Simple Models and Biased Forecasts

Pooya Molavi

July 12, 2022

NBER Summer Institute

Motivation

Rational expectations: agents forecast *as if* they knew the true model of the economy.

- But why can we assume that agents know the true model?
- Learning literature: let's "make the agents like econometricians."

Motivation

Rational expectations: agents forecast *as if* they knew the true model of the economy.

- But why can we assume that agents know the true model?
- Learning literature: let's "make the agents like econometricians."

This paper: agents as econometricians who use flexible but simple time-series models.

Motivation

Rational expectations: agents forecast *as if* they knew the true model of the economy.

- But why can we assume that agents know the true model?
- Learning literature: let's "make the agents like econometricians."

This paper: agents as econometricians who use flexible but simple time-series models.

(1) Agents entertain a large and flexible class of time-series models.

- all state-space models of a given dimension

Motivation

Rational expectations: agents forecast *as if* they knew the true model of the economy.

- But why can we assume that agents know the true model?
- Learning literature: let's "make the agents like econometricians."

This paper: agents as econometricians who use flexible but simple time-series models.

(1) Agents entertain a large and flexible class of time-series models.
  - all state-space models of a given dimension

(2) All models in the class are too simple relative to the truth, i.e., they are misspecified.
  - the models are low-dimensional

Motivation

Rational expectations: agents forecast *as if* they knew the true model of the economy.

- But why can we assume that agents know the true model?
- Learning literature: let's "make the agents like econometricians."

This paper: agents as econometricians who use flexible but simple time-series models.

(1) Agents entertain a large and flexible class of time-series models.
  - all state-space models of a given dimension

(2) All models in the class are too simple relative to the truth, i.e., they are misspecified.
  - the models are low-dimensional

(3) Study the long-run limit when learning is complete.
  - agents settle on pseudo-true models that approximate the true model

The Framework

Individual Problem

- Discrete-time economy with a single agent (for now).

Individual Problem

- Discrete-time economy with a single agent (for now).
- A sequence of observables $\{y_t\}_{t=-\infty}^{\infty}$.
  - $y_t \in \mathbb{R}^n$

Individual Problem

- Discrete-time economy with a single agent (for now).
- A sequence of observables $\{y_t\}_{t=-\infty}^{\infty}$.
  - $y_t \in \mathbb{R}^n$
- Observables are distributed according to some probability distribution $\mathbb{P}$.
  - for now, assume $\mathbb{P}$ is exogenous (will be endogenous in GE)
  - mean zero, stationary, and Gaussian
  - expectation operator $\mathbb{E}$

Individual Problem

- Discrete-time economy with a single agent (for now).
- A sequence of observables $\{y_t\}_{t=-\infty}^{\infty}$.
  - $y_t \in \mathbb{R}^n$
- Observables are distributed according to some probability distribution $\mathbb{P}$.
  - for now, assume $\mathbb{P}$ is exogenous (will be endogenous in GE)
  - mean zero, stationary, and Gaussian
  - expectation operator $\mathbb{E}$
- The agent is attempting to forecast future values of the observables.
  - time-t information set is $\{y_\tau\}_{\tau=-\infty}^{t}$
  - agent uses a model to map past observables to her forecasts:

$$\theta : \{y_\tau\}_{\tau=-\infty}^{t} \mapsto E_t[\cdot]$$

State-Space Models

Main assumption: the agent can only entertain stationary ergodic distributions P that can be represented by state-space models with at most d states.

$$
\begin{aligned}
z_t &= A z_{t-1} + w_t, \qquad & w_t &\sim \text{ i.i.d. } \mathcal{N}(0, Q) \\
y_t &= B' z_t + v_t, \qquad & v_t &\sim \text{ i.i.d. } \mathcal{N}(0, R)
\end{aligned}
$$

- $z \in \mathbb{R}^d$ is the set of subjective state variables.

