the annual salary for the best offer will be for the first year?*" The phrasing of the question is ambiguous and it is unclear what statistics are measured here. This is not too critical for our research question, as our analysis uses the follow-up question. This question only partitions the wages into different wage bins.

In the follow-up question, the survey elicits the belief distribution of receiving the best offer that belongs to a range that is determined by the subject's response in "OO2a2". Hereafter we will refer to wage "bins", by which we mean the ranges of wages from the above survey question.

A challenge working with this question is that the wage bins are different depending on the respondent's response to "OO2a2". To make the responses comparable between an individual's consecutive surveys, we fit various distributions to each individual's responses. For further discussion on the fitting process, refer to section 5.1.

The survey also included the question "OO2a" to measure an individual's expectation of the average wage. This question reads "*Think about the job offers that you may receive within the coming four months. Roughly speaking, what do you think the average annual salary for these offers will be for the first year*?" We do not use this question for the excess belief movement test since the follow-up question that elicits a probability distribution is necessary for the excess belief movement test. We do, however, use this response to test for asymmetric updating to determine how the expected wage offer changes depending on the wage offer they receive. Finally, we also have questions about the offers they received,[6] as well as a question on the number of offers expected.

[6]The survey asks individuals to report wages of their best three offers, but it also collects information on the number of offers expected and received. Hence, we can tell if an individual received more offers than they have space to report.

3.3 Data Structure and Summary Statistics

3.3.1 Timeline

To study belief updating, we require at least two responses from the same individual to observe the update in beliefs. Hence, we restrict our sample to only respondents who completed at least two surveys. Respondents can complete the survey three times, which allows us to observe at most two updates from each individual. The timeline is detailed in figure 2.

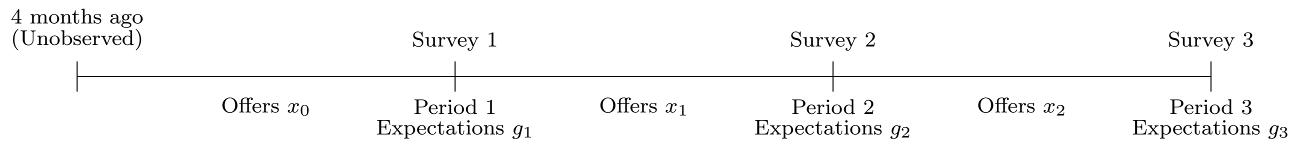

Figure 2: Timeline of Survey

In this paper, our main analyses are concerned with the *change* in beliefs between surveys. Therefore, our final dataset is organized into "updates," where each observation contains information from two consecutive labor market surveys for a single individual. Due to attrition, the number of updates will not be twice the number of individuals. However, for those who took all three surveys, there will be two separate updates: one for the update from period 1 to period 2, and one for the update from period 2 to period 3.[7]

3.3.2 Descriptive Statistics

The composition of our final dataset by employment status (defined as the employment status of the individual on each update's initial survey) is given in table 1. In addition to showing employment status for all individuals collectively, we also show it by whether they received at least one wage offer between the update's surveys, as well as by whether they searched for work in the month before the survey.

[7]This format allows us to analyze updates separately in section 6.4. For the analysis in section 6.1 which uses an individual's entire sequence of updates, we add the period 3 information to the observation covering the update from period 1 to 2 where relevant and instead use only one observation per individual.

Count	All Individuals	Got Offer?		Searched?		
		Yes	No	Yes	No	Unknown
Number Unemployed	127	37	90	71	51	5
Number Employed	3,038	542	2,496	742	2,017	279
Number Not in Labor Force	406	43	363	34	368	4
Missing Employment Status	36	4	32	17	16	3
All Observations	3,607	626	2,981	864	2,452	291
Unique Individuals	2,532	565	1,967	706	1,637	189
Data Range	3/2015-11/2019					

Table 1: Dataset composition. Observations consist of information from two consecutive surveys for a given individual. Observations from one of the surveys in 2020 were excluded to avoid measuring the effects of the pandemic on expectations.

For the unique individual counts, an individual was counted as having received an offer or having searched if they did so for at least one of their updates. Some individuals left the question blank for the search question. If the individual only answered the question once, this answer was used to proxy for whether the individual had searched or not. Observations reporting expected annual wages (using either question OO2a or OO2a2 for either period) below $10,000 were dropped, as it is suggestive that these individuals are not reporting their annual salary.[8]

In the following graph, we report the expected and realized (maximum) wage offers. One concern is that the beliefs from the survey are not incentivized and that this may affect the quality of the responses. In figure 3 below, we see that the responses are not entirely random and is fairly close to the distribution of offers that were reported in the survey. However, the survey respondents are slightly optimistic about their wage offers but this is consistent with the findings in existing literature (Krueger and Mueller, 2016). This suggests that people are not randomly responding to the survey questions.

[8] 618 observations were dropped this way, 257 of which had a reported wage or one of the expectation questions less than $100.

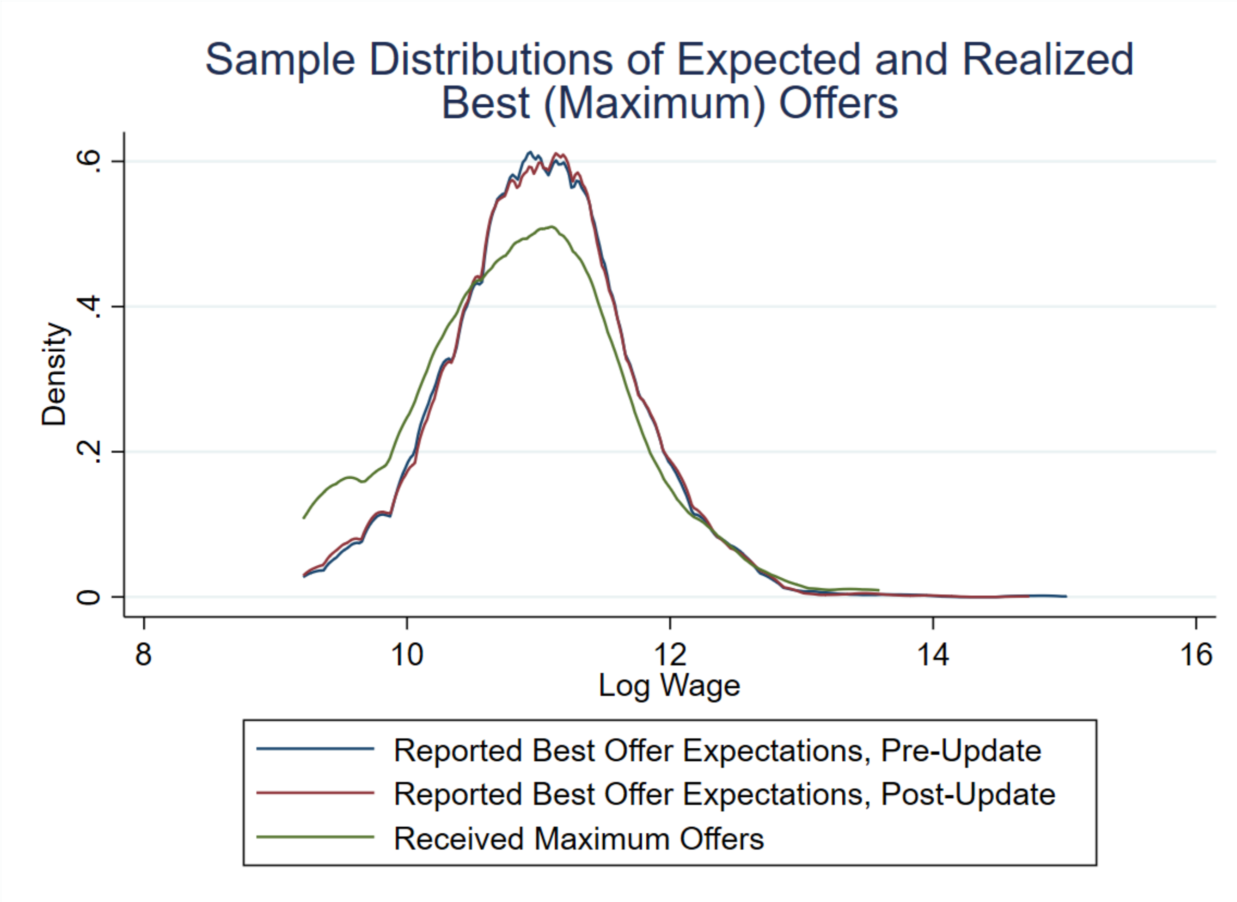

Figure 3: Distributions of each update's reported beliefs and offers. Reported expected wages are based on question OO2a2 and include all individuals, including those who did not receive an offer. The highest offers received between the update's two surveys were used to graph the third distribution.

4 Theoretical Framework

We will now provide a model of belief dynamics related to the job search setting we are studying. We will follow Li and Yu's (2018) method of modeling uncertainty. In their model, the agent believes there is a set of distributions the agent is drawing from, and the distributions from this set have the monotone likelihood ratio property. Unlike their model which is interested in obtaining comparative statics across different signal realizations, we are primarily interested in the belief dynamics, hence we do not require the assumption about MLRP.

This model also differs from the Augenblick and Rabin (2021) setting. In their setting, the state of the world is an outcome (e.g. Biden winning the election), while in our setting, the state of the world is a probability distribution. We assume that the agent's reported belief in the SCE survey is a mixture distribution. We show that their results generalize to this setting as well. We will need to first verify that for Bayesian agents, the reported beliefs from the mixture distribution have the martingale property to apply their test.

Suppose there is an agent who is searching for a job and believes there is a set of possible wage distributions, $\mathcal{F}$, that he is drawing his wages from. We will index the density functions in $\mathcal{F}$, by a parameter θ taking values in an ordered set Θ. The agent has a non-degenerate belief g_t over θ at time t. We will also assume that the agent places a non-zero probability weight on the true wage distribution, F, he is drawing from. The agent's belief about the wage he will receive is a mixture distribution of the possible wage distributions in $\mathcal{F}$.

In each period, the agent can observe a signal that reveals some information about the wage distribution he is drawing from. We let the set of possible signal realizations be $X \subseteq \mathbb{R}^n$; we can think of the signal realization $x \in X$ as a vector of wages of the job offer the agent received or any news that reveals information about his wage distribution. The conditional density function is denoted as $p(x_t|\theta)$, which reflects the likelihood of observing signal x at time t conditioned on the wage distribution being f_θ.

In the survey question "OO2b," the respondents provided a probability distribution about the best wage offer they will receive in the form of assigning probabilities to "binned" ranges of potential wages. We denote the probability of the wages being in bin $i \in \{1, 2, \ldots, n\}$ at time t as π_t^i.[9]

Under the model we have constructed, we will partition the wages into n wage bins $\{[a_0, a_1), [a_1, a_2), \ldots [a_n, a_{n+1})\}$. In the survey, $a_0 = 0$ and $a_{n+1} = \infty$. This partition covers the entire support of the wage distribution. We will assume that the agent's reported belief can be represented by the expression below

[9] In the survey, there are 6 bins, but we are keeping the model general.

$$\pi_t^i = \overbrace{\int_{\theta' \in \Theta} g_t(\theta') \underbrace{\int_{a_{i-1}}^{a_i} f(w|\theta')dw}_{\substack{\text{Probability of drawing} \\ \text{a wage within the bin} \\ \text{from distribution } \theta'}} d\theta'}^{\text{Averaged over beliefs of distributions}} \quad (1)$$

where w is the wage drawn from the distribution. The inner integral is the probability of receiving a wage offer within the wage bin from the distribution indexed by θ'. The outer integral integrates the agent's beliefs over which distribution he is drawing from.

Our primary test tests if people's belief follows the martingale property. Formally, the martingale property states that the expected posterior beliefs have to be equal to the prior belief, $\mathbb{E}(g_{t+1}(\theta|x)|g_t(\theta)) = g_t(\theta)$. Intuitively, this property states that the agent should not expect himself to change his belief before seeing the signal. If the agent expects himself to change his belief, he should have already done so which leads to inconsistency with the current prior.

If the agent's updating rule on g has the martingale property, the probabilities π^i will have the martingale property as well. We can easily verify that $\mathbb{E}(\pi_{t+1}^i|g_t)$ converges absolutely since it has to be less than 1 and the integrands are non-negative. We can apply Fubini's theorem and interchange the order of expectation and the integral.

$$\mathbb{E}(\pi_{t+1}^i|g_t) = \int_{\theta' \in \Theta} \mathbb{E}(g_{t+1}(\theta'|x_{t+1})|g_t) \int_{a_{i-1}}^{a_i} f(w|\theta')dw = \int_{\theta' \in \Theta} g_t(\theta') \int_{a_{i-1}}^{a_i} f(w|\theta')dw = \pi_t^i$$

This allows us to perform the excess belief movement test on the statistics reported to determine if people are updating their beliefs in a Bayesian manner.

