Testing Theories of Expectation Formation using Probabilistic Answers

Tao Wang

June 19, 2019

1 Introduction

The theories on how different agents form expectations have proliferated over the past decade. On one hand, these theories built upon different micro foundations produce some similar macro patterns. For instance, the notion of information rigidity, or expectation rigidity could be micro founded by many different theories, but mostly generate sluggish response to new information.

On the other hand, however, there are important and subtle differences in testable predictions from these theories for both individual forecasts and aggregate moments of the forecasts. My goal of this exercise is to derive the testable predictions based on various theories, and utilize the probablistic questions from both professional forecasters survey and household survey to test them.

The potential contribution lies in the extra insights I may get from using probablistic questions.[1] One obvious advantage of density expectations compared to point one is availability of higher moments, i.e. variance of the forecast indicates the dispersion of the forecasts. Along time dimension, the dynamics of conditional relative likelihood is a useful indicator of information gain or forecasting efficiency. Besides, as in a typical signal-extraction model, the prediction that new information reduce conditional variance can be tested only using probability distributions. Although the detailed implementation of these tests involves technical difficulties, the ideas are quite clear. Lastly, with second moments in individual level expectations, we are allowed to potentialy explore the heterogeneity across agents within one theory. For instance, sticky information typically assume a posson update rate for all agents in the economy. This is partly due to the fact that available data only allows us to recover single parameter of regidity instead of inter-group heterogeneity.

[1] For an insightful survey on the importance of measuring subjective expectations using probablistic surveys, see [Manski, 2004].

2 Theories

2.1 A Unified Framework

Agent i's h-period ahead density forecast of variable y_i at time t is the conditional density of $y_{i,t+h}$ given the information set $I_{i,t}$ available at time t.

$$\widehat{f}_{i,t}(y^i_{t+h}|I_{i,t})$$

Notice here the y_{t+h} has supscript i. This is to remind that in general the variables y_{t+h} can be either aggregate variables or idiosyncratic ones, in the former case, we could drop subscript i.

$I_{i,t}$ is the information set available at time t for individual i. The information set can be also agent specific, therefore, it has subscript i. The specific content contained in I_t varies from different models. Seemingly a simple and justified assumption is that $x_t \in I_t$. This, however, may not be ture for all models. For instance, sticky expectation and rational inattention literature all predict that agents are not able to update new information instataneously. Thus the information set may not winclude the recent realization of the variable itself.

Assumption also needs to be made about the underlying model that generates the conditional density of the variables by agent i.

Given the forecast density, the expectation of forecast at time t is.

$$E_{i,t}(y_{t+h}|I_{i,t}) = \int \widehat{f}_{i,t}(y_{t+h}|I_{i,t})y_{t+h}dy_{t+h} \equiv \bar{y}_{i,t,t+h}$$

The forecasting variance is

$$Var_{i,t}(y_{t+h}|I_{i,t}) = \int \widehat{f}_{i,t}(y_{t+h}|I_{i,t})(y_{t+h} - \bar{y}_{i,t,t+h})^2 dy_{t+h}$$

The true process of y_t depends on the variables. Here, we consider a simple case where true process of y_t is AR(1).

$$y_{t+1} = \rho y_t + \omega_t$$

$$\omega_t \sim N(0, \sigma^2_\omega)$$

2.2 Full-information rational expectation benchmark

2.2.1 Individual moments

Full information rational expectation implies $y_t \in I_t$

Expectation

$$E_{i,t}(y_{t+h}|y_t) = E_{i,t}(y_{t+h}|y_t) = \rho^h y_t$$

Variance

$$Var_{i,t}(y_{t+h}|y_t) = \sum_{s=1}^{h} \rho^{2s}\sigma^2_\omega$$

Ex post forecast error of individual i

$$FE_{i,t} = y_{t+h} - E_{i,t}(y_{t+h}|y_t) = \sum_{s=1}^{h} \rho^s \omega_{t+h-s}$$

has is mean zero and orthogonal to information at time t.

2.2.2 Population moments

A special case below.

