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

Sebastian Brown, Kenneth Chan

University of California, Santa Barbara

April 4, 2024

Motivation

- Providing different types of labor market information has been shown to alter important aspects of individual labor market behavior, affecting outcomes from how many jobs people find (as in Arni 2016) to where they apply (as in Gee 2018).
- The effectiveness of such policies is likely to depend on exactly *how* people update their beliefs about what opportunities exist for them in the labor market in response to new information they receive.

Motivation

- Many models of belief updating have been proposed which could be applied to understand the specifics of updating in a labor market context.
- However, empirical job search data sources with belief (expectations) questions remain rare, and it is only recently that these questions have begun to be used more often. (Mueller & Spinnewijn 2023)
- Unlike in the lab, we cannot directly control signal structure and priors, making it more difficult to evaluate these models.

Research Questions

- Main Question: How do people learn their own wage offer distributions over time?
  - We assume that individual wage offer distributions are static, but unknown.
  - Under this assumption, do individuals learn about wages in a biased (non-Bayesian) way?
  - If so, how strong/what type are their biases?
  - Which existing method of modeling belief updating is most consistent with empirical labor market data?

Methodology Overview

- We use a nationally-representative survey (the Survey of Consumer Expectations) that surveys each individual up to 3 times every 4 months.
- We test for non-Bayesian behavior by adapting a test from recent work by Augenblick & Rabin (2021).
- We also do some individual-level analyses and find results consistent with asymmetric updating.

Literature

- Theoretical Search models
  1. Model with known underlying distribution (Stigler 1961, McCall 1970, Weitzman 1979)
  2. Model with unknown underlying distribution (Rothschild 1978, Rosenfield & Shapiro 1981, Talmain 1992, Li & Yu 2018, Potter 2021)
- Empirical work has primarily focused on the first class of models due to insufficient data on job searchers' beliefs.
- Recent empirical work related to learning in job search: Conlon, Pilossoph, Wiswall & Zafar 2018 (estimate wage updating; key methodological differences discussed later after discussion of how our test works); Mueller, Spinnewijn and Topa 2021 (cross-sectional analysis of expectations of unemployed by unemployment length); Potter 2021 (a model with learning fits dynamics of time use data from Great Recession)
- Papers on non-Bayesian updating (Epstein et al. 2010, Grether 1980, Hagmann & Loewenstein 2017, Ba 2022, Ortoleva 2012)

Contributions to the Literature

- Provide an estimate of the prevalence of non-Bayesian updating in the data which has some methodological advantages over previous empirical work (fewer assumptions as explicit estimation of Bayesian posterior is unnecessary) and which detects a new feature of the non-Bayesian updating (movement and reduction).
- Provide individual-level analysis which suggests asymmetric wage updating responses to "happy" and "sad" news.
- Distinguish which (non-Bayesian) updating rules are consistent with the data.

Martingale Propery

- Let the state space be Θ and set of signals X
- **Martingale Property:** $\mathbb{E}_X(g_{t+1}(\theta|x)|g_t(\theta)) = g_t(\theta)$
- Idea: before you observe the signal, the expected posterior has to equal to prior

Competing Behavioral Models of Updating

1. Bayesian updating

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

2. Affine Transformation of Prior and Bayesian Belief (Epstein, Noor & Sandroni 2010)

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

3. Exponential Non-Bayesian updating (Grether 1980)

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

4. Affine Transformation of Reference Belief and Bayesian Belief (Hagmann & Loewenstein 2017)

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

Best offer estimate: Question text

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

Best offer distribution: Question text

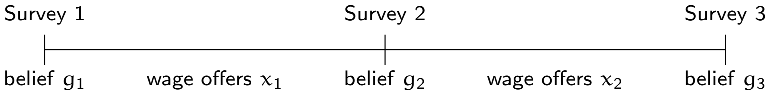
OO2b - OO2b (shown if OO2a2 > 0 each response is % of OO2a2 ranging from .8 to 1.2) **(Added November 2014)**