4.1 Updating Rules

We will be testing the data against some updating rules to determine which rules can best explain people's behavior.

4.1.1 Updating Rules with Martingale Property

1. Bayesian updating
$$g^{bayes}_{t+1}(\theta|x_{t+1}) = \frac{g_t(\theta)p(x_{t+1}|\theta)}{\int_{\theta'\in\Theta} g_t(\theta')p(x_{t+1}|\theta')} \tag{2}$$

Bayesian updating is the standard updating rule in microeconomic theory. This updating rule has many desired properties such as the martingale property, which makes models tractable.[10]

2. Affine Transformation of Bayesian Belief and prior (Epstein, Noor and Sandroni, 2010)[11]
$$g^{bias}_{t+1}(\theta|x_{t+1}) = (1-\lambda)g_t(\theta) + \lambda g^{bayes}_{t+1}(\theta|x) \tag{3}$$

The Epstein, Noor and Sandroni (2010) updating rule is attractive as it preserves the martingale property, and yet it accommodates over- and under-updating relative to the Bayesian benchmark. $\lambda \geq 0$ is the parameter that determines the degree of over and underreaction. When $\lambda < 1$ we have underreaction and $\lambda > 1$ we have overreaction.[12] When $\lambda = 1$ we have the standard Bayesian updating.

4.1.2 Updating Rules without Martingale Property

3. Exponential distortion to prior and conditional probabilities (Grether, 1980)

$$g^{bias}_{t+1}(\theta|x_{t+1}) = \frac{g_t(\theta)^a p(x_{t+1}|\theta)^b}{\int_{\theta'\in\Theta} g_t(\theta')^a p(x_{t+1}|\theta')^b} \tag{4}$$

This updating rule can accommodate several updating biases but it violates the martingale property. $a \geq 0$ is the weight the agent places on the prior; when $a < 1$, we have

[10] Another useful property is divisibility (Cripps, 2018), where updating the signals sequentially or simultaneously does not affect the posterior belief. In the ambiguity literature, dynamic consistency (Epstein and Le Breton, 1993), receiving more information about the state of the world does not change the agent's optimal contingent plan, leading to Bayesian updating under expected utility theory.

[11] This model nests cursed belief (Eyster and Rabin, 2005), where the updated belief is just a convex combination between the Bayesian posterior and the prior.

[12] There is an upper bound on λ to ensure that it is a valid probability measure.

base rate neglect and when $a > 1$, we have confirmation bias. $b \geq 0$ is the weight the agent places on the signal; when $b < 1$ we have underreaction to signals, and when $b > 1$ we have overreaction to signals. When $a = 1$ and $b = 1$ we have the standard Bayes' rule.

This model is widely used in the analysis of experimental data because the odds ratio can be log-linearized and estimated with a linear regression.

4. Convex combination of Bayesian belief and reference belief (Hagmann and Loewenstein, 2017)

$$g_{t+1}^{bias}(\theta|x_{t+1}) = (1-\lambda)\mu(\theta) + \lambda g_{t+1}^{bayes}(\theta|x) \tag{5}$$

In this model, μ is a reference belief which is a belief that the agent wants to have (for utility reasons), and $0 \leq \lambda \leq 1$ is a parameter that draws the updated belief towards the reference belief. This model has been use to explain motivated beliefs (Bénabou and Tirole, 2002; Eil and Rao, 2011). A key prediction from this model is the agent will update asymmetrically, the agent over-updates when the signal moves the prior towards the reference belief and under-updates when the signal moves the prior away from the reference belief.

For instance, if an individual is optimistic about the wage offers he can receive. We set μ to be an optimistic belief (more likely for high wages to be drawn), and when $\lambda \neq 1$, the agent's belief will update towards this reference belief.

5 Empirical Strategy

5.1 Distribution Fitting

One main challenge in working with this data is that the bins of the wage offer differ across each individual's survey responses. The bins of the wage offer are constructed using percentages of the expected best wage offers. If subjects report different best wage expectations in different survey responses, the wage bins will have different intervals across each response. For example, when an individual changes his wage expectation, the data is as follows. The wage bins are not comparable across time.

Initial Survey Belief

	$w <$ 52,000	$52,000 \leq w <$ 58,500	$58,500 \leq w <$ 65,000	$65,000 \leq w <$ 71,500	$71,500 \leq w <$ 78,000	$w \geq$ 78,000
$p(\cdot)$	0	0.2	0.3	0.4	0.1	0

Second Survey Belief

	$w <$ 56,800	$56,800 \leq w <$ 63,900	$63,900 \leq w <$ 71,000	$71,000 \leq w <$ 78,100	$78,100 \leq w <$ 85,200	$w \geq$ 85,200
$p(\cdot)$	0	0	0.7	0.2	0.1	0

Table 2: Example of a response from the SCE survey

To make the bins comparable across time, we fit a log-normal distribution to the second survey's belief. This allows us to estimate what the updated beliefs are in the initial survey's wage bins.

We first convert the survey responses to a CDF and we use the Simulated Method of Moments (SMM) to fit various distributions to the five points on individuals' updated beliefs from the survey questionnaire.[13] We try to minimize the "distance" between the estimated probability from the fitted distribution and the actual probability. Since this process gives us an entire distribution for the updated beliefs, we can directly estimate the probability weight placed over the ranges of the updated beliefs that align with the prior wage bins defined by the survey questions by directly estimating the weight over these bins from the simulated distribution.

The main reason we fit the updated beliefs is that unbiased measurement error in the updated beliefs does not affect the excess belief movement test statistics. We discuss this in greater detail in Section 6.2.1.

We present results from fitting a log-normal distribution as the main result since it has

[13]The sixth bin is unbounded hence it is not meaningful to include this point in the fitting process of the CDF.

the lowest mean square error among all the distributions that we have tried. These alternate methods, along with their fitting errors and their effects on our main result, are described in Appendix B. Overall, while the method used to fit the data sometimes affected the magnitude of our point estimates, it did not change the direction, significance, or general interpretation of our main results (the excess belief movement test results described in section 6.1).

5.2 Martingale Test: Excess Belief Movement

5.2.1 Test Description

We will be using the excess belief movement test Augenblick and Rabin (2021) to test for the martingale property. There are alternative tests, such as testing if the updated beliefs average out to be the prior belief. The main benefit of the excess belief movement test is that the test statistics are closely related to some of the more prominent belief updating biases, such as base rate neglect. This will give us insights into the type of belief updating bias the agent is exhibiting. Moreover, Augenblick and Rabin (2021) showed that their test has a higher power in detecting non-Bayesian updating compared to other existing tests.

The test involves computing two statistics: (1) belief movement, m_{t_1,t_2}, and (2) uncertainty reduction, r_{t_1,t_2}, as shown in equations 6 and 7 respectively. We will also assume $t_2 > t_1$. Let π_t^i denote the probability assigned to bin i at time t. With h bins, the two statistics are defined as

$$m_{t_1,t_2} \equiv \sum_{i=1}^{6} \sum_{\tau=t_1}^{t_2-1} (\pi_{\tau+1}^i - \pi_\tau^i)^2 \tag{6}$$

$$r_{t_1,t_2} \equiv \sum_{i=1}^{6} \sum_{\tau=t_1}^{t_2-1} \pi_\tau^i(1 - \pi_\tau^i) - \pi_{\tau+1}^i(1 - \pi_{\tau+1}^i) \tag{7}$$

Both statistics have an intuitive interpretation. The belief movement is the total squared difference between period 2's belief and period 1's belief. This captures how much belief is changing regardless of the direction of change. For uncertainty reduction, the statistics can be interpreted as a measurement of the "variance" of the belief. If we treat each bin like a Bernoulli distribution, the expression in the summation is the variance of the Bernoulli distribution in period 1 minus the variance of the Bernoulli distribution in period 2. This is summed across all the bins. This gives us a proxy of the amount of uncertainty in the belief distribution.

If the belief updating rule satisfies the martingale property, the expected belief movement will be equal to the expected uncertainty reduction.[14] This means that if the agent's belief is expected to move greatly, we will expect the agent to become more certain about the state.

The ideal test will require us to elicit the respondents' beliefs at every possible signal realization to compute the expected belief movement and uncertainty reduction. This will allow us to determine if an individual is Bayesian. Since we only observe a single updated belief profile based on the signals the respondent observed in the last four months, we can only test if the population is Bayesian. Given that there are n observations, we compute the average belief movement as $\overline{m}_{t_1,t_2} \equiv \frac{1}{n}\sum_{j=1}^{n} m^j_{t_1,t_2}$ and the average uncertainty reduction as $\overline{r}_{t_1,t_2} \equiv \frac{1}{n}\sum_{j=1}^{n} r^j_{t_1,t_2}$. We then compute the average excess belief movement statistic:

$$X = \overline{m}_{t_1,t_2} - \overline{r}_{t_1,t_2} \tag{8}$$

where "excess" refers to the amount of movement exceeding reduction. An excess belief movement has a different interpretation depending on the amount of uncertainty reduction. For instance, with an excess belief movement of 0.01, the agent is closer to Bayesian if the uncertainty reduction is large compared to the uncertainty reduction when it is small. To allow for comparability across different studies, we also compute the normalized excess movement, which can be interpreted as the percentage of excess belief movement relative to the amount of uncertainty reduction. Under the null hypothesis that people are Bayesian, we expect $X_{norm} = 1$.

$$X_{norm} = \frac{\overline{m}_{t_1,t_2}}{\overline{r}_{t_1,t_2}} = \frac{X}{\overline{r}_{t_1,t_2}} + 1 \tag{9}$$

[14]To show this we only need to apply the law of iterated expectations.

If we reject the null hypothesis, we reject that the martingale property is satisfied. There are two possible interpretations of this result. Firstly, the updating rule that people use does not satisfy the martingale property. Secondly, the updating rule obeys the martingale property but people have incorrect priors which causes the test to reject the martingale property.

To see why correct prior matters, consider a simple two-state model with a fully revealing signal. Suppose the correct prior that state 1 is drawn is 0.5 but the Bayesian agent holds an incorrect prior of 0.7. If this trial is repeated many times the agent will expect that 70% of the time, the posterior belief is 1 and 0 in the remaining 30% of the time. However, when the data is collected, half the time the posterior belief will be 1 and 0 otherwise, and the beliefs average to 0.5 instead of the agent's prior of 0.7. This shows that the martingale property is violated and we may incorrectly conclude that the agent is non-Bayesian even when the agent is Bayesian.[15]

5.3 Test Assumption: Stable Wage Distribution During Survey

The worker's wage distribution will change depending on labor market conditions and the phase of the worker's career. For the test to be valid we only require the worker's wage distribution to be stable for the survey duration. If the wage distribution changes during the the survey period, the equality of the expected belief movement and uncertainty reduction will not hold.

While we cannot test how each individual's offer distributions are changing, we can test whether the aggregate distribution of all reported wage offers changes from period to period for individuals who report offers. We perform a KL divergence test to determine if the distribution of the pooled wage offers received by all the survey respondents is the same across time. The KL-divergence test rejects that the distributions are the same at the 5%

[15] The average belief movement collected from the data in this example will be $\frac{1}{2}(0.7-1)^2 + \frac{1}{2}(0.7-0)^2 = 0.29$ while the average uncertainty reduction is $\frac{1}{2}(0.7)(1-0.7) + \frac{1}{2}(0.7)(1-0.7) = 0.21$. We can see that the excess belief movement does not equal zero.

level, but the magnitude of the difference is fairly small. The difference in the expected log likelihood ratio of the distributions of wage offers received before and after the initial survey is 0.0319. The graph below plots the wage offer distributions in the initial and follow-up surveys.[16]

Statistic	Value
Entropy	0.0319
Standard Error	0.0136
Reference Observations (Earlier Wage Offers)	537
Comparison Observations (Later Wage Offers)	537

Table 3: Kullback–Leibler divergence statistics for wage offer distributions relative to the survey date. For each of the comparisons, the number of histogram bins was 20.

[16] Since there are not many people who reported offers in the third survey, we pooled the offers in the third survey with the second survey.