2.3 Sticky Expectation

Agent does not update information instantaneously, instead at a Possion rate λ. Specificaly, at any point of time t, the agent learns about the up-to-date realization of y_t with probability of λ; otherwise, it forms the expectation based on the most recent up-to-date realization of $y_{t-\tau}$, where τ is the time experienced since previous update.

2.3.1 Individual moments

For an individual whose most recent update occurs at τ period before, $I_{i,t|t-\tau} = I_{i,t-\tau} = y_{t-\tau}$. Thus

$$E_{i,t|t-\tau}(y_{t+h}|I_{i,t}) = E_{i,t}(i, y_{t+h}|y_{t-\tau}) = \rho^{h+\tau} y_{t-\tau}$$

$$Var_{i,t|t-\tau}(y_{t+h}|I_{i,t}) = Var_{i,t}(y_{t+h}|y_{t-\tau}) = \sum_{s=0}^{h+\tau} \rho^{2s} \sigma^2_\omega$$

$Var_{i,t|t-\tau}(y_{t+h}|_{i,t})$ increases as τ increases. The model collapses to full-information rational expectation if $\tau = 0$ for all individuals.

Now, we turn to the case when agent updates infrequently.

- When the update happens, the variance responds substantially.

$$Var_{i,t}(y_{t+h}|y_t) - Var_{i,t|t-\tau}(y_{t+h}|y_{t-\tau}) = \sum_{s=0}^{h} \rho^{2s}\sigma^2_\omega - \sum_{s=0}^{h+\tau} \rho^{2s}\sigma^2_\omega = \sum_{s=0}^{\tau} \rho^{2s}\sigma^2_\omega$$

- When the update does not happen, the variance responds little.

$$Var_{i,t|t-\tau-1}(y_{t+h}|y_{t-\tau-1}) - Var_{i,t|\tau}(y_{t+h}|y_{t-\tau}) = \sigma^2_\omega$$

  At individual level, it is hard to recover the information rigidity parameter λ directly. One testable prediction of information rigidity is that the change in forecast variance varies substantially depending on if updated or not. The longer period for which the agent stays unupdated(greater τ), the bigger the change is in the forecasting variance.

However, the difference in average responses in variance to new information may speak to potential heterogeneity in information rigidity. According to the theory above, higher information rigidity implies high volatility of variance responses.

2.3.2 Population moments

1. Average forecast

   The mean forecast across population is a weighted average of past rational expectations

$$
\begin{aligned}
\bar{E}_t(y_{t+h}) &= \lambda \underbrace{E_t(y_{t+h})}_{\text{rational expectation at t}} + (1-\lambda) \underbrace{\bar{E}_{t-1} y_{t+h}}_{\text{average expectation at} t-1} \\
&= \lambda E_t(y_{t+h}) + (1-\lambda)(\lambda E_{t-1}(y_{t+h}) + (1-\lambda)\bar{E}_{t-2}(y_{t+h}))... \\
&= \lambda \sum_{s=0}^{\infty} (1-\lambda)^s E_{t-s}(y_{t+h}) \\
&= \lambda \sum_{s=0}^{\infty} (1-\lambda)^s \rho^{s+h} y_{t-s}
\end{aligned}
\tag{1}
$$

   The change in average forecast is

$$
\begin{aligned}
\Delta \bar{E}_t(y_{t+h}) &= (1-\lambda)\Delta \bar{E}_{t-1}(y_{t+h}) + \lambda(E_t(y_{t+h}) - E_{t-1}(y_{t+h})) \\
&= (1-\lambda)\Delta \bar{E}_{t-1}(y_{t+h}) + \lambda \rho^h \omega_t
\end{aligned}
\tag{2}
$$

   This implies the change in average forecast is serially correlated, depending on the information rigidity, i.e. lower λ implies higher serial correlation. Also lower λ implies the expectation underreact to the shocks at t.