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)

Survey Timeline



Overview of Data

Description	Total	Got offer?	
		Yes	No
Total observations (updates)	4,374	747	3,627
Number unemployed	159	46	113
Number employed	3,634	637	2,997
Number not in labor force	534	57	477
Missing employment status	47	7	40
Unique individuals	3,103	661	2,680
Date Range	3/2015-3/2020		

Table: Employment status is defined as the status in the "before" period of the before-after pair. Numbers of individuals in "Got offer?" categories don't exactly sum to totals because individuals could change statuses between periods; in such a case, an individual would be counted in both categories for the "unique individuals" measure, although this will not impact the other counts.

Data Structure: Simplified Example

- **Period 1:** Guess of average best wage offer is \$100

	$90 < w \leqslant 100$	$100 < w \leqslant 110$	$110 < w \leqslant 120$
$p(\cdot)$	0.4	0.2	0.4

Table: Period 1 Beliefs

- **Period 2:** Guess of average best wage offer is \$90

	$81 < w \leqslant 90$	$90 < w \leqslant 99$	$99 < w \leqslant 108$
$p(\cdot)$	0.3	0.4	0.3

Table: Period 2 Beliefs

Data Structure: Simplified Example

- **Period 1:** Guess of average best wage offer is \$100

	$w \leqslant 100$	$w \leqslant 110$	$w \leqslant 120$
$p(\cdot)$	0.4	0.6	1

Table: Period 1 Beliefs

- **Period 2:** Guess of average best wage offer is \$90

	$w \leqslant 90$	$w \leqslant 99$	$w \leqslant 108$
$p(\cdot)$	0.3	0.7	1

Table: Period 2 Beliefs

Data Structure: Simplified Example

- **Period 1:** Guess of average best wage offer is \$100

	$w \leqslant 90$	$w \leqslant 99$	$w \leqslant 100$	$w \leqslant 108$	$w \leqslant 110$	$w \leqslant 120$
$p(\cdot)$			0.4		0.6	1

Table: Period 1 Beliefs

- **Period 2:** Guess of average best wage offer is \$90

	$w \leqslant 90$	$w \leqslant 99$	$w \leqslant 100$	$w \leqslant 108$	$w \leqslant 110$	$w \leqslant 120$
$p(\cdot)$	0.3	0.7		1		

Table: Period 2 Beliefs

Data Structure: Simplified Example

- **Period 1:** Guess of average best wage offer is \$100

	$w \leqslant 90$	$w \leqslant 99$	$w \leqslant 100$	$w \leqslant 108$	$w \leqslant 110$	$w \leqslant 120$
$p(\cdot)$			0.4		0.6	1

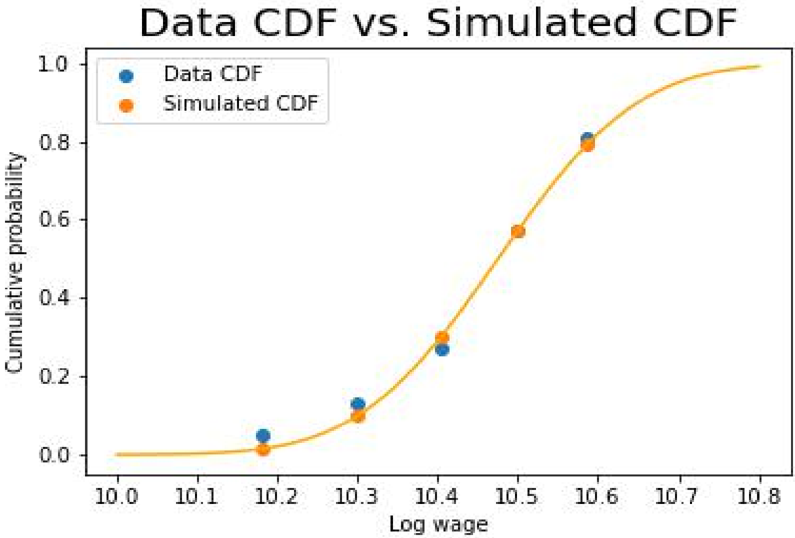
Table: Period 1 Beliefs

- **Period 2:** Guess of average best wage offer is \$90

	$w \leqslant 90$	$w \leqslant 99$	$w \leqslant 100$	$w \leqslant 108$	$w \leqslant 110$	$w \leqslant 120$
$p(\cdot)$	0.3	0.7	0.9	1	1	1

