

Simple Models and Biased Forecasts

Pooya Molavi

July 23, 2023

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?

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?

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

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?

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

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?

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

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?

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 where agents' estimates have settled.
 - 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 ergodic, and purely non-deterministic
 - finite second moment and entropy rate
 - 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 ergodic, and purely non-deterministic
 - finite second moment and entropy rate
 - 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:

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

State-Space Models

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

$$z_t = Az_{t-1} + w_t, \quad w_t \sim \text{i.i.d. } \mathcal{N}(0, Q)$$

$$y_t = B'z_t + v_t, \quad v_t \sim \text{i.i.d. } \mathcal{N}(0, R)$$

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

State-Space Models

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

$$z_t = Az_{t-1} + w_t, \quad w_t \sim \text{i.i.d. } \mathcal{N}(0, Q)$$

$$y_t = B'z_t + v_t, \quad v_t \sim \text{i.i.d. } \mathcal{N}(0, R)$$

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

State-Space Models

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

$$z_t = Az_{t-1} + w_t, \quad w_t \sim \text{i.i.d. } \mathcal{N}(0, Q)$$

$$y_t = B'z_t + v_t, \quad v_t \sim \text{i.i.d. } \mathcal{N}(0, R)$$

- $z \in \mathbb{R}^d$ is the set of subjective state variables.
- d captures the agent's sophistication.
 - large $d \rightarrow$ back to RE
 - small $d \rightarrow$ model misspecification
- d is the only free parameter.
- I take d to be (much) smaller than n .

State-Space Models

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

$$z_t = Az_{t-1} + w_t, \quad w_t \sim \text{i.i.d. } \mathcal{N}(0, Q)$$

$$y_t = B'z_t + v_t, \quad v_t \sim \text{i.i.d. } \mathcal{N}(0, R)$$

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

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.

Cross-Sectional and Times-Series Complexity

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.

A Complementary Interpretation: Limited Memory

Models with d running statistics

- Running statistics:

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

- The running statistics are updated linearly over time:

$$s_t = Ms_{t-1} + Ky_t$$

- 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

- Running statistics:

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

- The running statistics are updated linearly over time:

$$s_t = Ms_{t-1} + Ky_t$$

- 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

Endogenizing the Agent's Model

Two ways of specifying the agent's model $\theta \equiv (A, B, Q, R)$:

1. Determined by [maximum likelihood estimation](#) given data from the true DGP.

Endogenizing the Agent's Model

Two ways of specifying the agent's model $\theta \equiv (A, B, Q, R)$:

1. Determined by **maximum likelihood estimation** given data from the true DGP.
2. The agent starts with a prior over d -state models and updates it using **Bayes' rule** given data from the true DGP.

Endogenizing the Agent's Model

Two ways of specifying the agent's model $\theta \equiv (A, B, Q, R)$:

1. Determined by **maximum likelihood estimation** given data from the true DGP.
2. The agent starts with a prior over d -state models and updates it using **Bayes' rule** given data from the true DGP.

Focus on the limit where the agent has infinite data:

- The agent (generically) ends up with the same model using both approaches.
- The limiting model is a **pseudo-true model** that minimizes KLDR:

$$\lim_{t \rightarrow \infty} \frac{1}{t} \mathbb{E} \left[\log \left(\frac{f(y_1, \dots, y_t)}{f^\theta(y_1, \dots, y_t)} \right) \right]$$

The rest of the talk...

1. (Some) qualitative **features** of pseudo-true one-state models.
2. (Some) **implications** for agents' forecasts and actions.
3. Propagation of TFP shocks in the **RBC** model.
4. Propagation of separation and productivity shocks in the **DMP** model.

- **Misspecified learning:** Berk (1966), Huber (1967), White(1982, 1994), Shalizi (2009), Esponda–Pouzo (2016, 2021), ...
 - not about learning foundations
 - characterize properties of pseudo-true models

- **Misspecified learning:** Berk (1966), Huber (1967), White(1982, 1994), Shalizi (2009), Esponda–Pouzo (2016, 2021), ...
 - not about learning foundations
 - characterize properties of pseudo-true models
- **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 **cross-sectional** relationships (a result)
 - can only understand simple **time-series** relationships

- **Misspecified learning:** Berk (1966), Huber (1967), White(1982, 1994), Shalizi (2009), Esponda–Pouzo (2016, 2021), ...
 - not about learning foundations
 - characterize properties of pseudo-true models
- **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 **cross-sectional** relationships (a result)
 - can only understand simple **time-series** relationships
- **Bounded rationality with pseudo-true models:** Sawa (1978), Bray (1982), Bray–Savin (1986), Rabin–Vayanos (2010), Fuster–Laibson–Mendel (2010), Fuster–Hebert–Laibson (2012), Spiegel (2016), Levy–Razin–Young (2021), Molavi–Tahbaz-Salehi–Vedolin (2023), ...