2. Cross-sectional disagreements

   According to information rigity model, if everyone is instantaneously udpated, there should not be disagreements. In general, the dispersion in forecasting is non-zero because of different lags in updating. One can also derive variance of forecasts across agents at time t

$$
Var_t(E_{i,t}(y_{t+h})) = \lambda \sum_{\tau=0}^{\infty} (1-\lambda)^{\tau} (E_{t|t-\tau}(y_{t+h}) - \bar{E}_t(y_{t+h}))^2 \tag{3}
$$

From time t to $t+1$, the change in dispersion comes from two sources. One is newly realized shock at time $t+1$. The other component is from people who did not update at time t and update at time $t+1$.

$$\Delta Var_{t+1}(E_{i,t+1}(y_{t+h})) = \text{change due to new updaters} + \text{shock at time } t+1 \quad (4)$$

Notice the change is positive, meaning the dispersion rises in response to a shock. Importantly, the increase is the same regardless of the realization of the shock.

More generally, we can derive the impulse response of dispersion at time $t+j$ to a shock that realized at t.

$$\rho^{2(h+j)}(1-\lambda^{j+1})\lambda^{j+1}\omega_t^2 \quad (5)$$

A shock increases the disagreements across agents and then it gradually returns to its steady state level.

3. Average variance

   Since we have individual level variance, we can also derive average variance of the population. Taking the average of variance across individual agents at time t. Here I use the $\bar{}$ to represent population mean of variance.

$$\begin{aligned}\bar{V}ar_t(y_{t+h}) &= \sum_{\tau=0}^{+\infty} \underbrace{\lambda(1-\lambda)^\tau}_{\text{fraction who does not update until } t-\tau} \underbrace{Var_{t|t-\tau}(y_{t+h})}_{\text{Variance of most recent update at } t-\tau} \\ &= \sum_{\tau=0}^{+\infty} \lambda(1-\lambda)^\tau \sum_{s=0}^{h+\tau} \rho^{2s}\sigma_\omega^2\end{aligned} \quad (6)$$

Correspondingly, the change in average uncertainty will be

$$\Delta\bar{V}ar_t(y_{t+h}) = \sum_{\tau=0}^{+\infty} \lambda(1-\lambda)^\tau \sum_{s=0}^{h+\tau} \rho^{2s}\sigma_\omega^2 - \sum_{\tau=0}^{+\infty} \lambda(1-\lambda)^\tau \sum_{s=0}^{h+\tau} \rho^{2s}\sigma_\omega^2 = 0 \quad (7)$$

An information rigidity model based on fixed Poisson update rate predicts that the average variance does not change over time. This can be tested.

2.3.3 Summary of predictions of information rigidity

- Individual expectation may or may not change depend upon if updating.
- Individual variances changes non-monotonically depending on if updating. Always increase with arrival of new information.
- Population mean of forecast responds with lags. Change is serially correlated.
- Population dispersion of forecasts rise in response to new shocks and return to steady state level gradually.
- Population average variance does not change over time.

2.4 Noisy information

A class of so-called noisy information model describes the expectation formation as a process extracting or filtering true fundamental state variable y_t from a sequence of realized signals. The starting assumption is that agent cannot observe the true variable perfectly. Unlike information rigidity model, it is assumed that agents keep track of the realizations of the signals instantaneously all the time.

We assume agent i observe two signals s^{pb} and s_i^{pr}, with s^{pb} being public signal common to all agents, and s_i^{pr} private signals being individual specific. The generating process of two signals are

$$\begin{aligned} s_t^{pb} &= y_t + \epsilon_t, \quad \epsilon_t \sim N(0, \sigma_\epsilon^2) \\ s_{i,t}^{pr} &= y_t + \xi_{i,t} \quad \xi_{i,t} \sim N(0, \sigma_\epsilon^2) \end{aligned} \tag{8}$$

We can stack the two signals into one vector $s_{i,t} = [s_t^{pb}, s_{i,t}^{pr}]'$ and $v_{i,t} = [\epsilon_t, \xi_{i,t}]'$. So in a compact form, it can be written as

$$\begin{aligned} s_{i,t} &= H y_t + v_{i,t} \\ &\text{where } H = [1,1]' \end{aligned} \tag{9}$$

In our general framework, the noisy information implies that the information set $I_{i,t}$ available to individual i at time t only includes past and recent realizations of the signals. The individual density forecast of y_{t+h} is

$$\widehat{f}_{i,t}(y_{t+h}|I_{i,t}) = \widehat{f}_{i,t}(y_{t+h}|s_{i,t}, s_{i,t-1}...) \equiv \widehat{f}_{i,t|t}(y_{t+h})$$

We use $t|k$ to denote the moments at time t based on information(signals) till time k.