Table: Period 2 Beliefs

Matching Graphs: Example CDF



Method Wellness-of-Fit

- Have tried different distributions/fittings (Gumbel (extreme value) distribution, log normal, etc.)
- One we are using now fits the data the best on average from what we have tried so far, as measured by the average absolute error across all bins.

Simple Model

Red Die

Value of die roll	Tokens awarded
1	10
2	20
3	20
4	30
5	30
6	30

Blue Die

Value of die roll	Tokens awarded
1	10
2	10
3	10
4	20
5	20
6	30

- One of the die is selected randomly.
- You can only observe the tokens awarded, not the color of the die.
- Given a 50-50 prior, the expected value from rolling a die is 20.

Simple Model

Red Die

Value of die roll	Tokens awarded
1	10
2	20
3	20
4	30
5	30
6	30

Blue Die

Value of die roll	Tokens awarded
1	10
2	10
3	10
4	20
5	20
6	30

- One of the die is selected randomly.
- You can only observe the tokens awarded, not the color of the die.
- Given a 50-50 prior, the expected value from rolling a die is 20.
- How will your expectation change after observing a 10-token outcome?

Uncertainty in the Environment

- The agent believes there is a set of possible wage distributions, $\mathcal{F}$ (color of die)
- True wage distribution is in $\mathcal{F}$ which are indexed by an ordered set $\Theta \subset \mathbb{R}$
- The agent has belief g_t over θ at time t

Uncertainty in the Environment

- The agent believes there is a set of possible wage distributions, $\mathcal{F}$ (color of die)
- True wage distribution is in $\mathcal{F}$ which are indexed by an ordered set $\Theta \subset \mathbb{R}$
- The agent has belief g_t over θ at time t
- We partition the wages into n wage bins $\{[a_0, a_1), [a_1, a_2), \ldots [a_{n-1}, a_n)\}$

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

- If the updating rule used by g has the Martingale property, then so will π.

2 States (Augenblick and Rabin, 2021)

- Overview: Model of belief dynamics (multiple time periods)
- 2 states and state 1 occurs with probability π
- Belief movement

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

- Uncertainty Reduction

$$\begin{aligned} r_{t_1,t_2}(\pi) &\equiv \sum_{\tau=t_1}^{t_2-1} \pi_\tau(1-\pi_\tau) - \pi_{\tau+1}(1-\pi_{\tau+1}) \\ &= \pi_{t_1}(1-\pi_{t_1}) - \pi_{t_2}(1-\pi_{t_2}) \end{aligned}$$

Many States (Augenblick and Rabin, 2021)

- n states and state i occurs with probability π^i
- Belief movement

$$m_{t_1,t_2}(\pi) \equiv \sum_{i=1}^{n} \sum_{\tau=t_1}^{t_2-1} ({\pi^i}_{\tau+1} - {\pi^i}_{\tau})^2$$

- Uncertainty Reduction

$$r_{t_1,t_2}(\pi) \equiv \sum_{i=1}^{n} \sum_{\tau=t_1}^{t_2-1} {\pi^i}_{\tau}(1 - {\pi^i}_{\tau}) - {\pi^i}_{\tau+1}(1 - {\pi^i}_{\tau+1})$$

$$= \sum_{i=1}^{n} {\pi^i}_{t_1}(1 - {\pi^i}_{t_1}) - {\pi^i}_{t_2}(1 - {\pi^i}_{t_2})$$

