Learning and Subjective Expectation Formation: A Recurrent Neural Network Approach

Aditya Chaudhry

Chicago Booth

6th November 2020

Overview

- **Goal**: Develop an empirical framework to model relationship between macro signals and household expectations
- **Methodology**: Use recurrent neural networks (RNNs) to estimate this relationship and double machine learning (DML) to estimate marginal effects
- **Three new stylized facts:**
  1. Relationship between macro signals and expectations is nonlinear
  2. Marginal effects (ME) of signals on expectations are time-varying
  3. Signals about economic conditions explain majority of this time variation
- **Model implications:**
  1. New facts inconsistent with
     - 1.1 Full information rational expectaitons (FIRE) — need much time-variation in environment
     - 1.2 Noisy learning — linear parametrizations usually lead to constant ME
     - 1.3 Constant gain learning — ME are constant over time as more information observed
  2. Consistent with modification of standard rational inattention model

Recap: Empirical Framework

Goal 1: Estimate average structural function (ASF)

- Households report expectations $Y_{i,t+1|t}$ as function of:
  - History of signals: $\{Z_{i,\tau}\}_{\tau=0}^{t}$
  - Latent state variable: $\Theta_{i,t}$
  - History of idiosyncratic noise/private information: $\{\epsilon_{i,\tau}\}_{\tau=0}^{t}$
- We want "average mapping" $g : \left(\{Z_{i,\tau}\}_{\tau=0}^{t}, \Theta_{i,-1}\right) \to Y_{i,t+1|t}$

$$y_{i,t+1|t} \equiv \mathbb{E}_{\{\epsilon_{i,\tau}\}_{\tau=0}^{t}}\left[Y_{i,t+1|t}\right] \equiv g\left(\{Z_{i,\tau}\}_{\tau=0}^{t}, \Theta_{i,-1}\right)$$

- **ASF tells us how households map macro signals to expectations**

Recap: Empirical Framework

Goal 1: Estimate average structural function (ASF)

- Households report expectations $Y_{i,t+1|t}$ as function of:
  - History of signals: $\{Z_{i,\tau}\}_{\tau=0}^{t}$
  - Latent state variable: $\Theta_{i,t}$
  - History of idiosyncratic noise/private information: $\{\epsilon_{i,\tau}\}_{\tau=0}^{t}$
- We want "average mapping" $g : \left(\{Z_{i,\tau}\}_{\tau=0}^{t}, \Theta_{i,-1}\right) \to Y_{i,t+1|t}$

$$y_{i,t+1|t} \equiv \mathbb{E}_{\{\epsilon_{i,\tau}\}_{\tau=0}^{t}}\left[Y_{i,t+1|t}\right] \equiv g\left(\{Z_{i,\tau}\}_{\tau=0}^{t}, \Theta_{i,-1}\right)$$

- **ASF tells us how households map macro signals to expectations**

Goal 2: Estimate average marginal effects (AMEs)

- AME is partial derivative of $y_{i,t+1|t}$ with respect to some signal $Z_{i,t}^{j}$:

$$\beta^{j} = \mathbb{E}\left[\frac{\partial g\left(\{Z_{i,\tau}\}_{\tau=0}^{t}, \Theta_{i,-1}\right)}{\partial Z_{i,t}^{j}}\right]$$

- **AMEs shed light on households weigh different signals in expectations formation**

Recap: Methodology

Use RNNs to estimate ASF

- RNNs provide non-parametric estimator that accounts for time-dependency of expectations on historical signals
- Flexible functional form as opposed to strong parametric restrictions of previous work

Recap: Methodology

Use RNNs to estimate ASF

- RNNs provide non-parametric estimator that accounts for time-dependency of expectations on historical signals
- Flexible functional form as opposed to strong parametric restrictions of previous work

Use DML to estimate AMEs

- Want partial derivative of ASF wrt. signal
- RNN-estimated ASF is biased
- Naive AME estimates will be biased and not $\sqrt{n}$—consistent
- DML provides way to de-bias estimated AMEs

What are RNNs?