Then we are ready to apply Kalman Filter in this context. The posterior distribution of y_t after seeing all signals till t is

$$\widehat{f}_{i,t|t}(y_{t+h}) \sim N(E_{i,t|t}(y_{t+h}), Var_{i,t|t}(y_{t+h})) \tag{10}$$

where the expectation and variances are functions of noisiness of signals and fundamentals. The expectation also depends on the realized values. But this is not the case for variance.

2.4.1 Individual moments

1. Expectation

   Now any agent trying to forecast future variables will have to form her expectation of the contemporaneous state variable, $E_{i,t|t}(y_t)$. Then the best h-period ahead forecast is simply iterated h periods forward based on the AR(1) process.

   Thus, we first work out $E_{i,t|t}(y_t)$.

$$\begin{aligned} E_{i,t|t}(y_t) &= \underbrace{E_{i,t|t-1}(y_t)}_{\text{prior}} + P \underbrace{(s_{i,t|t} - s_{i,t|t-1})}_{\text{innovations to signals}} \\ &= (1-PH)E_{i,t|t-1}(y_t) + Ps_{i,t} \\ &= (1-PH)E_{i,t|t-1}(y_t) + PHy_{i,t} + Pv_{i,t} \end{aligned} \tag{11}$$

   where the Kalman gain $P = [P_\epsilon, P_\xi] = \Sigma^y_{i,t|t-1}H(H'\Sigma^y_{i,t|t-1}H + \Sigma^v)^{-1}$

   where $\Sigma^y_{i,t|t-1}$ is the variance of y_t based on prior belief

   and $\Sigma^v = \begin{bmatrix} \sigma^2_\epsilon & 0 \\ 0 & \sigma^2_\xi \end{bmatrix}$

   The h-period ahead forecast is

$$E_{i,t|t}(y_{t+h}) = \rho^h E_{i,t|t}(y_{t+h}) \tag{12}$$

   Individual forecast partially responds to new signals, i.e. $P < 1$. $P = 1$ is a special case when both signals are perfect thus $\Sigma^v = 0$, then the formula collapses to full rational expectation.

   Now, the rigidity parameter is governed by $1 - PH$ with multiple signals. It is a function of variance of y from the prior of previous period and noisiness of the signals. Therefore, it is time variant as the variance is updated by the agent each period.

   There are a few important distinctions between noisy information and sticky expectation.

   • First, the persistence of expectation exists at individual level. There is serially correlation between $E_{i|t|t}(y_t)$ and $E_{i,t|t-1}(y_t)$, or more generally, between $E_{i,t|t}(y_{t+h})$ and $E_{i,t|t-1}(y_{t+h})$. This pattern can be only observed from population moments according to sticky expectation models.

   To see this, the change in individual forecast from $t-1$ to t is

$$\Delta E_{i,t|t}(y_{t+h}) = \underbrace{\rho^h(1-PH)\Delta E_{i,t-1|t-1}(y_t)}_{\text{Lagged response}} + \underbrace{\rho^h PH\Delta y_{i,t} + \rho^h P\Delta v_{i,t}}_{\text{Shocks to signals}} \quad (13)$$

The serial correlation is $\rho^h(1-PH)$, it does not only depend on PH, but also the forecast horizon h. Therefore, one testable assumption is to see auto regression of change in forecast to see if the coefficient depends on horizon.

• Second, the expectation adjusts in each period as long as there is new information. In sticky expectation, however, the expectation adjusts only when the agent updates.

2. Variance

The posterior variance at time t is a linear function of prior variance and variance of signals.

$$\Sigma^y_{i,t|t} = \Sigma^y_{i,t|t-1} - \Sigma^y_{i,t|t-1}H'(H\Sigma^y_{i,t-1}H' + \Sigma^v)H\Sigma^y_{i,t|t-1} \quad (14)$$

There are a few important properties in the variance.