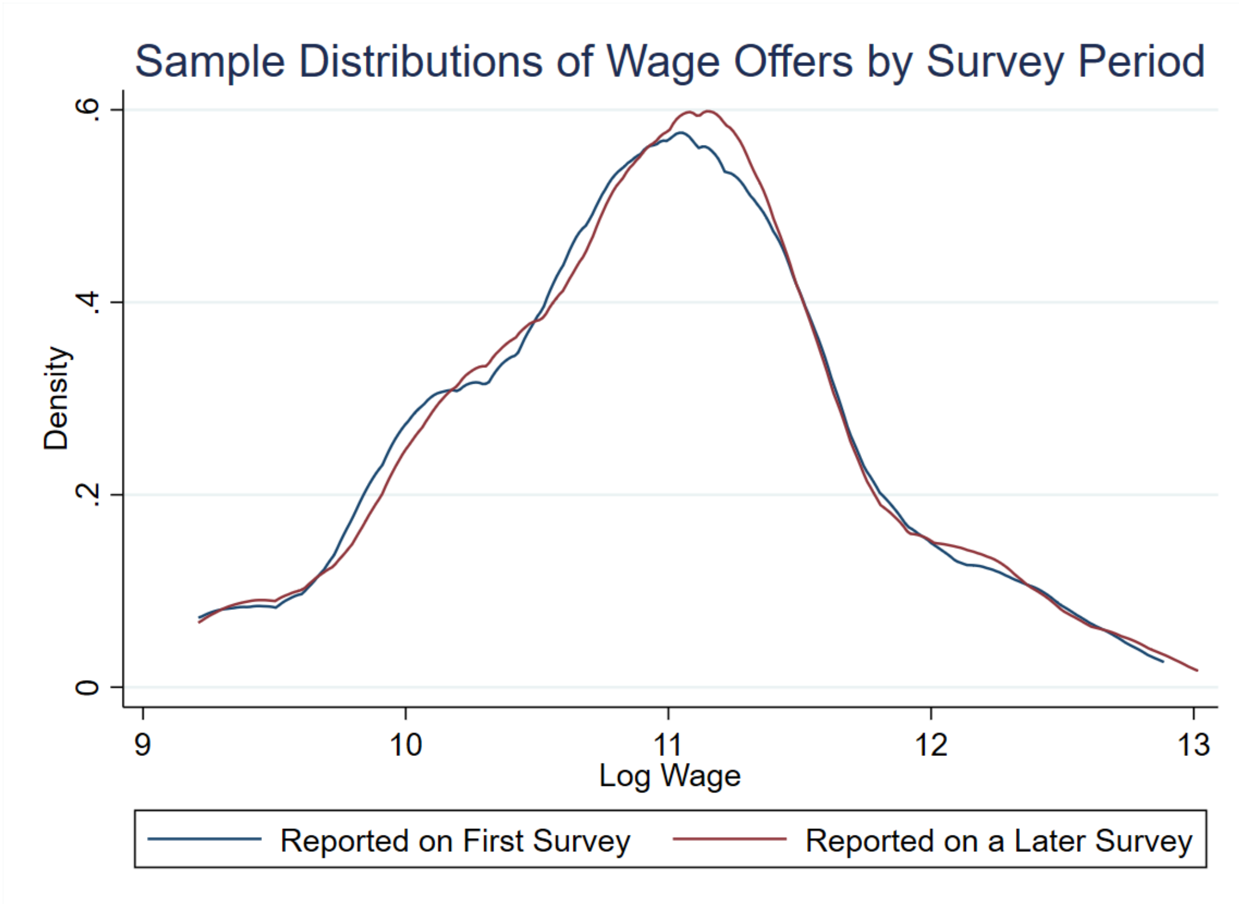

Figure 4: Wage offer distributions relative to the time of the survey. The first distribution measures offers in the period before the survey and the second distribution measures offers in the period after the survey. For individuals with two updates, offers were pooled to allow more observations in K-L and FE regression analyses.

We also perform an individual fixed effect regression to test if the average wage offer of survey respondents changed after completing a survey for those reporting offers. This analysis requires an individual to report receiving job offers in at least two surveys. It is also important to note that this test is necessary but not sufficient to show that the individual wage distribution did not change in the period of the study.

	Wage Offer
Post initial survey (=1)	262.1
	(1,815)
Constant	74,200
	(907.5)
Observations	1,074
Number of individuals	268

Table 4: Fixed effect regression of wage offers on an indicator for being in the period after the initial survey. Fixed effects are at the individual level. Standard errors are clustered at the state level.

As shown in table 4, we see that wage offers are not statistically different across the survey period. These tests do not find evidence of wage distribution instability during the survey period for each individual.

6 Results

6.1 Excess Belief Movement Test

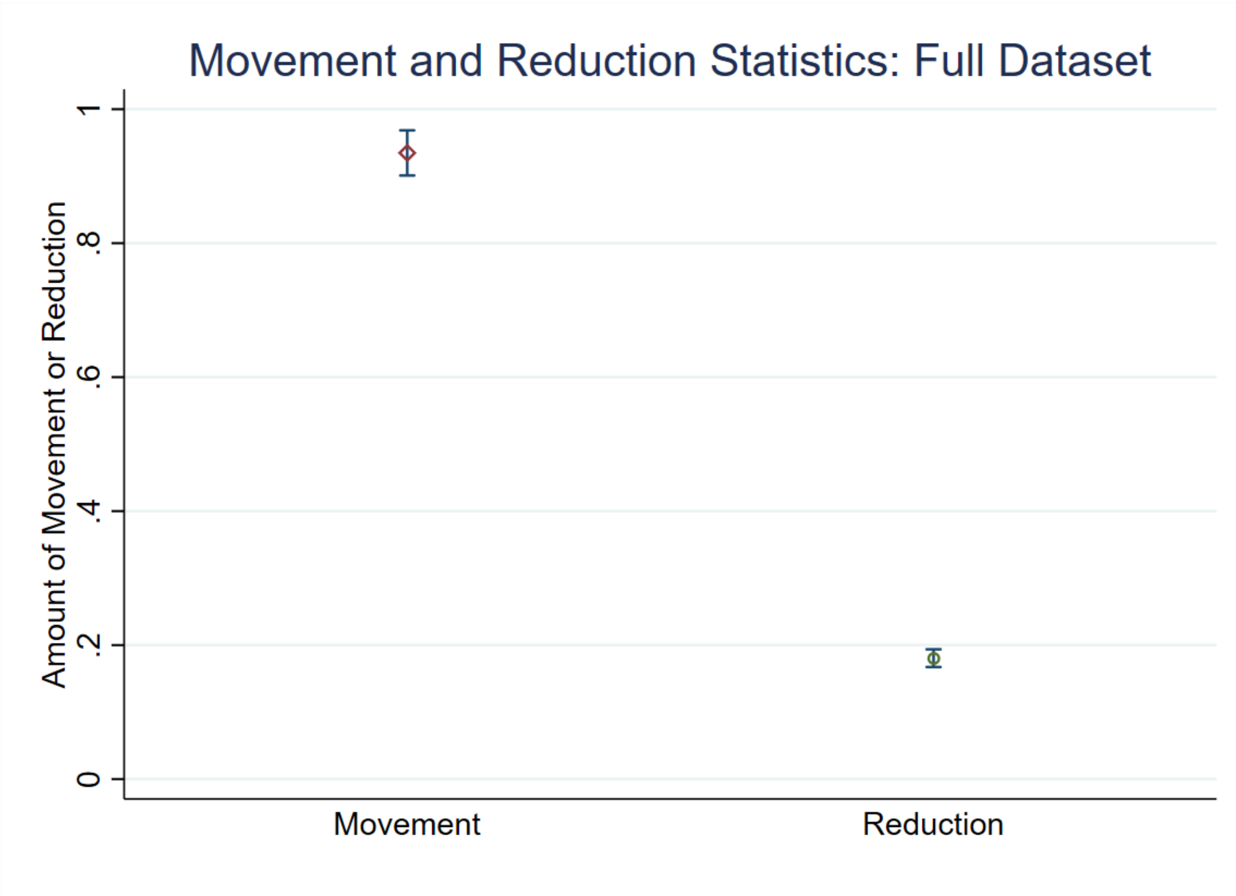

Figure 5: Movement and uncertainty reduction summary statistics. Error bars display 95% confidence intervals around the mean.

We first present our main result from the excess belief movement test. Overall we found excess belief movement relative to the amount of uncertainty reduction, as shown in column 1 of table 5 and the box plot in figure 5.[17] The normalized excess belief movement of $X_{norm} = 5.1744$ means that beliefs are moving 417% more than the amount of uncertainty

[17]The observation counts in table 5 are slightly less than in the summary statistics table because there were 3 individuals missing state (unable to assign to cluster) and 46 individuals with two updates for which we had enough information to use their second update, but not their first. In the latter case, we still used their second update in the analyses based on single updates, but not for the movement and reduction-based analyses. This is because the movement and reduction-based analyses are meant to use all measured periods for an individual.

reduction. Excess belief movement suggests that people are over-updating relative to the Bayesian benchmark, and this is consistent with updating biases such as base rate neglect and overreaction to signals.

Statistic	All Individuals	Got Offer?		Searched?		
		Yes	No	Yes	No	Unknown
$\overline{m}$	.9341	1.0499	.9011	1.0083	.9004	.9517
	(.0193)	(.0434)	(.0213)	(.0372)	(.0213)	(.0545)
$\overline{r}$	.1805	.1981	.1755	.1989	.1687	.2152
	(.0079)	(.0190)	(.0085)	(.0134)	(.0080)	(.0252)
X	.7536	.8518	.7256	.8094	.7317	.7365
	(.0207)	(.0472)	(.0248)	(.0358)	(.0224)	(.0539)
X_{norm}	5.1744	5.2999	5.134	5.0702	5.3369	4.4228
	(.2488)	(.5527)	(.3015)	(.3384)	(.2766)	(.5153)
p-value of t-test: $X = \overline{m} - \overline{r} = 0$	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Observations	2,489	552	1,937	691	1,613	185

Table 5: Excess movement statistics: Log normal-fitted results. Clustered errors by state in parentheses. Some of the subjects in our sample did not answer the question of whether they searched, but we still include them here as they are included in our full sample estimate.

We hypothesized that people who received a job offer would update in a more Bayesian manner because job offers are individualized and direct feedback about one's wage distribution, while other information like labor market news is not direct feedback on their wage distribution. But surprisingly, while those who got at least one offer have a slightly higher normalized statistic, the statistic is still very large and of a similar magnitude as the statistic for those who did not receive any offers.

We also examine the heterogeneous effect between respondents who are searching for a job and not searching. Intuitively, people who are actively searching for a job are likely to put more thought into forming their wage expectations. This could mean job searchers are more Bayesian than non-job searchers. We do observe that respondents who are searching for a job have slightly smaller normalized excess belief movement statistics on average. However, the normalized excess belief movement statistics are still large for respondents who are searching for a job.

Figure 6 below provides an alternative look into the excess belief movement test results. The provided box plot compares belief movement and uncertainty reduction measures of four different groups of individuals. The graph shows the surprising result that movement and reduction remain fairly similar across both categories, contrary to our prediction that those with personal wages and those who are searching would behave in a more Bayesian manner.

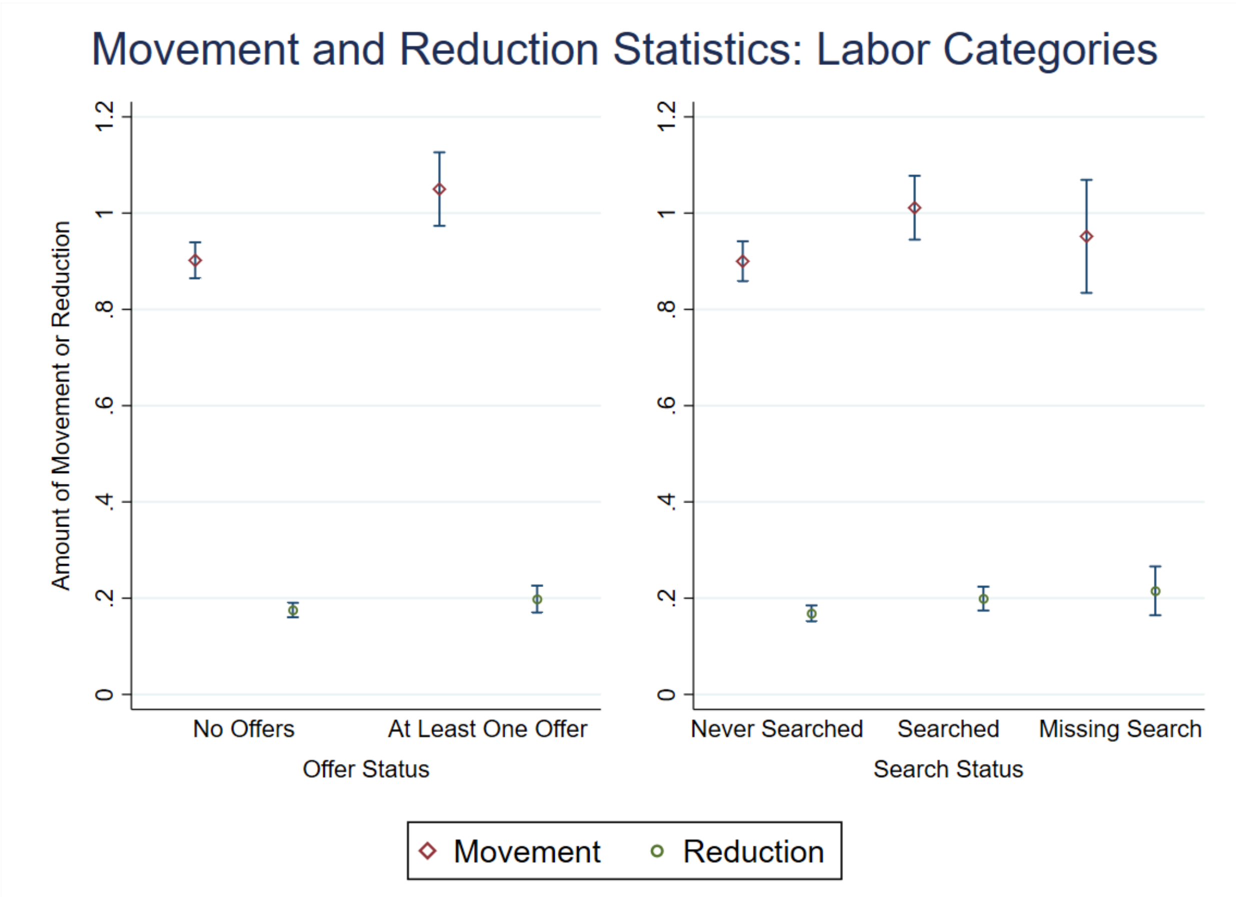

Figure 6: Movement and reduction summary statistics by whether individuals ever received an offer or ever searched. Error bars display 95% confidence intervals around the mean.

