Learning in Job Search: Insights on Belief Updating

Sebastian Brown, Kenneth Chan

September 3, 2023

Abstract

We use a nationally-representative survey (the labor supplement of the Survey of Consumer Expectations) to study individuals' behavioral biases in learning about their wage distributions. Using a recently-developed test for Bayesianism (Augenblick and Rabin, 2021), we find strong evidence of non-Bayesian learning. Specifically, among individuals who receive wage offers, we find an average movement in beliefs that is roughly 543% of the reduction in their beliefs' uncertainty, a result consistent with base-rate neglect and/or overreaction to signals. We further examine the heterogeneity in agents' learning by identifying specific non-Bayesian (and other) patterns, finding additional individual patterns that could be explained by Gambler's Fallacy and optimistic reasoning.

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 strategy 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] 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. Consequently, there has been a lack of research on how job seekers learn and update their beliefs in light of new information. Examining the process of learning and belief updating can provide valuable insights into the effectiveness of information provision policies. Our paper hopes to fill this gap and provide some insights into how job searchers incorporate new information into existing beliefs.

Providing information and correcting false beliefs is a potent policy tool for changing people's beliefs and behavior. In a series of recent field experiments, providing information has been shown to help students make more informed decisions about their college majors

[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.

(Wiswall and Zafar, 2015), improve people's knowledge about COVID-19 (Sadish et al., 2021), eliminate statistical 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.[2] The author argues this might be evidence that the information treatment caused 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. 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 beliefs and study their effects, 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 are representative and free of the cognitive influences of an artificial lab experiment environment, they also should serve as a more externally-relevant measure of how individuals process information in a true labor market context than could be collected in a controlled experiment environment.

In our first analysis, we employ a recently-developed test for Bayesian updating in beliefs from Augenblick and Rabin (2021). There are two statistics required for this test, the belief movement which is the squared difference of changes in belief and uncertainty reduction which is measured by how much variance has been reduced from belief updating. If the agent is Bayesian, 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 people overreact to signals and exhibit base rate neglect. Our remaining analysis attempts to examine specific individual behaviors that contribute to this aggregate result.

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 3 outlines a simple theoretical model of labor market updating and briefly describes the general theory of the Augenblick-Rabin test for Bayesianism. Section 5 explains our empirical strategy, detailing the steps that we took to apply the Augenblick-Rabin test to this dataset. Section 6 then lists our main results. Finally, section 7 concludes and provides suggestions for future work.

2 Literature Review

Our work is related to the literature on search theoretic models with unknown distributions. In these models, the underlying distribution from which offers are drawn, while fixed,

[2]We interpret this coaching program as a form of providing information to help the job searcher.

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 expectation 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 (Rothschild, 1978; Rosenfield and Shapiro, 1981; Talmain, 1992). This is because the Dirichlet distribution has a conjugacy property[4] and it helps them to solve for the model. Recently, Li and Yu (2018) deviated from the Dirichlet 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 allowed them to establish first-order stochastic dominance between distributions and posterior beliefs, which allows them to perform monotone comparative statics (Milgrom and Shannon, 1994) on the reservation value.

While these search models assume people are updating their beliefs in a Bayesian manner, behavioral economists have shown that people are not Bayesian. Some of the more well-known biases include base rate neglect (Kahneman and Tversky, 1972), 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 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 et al., 2018; Bordalo et al., 2019, 2020; Augenblick and Rabin, 2021).

There is a small, but growing literature on empirical evidence of learning in labor market beliefs. We cite a few key papers here. A first paper is Mueller et al. (2021), which uses data from the Survey of Consumer Expectations and Survey of Unemployed Workers in New Jersey. They found that unemployed workers do not sufficiently adjust their beliefs of the probability of finding an acceptable job downward over the unemployment spell, leading the long-term unemployed to display an optimistic bias in job finding. This is evidence for a lack of updating in the labor market, although it relates only to job finding, which is a function of both wage expectations and arrival rate expectations, and does not distinguish between these two potential sources of the studied beliefs.

Another relevant paper is Kudlyak et al. (2014), which studies panel data from a job posting website. Its authors find that job seekers first apply to jobs that are the first choices of other job seekers with similar education levels as themselves and that they then apply to the top choices of workers with less education than them within four to five weeks of

[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] For any signal structure, the posterior belief and the prior beliefs are from the same family of distribution.

continued unemployment. They argue that this is evidence that searching workers learn to adjust their expectations downward over the unemployment spell.

Two recent papers on labor market learning stand out as the most similar to ours. The first is Potter (2021), which fits a model of Bayesian updating in offer arrival rates to data from the Great Recession. We build on Potter's work in three key ways. First, our data allow us to directly observe subjects' beliefs, whereas Potter's model must infer these beliefs only from changes in time spent searching for work. Second, we focus on identifying key (non-)Bayesian learning patterns in the data, whereas Potter focuses only on showing whether there was evidence of learning at all. Finally, while Potter's data highlight learning in the Great Recession, our data allow us to examine learning over a longer and more recent time span, specifically from March 2015 until the present.[5]

The other paper is Conlon et al. (2018), which uses the Survey of Consumer Expectations data to study learning and heterogeneity in beliefs over wage offers. They find that people do not have the same systematically optimistic bias over wages as they do over the offer arrival rates. They also calibrate a simple model of learning over wages, and find that it better explains the data than a model without learning.[6] Their model finds that, on average, wage expectations increase by \$0.47 for every one dollar increase in observed wage offer. They find that this indicates that people update too much on average due to recent signals, since according to their analysis, Bayesian updating would instead imply increasing expectations by \$0.16 on average for every one dollar increase in observed wage offer. We build on this paper by conducting additional *individual-level* analyses that allow us to uncover additional details about labor market learning. Specifically, we measure and discuss the prevalence of various learning patterns in the data. Our work also allows us to test for Bayesianism while making fewer assumptions about an agent's individual wage distribution.[7]

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)).

[5] The survey we use started in December 2013, but it was not until March 2015 that all of the questions we use appeared in their current form in the data. Our dataset also only goes until March 2020, as this is the most recent data publicly available. However, the survey has continued to run beyond 2020, with more data is released periodically.

[6] They choose to focus on wage offers rather than arrival rates, deferring to analysis by an older version of Potter (2021) for a model of updating in arrival rate beliefs.

[7] Conlon et al. (2018) address non-Bayesian behavior in one of their appendices. While they do utilize the same primary survey question as we do for studying Bayesianism, they assume that all agents with a given set of covariates receive offers from one of two distributions - one for each of two unobservable "types" of worker. They also assume that all individuals' distributions share the same variance, which varies only on whether an individual has a college degree. By contrast, our martingale tests for Bayesianism assume individuals' wage beliefs are log-normal (like Conlon et al. (2018)), but don't require an assumption of shared variance, or an assumption that distributions only vary systematically in two ways outside of observed covariates. We also make no assumptions beyond log-normality when we identify counts of non-Bayesian patterns (overshooting and moving in the wrong direction) in individual behavior in figure 5.