Test For Bayesianism (Martingale)

- We denote M_{t_1,t_2} and R_{t_1,t_2} as the random variable for belief movement and uncertainty reduction respectively
- Statistical test: $\mathbb{E}M_{t_1,t_2} = \mathbb{E}R_{t_1,t_2}$
- Intuition: When there is small uncertainty reduction (not many or weak signals or a tight prior), belief movement should be small

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

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

$$X_{norm} = \frac{\overline{m}_{1,2}}{\bar{r}_{1,2}}$$

Differences from Most Similar Paper

- Conlon et al. (2018) also used the SCE data and has a similar research question and has found that job searchers are non-Bayesian
- Our method does not require explicitly estimating a Bayesian posterior, while theirs does.
- They do this by assuming that both prior and posterior beliefs over log wages follow a normal distribution with known variance σ^2, and apply the corresponding conjugate prior formula to get the Bayesian posterior:

$$E_{t+1}(\theta|x_{t+1}) = \frac{\sigma^2\mu + \sigma_0^2 x_{t+1}}{\sigma^2 + \sigma_0^2}$$

  where σ_0^2 can be estimated directly from the distributional question in the data once log normality is assumed.
- They allow σ^2 to vary by whether an individual has earned a college degree or not, and estimate it by estimating the variance of all reported wage offers in the sample.

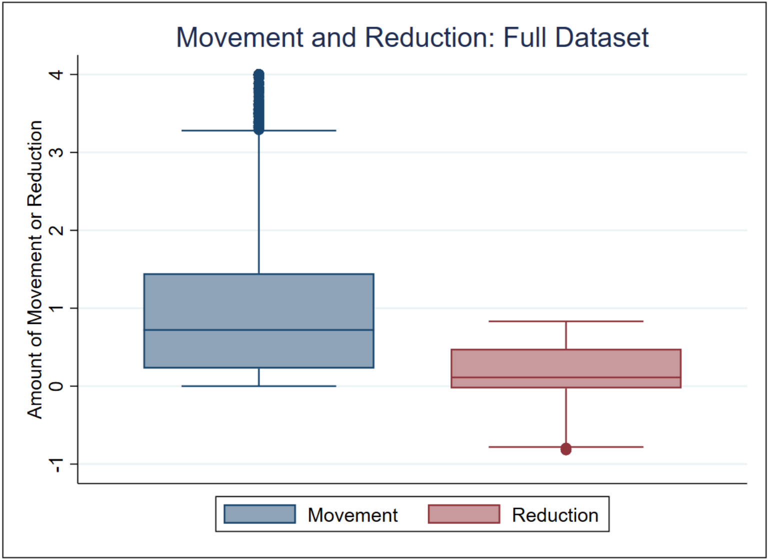
Estimate Contributions (Cont.)

- Even if this ended up being a good estimate of the variance of the wage distribution by education group, we must still assume that individuals use education group variance in their updating for the conjugate prior formulas to make sense.
- However, the individual wage distribution for a college graduate who is a CEO might be very different than a college graduate who is a high school teacher.
- Overall, there doesn't seem to be a good reason to expect individuals to individually update using such a broad group variance as a measure of signal informativeness (or to expect them to know it).

Estimate Contributions

- The main test we use in our paper sidesteps needing an explicit measure of beliefs over signal (wage) informativeness while still providing a test for whether individuals update in a Bayesian measure on average.
- It is also able to test for Bayesianism in individuals who did not receive a wage offer, which was not possible under the previous method.

Martingale Test



Test Results

Statistic	All Individuals	Got Offer?		Searched?	
		Yes	No	Yes	No
$\overline{m}$	.8480	.9060	.8357	.8494	.8336
	(.0129)	(.0305)	(.0142)	(.0257)	(.0156)
$\overline{r}$	.1449	.1376	.1464	.1521	.1324
	(.0058)	(.0137)	(.0064)	(.0114)	(.0070)
X	.7031	.7684	.6894	.6973	.7013
	(.0138)	(.0330)	(.0151)	(.0276)	(.0167)
X_{norm}	5.8533	6.5831	5.7089	5.5831	6.2979
Observations	3,060	532	2,528	740	2,083

Table: Excess movement statistics: Log normal-fitted results. Standard errors in parentheses.