State-Space Models

Main assumption: the agent can only entertain stationary ergodic distributions P that can be represented by state-space models with at most d states.

$$
\begin{aligned}
z_t &= A z_{t-1} + w_t, & w_t &\sim \text{ i.i.d. } \mathcal{N}(0, Q) \\
y_t &= B' z_t + v_t, & v_t &\sim \text{ i.i.d. } \mathcal{N}(0, R)
\end{aligned}
$$

- $z \in \mathbb{R}^d$ is the set of subjective state variables.
- d captures the agent's sophistication.
  - large $d \longrightarrow$ back to RE
  - small $d \longrightarrow$ model misspecification

State-Space Models

Main assumption: the agent can only entertain stationary ergodic distributions P that can be represented by state-space models with at most d states.

$$
\begin{aligned}
z_t &= A z_{t-1} + w_t, \qquad & w_t &\sim \text{ i.i.d. } \mathcal{N}(0, Q) \\
y_t &= B' z_t + v_t, \qquad & v_t &\sim \text{ i.i.d. } \mathcal{N}(0, R)
\end{aligned}
$$

- $z \in \mathbb{R}^d$ is the set of subjective state variables.
- d captures the agent's sophistication.
  - large $d \longrightarrow$ back to RE
  - small $d \longrightarrow$ model misspecification
- d is the only free parameter.
- $\theta \equiv (A, B, Q, R)$ is estimated endogenously by the agent.

A Dichotomy

A dichotomy: model θ is unconstrained other than the constraint on d.

- The agent can entertain *any* linear cross-sectional relationship between variables.
- But is constrained in the types of time-series relationships she can perceive.

A Dichotomy

A dichotomy: model θ is unconstrained other than the constraint on d.

- The agent can entertain *any* linear cross-sectional relationship between variables.
- But is constrained in the types of time-series relationships she can perceive.

Stark assumption, but...

- It allows me to focus on the difficulty of dealing with time-series complexity.
  - Cross-sectional complexity is the focus of rational inattention, sparsity, etc.
- It leads to a useful linear invariance property for expectations.

▹ limited memory interpretation

Pseudo-True Simple Models

Goodness-of-fit measure: Kullback–Leibler Divergence Rate

$$\mathrm{KLDR}(\theta) \equiv \lim_{t\to\infty} \frac{1}{t} \mathbb{E}\left[\log\left(\frac{\mathbb{f}(y_1,\ldots,y_t)}{f^{\theta}(y_1,\ldots,y_t)}\right)\right]$$

- f^{θ} is the agent's subjective density of $\{y_t\}_t$ under model θ.
- $\mathbb{f}$ is the density and $\mathbb{E}[\cdot]$ is the expectation under the true DGP.

Pseudo-True Simple Models

Goodness-of-fit measure: Kullback–Leibler Divergence Rate

$$\mathrm{KLDR}(\theta) \equiv \lim_{t\to\infty} \frac{1}{t} \mathbb{E}\left[\log\left(\frac{\mathbb{f}(y_1,\ldots,y_t)}{f^{\theta}(y_1,\ldots,y_t)}\right)\right]$$

Definition (Pseudo-True d-State Models)

θ^* is a *pseudo-true d-state model* if

$$\theta^* \in \arg\min_{\theta\in\Theta^d} \mathrm{KLDR}(\theta)$$

where

$$\Theta^d \equiv \{\text{all } d\text{-state models } \theta = (A, B, Q, R)\}$$

▹ learning foundations

Pseudo-True Simple Models

Definition (Pseudo-True d-State Models)

θ^* is a *pseudo-true d-state model* if

$$\theta^* \in \arg\min_{\theta \in \Theta^d} \mathrm{KLDR}(\theta)$$

where

$$\Theta^d \equiv \{\text{all } d\text{-state models } \theta = (A, B, Q, R)\}$$

Agent recovers the *true* model if we replace $\theta \in \Theta^d$ with $\theta \in \bigcup_{d=0}^{\infty} \Theta^d$.