3 Data

3.1 Data Source Description

Our primary data source is 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 remain in the survey for up to a year, and are replaced on a rolling basis. Therefore, we are able to observe the same subject for up to a maximum of three times for the labor force questions and twelve times for the core questions. 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 (see, for example, Potter (2021), Spinnewijn (2015)). Another benefit is its panel structure, which, while short, still does allow for within-subject observations of the evolution in beliefs over time. A third benefit is that it features mostly employed individuals, allowing study of a previously-understudied class of workers (as highlighted in Faberman et al. (2022)).

This last benefit, however, comes at the cost of including fewer unemployed individuals. While the SCE has been included in analyses of unemployed workers (e.g. Mueller et al. (2021); Conlon et al. (2018); Faberman et al. (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, only 3.9% of our final sample is unemployed.[8]

3.2 Survey Question Description

Our analysis uses a few survey questions which merit more detailed explanations. First, we use the distributional question labeled "OO2b" as shown in Figure 1 to measure the mean and variance of an individual's prior.

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?*" Therefore, the categories of this question may be taken to mean, "What is the probability your maximum offer will be 80% or less of your guess of the average maximum offer?", "What is the probability your maximum offer will be 80% to 90% of your guess of the average maximum offer?", and so on. Since the question is about the maximum wage offer, not the average, we must take steps to recover the single wage offer distribution from this question. This process is outlined in section 5.1.

Hereafter we will refer to wage "bins," by which we mean the ranges of wages implied by the survey questions. These range from bin 1, the lowest bin, consisting of all wages 80%

[8] It is therefore not surprising that Mueller et al. (2021) augments the SCE by including a second dataset, the Survey of Unemployed Workers in New Jersey, which as its name suggests, focuses entirely on unemployed individuals.

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)

Figure 1: Survey Question OO2b from SCE

times the expected best offer and below, to bin 6, the highest bin, consisting of all wages 120% times the expected best offer and above. A challenge working with this question is that the bins of wages are redefined with each new survey taken, including within subjects. The size of the wage bins is determined by their response to "OO2a2," which differs across every response to the survey. To make the responses comparable between an individual's consecutive surveys, we fit a log-normal distribution to each individual's responses, after making the adjustments discussed in section 5.1. This fitting also gives us the mean and variance used in our principal analysis for each individual distribution. For an explanation of this fitting process, see section 5.2.

While our main analysis uses the mean from the fitted log-normal distribution, we also use the question "OO2a" to measure an individual's expectation of the average wage for a given period to graph the sample distribution of expected offers in figure 1 in the next section. 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 our main analysis since unlike OO2a2, it does not have a corresponding question (OO2b) to elicit a probability distribution. However, its responses do allow us to see how closely our fitted means follow the directly-elicited means, as highlighted in the next section.

Finally, we also use questions about the wages and number of offers received, as well as a question on the number of offers expected. The survey asks individuals to report their three best 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. Further explanation of how we incorporate the expected number of offers to recover the individual wage distribution is given in section 5.1.

3.3 Data Structure and Summary Statistics

Our final dataset is organized such that each observation contains information from two consecutive labor market surveys for a single individual. While most individuals in the dataset appear only twice, for those that appear three times, there will therefore be two separate observations: one for the update from period 1 to period 2, and one for the update from period 2 to period 3.

The composition of our final dataset by employment status (defined as the employment status of the individual in each before-after pair's first period) and whether they received at least one wage offer is given Table 1:

Description	Count	
	At least one reported wage offer	No reported wage offers
Total observations	978	3,890
Unique individuals	847	2,883
Number unemployed	59	131
Number employed	804	3,151
Number not in labor force	104	562
Missing employment status	11	46
Data Range	3/2015-3/2020	

Table 1: Dataset composition. Observations consist of information from two consecutive surveys for a given individual.

In the following graphs, we report on the expected and realized wage offers, both in the wage amount and number of 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 1 below, we see that most of the responses are close to the average offers all the survey respondent has reported receiving. This suggests that at the aggregate level, people are not randomly responding to the survey questions. We can also see that the aggregate log distributions for expected offers and reported offers seem to follow a bell curve, consistent with our assumptions of log normality for individual distributions.