Result Interpretation

- The test results are consistent with overreacting to signals and base rate neglect in the Grether model.
- Statistics have similar magnitude whether pooling, separating on whether or not individual reported an offer, or separating on whether individual searched within four weeks prior to survey.
- This also remains when splitting the sample by age, education, gender, and race groups.

Measurement Errors

- Measurement Error can potentially explain excess belief movement even when agent is Bayesian (Augenblick & Rabin 2021)
- Consider that the agent has a true belief of π_t^i but reports a distorted belief $\hat{\pi}_t^i = \pi_t^i + \epsilon_t^i$, where ϵ_t^i is the measurement error.
- Assume that the measurement error term is mean zero with variance ${\sigma_{\epsilon}^{i}}^{2}$ and uncorrelated with recent belief and error realizations $(\mathbb{E}(\epsilon_t^i \pi_t^i) = \mathbb{E}(\epsilon_t^i \pi_{t-1}^i) = \mathbb{E}(\epsilon_t^i \epsilon_{t-1}^i) = 0)$
- We show in our appendix that expected excess belief movement will be equal to $\sum_{i=1}^{n} 2\sigma^2_{\epsilon_t^i} \neq 0$

Measurement Errors

- Monte Carlo Simulation: 10000 simulations
- Same number of observations as in our dataset.
- 6 states.
- Prior is uniform distribution.
- We pick a uniform distribution of symmetric posterior distributions that is Bayes' plausible to match the uncertainty reduction.
- Measurement error is simulated by drawing from a Dirichlet distribution with mean centered around the correct beliefs.
- $\Delta \equiv \sum_{i=1}^{6} |\pi_1^i - \pi_{prior}^i|$

Measurement Errors

Statistic	Uniform Prior with Matched Belief Movement
X	0.5319 [0.5196, 0.5441]
X_{norm}	5.3506 [5.1149, 5.6021]
Δ	0.9871 [0.9800, 0.9944]

Table: Monte Carlo Simulation Results

Competing Behavioral Models of Updating

We can reject updating rules with Martingale property

1. ~~Bayesian updating~~

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

2. ~~Affine Transformation of Prior and Bayesian Belief~~ (Epstein et al. 2010)

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

Updating Rules (Cont.)

3. Exponential Non-Bayesian updating (Grether 1980)

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

4. Affine Transformation of Reference Belief and Bayesian Belief (Hagmann & Loewenstein 2017)

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

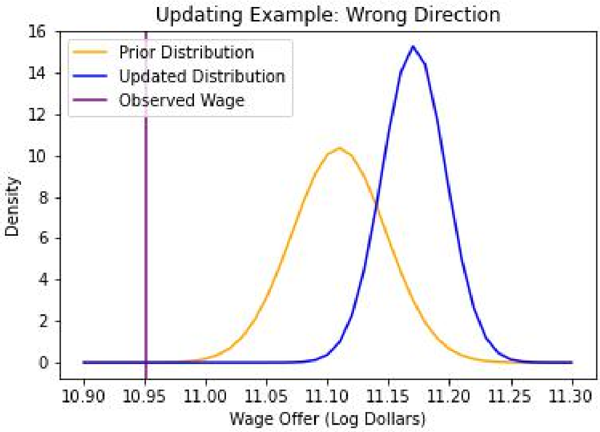
Directional Patterns

- Under a standard Bayesian updating job search model in a Gaussian updating framework, the mean of the posterior belief should always be a convex combination between the prior mean and the wage offer.
- In this case, a Bayesian agent will always update her prior mean in the direction of an observed signal.