Relation to the Literature

- **Factor analysis of business-cycle:** Sargent-Sims (1977), Watson (2004), Angeletos-Collard-Dellas (2020)
    - endogenizes the “main business-cycle shock” of Angeletos et al.
- **Noisy information/rational inattention/sparsity:** Mankiw-Reis (2002), Sims (2003), Woodford (2003), Gabaix (2014), Angeletos-Lian (2018), ...
    - perfect knowledge of current variables
    - perfect understanding of intratemporal relationships
    - can only understand simple intertemporal relationships
- **Learning models in macro:** Marcet-Sargent (1989), Evans-Honkapohja (1995), Adam-Marcet (2011), ...
    - focus on the asymptotics of learning
    - prior rules out the true model
- **Misspecified learning:** Berk(1966), Esponda-Pouzo (2016, 2021), Molavi (2019), ...

The Plan

The rest of the talk...

1. Characterization of pseudo-true 1-state models.
2. (Some) implications for agents' forecasts and actions.
3. Impulse and propagation: the TFP shock in the RBC model.
4. Application to forward guidance in the NK model.

In the paper (but not the talk)...

1. Generalization to the $d > 1$ case.
2. Additional implications.
3. Propagation of productivity and separation shocks in the DMP model.

Pseudo-True 1-State Models

An Invariance Result

Theorem (Linear Invariance)

Consider two agents:

- *Agent i observes y_t with distribution $\mathbb{P}$ and uses a pseudo-true model given $\mathbb{P}$.*
- *Agent j observes $\tilde{y}_t = T y_t$ with distribution $\tilde{\mathbb{P}}$ and uses a pseudo-true model given $\tilde{\mathbb{P}}$.*

An Invariance Result

Theorem (Linear Invariance)

Consider two agents:

- *Agent i observes y_t with distribution $\mathbb{P}$ and uses a pseudo-true model given $\mathbb{P}$.*
- *Agent j observes $\tilde{y}_t = T y_t$ with distribution $\tilde{\mathbb{P}}$ and uses a pseudo-true model given $\tilde{\mathbb{P}}$.*

For any full-rank matrix T,

$$E^*_{jt}[\tilde{y}_{t+s}] = T E^*_{it}[y_{t+s}]$$

An Invariance Result

Theorem (Linear Invariance)

Consider two agents:

- *Agent i observes y_t with distribution $\mathbb{P}$ and uses a pseudo-true model given $\mathbb{P}$.*
- *Agent j observes $\tilde{y}_t = T y_t$ with distribution $\tilde{\mathbb{P}}$ and uses a pseudo-true model given $\tilde{\mathbb{P}}$.*

For any full-rank matrix T,

$$E^*_{jt}[\tilde{y}_{t+s}] = T E^*_{it}[y_{t+s}]$$

Pseudo-true simple models respect *all* linear intratemporal relationships...

$$E^*_t[\alpha y_{1,t+s} + \beta y_{2,t+s}] = \alpha E^*_t[y_{1,t+s}] + \beta E^*_t[y_{2,t+s}]$$

Autocovariances and Autocorrelations

The 1-state pseudo-true model turns out to depend on the true DGP *only* via the true autocorrelations.

- True autocovariance matrices (standard definition):

$$\Gamma_l \equiv \mathbb{E}[y_t y_{t-l}']$$

Autocovariances and Autocorrelations

The 1-state pseudo-true model turns out to depend on the true DGP *only* via the true autocorrelations.

- True autocovariance matrices (standard definition):

$$\Gamma_l \equiv \mathbb{E}[y_t y'_{t-l}]$$

- True autocorrelation matrices (not a standard definition):

$$C_l \equiv \Gamma_0^{-1} \left(\frac{\Gamma_l + \Gamma'_l}{2} \right)$$

  - products of two symmetric matrices $\implies$ real eigenvalues
  - reduce to the usual autocorrelations when $n = 1$
  - by linear invariance, can assume without loss that Γ_0 is invertible

Main Characterization Result

Theorem

Let λ denote the eigenvalue of C_1 largest in magnitude.

Let p and q denote the corresponding right and left eigenvectors (normalized: $q'p = 1$).

Main Characterization Result

Theorem

Let λ denote the eigenvalue of C_1 largest in magnitude.

Let p and q denote the corresponding right and left eigenvectors (normalized: $q'p = 1$).