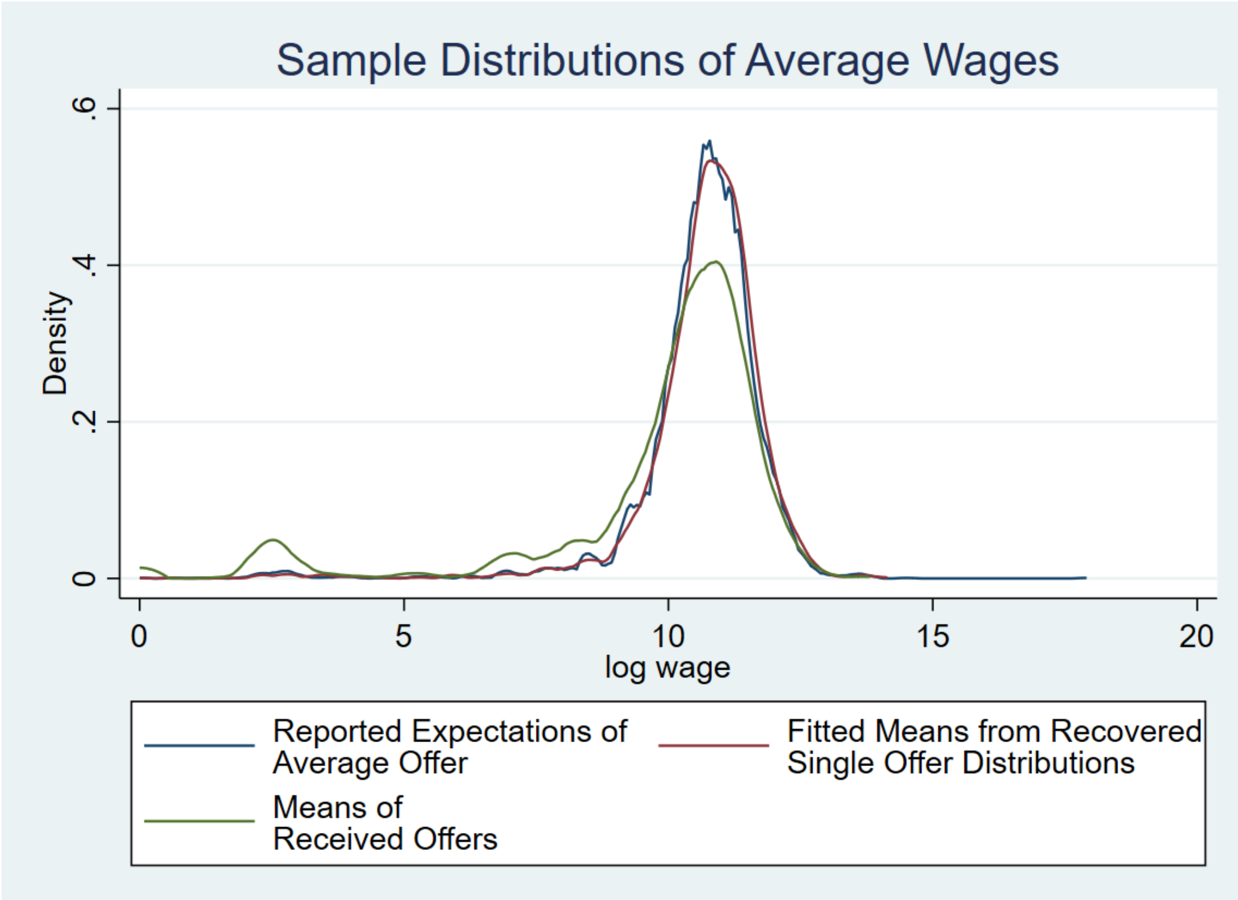

Figure 2: Distribution of reported expected wages, fitted expected wages, and received wages. Reported expected wages and fitted expected wages include all individuals, including those who did not receive an offer. Fitted expected wages are the means of the fitted distributions using the process discussed in section 5.2.

Figure 2 (below) shows that individuals tend to overestimate how many offers they will receive. This graph groups observations with more than ten offers into one category, of which there are very few. In addition to the information on the number of offers presented in the graph, we also note that the number of individuals who are unable to report all received offers is relatively small: Of the 978 observations in our final dataset with realized offers, 601 reported one offer, 230 reported two offers, and 147 reported three offers. Of those 147, only 70 reported receiving more than three offers (the maximum number of wages that can be reported).

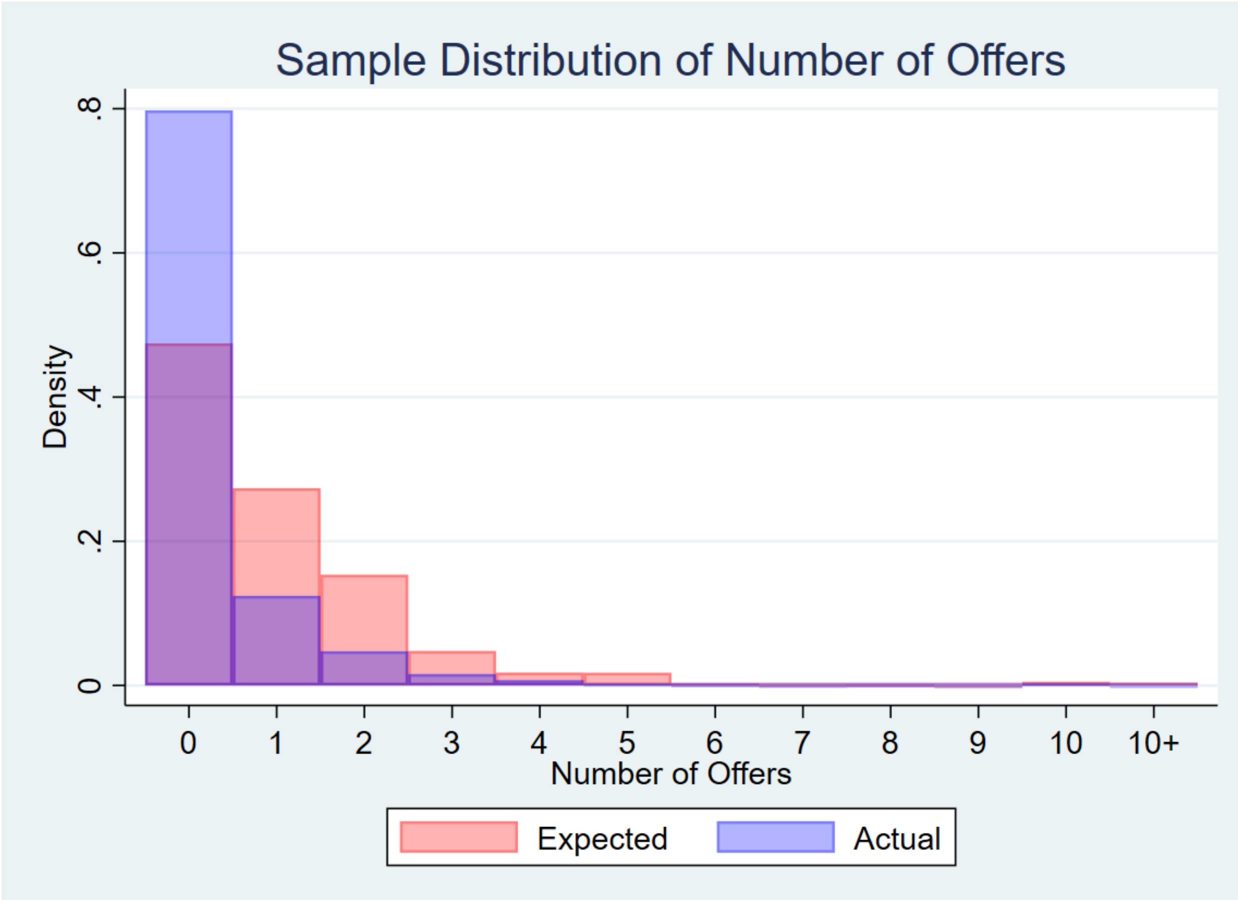

Figure 3: Distribution of expected and realized wage offers. Expected offers include the expectations of all individuals, whether or not they received an offer. There were only 19 observations in the final sample for which individuals expected more than ten offers, and only one observation in the sample reported receiving more than ten offers.

4 Theoretical Model

We will now write down a model of belief dynamics in the job search setting we are studying. We will follow Li and Yu's (2018) method of modeling uncertainty. Unlike their model that is interested in obtaining comparative statics across different signal realizations, we are primarily interested in the belief dynamics, hence we will not need to assume MLRP in the set of distributions. This model also differs from Augenblick and Rabin (2021) where the state of the world is a probability distribution (not an outcome) and the reported beliefs is a mixture distribution. We will need to first verify that for Bayesian agents, the reported beliefs have the martingale property.