- First, it does not depend on the realizations of the signal.

- Second, it decreases unambigously from $t-1$ to t. To see this

$$\Sigma^y_{i,t|t} - \Sigma^y_{i,t|t-1} = -\Sigma^y_{i,t|t-1}H'(H\Sigma^y_{i,t-1}H' + \Sigma^v)H\Sigma^y_{i,t|t-1} < 0 \quad (15)$$

These two properties carry through to the h-period ahead forecast as well. As the forecast variance is the following

$$Var_{i,t|t}(y_{t+h}) = \rho^{2h}\underbrace{Var_{i,t}(y_t)}_{\Sigma_{i,t|t}} + \sum_{s=0}^{h}\rho^{2s}\sigma^2_\omega \quad (16)$$

$$\Delta Var_{i,t|t}(y_{t+h}) = \rho^{2h}\Delta\Sigma_{i,t|t} - \rho^{2h}\sigma^2_\omega \quad (17)$$

From t to $t+1$, when $h \geq 1$, the decline in variance come from two sources. The first source is the pure gain from the new signals, i.e. $\Delta\Sigma^y_{i,t|t}$. It is scalled by the factor ρ^{2h}. The second source is present in full information rational expectation model: as time goes from $t-1$ to t, there is a reduction of uncertainty about ω_t.

2.4.2 Population moments

1. Average forecast

$$
\begin{aligned}
\bar{E}_{t|t}(y_{t+h}) &= \rho^h[(1-PH)\underbrace{\bar{E}_{t-1}(y_{t+h})}_{\text{Average prior}} + P \underbrace{\bar{s}_t}_{\text{Average Signals}}] \\
&= (1-PH)\bar{E}_{t-1}(y_{t+h}) + P[\epsilon_t, 0]' \\
&= (1-PH)\bar{E}_{t-1}(y_{t+h}) + P\epsilon_t
\end{aligned} \tag{18}
$$

2. Change in average forecast

$$
\Delta \bar{E}_t|t(y_{t+h}) = \rho^h(1-PH)\Delta\bar{E}_{t-1}(y_{t+h}) + \rho^h P \Delta\epsilon_t \tag{19}
$$

   Same to the individual forecast, the change in average forecasts has serial correlation with the same auto regression parameter $\rho^h(1-PH)$.

3. Cross-sectional disagreements

   In this model, the only disagreements across agents come from the difference in realized private signals. Therefore, in short-cut, the disagreements are

$$
\begin{aligned}
Var_t(y_{t+h}) &= E((E_{i,t|t}(y_{t+h}) - \bar{E}_t(y_{t+h}))^2) \\
&= \rho^{2h}P_\xi^2\sigma_\xi^2
\end{aligned} \tag{20}
$$

   Several properties.

   • First, the disagreements increase with the forecast horizon. • Second, the disagreements depends on noisiness private signals, but not on that of public signals and the variance of the true variable y. • Third, similar to sticky expectation model, the disagreements also increase with the rigidity parameter P in this model.

4. Change in disagreements

$$
\Delta Var_t(y_{t+h}) = \rho^{2h}(1-\rho^2)P_\xi^2\sigma_\xi^2 > 0 \tag{21}
$$

   The disagreements increase as time goes from $t-1$ to t. Also, as the time approaches $t+h$, the disagreements increase. This seems counterintuitive. But the reason is that here the disagreenments always exist simply because agents receive private signals, this disagreements is actually amplified as time goes forward.

5. Average variance

   Since the variance does not depend on signals and the precision is the same aross the agents, average variance is equal to the variance of each individual.

   $$\bar{V}ar_t(y_{t+h}) = \bar{\Sigma}^y_t \tag{22}$$

   Also, same as the individual variance, the variance unambiguiously drop as time goes by.

   $$\Delta Var_t(y_{t+h}) < 0 \tag{23}$$

2.4.3 Summary of predictions from noisy information

• Invididual expectation adjusts in each period, but only partially adjusts to new information.
• Unlike sticky expectation, slugishness in adjustment or serial correlation of adjustment exists in individual level. The correlation parameter decreases with forecast horizon, which is not the case in sticky expectation.
• Individual variance unambiguously drops each period as one approaches the period of realization. In sticky expectation, it increases regardless of updating or not.
• Population average forecast partially adjusts to news and has serial correlation as the individual level.
• Population disagreements rise in each period as time approaches the period of realization. Disagreetments will never be zero.
• Average variance declines unambiguously each period.