Updating in Wrong Direction

- We find two patterns in the data which do not follow this idea.
- The first we call "wrong direction."
- In this case, the agent updates the mean away from the observed wage signal(s).
- We have some people who increase their reported (not fitted) wage expectation after seeing a low-wage offer

Example Wrong Direction Update

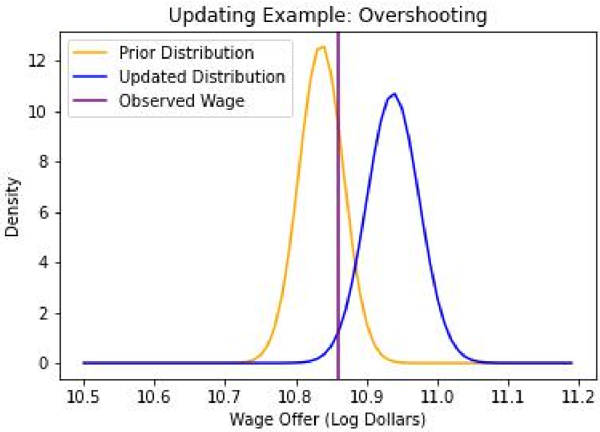


- Fitted distributions displayed for illustrative purposes only; measurement of the pattern uses only reported means, not fitted means.

Overshooting

- The second pattern is called "overshooting."
- In this case, the individual updates the posterior mean *beyond* the observed signal.

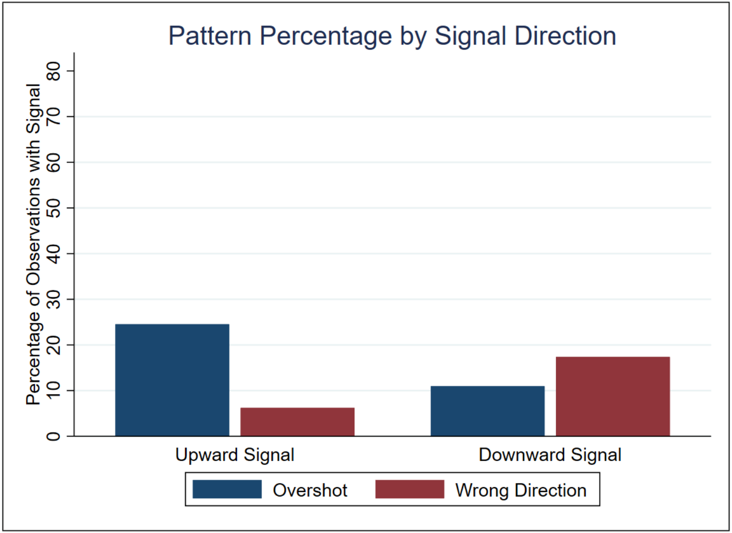
Observed Patterns



Pattern Counts

- We find evidence for both patterns among those with offers.
- About 15% of observations display overshooting, while about 11% of observations display moving in the wrong direction (note also that the categories are mutually exclusive).
- However, we also found another pattern among these results.

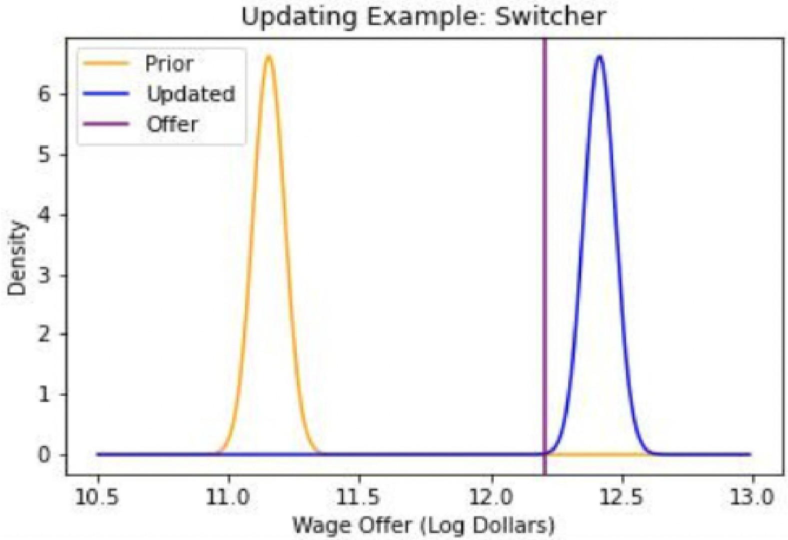
Asymmetric Updating: Good vs. Bad News “Mistake” Rates