6.1.1 Updating Rules that Produce Excess Belief Movement

Given the result that people are non-Bayesian, what are some updating rules that can generate this data? Firstly, we can reject all updating rules with the martingale property, which includes Bayesian updating and the Epstein, Noor and Sandroni (2010) model.

The Grether (1980) model as described in equation 4 can be consistent with the results obtained from this test. If we set $a < 1$ and/or $b > 1$, which suggests the agent exhibits base rate neglect and overreaction to signals, we can get excess belief movement. The intuition is that if people overweight the signal they observe or underweight their prior beliefs, this will result in over-updating in the beliefs compared to the Bayesian benchmark which results in excess belief movement.

The Hagmann and Loewenstein (2017) model can produce excess belief movement, but the conditions to produce the excess belief movement are more complicated than the Grether (1980) model in a setting with multiple states or wage bins. From Augenblick and Rabin (2021), we know that the excess belief movement statistics has the following formula[18]

$$X = \sum_{i=1}^{6} \mathbb{E}[(2\pi_t^i - 1)(\pi_t^i - \pi_{t+1}^i)]. \tag{10}$$

In this model $\mathbb{E}\pi_{t+1}^i = (1-\lambda)\pi_{ref}^i + \lambda\pi_t^i$, where π_{ref}^i is the reference belief and $0 \leq \lambda \leq 1$. If we expand the above expression, we get

$$X = (1-\lambda)\sum_{i=1}^{6}(2\pi_t^i - 1)(\pi_t^i - \pi_{ref}^i). \tag{11}$$

There are various combinations of priors and reference beliefs where we can obtain excess belief movement from this model.[19] The sign of the test statistics depends on the summation. However, the parameter λ affects the magnitude of the test statistics. If λ is small, the agent puts more weight on the reference belief and less weight on the Bayesian beliefs. This makes the agent more non-Bayesian, resulting in an excess movement statistic that is larger in magnitude.

[18] We will use a two-period model to provide intuition, the result generalizes to more than two periods.

[19] In a two-state case, $X = (1-\lambda)(2\pi_t - 1)(\pi_t - \pi_{ref})$. We can see that we get excess belief movement in two cases: (1) $\pi_t < 0.5$ and $\pi_{ref} > \pi_t$, (2) $\pi_t > 0.5$ and $\pi_{ref} < \pi_t$. The idea is that if the expected updated belief is biased towards 0.5, we will get excess belief movement. This is consistent with the result from Augenblick and Rabin (2021).

6.2 Robustness Checks

6.2.1 Measurement Error

In surveys, it is possible to potentially misreport their beliefs or round off some of the probabilities. In this section, we address the concern that our primary result might be driven by this measurement error.

Consider that the theoretical framework where the agent has a true belief of π_t but reports a distorted $\hat{\pi}_t = \pi_t + \epsilon_t$, where ϵ_t is the measurement error. In a two-state model, assuming that the measurement error term is mean zero with variance σ_ϵ^2 and uncorrelated with recent belief and error realizations ($\mathbb{E}(\epsilon_t \pi_t) = \mathbb{E}(\epsilon_t \pi_{t-1}) = \mathbb{E}(\epsilon_t \epsilon_{t-1})$), Augenblick and Rabin (2021) showed that the excess belief movement will be equal to $2\sigma_{\epsilon_t}^2$.

Generalizing this to n states, we show that the excess belief movement will be equal to $\sum_{i=1}^{n} 2\sigma_{\epsilon_t^i}^2$.[20] With measurement error, the excess belief movement is equal to the total variance of the measurement error in the prior belief multiplied by 2. One interesting feature of this test is only measurement error in the prior affects the test statistics. This motivates our method of fitting a log-normal distribution to the posterior to compare it to the prior beliefs, which minimizes the error term in the prior.

This result tells us that only the measurement error in the first period will affect the excess belief movement statistics. In the calibration exercise, we only have to introduce the measurement error to the prior belief.

We also perform a Monte Carlo simulation to determine how much measurement error is required to produce the result we obtained. The Monte Carlo simulation is designed to mirror the setting in our data as closely as possible. Since there are 6 wage bins and most subjects responded to the survey only twice, we have 6 states and with only 2 periods we will have only 1 signal realization and 1 update. Since we have 2,489 individuals for which we can calculate movement and reduction statistics, we generated 2,489 pairs of priors and posteriors in each simulation.

[20] Refer to appendix A.1 for the poof.

Since our statistics of interest only depend on measurement errors in the prior, we only introduce measurement errors to the prior beliefs. We draw our simulated prior data from a symmetric Dirichlet distribution that is centered around a uniform prior distribution.[21] Since we are working with uniform priors, the parameters of the Dirichlet distribution are equal. We then scale the parameters to adjust the variance of the distribution to match the normalized excess belief movement that we obtained.

Assuming that only measurement error is driving the results we observe, we can compute the supposed Bayesian belief movement and uncertainty reduction statistics. Based on the result we obtained, the supposed Bayesian belief movement and uncertainty reduction are 0.9341 and 0.1805, respectively, and we try to match these statistics in our calibration exercise.

It is important to note that there are infinitely many possible combinations of prior beliefs and distributions of posterior beliefs that can produce the same desired belief movement statistics. For this calibration exercise, we assume a uniform prior. The main reason for this assumption is that we can construct a uniform distribution of the posterior beliefs with symmetric posteriors, which is easier to work with.

The posterior belief is then selected to match the supposed Bayesian belief movement and reduction statistics. One of the possible posterior beliefs that are selected in this calibration exercise is $(0, 0.8346, 0.1654, 0, 0, 0)$. To obtain the full distribution of posterior beliefs, we can permute the order of the probabilities and each of these permutations will be realized with equal probability.

Finally, the variance of the error term is not an intuitive measure of the amount of measurement error. To quantify measurement error more intuitively, we define $\hat{\pi}_1^i$ as the prior beliefs drawn from the Dirichlet distribution. The measurement error, $\Delta \equiv \sum_{i=1}^{6} |\hat{\pi}_1^i - \pi_{prior}^i|$, is the total "distance" between the beliefs drawn from the Dirichlet distribution and true

[21] The draws from the Dirichlet distributions are a valid probability distribution that has to sum to 1. This approach is better than adding an error that is normally distributed which may cause the probability measure to be negative or exceed 1.

prior. Δ is bounded above by 2, which is a complete misreporting of beliefs.

We simulate this process 10,000 times. In Table 6, we report the average statistics computed from the 10,000 Monte Carlo simulations and the interval that contains 95% of our simulation result.

We see that to obtain an excess belief movement from our dataset assuming the respondents are Bayesian, we need the survey respondents to misreport their prior beliefs by a total of 116 percentage points across the six bins. This means that survey respondents who have a belief of 70 in one of the wage bins have to misreport the belief as 12, with the misassigned weight redistributed across the remaining wage bins. The measurement error required to rationalize the behavior is very large and is unlikely to be the sole explanation for the result we obtained.

Simulated Statistic	Uniform Prior with Matched Parameters	Target Value from excess belief movement test
$\overline{m}$	0.9341 [0.9163, 0.9521]	0.9341
$\overline{r}$	0.1792 [-0.2344, 0.4562]	0.1805
X	0.7535 [0.7316, 0.7756]	0.7536
X_{norm}	5.1770 [4.9170, 5.4573]	5.1744
Δ	1.1618 [1.1522, 1.1716]	

Table 6: Measurement error calibration statistics. The measurement error is generated by a Dirichlet prior and we pick Bayes' plausible distribution of posterior beliefs in order match the belief movement and uncertainty reduction statistics.

The analysis here assumes all survey respondents have the same measurement error in reporting their beliefs. This assumption can be relaxed and we require the average total variance in the error to be equal to half of the excess belief movement statistics. However, Δ will be lower in the case of heterogeneous measurement error due to the concave relationship between Δ and excess belief movement. Figure 7 plots the relationship between Δ and the excess belief movement statistics from the simulation exercise that is described above.

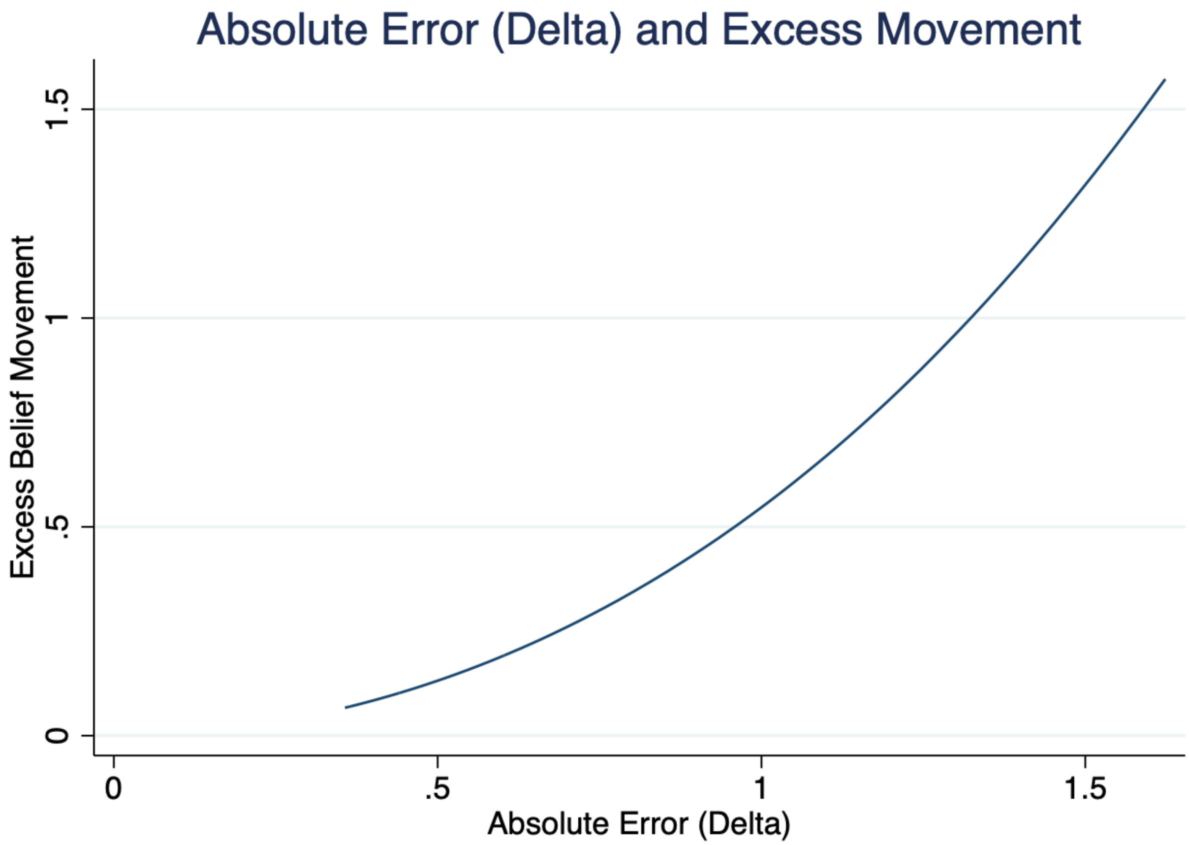

Figure 7: Relationship between Δ for a Dirichlet distribution centered around the uniform distribution and excess belief movement statistics.