If an ergodicity assumption is satisfied, then given any pseudo-true 1-state model

$$E_t^*[z_t] = p' y_t$$

$$E_t^*[z_{t+s}] = \lambda^s E_t^*[z_t]$$

$$E_t^*[y_{t+s}] = q E_t^*[z_{t+s}]$$

▹ ergodicity assumption

Main Characterization Result

Theorem

Let λ denote the eigenvalue of C_1 largest in magnitude.

Let p and q denote the corresponding right and left eigenvectors (normalized: $q'p = 1$).

If an ergodicity assumption is satisfied, then given any pseudo-true 1-state model

$$E_t^*[z_t] = p'y_t$$

$$E_t^*[z_{t+s}] = \lambda^s E_t^*[z_t]$$

$$E_t^*[y_{t+s}] = qE_t^*[z_{t+s}]$$

▹ ergodicity assumption

Main Characterization Result

Theorem

Let λ denote the eigenvalue of C_1 largest in magnitude.

Let p and q denote the corresponding right and left eigenvectors (normalized: $q'p = 1$).

If an ergodicity assumption is satisfied, then given any pseudo-true 1-state model

$$E_t^*[z_t] = p'y_t$$

$$E_t^*[z_{t+s}] = \lambda^s E_t^*[z_t]$$

$$E_t^*[y_{t+s}] = qE_t^*[z_{t+s}]$$

▹ ergodicity assumption

Main Characterization Result

Theorem

Let λ denote the eigenvalue of C_1 largest in magnitude.

Let p and q denote the corresponding right and left eigenvectors (normalized: $q'p = 1$).

If an ergodicity assumption is satisfied, then given any pseudo-true 1-state model

$$E_t^* [y_{t+s}] = \lambda^s q p' y_t$$

▹ ergodicity assumption

A Diagonal Example

- True data-generating process:

$$y_t = \begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_n \end{pmatrix} y_{t-1} + \begin{pmatrix} b_1 & 0 & \dots & 0 \\ 0 & b_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & b_n \end{pmatrix} \epsilon_t, \qquad \epsilon_t \sim \text{i.i.d. } \mathcal{N}(0, I)$$

  where

$$|\alpha_1| > |\alpha_2| > \cdots > |\alpha_n|$$

A Diagonal Example

- True data-generating process:

$$y_t = \begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_n \end{pmatrix} y_{t-1} + \begin{pmatrix} b_1 & 0 & \dots & 0 \\ 0 & b_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & b_n \end{pmatrix} \epsilon_t, \qquad \epsilon_t \sim \text{i.i.d. } \mathcal{N}(0, I)$$

  where

$$|\alpha_1| > |\alpha_2| > \cdots > |\alpha_n|$$

- The autocorrelation matrix at lag $l = 1$:

$$C_1 = \begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_n \end{pmatrix}$$

Pseudo-True Model in the Diagonal Example

- Eigenvalue and eigenvectors...

$$\lambda = \alpha_1, \qquad q = p = (1, 0, \ldots, 0)'$$

Pseudo-True Model in the Diagonal Example

- Eigenvalue and eigenvectors...

$$\lambda = \alpha_1, \qquad q = p = (1, 0, \ldots, 0)'$$

- Forecasts:

$$\begin{aligned} E_t^*[y_{1,t+s}] &= \alpha_1^s y_{1t} \\ E_t^*[y_{j,t+s}] &= 0, \qquad \forall j \neq 1 \end{aligned}$$

Pseudo-True Model in the Diagonal Example

- Eigenvalue and eigenvectors...

$$\lambda = \alpha_1, \qquad q = p = (1, 0, \ldots, 0)'$$

- Forecasts:

$$E_t^*[y_{1,t+s}] = \alpha_1^s y_{1t}$$
$$E_t^*[y_{j,t+s}] = 0, \qquad \forall j \neq 1$$

- Persistence bias: forecasts are anchored to the most persistent observable.
  - Forecasts of y_1 coincide with the RE.
  - Forecast y_j for $j \neq 1$ as if i.i.d.