Suppose there is an agent who is searching for a job believes there is a set of possible wage distributions, $\mathcal{F}$, that he is drawing his wages from.[9] We will index the density functions in

[9]Another way to model this would be to assume that the agent has a Dirichlet prior over the wage distribution. But the main drawback of this approach is we have to assume the wage realization is finite.

$\mathcal{F}$, by a parameter θ taking values in the 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 and news related to the labor market. 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.

For a Bayesian agent, the agent will update his belief using Bayes' rule:

$$g_{t+1}(\theta|x) = \frac{g_t(\theta)p(x_{t+1}|\theta)}{\int_{\theta' \in \Theta} g_t(\theta')p(x_{t+1}|\theta')} \tag{1}$$

With Bayesian updating, we expect the belief updating rule to have the martingale property, $\mathbb{E}(g_{t+1}(\theta|x_{t+1})|g_t(\theta)) = g_t(\theta)$.

In the survey questionnaire "OO2b," the respondent provided a probability distribution about the best wage offer they will receive. We can recover the single wage offer following the procedure in section 5.1, which gives us data about the respondent's probability distribution about the wage offer they will receive as probabilities over several wage bins. We denote the probability of the wages being in bin $i \in \{1, 2, \dots, n\}$ at time t as π_t^i. [10]

Under the model we have constructed, we will partition the wages into n wage bins $\{[a_0, a_1), [a_1, a_2), \dots [a_n, a_{n+1})\}$. 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

$$\pi_t^i = \int_{\theta' \in \Theta} g_t(\theta') \int_{a_{i-1}}^{a_i} f(w|\theta')dw \tag{2}$$

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 over the agent's beliefs about the distribution he is drawing from. The probabilities π^i will follow the martingale property as well.[11]

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

If the respondent is Bayesian, the reported probabilities in the survey will have the martingale property. We will be testing the martingale property to determine if people are updating their beliefs in a Bayesian manner.

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

[11] 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.

5 Empirical Strategy

There are two main challenges with this data. Firstly, the survey measured the probability distribution of the best wage offer, not the individual wage offer. The distribution of the best wage offer is a function of the arrival rate of job offers which may change over time as well. As such, when the respondent's belief changes, we can't attribute it completely to the respondent updating his belief about the wage distribution. Secondly, the bins of the wage offer differ across every individual and time period. 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 time periods, the wage bins will have different intervals across each response. To address these problems we first have to recover the individual wage distribution and then estimate the individual wage distribution.

5.1 Recovering the Single Wage Offer from the Maximum

There are two reasons why we want to work with a single-wage offer distribution instead of the best wage offer distribution directly. Firstly, 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 search effort differs across time. To alleviate this concern, we estimate their single-wage distribution using the data.

Secondly, given the structure of the data, we have to make some parametric assumptions to fit a distribution of the data to the responses and make them comparable across time. While it is well known that the single wage distribution can be modeled well by a log-normal distribution (e.g. Clementi and Gallegati (2005)), it is challenging to justify the family of distributions to fit the best (maximum) wage offer distribution.

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.

5.2 Distribution Fitting

We use the Simulated Method of Moments (SMM) to fit log-normal distributions to the five points on individuals' prior cumulative density functions recorded in question OO2b. The horizontal values of these five points are the inverse CDF values located at the wage bin thresholds (e.g. 80% of OO2a2, 90% of OO2a2, and so on), and the vertical values are survey responses of OO2b with the adjustments described in 5.1. We find the mean and standard deviation which minimize the horizontal distance in log wages (using the L2 norm) between points on the simulated CDFs with the same CDF values as the five points and the five points themselves. An example is given in figure 3, where the blue points are points on the CDF implied by survey responses, and the orange points are points that minimize the horizontal distance between the orange curve and blue points.

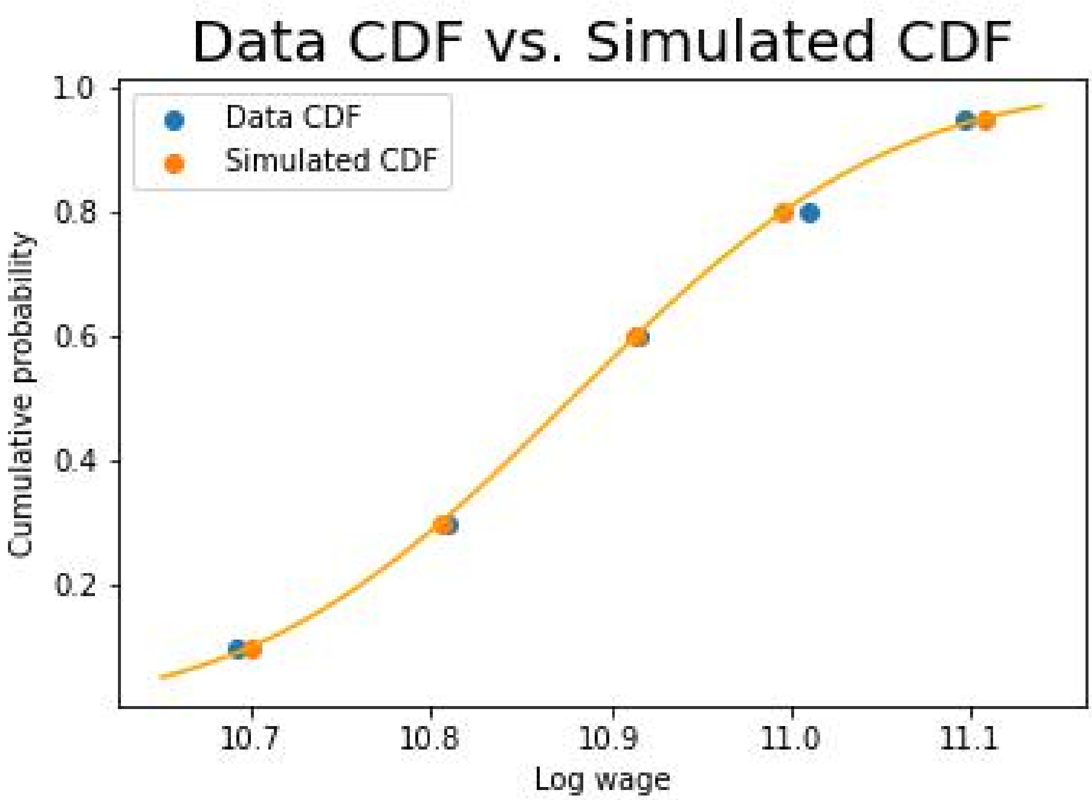

Figure 4: Example fitted distribution.