We can see that if we allow for heterogeneous measurement error and fix the average excess movement statistics, the average Δ required to rationalize the statistics will be lower. Assuming homogeneous measurement error gives us an upper bound for Δ.

6.2.2 Two Wage Bins

Because of how the excess belief movement statistic is constructed, it is possible that having six different wage bins could inflate the excess belief movement statistics. Moreover, since most of the other studies that use the excess movement test have binary states (Augenblick and Rabin, 2021; Augenblick, Lazarus and Thaler, 2023).[22] For comparability with other studies, we also perform a robustness test by collapsing the six wage bins into just two bins. We combine the first 3 wage bins, beliefs that the maximum wage offer is lower than the response in OO2a2, and the last 3 wage bins, beliefs that the maximum wage offer is greater than the response in OO2a2, into a single wage bin.

The excess belief movement statistics in our study are still significantly larger than other studies. In Augenblick and Rabin (2021) the largest normalized excess belief movement was 1.2, while for Augenblick, Lazarus and Thaler (2023) it is about 1.24 as computed from their regression estimates in Table 2 for the basketball game, the statistics from financial data is 1.46. As shown in table 7, we see that the statistics that we got are still significantly larger than the other studies. In section 6.3, we provide some explanation for the large excess movement statistics that are observed in our dataset. To understand why these statistics are larger, we further examine the survey respondents' updating patterns.

Statistic	All Individuals	Got Offer?		Searched?		
		Yes	No	Yes	No	Unknown
$\overline{m}$	.6058	.6637	.5892	.6374	.5876	.6457
	(.0148)	(.0401)	.0147	(.0321)	(.0172)	(.0619)
$\overline{r}$	.1072	.1221	.1029	.1156	.1044	.0998
	(.0064)	(.0111)	(.0060)	(.0111)	(.0066)	(.0177)
X	.4986	.5417	.4863	.5218	.4832	.5459
	(.0149)	(.0413)	(.0156)	(.0299)	(.0174)	(.0622)
X_{norm}	5.6524	5.437	5.7252	5.0702	5.3369	4.4228
	(.3383)	(.5871)	(.3574)	(.3384)	(.2766)	(.5153)
p-value of t-test: $X = \overline{m} - \overline{r} = 0$	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001	< 0.001
Observations	2,489	552	1937	691	1,613	185

Table 7: Excess movement statistics: Two-bin results. Clustered errors by state in parentheses.

[22] Some of the other settings include forecasters predicting the likelihood of geopolitical events, as well as beliefs in a British prediction market

6.2.3 Incorrect Priors

Our excess belief movement test results could also be influenced by the level of bias in agents' priors. In this section, we discuss why we do not believe incorrect priors can fully explain our results. First, we do not find evidence of convergence to correct beliefs in table 8 of section 6.4.1. If individuals only had biased starting beliefs, but still employed unbiased updating, this would eventually lead to belief convergence.

Second, we present the following analysis which estimates the minimum amount of error in priors needed to explain our observed excess movement statistic, based on the data. We find that this minimum bound is much higher than our best estimate of the actual level of error in priors in our data. For tractability, we apply this analysis to the two-state case used in our analysis from section 6.2.2.

Consider the two-bin analogue of equation 10:

$$X = \mathbb{E}[(2\pi_t - 1)(\pi_t - \pi_{t+1})] \tag{12}$$

Suppose individuals all perfectly update to correct priors after one period, but begin with biased beliefs. Then the most incorrect priors could contribute to our observed positive result would occur if people who believed it was more likely to get a higher wage (relative to their guess of the average best wage) were as overly-optimistic as possible and if people who believed it was more likely to get a lower wage were as overly-pessimistic as possible. This is because X is maximized when individuals move towards $\pi_t = \frac{1}{2}$, the point of greatest uncertainty.

To construct a bound for the minimum level of bias in priors needed to produce our results, we first assume that signals are fully revealing so that individuals fully update to the correct prior after one period. Next, we further assume that each individual has the direction of bias in their prior that will increase movement as much as possible when updating to this correct prior. Finally, we assume that each person has the same absolute magnitude of error in priors, $b > 0$, for tractability.

The sample estimator for X can then be written as

$$\hat{X} = \frac{1}{n}\sum_{i=0}^{n}[(2\pi_{t,i} - 1)k_j] \tag{13}$$

where $k_j = b$ if $\pi_t > \frac{1}{2}$, $k_j = -b$ if $\pi_t < \frac{1}{2}$, i is subject number, and n is sample size.

The figure below plots the level of bias in priors against the value of the excess movement in the sample, calculated using equation 13.

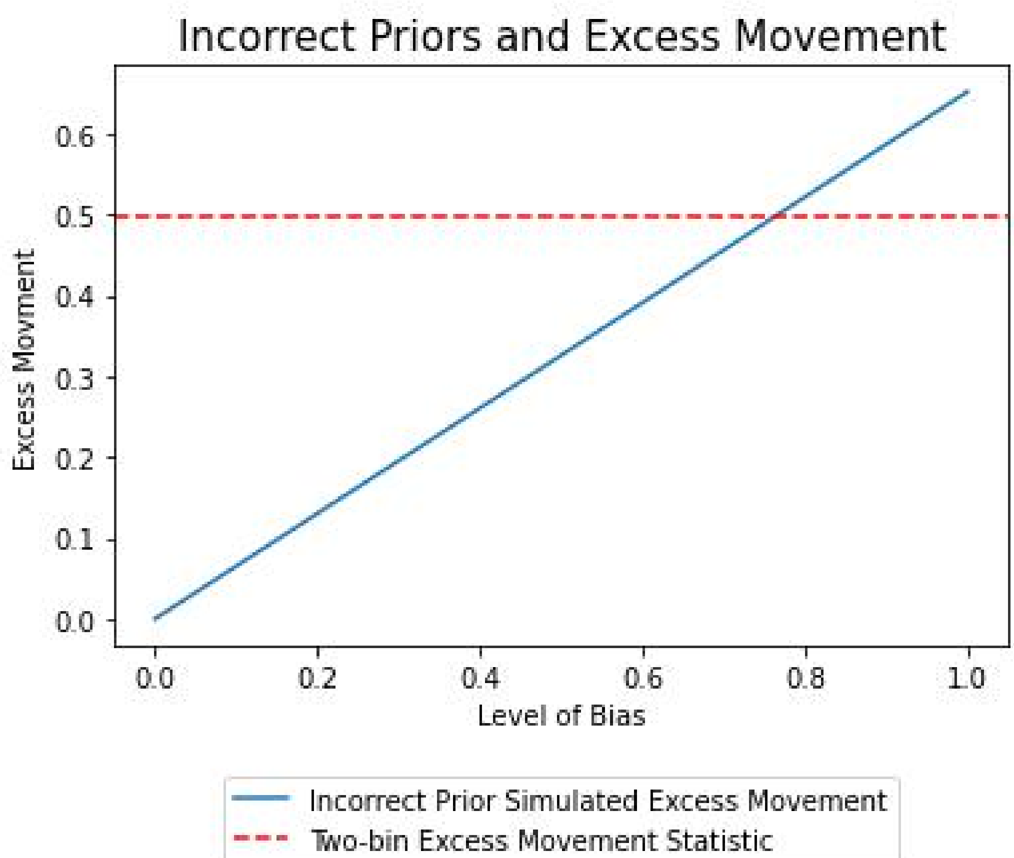

Figure 8: Simulation of minimum prior bias needed to observe different levels of excess movement. This simulation assumes all subjects are equally biased in their priors in a way that maximizes the excess movement from the shared bias.

As indicated by the graph, for our two-bin result of $\hat{X} = 0.4986$ to be fully explainable by incorrect priors in the theoretical case where incorrect priors would have the greatest impact, prior beliefs would have to differ from correct beliefs by about 0.7650.

To see if this is a reasonable amount of error given our data, we estimate correct priors for each individual using data from the monthly Current Population Survey (CPS). Specifically, we fit normal distributions to log wages reported by CPS respondents during the same period covered by our sample. The distributions are allowed to vary by race, gender, age group, whether respondents hold a college degree.[23] We then estimate correct priors by summing the weight for the matched distributions according to the two-state version of the wage bin definitions implied by survey question OO2b for each SCE subject.

When calculated this way, we find that the average absolute value difference between prior beliefs on the CPS and those on the SCE is about 0.3178, with this difference having a standard error of 0.0049. This is far below the necessary bound needed for incorrect priors needed to fully explain our results.

6.3 Updating Unlikely Events

In section 6.2.2, we showed that the excess belief movement we obtained is significantly larger than other studies even when we only have two wage bins. The main reason is wage offers that were thought to be unlikely in the initial survey are updated to likely offers in subsequent surveys, as shown in figure 9 where the density function shifted to a region that has a low probability satisfied. In this example, this update is likely to be caused by receiving a "surprising" offer that is highly unlikely according to the prior which causes a large change in beliefs. This type of update is possible for a Bayesian agent when the agent receives a very informative signal. But such updates should be rare if the martingale property is satisfied.

[23] The chosen age groups used are 18-40 (early career), 40-65 (late career), and 65+ (retirement age). These age groups roughly mimic those pre-existing in the SCE data, although we do not rely on these categorical measures directly when matching on subject age.

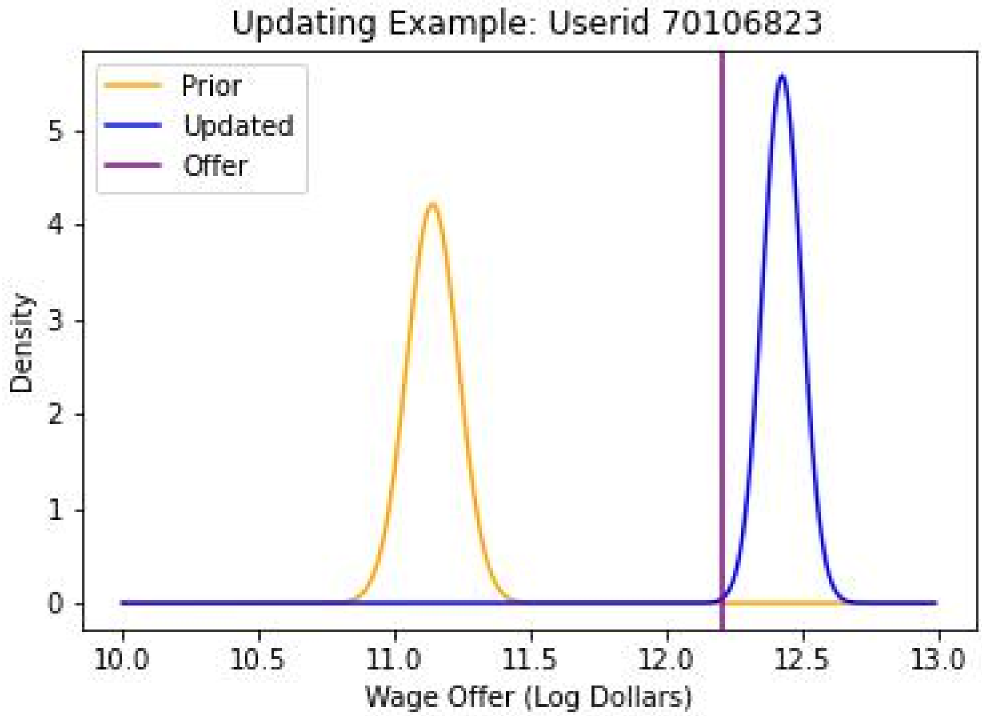

Figure 9: Example of a survey response where beliefs are fitted with a log-normal distribution.

In this section, we estimate the number of instances where an unlikely wage offer is updated to a likely outcome. First, we identify wage bins in the initial survey where the survey respondent indicated zero probability of receiving a wage offer from that wage bin. We interpret a zero probability response as the survey respondent thinking that the offer is unlikely instead of an impossible offer. We then fit a log-normal distribution to the updated beliefs as discussed in the earlier section, and set various thresholds for probability weights that are considered likely. Figure 10 shows the relationship between the proportion of updates and the different thresholds used. For thresholds between 10 and 50%, we estimate that about 23-36% of observations update an unlikely wage offer in the initial survey into a likely wage offer in the follow-up survey.