Pseudo-True One-State Models

An Invariance Result

Theorem (linear invariance)

Consider two agents:

- *Agent i observes y_t and uses pseudo-true model θ .*
- *Agent j observes $\tilde{y}_t = Ty_t$ and uses pseudo-true model $\tilde{\theta}$.*

An Invariance Result

Theorem (linear invariance)

Consider two agents:

- Agent i observes y_t and uses pseudo-true model θ .
- Agent j observes $\tilde{y}_t = Ty_t$ and uses pseudo-true model $\tilde{\theta}$.

For any full-rank matrix T ,

$$E_{j,t}^{\tilde{\theta}}[\tilde{y}_{t+s}] = TE_{i,t}^{\theta}[y_{t+s}]$$

An Invariance Result

Theorem (linear invariance)

Consider two agents:

- Agent i observes y_t and uses pseudo-true model θ .
- Agent j observes $\tilde{y}_t = Ty_t$ and uses pseudo-true model $\tilde{\theta}$.

For any full-rank matrix T ,

$$E_{jt}^{\tilde{\theta}}[\tilde{y}_{t+s}] = TE_{it}^{\theta}[y_{t+s}]$$

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

$$E_t^{\theta}[\alpha y_{1,t+s} + \beta y_{2,t+s}] = \alpha E_t^{\theta}[y_{1,t+s}] + \beta E_t^{\theta}[y_{2,t+s}]$$

- Not always true in non-RE models.

No Misperception in the Cross Section

Theorem

Under any pseudo-true model θ ,

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

- The agent correctly perceived *all* the cross-sectional correlations.
- She only misperceives the serial correlations.

No Misperception in the Cross Section

Theorem

Under any pseudo-true model θ ,

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

- The agent correctly perceived *all* the **cross-sectional** correlations.
- She only misperceives the **serial correlations**.
- Different from rational inattention, noisy information, sparsity, etc.

No Misperception in the Cross Section

Theorem

Under any pseudo-true model θ ,

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

- The agent correctly perceived *all* the **cross-sectional** correlations.
- She only misperceives the **serial correlations**.
- Different from rational inattention, noisy information, sparsity, etc.
- **Intuition:** The agent can always match the cross-sectional correlations by an appropriate choice of $\theta = (A, B, Q, R)$. MLE/Bayesian learning leads her to do so.

- Under any pseudo-true one-state model θ ,

$$E_t^\theta[y_{t+s}] = a^s(1 - \eta)qp' \sum_{\tau=0}^{\infty} a^\tau \eta^\tau y_{t-\tau}$$

- Under any pseudo-true one-state model θ ,

$$E_t^\theta[y_{t+s}] = a^s(1 - \eta)qp' \sum_{\tau=0}^{\infty} a^\tau \eta^\tau y_{t-\tau}$$

- The tuple (a, η, p, q) depends on the true autocovariance matrices $\{\Gamma_l\}_{l=0}^{\infty}$, where

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

- Under any pseudo-true one-state model θ ,

$$E_t^\theta[y_{t+s}] = a^s(1 - \eta)qp' \sum_{\tau=0}^{\infty} a^\tau \eta^\tau y_{t-\tau}$$

- The tuple (a, η, p, q) depends on the true autocovariance matrices $\{\Gamma_l\}_{l=0}^{\infty}$, where

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

- (a, η, p, q) is the solution to a non-convex optimization problem.

Exponential Ergodicity and Markovian Models

Definition (exponential ergodicity)

The true process is **exponentially ergodic** if

$$\rho\left(\left(\frac{\Gamma_l + \Gamma'_l}{2}\right) \Gamma_0^\dagger\right) \leq \rho\left(\left(\frac{\Gamma_1 + \Gamma'_1}{2}\right) \Gamma_0^\dagger\right)^l$$

where $\rho(\cdot)$ denotes the spectral radius.

Exponential Ergodicity and Markovian Models

Definition (exponential ergodicity)

The true process is **exponentially ergodic** if

$$\rho \left(\left(\frac{\Gamma_l + \Gamma'_l}{2} \right) \Gamma_0^\dagger \right) \leq \rho \left(\left(\frac{\Gamma_1 + \Gamma'_1}{2} \right) \Gamma_0^\dagger \right)^l$$

where $\rho(\cdot)$ denotes the spectral radius.

Theorem

If the true process is exponentially ergodic, then the pseudo-true one-state model:

- (a) is **Markovian**
- (b) only depends on Γ_0 and Γ_1

So, Which Processes are Exponentially Ergodic?

- A spanning linear combination of independent AR(1) processes:

$$f_{it} = \alpha_i f_{i,t-1} + \epsilon_{it}, \quad i = 1, \dots, m$$

$$y_t = H' f_t, \quad H \text{ rank } m$$

So, Which Processes are Exponentially Ergodic?