This is done for both periods in an observational pair, which allows us to estimate the probability over bins using the period 1 OO2b responses for period 2; this allows us to overcome the obstacle of period 2 using different wage bins than period 1. Note here than we use the CDF to do the fitting rather than the PDF. This is because there is no bijection between a log-normal PDF and its inverse and because the survey responses are binned rather than for individual wages.

5.3 Martingale Tests

Bayesian updating is one of the updating rules that must satisfy the martingale property.[12] 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.

[12] A non-Bayesian updating rule that satisfies the martingale property is the Epstein, Noor and Sandroni (2010) model where the updated belief is an affine transformation of the Bayesian posterior with the prior.

The martingale property can be tested directly, and we will be using a test proposed by Augenblick and Rabin (2021) to do so. There are alternative tests, such as testing if the updated beliefs average out to be the prior belief. The main benefit of the Augenblick and Rabin (2021) test is that it is closely related to some of the more prominent belief-updating biases, such as base-rate neglect and underreaction to signals. This will give us insights into the type of belief updating bias the agent is exhibiting. Moreover, we do not need to know the Bayesian beliefs or the conditional probabilities to determine if the survey respondents are Bayesian. This allows us to make minimal assumptions for the analysis.

The test involves computing two statistics: (1) belief movement, $m_{1,2}$, and (2) uncertainty reduction, $r_{1,2}$, as shown in equations 3 and 4 respectively. Let π_t^i denote the probability assigned to bin i at time t. With h bins, the two statistics are defined as

$$m_{1,2} \equiv \sum_{i=1}^{h} (\pi_2^i - \pi_1^i)^2 \tag{3}$$

$$r_{1,2} \equiv \sum_{i=1}^{h} \pi_1^i(1 - \pi_1^i) - \pi_2^i(1 - \pi_2^i) \tag{4}$$

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. 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 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. 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.[13]

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 single updated belief profile based on the signals the respondent observed in the last four months, we can only determine if the population is Bayesian at the aggregate level. Given that there are n observations, we compute the average belief movement as $\overline{m}_{1,2} \equiv \frac{1}{n}\sum_{j=1}^{n} m_{1,2}^j$ and the average uncertainty reduction as $\overline{r}_{1,2} \equiv \frac{1}{n}\sum_{j=1}^{n} r_{1,2}^j$. We then compute the average excess belief movement statistic:

$$X = \overline{m}_{1,2} - \overline{r}_{1,2} \tag{5}$$

where "excess" refers to the amount of movement exceeding reduction.

The standard error for this test is

$$s_{1,2} = \sqrt{\frac{1}{n}\sum \left(m_{1,2}^j - r_{1,2}^j - (\overline{m}_{1,2} - \overline{r}_{1,2})\right)^2} \tag{6}$$

[13] To show this, we have to expand the terms and apply the law of iterated expectations. With the martingale property, the terms will cancel out nicely.

Under the null hypothesis that the population is Bayesian, we expect $X = 0$. With the central limit theorem, we can compute the Z-statistic that converges to a standard normal distribution:

$$Z \equiv \frac{\sqrt{n}}{s_{1,2}}(\overline{m}_{1,2} - \overline{r}_{1,2}) \xrightarrow{n\to\infty} N(0,1) \tag{7}$$

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 address this concern, 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 $X_{norm} = 1$.

$$X_{norm} = \frac{\overline{m}_{1,2}}{\overline{r}_{1,2}} = \frac{X}{\overline{r}_{1,2}} + 1 \tag{8}$$

As proposed by Augenblick and Rabin (2021), we will report the excess belief movement, the Z statistics as well as the normalized belief movement in our results.

6 Results

6.1 Martingale Test Statistics

We report the martingale test statistics in table 2 below. Our main result is in column 4, where a statistic of $X_{norm} = 5.4344$ indicates an average movement among those with offers that is 5.4344 times the average reduction among the same observations.

For Bayesian updating under the Gaussian framework, the beliefs have two distinct characteristics. Firstly, the expected wage offer from the posterior wage distribution has to be a convex combination between the expected wage offer from the prior distribution and the wages that were observed. Secondly, the variance of the wage distribution should be non-increasing. Bayesian updating in this environment cannot rationalize any of these behaviors. Column 5 drops observations that exhibit these behaviors.[14] Excess movement decreases somewhat (though it remains very large) moving from column 4 to 5, as might be expected given that one of the categories of non-Bayesian behavior ("overshooting") involves moving more than can be justified by any observed wage.

[14] The measurement of these behaviors is discussed further in section 6.3.

Statistic	(1) All Individuals	(2) Without Offers	(3) With Offers Only	(4) With Offers Excluding Identified Non-Bayesians
X	.5314	.5161	.5924	.5056
Z^1	58.4621	51.4491	27.8964	16.0148
X_{norm}	5.2194	5.1611	5.4344	4.4187
Observations	4,866	3,888	978	454

Table 2: Excess movement statistics.

[1] Here, we actually do a paired t test, not a Z test, since we do not know the population standard deviation. However, we use "Z" to remain consistent with the notation in Augenblick and Rabin (2021)

However, it is unclear from this analysis how much of our main result can be attributed directly to the effect of wages: Column 3 also shows a similarly large effect, even among individuals with no reported wage offers.

Figure 4 below provides an alternative look into the martingale test results. The provided box plot compares summary statistics for the individual movement, uncertainty reduction, and excess movement measures of three different groups of individuals. The first group ("Possibly Bayesian") includes those with reported offers not identified as explicitly non-Bayesian by the process outlined in section 6.3, as well as those without realized wage offers. The second group ("Non-Bayesian") includes those with reported offers identified as non-Bayesian in section 6.3. The third group ("All") includes all individuals in the sample, with or without realized wage offers.