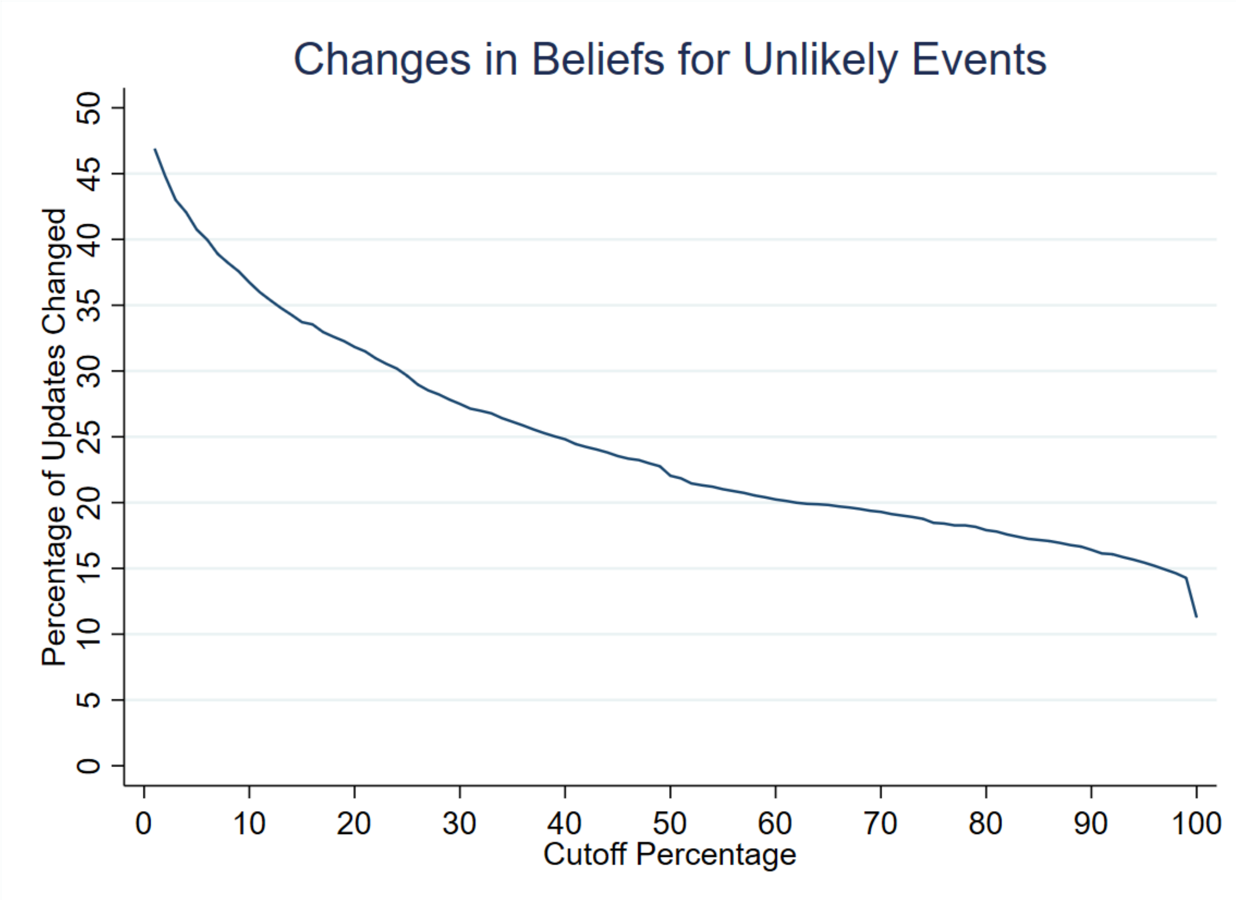

Figure 10: Fraction of observations that updates unlikely wage offers to likely offers. The y-axis shows the proportion of updates and the x-axis shows the different cutoffs.

For updating rules that multiply the conditional probabilities with the prior, like Bayes' rule and Grether (1980), we require a very informative signal or extreme biases in the latter model to update unlikely events in the prior into a likely event. These models may not be useful if we want to accommodate such updating behaviors. Alternatively, some non-Bayesian updating rules accommodate updating zero probability events or unlikely events in the prior into likely events and highly possible events in the updated beliefs (Ortoleva, 2012; Hagmann and Loewenstein, 2017; Ba, 2022). These updating rules may be more appropriate to explain these behaviors.[24]

[24]It is important to acknowledge that these models have several free parameters and it is difficult to falsify these models.

6.3.1 Overreaction and Base Rate Neglect

As discussed previously in section 6.1.1, we see that a Grether (1980) model with base rate neglect or overreaction to signals can produce excess belief movement. However, these two biases have different implications in the long-run beliefs.

Figure 11 presents a simulation result for the dynamics of beliefs for a sequence of signals that are drawn randomly. In this simulation, there are two possible states $\{H, L\}$ and the agent began with a prior of 0.5 for state H in the baseline case. In each period, the agent observes an independently drawn signal that predicts the state accurately 75% of the time. Suppose the drawn state is H. We can see that a Bayesian agent's belief will converge to the real state. An agent with overreaction will converge faster than the Bayesian agent, while the beliefs of an agent with base rate neglect will not converge. The intuition is that a base rate neglect agent "forgets" the prior beliefs and old signals which prevent beliefs from converging. Additionally, if the agent starts with a prior of 0.25 instead, she will still converge to the real state, although it will take her more periods to do so.

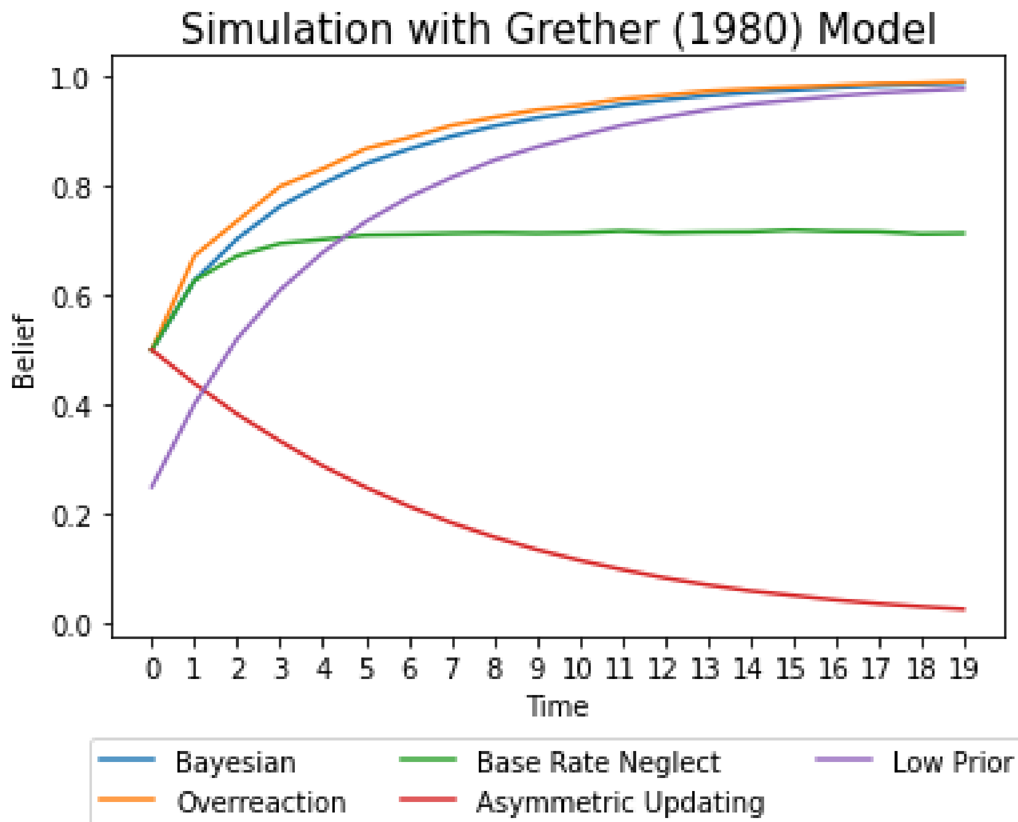

Figure 11: Simulation result showing the dynamics of beliefs. Overreaction and base rate neglect is modeled with Grether (1980) model and we choose $a = 0.5$ for the agent with base rate neglect and $b = 1.5$ for the agent with overreaction. Graphed points are averages from 10,000 sequences drawn Monte Carlo.

It is crucial to differentiate between these biases because they carry distinct implications for information provision policies. If overreaction is causing the excess belief movement, there might not be a need for an intervention in an information-rich environment as individuals will eventually learn about their wage distribution. Similarly, if the excess belief movement is caused by incorrect prior beliefs due to overoptimism, people's belief will eventually converge to the true distribution if they are Bayesian or exhibit overreaction. On the other hand, if base rate neglect is causing the excess belief movement, individuals may never learn about their wage distribution which provides a stronger justification for a policy intervention. Moreover, we will need to consistently provide information to the individuals as old information will be "neglected" or "forgotten."

Unfortunately, the SCE dataset is unable to convincingly distinguish between these two biases as we only observe at most two updates. The ideal test would be to observe the same individual over a longer time horizon with more frequent surveys and test if the beliefs converge over time.

To test for convergence of beliefs, we focus on individuals who completed three surveys (had two updates). If beliefs are converging, we would expect the belief movement and uncertainty reduction to decrease over time.[25] To test this we run a fixed effect regression and find that belief movement and uncertainty reduction are not significantly different across periods of survey. This suggests that base rate neglect is a likely cause of the excess belief movement. This finding supports the robust finding from lab experiments where people neglect their priors (Kahneman and Tversky, 1972; Benjamin, 2019; Esponda, Vespa and Yuksel, 2024).

	Movement	Reduction
Second Update (=1)	0.0121	-0.0131
	(0.0196)	(0.0156)
Constant	0.624	0.171
	(0.00979)	(0.00780)
Updates	2,146	2,146
R-squared	0.000	0.001
Individuals	1,073	1,073

Table 8: Individual fixed effect regressions of movement and reduction on sequential update number for individuals with two updates. Standard errors clustered by state.

Moreover, lab experiments have found that the main belief updating biases in the lab are base rate neglect and conservatism bias, where people underact to new information Benjamin (2019). If the results from the lab can be generalized beyond the lab, then it is unlikely that overreaction is causing the excess belief movement since conservatism bias is the main bias from lab experiments.

[25] Our main test calculates movement and uncertainty reduction measures using an individual's entire sequence of updates together, but for this test, we compare movement and reduction from one update to the next.

6.3.2 Asymmetric Updating Patterns

Another justification for a policy intervention is asymmetric updating. Consider a worker who responds over-updates when positive signals are received and under-updates when negative signals are received. If the environment is informative and provides enough bad signals for the agent, the worker's belief will converge to the true distribution at a slower rate compared to Bayesian updating. If the environment is not informative enough or in an extreme case of asymmetric updating,[26] the worker's belief may converge to the wrong distribution.

Among the 3,607 updates in our dataset, 626 updates, approximately 17% of the sample, reported receiving at least one job offer between periods. For individuals who reported receiving a job offer, we have some information about the type of signals they have received during the job search. This allows us to further examine how they are updating their beliefs. We observe that there is asymmetric updating which is consistent with motivated beliefs (Bénabou and Tirole, 2002; Eil and Rao, 2011; Hagmann and Loewenstein, 2017).

To identify this asymmetric updating pattern, we use the response about the average wage offer. We construct the following statistic, which we will refer to as the normalized change in expected wage.

$$\frac{y_2 - y_1}{|\overline{x}_2 - y_1|}$$

Here y_i is the subject's guess of the average offer they could earn over the next four months on survey i (that is, the response to survey question "OO2a"). $\overline{x}_i$ is the average of the offers reported by the subject on survey i. In the numerator, we have the change in the respondents' beliefs. In the denominator, we have the absolute difference between the average offers the survey respondent received and the initial beliefs to control for the magnitude of the update.

[26] In the most extreme case consider a worker who only updates when positive signals are observed and ignores all negative signals. The belief will converge to the wrong wage distribution.