Grether Model Implications

- Grether (1980) model predicts that if 0 probability weight is assigned in the prior, the posterior belief cannot be positive unless $\alpha = 0$ (total base rate neglect).
- We have people who started with 0 probability on some wage values but have a non-zero probability in the posterior.

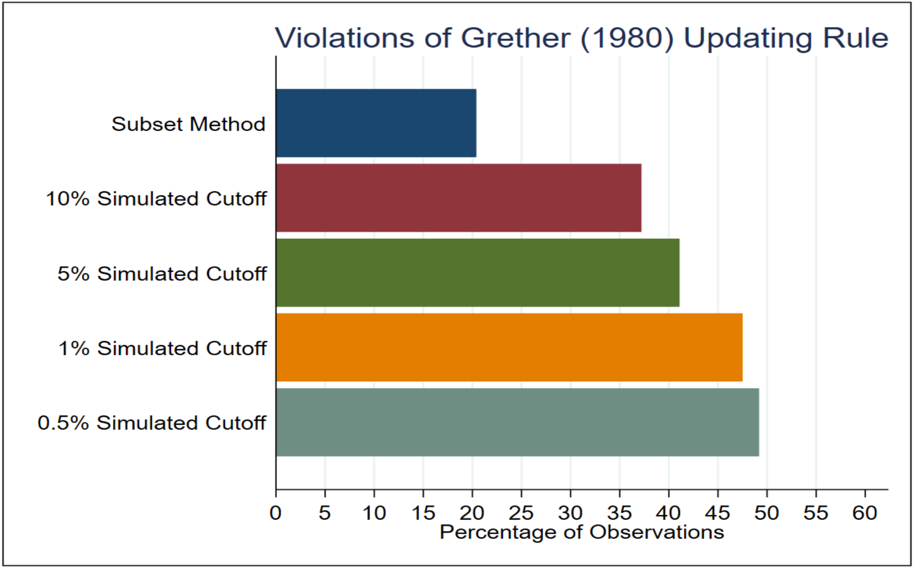
Example: Zero to Positive Belief



Description of Zero-Probability Shift Identification Methods

1. **Subset method (nonparametric, lower bound):**
   - The individual has a prior bin with 0 probability weight.
   - The individual has at least one posterior bin with positive weight contained within the aforementioned prior bin.
2. **Simulated method with threshold (parametric):**
   - Use fitted posterior instead of reported posterior so bin definitions are exactly the same between prior and posterior.
   - Since simulated posterior is log normal distribution, no range will have exactly zero probability weight.
   - A threshold (e.g. 1%, 5%) is used to determine whether the simulated posterior bin has non-zero probability weight.

Counts by Method



Zero-Probability Shift Results

- A significant portion of individuals are identified as updating a zero-probability belief to a positive-probability posterior, regardless of method used.

Conclusion

- To the extent that individuals update wage expectations, they seem much more willing to change the value they think their next offer will be than to change how sure they are that the next offer will be that value.
- This result is inconsistent with updating rules which rely on the Martingale property such as Bayesian updating and linearly combining the prior and Bayesian posterior.
- Among individuals who received offers in our data, individuals who made updating "mistakes" tended to do so in a way that favored more positive posterior wage expectations[1]
- Therefore, non-Bayesian updating models which allow motivated updates, such as Hagmann & Loewenstein (2017), may be useful in modeling labor market updating.

[1] Assuming no systematic differences in non-wage news among such individuals.