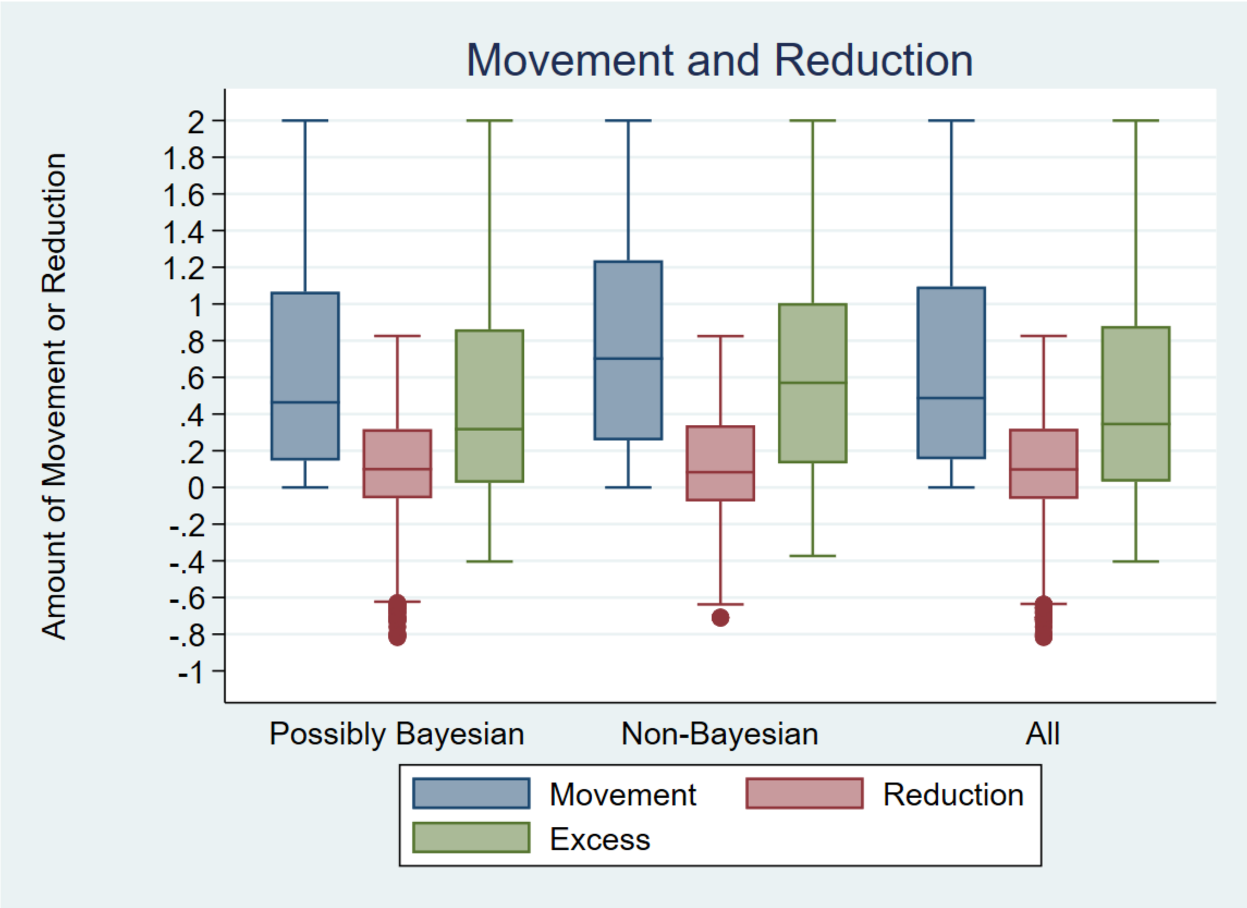

Figure 5: The "Non-Bayesian" category includes individuals whose updated mean overshot all observed wage offers or moved in the wrong direction. The "Possibly Bayesian" and "All" categories include individuals without realized wage offers in addition to those outside of the "Non-Bayesian" category with offers.

While the "Non-Bayesian" category has a slightly higher movement box and a slightly shorter (though not clearly lower) reduction box than the "Possibly Bayesian" category, the results for movement and reduction remain fairly similar across all three categories, again adding reason to caution our interpretation of results as coming from learning about received wages (while all in the "Non-Bayesian" category received wage offers, most in the "Possibly Bayesian" category did not).

6.1.1 Possible Updating Rules

Given the result that people are non-Bayesian, what are some updating rules that can generate this data? The first model is Grether (1980) whose updating rule generalized version of Bayes' rule that accommodate various belief-updating biases. Using the same notation as our theoretical model, the updating can be written as

$$g_{t+1}^{bias}(\theta|x) = \frac{g_t(\theta)^a p(x_{t+1}|\theta)^b}{\int_{\theta' \in \Theta} g_t(\theta')^a p(x_{t+1}|\theta')^b} \tag{9}$$

$a > 0$ is the weight the agent places on the signal; when $a < 1$, we have base rate neglect and when $a > 1$, we have confirmation bias. $b > 0$ 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 updating rule violates the martingale property and if we set $a > 1$ and/or $b < 1$, 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.

It is also important to note that there are ways to model over-updating without excess belief movement. For instance, consider the Epstein, Noor and Sandroni (2010) model where the non-Bayesian belief is an affine transformation of the Bayesian belief and the prior.[15]

$$g_{t+1}^{bias}(\theta|x) = (1-\lambda)g_t(\theta) + \lambda g_{t+1}^{bayes}(\theta|x) \tag{10}$$

where $\lambda > 0$ is the parameter that determines the degree of over and underreaction. When $\lambda < 1$ we have underreaction and $\lambda > 1$ we have overreaction.[16]. It is also easy to verify that the martingale property holds for this updating rule. With this updating rule, we expect the belief movement to be equal to the uncertainty reduction. Based on the result we obtain, we can also reject the Epstein, Noor and Sandroni (2010) model as well.

Another possible class of model that can potentially explain excess belief movements is multiple prior models where the agent can switch between prior beliefs (Ortoleva, 2012; Ba, 2022). These models are typically used in misspecified learning models, where an agent who incorrectly believes that some states are impossible can correct their mistaken beliefs. Under Bayesian updating, once zero probability weight is assigned to a state, there exists no signal that can cause the agent to assign positive probability weight in the posterior belief to this state.

To see how these models can result in excess belief movement, consider an agent who is predicting a binary state outcome $\{\theta_1, \theta_2\}$ and has a prior belief of $\pi_0^a = 0.25$ but entertains an alternative prior belief of $\pi_0^b = 0.75$. Let there be two possible signal realizations and the conditional probability is as follows $p(x_1|\theta_1) = p(x_2|\theta_2) = 0.75$. If the agent observes x_1, he updates his belief using π_0^a with Bayes' rule, if the agent observes x_1 he switches his prior to π_0^b and updates it with Bayes' rule. Under the initial prior, signal x_1 will be observed with $\frac{3}{8}$ probability. The expected excess belief movement can be written as $\mathbb{E}[(2\pi_t - 1)(\pi_t - \pi_{t+1})]$, and if we compute the excess belief movement statistics we have $X = 0.175$.

6.2 Regression Analysis

We also regress the expected wage offer and the actual wage offer they receive to examine the relationship between these two variables, which is presented in table 3.

[15]This is a generalization of cursed inference (Eyster and Rabin, 2005) where the biased belief is a convex combination of the prior and bayesian belief.