Second Moments

Theorem

The subjective variance under the pseudo-true model coincides with the true variance:

$$E^*[y_t y_t'] = \mathbb{E}[y_t y_t']$$

Second Moments

Theorem

The subjective variance under the pseudo-true model coincides with the true variance:

$$E^*[y_t y_t'] = \mathbb{E}[y_t y_t']$$

- PCA:
  - Project onto the dominant eigenvectors of the variance-covariance matrix, Γ_0.
  - Purely cross-sectional; uses no information about the serial correlations.
- Pseudo-true simple models:
  - Project onto the dominant eigenvectors of the first autocorrelation matrix, C_1.
  - No simplification in the cross section; perfectly matches Γ_0.

Some Implications

Unidimensional Dynamics and a Main Shock

- Recall that...

$$E_t^*\left[y_{t+s}\right] = \lambda^s q p' y_t$$

Unidimensional Dynamics and a Main Shock

- Recall that...

$$E_t^*[y_{t+s}] = \lambda^s q p' y_t$$

- The agent decomposes the observed y_t into the sum of two terms:

$$y_t = \underbrace{y_t^{\parallel}}_{\text{parallel to } p} + \underbrace{y_t^{\perp}}_{\text{orthogonal to } p}$$

Unidimensional Dynamics and a Main Shock

- Recall that...

$$E_t^*[y_{t+s}] = \lambda^s q p' y_t$$

- The agent decomposes the observed y_t into the sum of two terms:

$$y_t = \underbrace{y_t^{\parallel}}_{\text{parallel to } p} + \underbrace{y_t^{\perp}}_{\text{orthogonal to } p}$$

- Update expectations in response to $y_t^{\parallel}$.

Unidimensional Dynamics and a Main Shock

- Recall that...

$$E_t^*[y_{t+s}] = \lambda^s q p' y_t$$

- The agent decomposes the observed y_t into the sum of two terms:

$$y_t = \underbrace{y_t^{\|}}_{\text{parallel to } p} + \underbrace{y_t^{\perp}}_{\text{orthogonal to } p}$$

- Update expectations in response to $y_t^{\|}$.
- Does not update expectations *at all* in response to $y_t^{\perp}$.

Unidimensional Dynamics and a Main Shock

- Recall that...

$$E_t^*[y_{t+s}] = \lambda^s q p' y_t$$

- The agent decomposes the observed y_t into the sum of two terms:

$$y_t = \underbrace{y_t^{\|}}_{\text{parallel to } p} + \underbrace{y_t^{\perp}}_{\text{orthogonal to } p}$$

- Update expectations in response to $y_t^{\|}$.
- Does not update expectations *at all* in response to $y_t^{\perp}$.
- Forecasts behave as if the economy is driven by a single "main shock."

Linear best responses:

$$x_{jt} = E_t\left[\sum_{s=1}^{\infty} \beta_j^s c_j' y_{t+s}\right]$$

Comovement

Linear best responses:

$$x_{jt} = E_t \left[\sum_{s=1}^{\infty} \beta_j^s c_j' y_{t+s} \right]$$

Result: for any two forward-looking choices j and k

$$1 = \left| \mathbb{corr} \left(x_{jt}^*, x_{kt}^* \right) \right| \geq \left| \mathbb{corr} \left(x_{jt}^{\mathrm{RE}}, x_{kt}^{\mathrm{RE}} \right) \right|$$

Intuition:

- Expectations are unresponsive to $y_t^{\perp}$.
- It is as if there is a single shock $y_t^{\parallel}$ driving everything.
- This increases the comovement of different choices.

TFP Shocks in the RBC Model

The Loglinearized RBC Model

- TFP:

$$a_t = \rho a_{t-1} + \epsilon_t$$

- Capital:

$$k_t = (1-\delta)k_{t-1} + \delta i_{t-1}$$

- Consumption (permanent income hypothesis):

$$c_t = \frac{\chi}{\beta} k_t + \chi r_t + \chi \zeta w_t + (\chi - \beta\sigma) \sum_{s=1}^{\infty} \beta^s E_t [r_{t+s}] + \chi \zeta \sum_{s=1}^{\infty} \beta^s E_t [w_{t+s}]$$

- True for arbitrary expectations that satisfy the LIE.
    - the aggregate Euler equation may *not* be valid away from RE (Preston, 2005)

- r_t, w_t, and i_t are linear functions of k_t, a_t, and c_t.