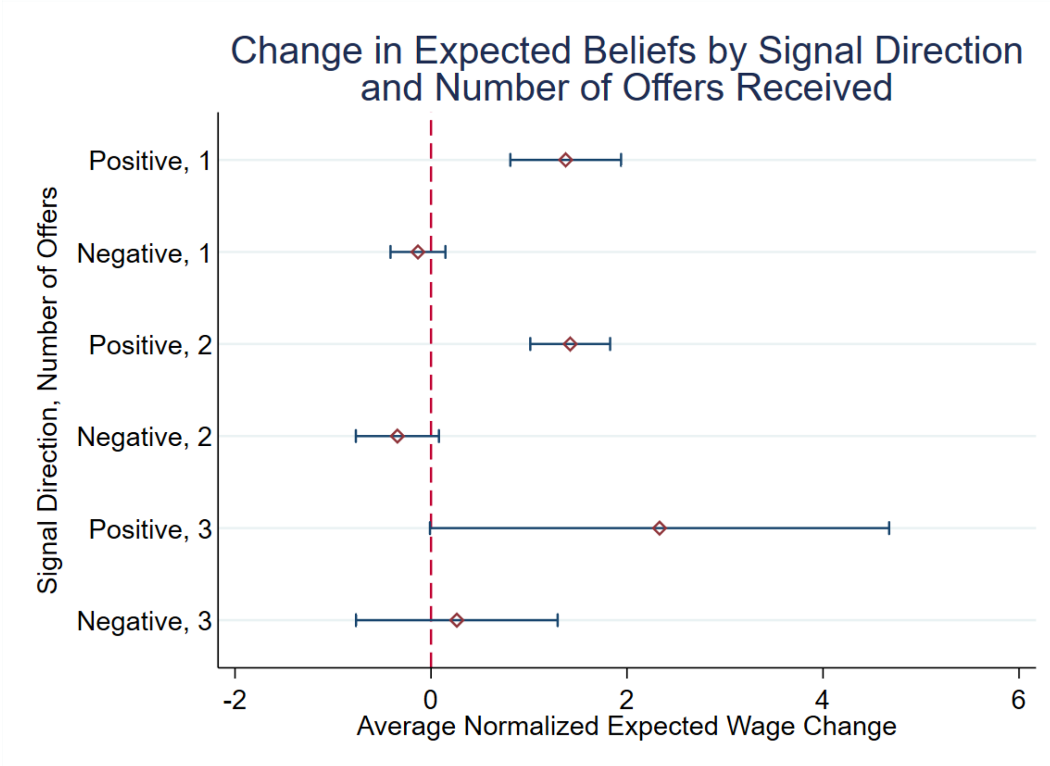

Figure 12: Normalized Expected Wage Changes by Signal Direction. Belief variable here is the average annual salary the survey respondent expects. Error bars display 95% confidence intervals around the mean.

As shown in figure 12, we see that the normalized expected wage change is larger in magnitude when the survey respondent receives a positive signal compared to a negative signal. This pattern is robust to the number of job offers they receive. The updating pattern is consistent with motivated beliefs (Bénabou and Tirole, 2002; Eil and Rao, 2011), where individuals want to have an optimistic outlook on their job prospects. In a theoretical model like Hagmann and Loewenstein (2017), we set the reference belief to an "optimistic" belief and we can obtain this asymmetric updating pattern where beliefs would move upwards. If we set λ to be large, people can update in the wrong direction, as well.

Asymmetric updating provides a potential explanation to reconcile the seemingly disparate findings of beliefs over periods of unemployment (Krueger and Mueller, 2016; Mueller, Spinnewijn and Topa, 2021), while Conlon, Pilossoph, Wiswall and Zafar (2018) and our paper found general over-updating pattern behavior.[27] Studies that found under-reaction typically requires downward adjustments in beliefs or reservation wage, this is consistent with our result where people are not adjusting their beliefs as much when they should update their beliefs downwards.

Finally, we do observe a surprising result in figure 12, where the normalized expected wages change are not too different across the number of offers received. Receiving multiple offers is more informative than receiving a single, we should expect more updating towards the average wage offers. One possible explanation is the "law of small numbers" (Rabin, 2002; Benjamin, 2019), where people hastily draw conclusions from a small sample size.

7 Discussion

In the excess belief movement test (Augenblick and Rabin, 2021), we find significantly more movement than uncertainty reduction in our data. This allows us to reject updating rules that have the martingale property such as Bayesian updating and Epstein, Noor and Sandroni (2010) updating rule. Theoretical models that want to incorporate non-Bayesian updating to job search model, should avoid using updating rules with the martingale property as such models do not describe how people update their beliefs.

The updating patterns observed are consistent with base rate neglect and overreaction to signals in the Grether (1980) model. Our findings suggest that information intervention is a powerful tool for influencing people's beliefs in a job search setting, due to over-updating relative to the Bayesian benchmark. To distinguish between base rate neglect and overreaction, we look at patterns of convergence in beliefs. While the dataset we use is limited in that we can only observe at most three responses from the survey respondents, we still found suggestive evidence that subject beliefs are not converging. This finding is suggestive of the presence of base rate neglect in this environment.

Moreover, we detected evidence of asymmetric updating, which can be explained by

[27] It is important to note that Conlon, Pilossoph, Wiswall and Zafar (2018) and our paper use the same dataset.

the Hagmann and Loewenstein (2017) updating model. This updating pattern may hinder individuals from learning about their wage distribution. There are also policy implications on how good news and bad news should be conveyed to individual due to the asymmetric patterns in updating. When providing information that can cause individuals to adjust their wage expectations downwards, we may need a stronger or more informative signal.

Our results are limited in a two important ways that leave scope for future work. Firstly, the current dataset does not allow us to effectively distinguish between overreaction and base rate neglect. These two updating biases have different asymptotic properties which will result in different long run behavior. If individuals exhibit overreaction, it is possible for them to eventually learn about the wage distribution they are drawing from. Overreaction does not pose a problem in the long-run if there is enough feedback and information to allow their beliefs to converge. On the other hand, if individuals exhibit base rate neglect, it may be impossible for them to learn about the distribution they are drawing their wage distribution. In this case, there may be a stronger need for intervention to correct people's beliefs as it is unlikely they will ever learn about their underlying wage distribution through feedback alone. Future analyses could collect higher-frequency data, allowing for a better study of how beliefs changes over time to detect patters of convergence.

Secondly, our sample chiefly includes employed individuals, so it is difficult for us to comment on the differences between employed and unemployed learning, with our analysis mostly applying to learning among employed individuals. The beliefs of unemployed individuals are likely to have the largest effect on their welfare as their beliefs will affect their decision to accept a job offer. It will be useful to have a survey where we can primarily focus on unemployed individuals. This will help us to design more effective policy.

References

Ahn, Hie Joo and Ling Shao, "The Cyclicality of On-the-Job Search Effort," *The B.E. Journal of Macroeconomics*, 2021, *21* (1), 185–220.

Armantier, Olivier, Giorgio Topa, Wilbert Van der Klaauw, and Basit Zafar, "An overview of the survey of consumer expectations," *Economic Policy Review*, 2017, pp. 51–72.

Arni, Patrick P, "Opening the Blackbox: How Does Labor Market Policy Affect the Job Seekers' Behavior? A Field Experiment," *IZA Discussion Paper Series*, 2016.

Athey, Susan, "Monotone comparative statics under uncertainty," *The Quarterly Journal of Economics*, 2002, *117* (1), 187–223.

Augenblick, Ned and Matthew Rabin, "Belief movement, uncertainty reduction, and rational updating," *The Quarterly Journal of Economics*, 2021, *136* (2), 933–985.

_ , Eben Lazarus, and Michael Thaler, "Overinference from weak signals and underinference from strong signals," *arXiv preprint arXiv:2109.09871*, 2023.

Ba, Cuimin, "Robust Misspecified Models and Paradigm Shift," 2022.

Benjamin, Daniel J, "Errors in probabilistic reasoning and judgment biases," *Handbook of Behavioral Economics: Applications and Foundations 1*, 2019, *2*, 69–186.

Bordalo, Pedro, Nicola Gennaioli, Rafael La Porta, and Andrei Shleifer, "Diagnostic expectations and stock returns," *The Journal of Finance*, 2019, *74* (6), 2839–2874.

_ , _ , Yueran Ma, and Andrei Shleifer, "Overreaction in macroeconomic expectations," *American Economic Review*, 2020, *110* (9), 2748–82.

Bénabou, Roland and Jean Tirole, "Self-Confidence and Personal Motivation," *The Quarterly Journal of Economics*, 2002, *117* (3), 871–915.

Chan, Alex, "Discrimination Against Doctors: A Field Experiment," 2022. Unpublished.

Conlon, John J, Laura Pilossoph, Matthew Wiswall, and Basit Zafar, "Labor market search with imperfect information and learning," Technical Report, National Bureau of Economic Research 2018.

Cooper, Michael and Peter Kuhn, "Behavioral job search," *Handbook of Labor, Human Resources and Population Economics*, 2020, pp. 1–22.

Cripps, Martin W, "Divisible updating," *Manuscript, UCL*, 2018.

DellaVigna, Stefano, Attila Lindner, Balázs Reizer, and Johannes F. Schmieder, "Reference-Dependent Job Search: Evidence from Hungary*," *The Quarterly Journal of Economics*, 05 2017, *132* (4), 1969–2018.

Eil, David and Justin M. Rao, "The Good News-Bad News Effect: Asymmetric Processing of Objective Information about Yourself," *American Economic Journal: Microeconomics*, 2011, *3* (2), 114–138.

Epstein, Larry G and Michel Le Breton, "Dynamically consistent beliefs must be Bayesian," *Journal of Economic theory*, 1993, *61* (1), 1–22.

_ , Jawwad Noor, and Alvaro Sandroni, "Non-bayesian learning," *The BE Journal of Theoretical Economics*, 2010, *10* (1).

Esponda, Ignacio, Emanuel Vespa, and Sevgi Yuksel, "Mental Models and Learning: The Case of Base-Rate Neglect," *American Economic Review*, 2024, *114* (3), 752–782.

Eyster, Erik and Matthew Rabin, "Cursed equilibrium," *Econometrica*, 2005, *73* (5), 1623–1672.

Faberman, R. Jason, Andreas I. Mueller, Ayşegül Şahin, and Giorgio Topa, "Job Search Behavior Among the Employed and Non-Employed," *Econometrica*, 2022, *90* (4), 1743–1779.

Gee, Laura K, "The more you know: Information effects on job application rates in a large field experiment," *Management Science*, 2018, *65* (5), 2077–2094.

Grether, David M, "Bayes rule as a descriptive model: The representativeness heuristic," *The Quarterly journal of economics*, 1980, *95* (3), 537–557.

Haaland, Ingar and Christopher Roth, "Beliefs about racial discrimination and support for pro-black policies," *Review of Economics and Statistics*, 2023, *105* (1), 40–53.

Hagmann, David and George Loewenstein, "Persuasion with motivated beliefs," in "Opinion Dynamics & Collective Decisions Workshop" 2017.

Jiang, Michelle and Kai Zen, "Information Asymmetry in Job Search," 2023.

Jin, Ginger Zhe, Michael Luca, and Daniel Martin, "Is no news (perceived as) bad news? An experimental investigation of information disclosure," *American Economic Journal: Microeconomics*, 2021, *13* (2), 141–173.

Kahneman, Daniel and Amos Tversky, "On prediction and judgement," *ORI Research monograph*, 1972, *1* (4).

Krueger, Alan B and Andreas I Mueller, "A contribution to the empirics of reservation wages," *American Economic Journal: Economic Policy*, 2016, *8* (1), 142–79.

Kudlyak, Marianna, Damba Lkhagvasuren, and Roman Sysuyev, "Systematic Job Search: New Evidence from Individual Job Application Data," Technical Report 12-03r, Federal Reserve Bank of Richmond 2014.

Li, Shengwu and Ning Neil Yu, "Context-dependent choice as explained by foraging theory," *Journal of Economic Theory*, 2018, *175*, 159–177.

McCall, John Joseph, "Economics of information and job search," *The Quarterly Journal of Economics*, 1970, pp. 113–126.

Milgrom, Paul and Chris Shannon, "Monotone comparative statics," *Econometrica: Journal of the Econometric Society*, 1994, pp. 157–180.

Milgrom, Paul R, "Good news and bad news: Representation theorems and applications," *The Bell Journal of Economics*, 1981, pp. 380–391.

Mobius, Markus and Tanya Rosenblat, "Social learning in economics," *Annu. Rev. Econ.*, 2014, *6* (1), 827–847.

Mortensen, Dale T., "Job Search, the Duration of Unemployment, and the Phillips Curve," *The American Economic Review*, 1970, *60* (5), 847–862.

Mueller, Andreas I, Johannes Spinnewijn, and Giorgio Topa, "Job seekers' perceptions and employment prospects: Heterogeneity, duration dependence, and bias," *American Economic Review*, 2021, *111* (1), 324–363.

Ortoleva, Pietro, "Modeling the change of paradigm: Non-Bayesian reactions to unexpected news," *American Economic Review*, 2012, *102* (6), 2410–2436.

Phillips, Lawrence D and Ward Edwards, "Conservatism in a simple probability inference task.," *Journal of experimental psychology*, 1966, *72* (3), 346.