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

Variable	(1)	(2)	(3)
Maximum Wage Offer	0.677***	0.657***	0.609***
	(0.0967)	(0.101)	(0.108)
White		-1,578	1,619
		(1,571)	(2,231)
Female		-7,821**	-7,782*
		(3,002)	(4,030)
Some College		17,046**	17,387**
		(7,912)	(7,150)
Constant	20,116***	-6,579	40,597**
	(4,641)	(10,186)	(17,556)
Observations	978	978	834
R-squared	0.696	0.706	0.731

Table 3: Wage offers and wage offer expectations.
Dependent variable is updated average wage offer expectation (in dollars, not logs). Robust standard errors are in parentheses. Column 3 includes fixed effect dummies for industry, state, and year. Three asterisks represent significance at the 1% level.

The results of the regression show a significant positive relationship between the maximum wage offer and expectations of the average wage, as expected given the previous literature. This is also consistent with the theory, where receiving a higher wage offer should cause the expected wage offer to increase. The inclusion of fixed effects for year, state, and industry improve the fit of the data but does not change the coefficients very much besides the constant.

6.3 Type Analysis

In addition to testing for non-Bayesian behavior as a whole, we also identify subjects exhibiting individual patterns that cannot be justified by Bayesian updating. Specifically, we detect individuals whose mean updated towards but beyond every observed wage offer ("overshooting") and whose mean moved in the opposite direction of any wage offer ("wrong direction"). For individuals who received at least one offer, we found overshooting in 276 (28.22%) of the updates and moving in the wrong direction in 248 (25.36%) of the updates. These observations are those considered by our analysis to display clearly-observable non-Bayesian behavior, and are given their own category in table 2 and figure 4.

One explanation for the observed non-Bayesian patterns is the Gambler's Fallacy. As applied to this case, this would have workers retain or even increase their wage expectations upon receiving poor offers, with the belief that a good offer will be more likely to come next due to having already observed several bad offers. This would be consistent with the existence of updating in the wrong direction, but would not help explain overshooting. A possible explanation for overshooting, to the extent non-Bayesian movement is positive, is motivated reasoning.Under this explanation, individuals witnessing a wage offer above their prior distribution might go too far because they really want there to be higher offers, leading them to hope that there are even better offers available than those they have observed.

In addition to these two behaviors, we also measure whether a subject had such a major shift in distributions that it could be considered switching to a different distribution entirely. For this measure, we consider an agent to have switched distributions as an update if either:

1. The upper bound of the lowest wage bin of the prior distribution is above the lower bound of the highest wage bin of the updated distribution, or

2. The lower bound of the highest wage bin of the prior distribution is below the upper bound of the lowest wage bin of the updated distribution.

In the first case, the updated distribution can be considered to have moved almost entirely to the left of the prior distribution, and in the second case, the updated distribution can be considered to have moved almost entirely to the right of the prior distribution.

We found that there were 958 (19.68%) switched distributions in our sample. Since this measure relies only on reported beliefs, not on receiving an offer, it is not necessarily restricted to those who receive offers. When restricting to only those who could be updating due to received wages, we found that there were 251 (25.66%) switched distributions.

Finally, we measure whether all of a subject's offers for a period are in the outer parts of their fitted belief distribution (both for the prior and updated distributions). Similar to the switching measure, an offer is considered in the "outer" region of a distribution if it is in the lowest wage bin or the highest wage bin using question OO2b. While this phenomenon is not inherently non-Bayesian, it could indicate that individuals do not consider such offers informative or legitimate work opportunities. A high presence of such situations could indicate that a lack of updating is not necessarily non-Bayesian behavior (since reported offers are less likely to be informative signals), even though standard Bayesian metrics applied in such cases would indicate the opposite (since observed offer wages would be far away from the mean of the prior). We found that receiving offers "outside" of one's prior belief distribution was very common, with it happening for 508 (51.94%) of the updates. We also found that this number decreased to 392 (40.08%) when calculating the same measure for the updated distributions, which is evidence of there being at least some learning.

The results of this section are summarized in figure 5 below.

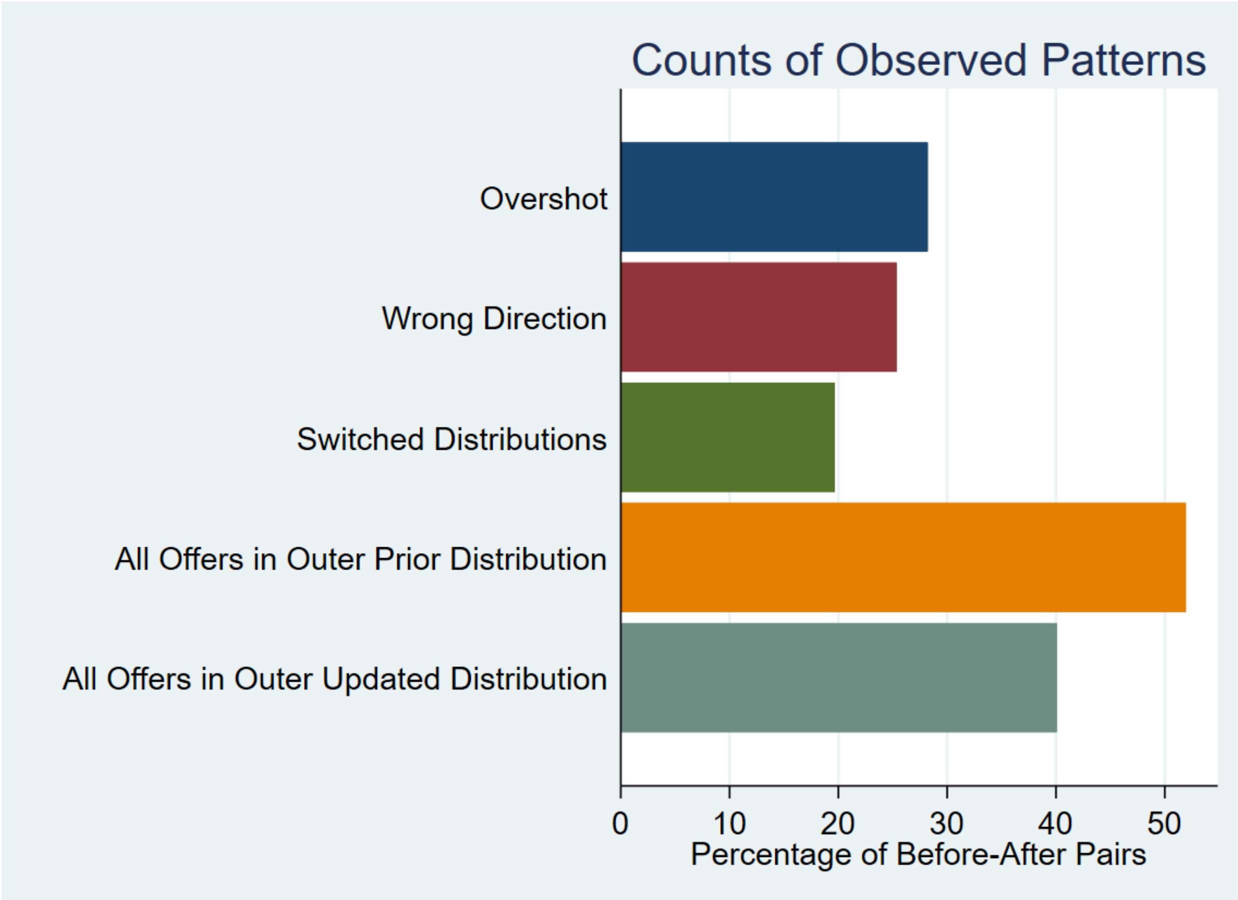

Figure 6: Counts of specific patterns in data as percentages. For "Switched Distributions," the percentage given is for the entire sample. For all other bars, the percentage given is for all observations with at least one reported offer.