Calibration and Agents' Pseudo-True Model

- Textbook calibration of primitives.
- Persistence of TFP shock: $\rho = 0.95$. The variance is irrelevant.
- Agents have full information.

Calibration and Agents' Pseudo-True Model

- Textbook calibration of primitives.
- Persistence of TFP shock: $\rho = 0.95$. The variance is irrelevant.
- Agents have full information.

- d is the *only* additional free parameter.
- $d \geq 2 \implies$ REE
- So, we only need to consider $d = 1$.

Calibration and Agents' Pseudo-True Model

- Textbook calibration of primitives.
- Persistence of TFP shock: $\rho = 0.95$. The variance is irrelevant.
- Agents have full information.

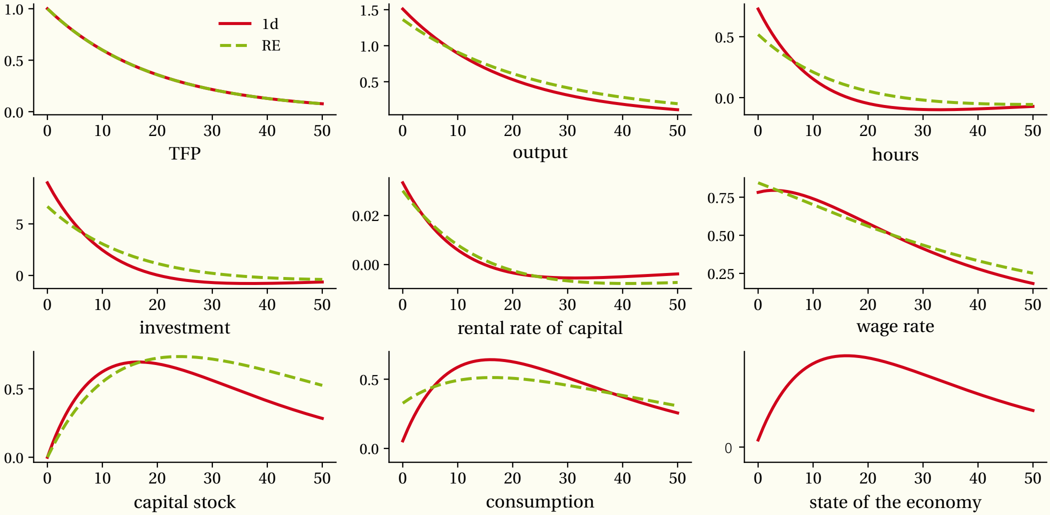
- d is the *only* additional free parameter.
- $d \geq 2 \implies$ REE
- So, we only need to consider $d = 1$.

- Agents' nowcast of the subjective state

$$E_t^*[z_t] = p' y_t = 0.947 k_t + 0.053 a_t$$

- Expectations move (almost) one-for-one with changes in capital stock.
    - almost no (direct) response to changes in TFP
    - persistence bias

Impulse Response Functions



Forward Guidance

The Two-Equation New-Keynesian Model

- Dynamic IS curve:

$$\hat{x}_t = -\sigma\left(\hat{i}_t - r_t^n\right) + E_t^h\left[\sum_{s=1}^{\infty}\beta^s\left(\frac{1-\beta}{\beta}\hat{x}_{t+s} - \sigma\left(\hat{i}_{t+s} - r_{t+s}^n\right) - \frac{\sigma}{\beta}\hat{\pi}_{t+s}\right)\right]$$

- NK Phillips curve:

$$\hat{\pi}_t = \kappa\hat{x}_t + \mu_t + E_t^f\left[\sum_{s=1}^{\infty}(\beta\delta)^s\left(\kappa\hat{x}_{t+s} + \frac{1-\delta}{\delta}\hat{\pi}_{t+s} + \mu_{t+s}\right)\right]$$

- No need for the Taylor principle.
  - assume observables *cannot* include sunspots
  - equilibrium is determinate as long as actions are measurable with respect to observables
- Natural rate, cost push-up, and interest-rate shocks.

- Textbook calibration of primitives: $\beta = 0.99$, $\sigma = 1$, $\delta = 3/4$, and $\kappa = 0.172$.