Potter, Tristan, "Learning and job search dynamics during the Great Recession," *Journal of Monetary Economics*, 2021, *117*, 706–722.

Rabin, Matthew, "Inference by believers in the law of small numbers," *The Quarterly Journal of Economics*, 2002, *117* (3), 775–816.

Rosenfield, Donald B and Roy D Shapiro, "Optimal adaptive price search," *Journal of Economic Theory*, 1981, *25* (1), 1–20.

Rothschild, Michael, "Searching for the lowest price when the distribution of prices is unknown," in "Uncertainty in Economics," Elsevier, 1978, pp. 425–454.

Sadish, D, Achyuta Adhvaryu, and Anant Nyshadham, "(Mis) information and anxiety: Evidence from a randomized Covid-19 information campaign," *Journal of Development Economics*, 2021, *152*, 102699.

Spinnewijn, Johannes, "UNEMPLOYED BUT OPTIMISTIC: OPTIMAL INSURANCE DESIGN WITH BIASED BELIEFS," *Journal of the European Economic Association*, 2015, *13* (1), 130–167.

Talmain, Gabriel, "Search from an unkown distribution an explicit solution," *Journal of Economic Theory*, 1992, *57* (1), 141–157.

Thaler, Richard H. and Cass R. Sunstein, *Nudge*, New Haven, CT and London: Yale University Press, 2008.

Weitzman, Martin L, "Optimal search for the best alternative," *Econometrica: Journal of the Econometric Society*, 1979, pp. 641–654.

Wiswall, Matthew and Basit Zafar, "Determinants of college major choice: Identification using an information experiment," *The Review of Economic Studies*, 2015, *82* (2), 791–824.

Appendix A Proofs

A.1 Measurement Error

Suppose the survey respondent report a distorted prior of $\hat{\pi}_t = \pi_t + \epsilon_t$, where ϵ_t is the measurement error. In a two-state model, assuming that the measurement error term is mean zero with variance σ_ϵ^2 and uncorrelated with beliefs and past errors ($\mathbb{E}(\epsilon_t \pi_t) = \mathbb{E}(\epsilon_t \pi_{t-1}) = \mathbb{E}(\epsilon_t \epsilon_{t-1})$). To simplify our notation, we will use $\mathbb{E}$ to represent the expectation of the updated beliefs and the measurement error.

The belief movement will be larger by σ_ϵ^2

$$
\begin{aligned}
\mathbb{E}(M_{t,t+1}) &= \mathbb{E}(\pi_{t+1} + \epsilon_{t+1} - \pi_t - \epsilon_t)^2 \\
&= \mathbb{E}[(\pi_{t+1} - \pi_t)^2 + 2(\pi_{t+1} - \pi_t)(\epsilon_{t+1} - \epsilon_t) + (\epsilon_{t+1} - \epsilon_t)^2] \\
&= \mathbb{E}[(\pi_{t+1} - \pi_t)^2 + \epsilon_{t+1}^2 - 2\epsilon_{t+1}\epsilon_t + \epsilon_{t+1}^2] \\
&= \mathbb{E}[(\pi_{t+1} - \pi_t)^2] + \sigma_{\epsilon_t}^2 + \sigma_{\epsilon_{t+1}}^2
\end{aligned}
$$

The uncertainty reduction will be smaller by σ_ϵ^2

$$
\begin{aligned}
\mathbb{E}(R_{t,t+1}) &= \mathbb{E}[(\pi_t + \epsilon_t)(1 - \pi_t - \epsilon_t) - (\pi_{t+1} + \epsilon_{t+1})(1 - \pi_{t+1} - \epsilon_{t+1})] \\
&= \mathbb{E}[(\pi_t)(1 - \pi_t - \epsilon_t) + \epsilon_t(1 - \pi_t - \epsilon_t) - (\pi_{t+1})(1 - \pi_{t+1} - \epsilon_{t+1}) + \epsilon_{t+1}(1 - \pi_{t+1} - \epsilon_{t+1})] \\
&= \mathbb{E}[(\pi_t)(1 - \pi_t) - \pi_{t+1}(1 - \pi_{t+1})] - \sigma_{\epsilon_t}^2 + \sigma_{\epsilon_{t+1}}^2
\end{aligned}
$$

The expected excess belief movement statistics for a Bayesian agent with measurement error is

$$
\mathbb{E}(M_{t,t+1}) - \mathbb{E}(R_{t,t+1}) = 2\sigma_{\epsilon_t}^2
$$

If we generalize this to n states,

$$
\begin{aligned}
\mathbb{E}(M_{t,t+1}) &= \mathbb{E}\left[\sum_{i=1}^{n}(\pi_{t+1}^{i}+\epsilon_{t+1}^{i}-\pi_{t}^{t}-\epsilon_{t}^{i})^{2}\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{n}(\pi_{t+1}^{i}-\pi_{t}^{i})^{2}+2(\pi_{t+1}^{i}-\pi_{t}^{i})(\epsilon_{t+1}^{i}-\epsilon_{t}^{i})+(\epsilon_{t+1}^{i}-\epsilon_{t}^{i})^{2}\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{n}(\pi_{t+1}^{i}-\pi_{t}^{i})^{2}\right]+\sum_{i=1}^{n}(\sigma_{\epsilon_t}^{i})^{2}+(\sigma_{\epsilon_{t+1}}^{i})^{2}
\end{aligned}
$$

$$
\begin{aligned}
\mathbb{E}(R_{t,t+1}) &= \mathbb{E}\left[\sum_{i=1}^{n}(\pi_{t}^{i}+\epsilon_{t}^{i})(1-\pi_{t}^{i}-\epsilon_{t}^{i})+(\pi_{t+1}^{i}+\epsilon_{t+1}^{i})(1-\pi_{t}^{i}-\epsilon_{t+1}^{i})\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{n}(\pi_{t}^{i})(1-\pi_{t}^{i}-\epsilon_{t}^{i})+\epsilon_{t}^{i}(1-\pi_{t}^{i}-\epsilon_{t}^{i})-(\pi_{t+1}^{i})(1-\pi_{t+1}^{i}-\epsilon_{t+1}^{i})+\epsilon_{t+1}^{i}(1-\pi_{t+1}^{i}-\epsilon_{t+1}^{i})\right] \\
&= \mathbb{E}\left[\sum_{i=1}^{n}(\pi_{t}^{i})(1-\pi_{t}^{i})-\pi_{t+1}^{i}(1-\pi_{t+1}^{i})\right]+\sum_{i=1}^{n}(\sigma_{\epsilon_{t+1}}^{i})^{2}-(\sigma_{\epsilon_t}^{i})^{2}
\end{aligned}
$$

The excess belief movement is

$$
\mathbb{E}(M_{t,t+1})-\mathbb{E}(R_{t,t+1})=\sum_{i=1}^{n}2(\sigma_{\epsilon_t}^{i})^{2}
$$

Appendix B Alternative Posterior Fittings

B.1 Fitting Method Descriptions

In this section, we describe other methods we tried to fit the posteriors to the data besides fitting log normal distributions directly to the OO2b responses. The effects of these alternate fitting methods are given in the section after.

B.1.1 Log Normal Fitting, Recovered Individual Offer Distribution

The first other way we tried was to back out the offer distribution for individual wage offers rather than directly fitting the distribution implied by OO2b. In the main paper, we fit the data with a log normal distribution because observed wages have been found to be generally log normal. Since the question asks about offers the agent is most likely to accept, beliefs about these offers are likely to more closely reflect beliefs about accepted wages and thus also follow a log normal distribution. However, the best wage offer distribution is a function of the individual wage distribution and the number of offers the respondent expects. The latter is likely to be a function of the respondent's search effort. For instance, if the respondent is actively searching for a job, the worker may expect to receive more job offers, and the respondent will report a "better" best-wage distribution. This makes it challenging to deduce if the respondents are Bayesian if the search effort differs across time. To alleviate this concern, we estimate their single-wage distribution using the data and fit it as a robustness check.

We recover the single-wage offer using the following procedure. Assuming that all wages are drawn from the same single wage offer distribution independently, we let the CDF from a single wage offer be $F(w)$, the CDF of the maximum wage distribution from n offer is $F^n(w)$. From the maximum wage distribution given in "OO2b," we take the nth root to obtain the CDF of the individual wage offer. For individuals who expect to receive zero wage offers, we assume the distribution they report in this question is the distribution of the single wage offer distribution.

B.1.2 Extreme Value Fitting

We next tried to fit an extreme value distribution to the data. The motivation for this was that question OO2a2 was about the "best" offer an individual received. Our intuition is that this would be the highest (maximum) offer, on average. We therefore used SMM to fit location and scale parameters for a Gumbel distribution to the data, similar to how we fit mean and variance parameters for the log normal distributions.

B.1.3 Kernel-Fitted Posterior

Finally, we also estimated the probability density function of the posterior by considering the midpoint of each bin as representative of samples drawn from the bin. The reported bin probability from question OO2b was considered as if it were the percentage of samples drawn from the bin. The kernel density estimator was then applied to logged values of the six bin midpoints. That is, for a value x on the logged posterior distribution corresponding to question OO2b, probability density was estimated as

$$f_h(x) = \sum_{i=1}^{6} w_i K_h(x - m_i)$$

where w_i was the probability assigned to each of the six bins, m_i was the logged value of the i^{th} bin's midpoint, and h was the bandwidth. h was selected for each individual using the simulated method of moments to minimize the error between the cumulative fitted density over the bin and reported density over the bin. The Gaussian kernel was used.

Note, however, that the top bin is unbounded and that the bottom bin is much larger than the others in the survey question. To allow for assigning the top bin a point for use in the density estimation, the top bin was assumed to be from 120% to 130% of the response of question OO2a2, rather than from 120% to infinity as on the actual questionnaire. This let the top bin cover as much of the distribution as the other bins.

For the bottom bin, we tried two different specifications. In the "restricted" method, we assumed the bottom bin was 70% to 80% of the OO2a2. This meant that each bin would have equal width (10%) and that the midpoint for the bottom bin would be assigned to 75% of OO2a2, something we expected would be closer to where individuals would place the actual weight of the distribution (since 40% offers could be very low or unrealistic, depending on the value of OO2a2). In the "unrestricted" method, we assumed the bottom bin was between 0% and 80% as explicitly defined on the survey, so that the midpoint of the bottom bin would be 40% of OO2a2.

B.2 Model Fit and Estimate Results by Method

The table below gives the average fitting error (mean squared error) for each method described in the previous section, as well as for the main method used in the paper. The log normal distributions fit the best, but each of the fittings not based on kernel estimation is very close. The kernel-based methods have worse fits than the other methods by a noticeable margin. Note also that fitting errors are very similar whether we consider first updates or second updates.[28] Therefore, our ability to fit posteriors seems about the same for both second and third surveys.

Distribution	Update	MSE Average	MSE Standard Deviation
Log normal, Direct Survey Responses	1	.0052	.0104
	2	.0053	.0106
Log normal, Recovered Individual Offer	1	.0049	.0099
	2	.0050	.0101
Gumbel	1	.0054	.0127
	2	.0053	.0131
Midpoint Kernel, Restricted	1	.0730	.0427
	2	.0720	.0412
Midpoint Kernel, Unrestricted	1	.0642	.0463
	2	.0637	.0446

Table 9: Average mean squared error by data fitting method and update number.

In the next table, we see whether changing the fitting method impacts our main results. We find large normalized statistics for all of the non-kernel methods, and the other statistics seem fairly close to each other, regardless of fitting method. The non-kernel methods have much lower normalized statistics, but it should be noted that they also have much worse fit. Therefore, these results suggest that our results are not very sensitive to which distribution we fit. Finally, we note that the main results we get are very similar whether we use direct survey responses from OO2b or recover the individual wage distribution. Thus, our result of non-Bayesianism appears robust to concerns over the number of offers expected by subjects in our sample.

[28] First updates include individuals without second updates.

Statistic	Log Normal, Direct Survey Responses	Log Normal, Recovered Individual Offer	Gumbel	Midpoint Kernel, Restricted	Midpoint Kernel, Unrestricted
$\overline{m}$	.9341	.9316	.9399	.4237	.4719
	(.0193)	(.0215)	(.0203)	(.0055)	(.0067)
$\overline{r}$	.1805	.1657	.1895	.3131	.2977
	(.0079)	(.0077)	(.0081)	(.0047)	(.0055)
X	.7536	.7658	.7504	.1106	.1742
	(.0207)	(.0222)	(.0209)	(.0101)	(.0116)
X_{norm}	5.1744	5.6205	4.9604	1.3534	1.5852
	(.2488)	(.2818)	(.2268)	(.0374)	(.0494)

Table 10: excess belief movement test statistics using alternate distribution fittings. Standard errors clustered by state are in parentheses.