7 Robustness Tests

7.1 Measurement Error

7.1.1 Theory

The normalized excess belief movement statistics is significantly larger than the estimates Augenblick and Rabin (2021) got in their empirical application in their paper. To further understand why we have such a large estimate we perform a Monte Carlo Simulation where we allow for measurement error.

Consider that 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 the excess belief movement will be equal to $\sum_{i=1}^{n} 2\sigma^2_{\epsilon^i_t}$.[17] With measurement error, the excess belief movement to be equal to the total variance of the measurement error in the prior belief multiplied by 2.

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. Moreover, given that we have an excess belief movement of more than 0.5, this suggests that the total variance of the measurement error has to be at least 0.25 to explain the result we are observing.

7.1.2 Monte Carlo Simulation

We 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 as closely as possible. Since there are 6 wage bins and most subjects responded to the survey twice, we have 6 states and with only 2 periods we will have only 1 signal realization and 1 update. With 4866 respondents in our dataset, we generated 5000 pairs of priors and posteriors in each simulation.

Since our statistics of interest only depends on measurement error in the prior, we only introduce measurement errors to the prior beliefs. We draw our simulated prior data from a Dirichlet distribution that is centered around a uniform prior distribution.[18] 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.

With our assumptions on the measurement error as discussed in the earlier subsection, we can show that the belief movement will be larger by $\sum_{i=1}^{n} \sigma^2_{\epsilon^i_t}$ and the uncertainty reduction is smaller by $\sum_{i=1}^{n} \sigma^2_{\epsilon^i_t}$ compared to the scenario without measurement.[19] 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 is 0.392, and we try to match this statistics in our calibration exercise.

It is important to note that, there is infinitely many possible combinations of prior beliefs and distribution 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 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 of 0.392. One of the possible posterior beliefs that is selected in this calibration exercise is $(0.75, 0.2, 0.05, 0, 0, 0)$. To obtain the full distribution of posterior beliefs, we can permutate the order of the probabilities and each of these permutations will be realized with equal probabilities.

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

[18] 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.

[19] This is only true if there is only measurement error in the prior.

Finally, the variance of the error term is not an intuitive measure of the amount of measurement error. To quantify measurement error in a more intuitive manner, 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 prior.

We are interested in the normalized excess movement and Δ, this process is simulated 10,000 times. In Table 4, we report the average statistics computed from the 10,000 Monte Carlo simulation and the interval that contains 95% of our simulation result.

We see that in order to obtain a excess belief movement from our dataset assuming the respondents are Bayesian, we need the survey respondents to misreport their prior beliefs by about 98 percentage point. The measurement error required to rationalize the behavior is too large and is unlikely to explain the result we have obtained.

Statistic	(1) Uniform Prior with Matched Belief Movement	(2) Balls and Urn with Matched Normalized Movement
X	0.5275 [0.5158, 0.5388]	0.1054 [0.1029, 0.1079]
X_{norm}	5.2349 [5.0219, 5.4579]	5.3463 [5.1014, 5.6090]
Δ	0.9831 [0.9765, 0.9899]	0.4477 [0.4437, 0.4516]

Table 4: Monte Carlo Simulation Results

In column 2 we also report the calibration results from a balls and urn thought experiment where we attempt to match the normalized belief movement. Suppose there are 6 urns with 100 black and white balls each. The proportion of black and white balls differs across each urn and the number of black balls is 0, 20, 40, 60, 80, and 100 respectively. There are 5000 individuals and for every individual, one of these urns is randomly selected with equal probability. These individuals will start with a uniform prior belief and each individual will independently observe one ball that is drawn from their selected urn. With this setup, we can compute the distribution of the Bayesian posterior beliefs.

From this calibration exercise, we see that we cannot naively attempt to match the normalized excess belief movement. Under both calibration exercises, we obtain similar normalized belief movement, but the amount of measurement error required is different due to different amount of uncertainty reduction.

8 Discussion

We found significantly more movement than uncertainty reduction in our data. This is consistent with base-rate neglect and overreaction to signals. We find that roughly 54%

of observations with realized wage offers in our dataset can be classified as overtly non-Bayesian in additional ways, with roughly a fourth of observations with offers updating beyond any offer received and another fourth updating away from offers received. Two possible explanations for the prevalence of these specific non-Bayesian updating behaviors are Gambler's Fallacy and motivated beliefs (optimistic reasoning).

Our findings strongly indicate that information intervention is a powerful tool for influencing people's beliefs, due to the excess belief movement. In our information-rich world, where individuals are continuously exposed to a stream of new information, it is essential to recognize that the effect of providing a single piece of information may be short-lived as people may overreact to the new information they receive or neglect the older information. To ensure a more lasting impact on beliefs, it might be necessary to adopt a strategy of consistently providing new and relevant information over time.

Our results are limited in a few important ways that leave scope for future work. First, 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. Second, we find a large effect among individuals with no wage offers, implying that our main result may be the product of unobservable factors or survey design. Future analyses could test the sensitivity of question phrasing and propose possibly more effective survey techniques. Another limitation is that we observe most individuals for only two periods, which limits our ability to study learning over extended time periods. Future analyses could collect higher-frequency data, allowing for a better study of the change in learning biases over time and across different time horizons.

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^2_ϵ and uncorrelated with recent belief and error realizations ($\mathbb{E}(\epsilon_t \pi_t) = \mathbb{E}(\epsilon_t \pi_{t-1}) = \mathbb{E}(\epsilon_t \epsilon_{t-1})$).

The belief movement will be larger by σ^2_ϵ

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

The uncertainty reduction will be smaller by σ^2_ϵ

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

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

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

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.

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

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.

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

Clementi, Fabio and Mauro Gallegati, "Pareto's Law of Income Distribution: Evidence for Germany, the United Kingdom, and the United States," *Public Economics*, 2005, pp. 3–14.

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.

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.

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

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.

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

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.

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.

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.

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.