Information, Calibration, and Estimation

- Textbook calibration of primitives: $\beta = 0.99$, $\sigma = 1$, $\delta = 3/4$, and $\kappa = 0.172$.

- d is the *only* additional free parameter.
- Assume $d = 1$.
- Assume agents have full information.

Information, Calibration, and Estimation

- Textbook calibration of primitives: $\beta = 0.99$, $\sigma = 1$, $\delta = 3/4$, and $\kappa = 0.172$.

- d is the *only* additional free parameter.
- Assume $d = 1$.
- Assume agents have full information.

- Choose the DGP for $(r_t^n, \mu_t, \hat{i}_t)$ to target the autocovariance of $(\hat{x}_t, \hat{\pi}_t, \hat{i}_t)$ at lags 0, 1.
  - only identifies the autocovariance of $(r_t^n, \mu_t, \hat{i}_t)$ at lags 0, 1
  - sufficient for my analysis
  - can hit the target perfectly

Pseudo-True 1-State Model

Agents' nowcast of the subjective state:

$$E_t^*[z_t] = p'y_t = 0.022\hat{x}_t - 0.42\hat{\pi}_t - 0.014\hat{i}$$

- High inflation and high output gap have opposite effects on the agents' nowcast.
- Expectations respond a lot to inflation, not so much to the nominal rate.

Pseudo-True 1-State Model

Agents' nowcast of the subjective state:

$$E_t^*[z_t] = p' y_t = 0.022\hat{x}_t - 0.42\hat{\pi}_t - 0.014\hat{i}$$

- High inflation and high output gap have opposite effects on the agents' nowcast.
- Expectations respond a lot to inflation, not so much to the nominal rate.

Persistence of the subjective state:

$$E_t^*[z_{t+1}] = 0.985 E_t^*[z_t]$$

- Larger than the estimated persistence of any of the shocks.
- But *not* unit root.

Forward Guidance

- A credible commitment by monetary authority at time t to

$$\{\hat{i}_{t+1}, \hat{i}_{t+2}, \ldots, \hat{i}_{t+T}\}$$

Forward Guidance

- A credible commitment by monetary authority at time t to

$$\{\hat{i}_{t+1}, \hat{i}_{t+2}, \ldots, \hat{i}_{t+T}\}$$

- Agents' information set at time t under forward guidance:

$$\omega_T \equiv \{\ldots, y_{t-1}, y_t, \hat{i}_{t+2}, \ldots, \hat{i}_{t+T}\}$$

Forward Guidance

- A credible commitment by monetary authority at time t to

$$\{\hat{i}_{t+1}, \hat{i}_{t+2}, \ldots, \hat{i}_{t+T}\}$$

- Agents' information set at time t under forward guidance:

$$\omega_T \equiv \{\ldots, y_{t-1}, y_t, \hat{i}_{t+2}, \ldots, \hat{i}_{t+T}\}$$

- Subjective expectations:

$$E_t^{*-\mathrm{FG}}[\cdot] = E^*[\cdot|\omega_T]$$

Forward Guidance

- A credible commitment by monetary authority at time t to

$$\{\hat{i}_{t+1}, \hat{i}_{t+2}, \ldots, \hat{i}_{t+T}\}$$

- Agents' information set at time t under forward guidance:

$$\omega_T \equiv \{\ldots, y_{t-1}, y_t, \hat{i}_{t+2}, \ldots, \hat{i}_{t+T}\}$$

- Subjective expectations:

$$E_t^{*-\mathrm{FG}}[\cdot] = E^*[\cdot|\omega_T]$$

- Agents' subjective model is Gaussian. $\implies$ Conditioning is easy!

$$E^*[\zeta_{t+s}|\omega_T] = \Sigma_{\zeta_s \omega_T} \Sigma^{-1}_{\omega_T \omega_T} \omega_T$$