2.5 Other Theories

1. Rational inattention.

   - Chris Sim's model.[Sims, 2003] Information serves the role of uncertainty reduction measured by relative entropy. Agents optimally trade off the fixed cost of being attentative versus the gain from uncertainty gain.
   - Ricardo Reis's model. [Reis, 2006]
   - Xavier Gabaix's sparse matrix model. [Gabaix, 2014]

2. Epidemiologic view. [Carroll, 2003] Regardless of the microfoundations of the information rigidity, households infrequently get access to more rational-and-up-to-date expectations from professional forecasters with a poisson process.

Table 1: Information of Data

	SCE	SPF
Time period	2013-present	2007-present
Frequency	Monthly	Quarterly
Sample Size	1,300	30-50
Aggregate Var in Density	1-yr and 3-yr ahead inflation	1-yr CPI and PCE
Individual Var in Density	1-yr earning growth	No
Pannel Structure	stay up to 12 months	average stay for 5 years
Demographic Info	Education, Income, Age, Location	Industry

3. Strategic behaviors. Second order belief, i.e., what you believe of what others believe. [Angeletos and Jennifer, 2009]

3 Empirical Results

3.1 Data

- **New York Fed Survey of Consumer Expectation**
  - Include not only perceived probabilities of binary event as in Michigan Survey does, but also elicit densitity forecasts.
- **Professional Forecasters**

A summary of the data information is as below.

3.2 Density Estimates

3.3 Test of Null Hypothesis of Rational Expectation

3.3.1 Replicating [?]

3.4 Professional Forecasters as a Benchmark

3.4.1 Replicating [Coibion and Gorodnichenko, 2012]

3.4.2 Additional Evidence from Uncertainty

3.5 Evidence from Households

4 Conclusion

References

[Angeletos and Jennifer, 2009] Angeletos, G.-M. and Jennifer, L. (2009). Incomplete information, higher-order beliefs and price inertia. *Journal of Monetary Economics*, 56:S19–S37.

[Bordalo et al., 2018] Bordalo, P., Gennaioli, N., and Shleifer, A. (2018). Diagnostic expectations and credit cycles. *The Journal of Finance*, 73(1):199–227.

[Carroll, 2003] Carroll, C. D. (2003). Macroeconomic expectations of households and professional forecasters. *the Quarterly Journal of economics*, 118(1):269–298.

[Coibion and Gorodnichenko, 2012] Coibion, O. and Gorodnichenko, Y. (2012). What can survey forecasts tell us about information rigidities? *Journal of Political Economy*, 120(1):116–159.

[Evans and Honkapohja, 2012] Evans, G. W. and Honkapohja, S. (2012). *Learning and expectations in macroeconomics*. Princeton University Press.

[Gabaix, 2014] Gabaix, X. (2014). A sparsity-based model of bounded rationality. *The Quarterly Journal of Economics*, 129(4):1661–1710.

[Hansen and Sargent, 2001] Hansen, L. and Sargent, T. J. (2001). Robust control and model uncertainty. *American Economic Review*, 91(2):60–66.

[Malmendier and Nagel, 2015] Malmendier, U. and Nagel, S. (2015). Learning from inflation experiences. *The Quarterly Journal of Economics*, 131(1):53–87.

[Mankiw et al., 2003] Mankiw, N. G., Reis, R., and Wolfers, J. (2003). Disagreement about inflation expectations. *NBER macroeconomics annual*, 18:209–248.

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

[Reis, 2006] Reis, R. (2006). Inattentive consumers. *Journal of monetary Economics*, 53(8):1761–1800.

[Sims, 2003] Sims, C. A. (2003). Implications of rational inattention. *Journal of monetary Economics*, 50(3):665–690.