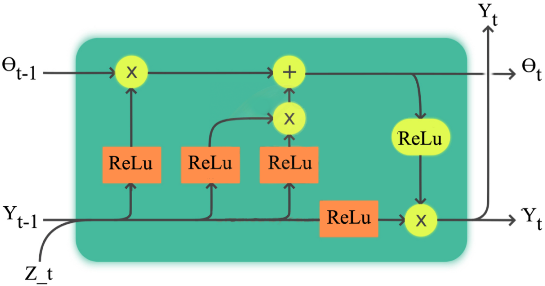
RNNs are a class of neural networks that use previous output as input

- Allow us to non-parametrically estimate ASF:

$$y_{i,t+1|t} \equiv g\left(\{Z_{i,\tau}\}_{\tau=0}^{t}, \Theta_{i,-1}\right)$$

This paper: Long short term memory networks (LSTMs)

- Z_t is signal, Y_t is fitted expectation
- Θ_t is "cell state" = keeps track of "long-term memory"
- Y_t, Θ_t are non-linear functions of $(Z_t, Y_{t-1}, \Theta_{t-1})$



Recap: Empirical Results

ASF is nonlinear

- Asymmetry: Expectations respond more strongly to bad news than good
- Non-linearity strongest for unemployment, business conditions.

Time-varying AMEs

- Countercyclical AMEs for professional forecasts, procyclical AMEs for contemporaneous signals
- Households learn adaptively in good times, become forward-looking in bad times

Economic conditions signals drive AME time-variation

- Signals of economic conditions explain majority of variance in AMEs of both current and forward-looking signals
- Households endogenously change how they weigh information across business cycle

Recap: Economic Content

New stylized facts discipline macro learning models

- FIRE would need alot of known time-variation in underlying environment to generate this variation in AMEs
- Noisy & constant gain learning usually assume linear expectations formation that leads to constant AMEs

Variant on standard rational inattention model is consistent with these facts

- In rational inattention, information structure is endogenous
- Agents can pay to acquire information that can affect their consumption decision

What's new here?

- Agent's can learn about current and future state of economy via two different signals
- Capture nonlinearity in ASF by solving numerically instead of by linear-quadratic approximation

Comments & Questions

Innovative methodology enables better measurement

- RNNs provide good way to loosen standard functional forms to measure empirical ASF
- DML assuages concerns about estimation efficiency

Measured ASF yields useful economic restrictions

- New stylized facts help discipline learning models
- Discussion of how and why rational inattention matches facts is helpful

Comments & Questions

Can you do more to demonstrate usefulness of RNN?

- Does RNN outperform standard parametric models in-sample?
  - Show us RNN approach fits expectations data better than parametric approaches

$$\hat{g}_{rnn} := \underset{g_w \in \mathcal{G}_{foh}^{RNN}}{\arg\min} \sum_{i,t} \frac{1}{2} \left(y_{i,t+1|t} - g_w \left(\{Z_{i,\tau}\}_{\tau=0}^{t} \right) \right)^2$$

- Does RNN outperform out-of-sample?
  - Convince us RNN is not overfitting to noise by showing it matches expectations out-of-sample

Can you do more to eliminate alternative models?

- Use future macro variables as RNN estimation target (i.e. use RNN to forecast)
- Compare AMEs for future macro data vs. AMEs for expectations data
- Wedge between these AMEs could shed further light on deviations from rational expectations

Can you go higher-dimensional?

- Use more signals to better approximate household information set
- E.g. All of FRED-MD, text data, etc.

Other comments

- Discuss in more detail how RNN and DML work as well as the intuition for these methods
    - Current exposition requires much pre-existing knowledge
    - Fleshing these details out would make paper more self-contained
- Give more implementation details:
    - How exactly do you structure the LSTM?
    - How do perform sample splitting for DML?
    - How do you perform the variance decomposition in Table 4?
- Simplify figures and tables:
    - E.g. Figure 4 is hard to interpret
- Simplify notation