Forward Guidance

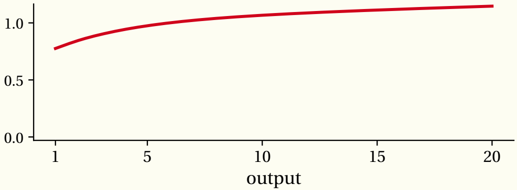
$$\hat{x}_t = \nu_{xi}^{(T)} \hat{i}_t + \nu_{xn}^{(T)} r_t^n + \nu_{x\mu}^{(T)} \mu_t + \sum_{s=1}^{T} \nu_{xi_s}^{(T)} \hat{i}_{t+s}$$

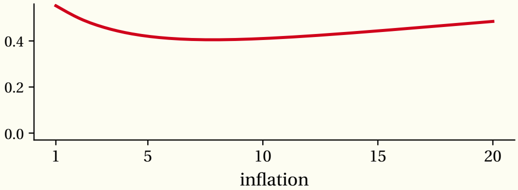
$$\hat{\pi}_t = \nu_{\pi i}^{(T)} \hat{i}_t + \nu_{\pi n}^{(T)} r_t^n + \nu_{\pi\mu}^{(T)} \mu_t + \sum_{s=1}^{T} \nu_{\pi i_s}^{(T)} \hat{i}_{t+s}$$

Forward Guidance

$$\hat{x}_t = \nu_{xi}^{(T)} \hat{i}_t + \nu_{xn}^{(T)} r_t^n + \nu_{x\mu}^{(T)} \mu_t + \sum_{s=1}^{T} \nu_{xi_s}^{(T)} \hat{i}_{t+s}$$

$$\hat{\pi}_t = \nu_{\pi i}^{(T)} \hat{i}_t + \nu_{\pi n}^{(T)} r_t^n + \nu_{\pi\mu}^{(T)} \mu_t + \sum_{s=1}^{T} \nu_{\pi i_s}^{(T)} \hat{i}_{t+s}$$





Concluding Remarks

- I assume agents forecast using simple models that are fit to the DGP.
- Illustration in the context of the RBC and NK models.
- The framework can be embedded in workhorse macro models.

Concluding Remarks

- I assume agents forecast using simple models that are fit to the DGP.
- Illustration in the context of the RBC and NK models.
- The framework can be embedded in workhorse macro models.

- The Julia code is available on my website.
- Paper with Alireza Tahbaz-Salehi and Andrea Vedolin: asset-pricing implications.

A Complementary Interpretation: Limited Memory

Models with d running statistics

- The running statistics

$$s_t \in \mathbb{R}^d$$

- The running statistics are updated linearly over time:

$$s_t = M s_{t-1} + K y_t$$

- Forecasts are linear functions of the running statistics:

$$E_t [y_{t+\tau}] = v'_\tau s_t$$

A Complementary Interpretation: Limited Memory

Models with d running statistics

- The running statistics

$$s_t \in \mathbb{R}^d$$

- The running statistics are updated linearly over time:

$$s_t = M s_{t-1} + K y_t$$

- Forecasts are linear functions of the running statistics:

$$E_t [y_{t+\tau}] = v_\tau' s_t$$

An equivalence result: d-state models $\approx$ models with d running statistics

▹ back

Theorem (Huber, White, Douc-Moulines, ...)

Assume the agent estimates the parameters of the model $\theta \equiv (A, B, Q, R)$ *using (quasi-)MLE. Then asymptotically her point estimate converges to a pseudo-true model.*

Learning Foundations

Theorem (Huber, White, Douc-Moulines, ...)

Assume the agent estimates the parameters of the model $\theta \equiv (A, B, Q, R)$ using (quasi-)MLE. Then asymptotically her point estimate converges to a pseudo-true model.

Theorem (Berk, Bunke-Milhaud, Shalizi, ...)

Assume the agent starts with a full-support prior over the set of d-state models and updates her prior over time using Bayes' rule. Asymptotically the agent's belief concentrates over the set of pseudo-true models.

▹ back

Ergodicity Assumption

Assumption

For all $l \geq 1$

$$\rho(C_l) \leq \rho(C_1)^l$$

where

$$\rho(C_l) = \max\{|\lambda| : \lambda \text{ is an eigenvalue of } C_l\}$$

- Requires autocorrelations to decay sufficiently fast.
- Satisfied for many commonly used specifications.
- Satisfied in all the applications studied in this talk.

▹ back