- A spanning linear combination of independent AR(1) processes:

$$f_{it} = \alpha_i f_{i,t-1} + \epsilon_{it}, \quad i = 1, \dots, m$$

$$y_t = H' f_t, \quad H \text{ rank } m$$

- The equilibrium processes arising in *all* three macro applications in the paper.

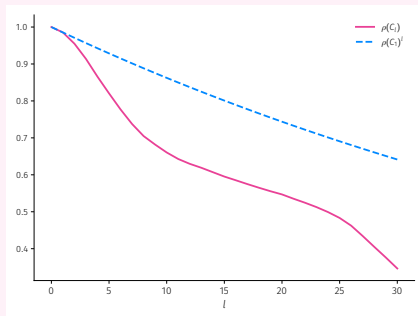
So, Which Processes are Exponentially Ergodic?

- A spanning linear combination of independent AR(1) processes:

$$f_{it} = \alpha_i f_{i,t-1} + \epsilon_{it}, \quad i = 1, \dots, m$$

$$y_t = H' f_t, \quad H \text{ rank } m$$

- The equilibrium processes arising in *all* three macro applications in the paper.
- Estimated (joint) time-series for output gap, inflation rate, and interest rate:



Intuition (1/2)

- Suppose $n = 1$ and the true process is invertible.
- First, suppose $d = \infty$.
- (Pseudo-)true forecasts are given by

$$\mathbb{E}_t[y_{t+1}] = \sum_{\tau=1}^{\infty} \phi_{\tau} y_{t+1-\tau}$$

Intuition (1/2)

- Suppose $n = 1$ and the true process is invertible.
- First, suppose $d = \infty$.
- (Pseudo-)true forecasts are given by

$$\mathbb{E}_t[y_{t+1}] = \sum_{\tau=1}^{\infty} \phi_{\tau} y_{t+1-\tau}$$

- Coefficients $\{\phi_{\tau}\}_{\tau=1}^{\infty}$ are all generically non-zero.
- But the optimal $\{\phi_{\tau}\}_{\tau=1}^{\infty}$ are **fine-tuned** to the true autocorrelation function (ACF).
 - a one-to-one mapping between the two
 - Yule-Walker equations:

$$\text{the ACF} \longleftrightarrow \{\phi_{\tau}\}_{\tau=1}^{\infty}$$

Intuition (2/2)

- Next, suppose $d = 1$.
- Pseudo-true forecasts are still given by

$$\mathbb{E}_t[y_{t+1}] = \sum_{\tau=1}^{\infty} \alpha_{\tau} y_{t+1-\tau}$$

Intuition (2/2)

- Next, suppose $d = 1$.
- Pseudo-true forecasts are still given by

$$\mathbb{E}_t[y_{t+1}] = \sum_{\tau=1}^{\infty} \alpha_{\tau} y_{t+1-\tau}$$

- But setting $d = 1$ imposes the constraint:

$$\alpha_{\tau} = (1 - \eta) a^{\tau} \eta^{\tau-1}$$

- Only two degrees of freedom. \implies **Cannot fine-tune** α_{τ} to fit the true ACF.

Intuition (2/2)

- Next, suppose $d = 1$.
- Pseudo-true forecasts are still given by

$$\mathbb{E}_t[y_{t+1}] = \sum_{\tau=1}^{\infty} \alpha_{\tau} y_{t+1-\tau}$$

- But setting $d = 1$ imposes the constraint:

$$\alpha_{\tau} = (1 - \eta) a^{\tau} \eta^{\tau-1}$$

- Only two degrees of freedom. \implies **Cannot fine-tune** α_{τ} to fit the true ACF.
- With exp. ergodicity, y_t is much more informative (than $y_{t+1-\tau}$ for $\tau > 1$) about y_{t+1} .
- Optimal to not distort how one uses y_t . Instead set $\eta = 0$, and so, $\alpha_{\tau} = 0$ for $\tau > 1$.

Theorem (persistence bias)

If the true process is exponentially ergodic, under any pseudo-true model θ ,

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

*a is the **top eigenvalue** and p and q are the corresponding left and right **eigenvectors** of*

$$\left(\frac{\Gamma_1 + \Gamma'_1}{2} \right) \Gamma_0^\dagger$$

Theorem (persistence bias)

If the true process is exponentially ergodic, under any pseudo-true model θ ,

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

*a is the **top eigenvalue** and p and q are the corresponding left and right **eigenvectors** of*

$$\left(\frac{\Gamma_1 + \Gamma_1'}{2} \right) \Gamma_0^\dagger$$

- The agent's forecasts only depend on the **most persistent component** of y_t .

A Diagonal Example

- True data-generating process:

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

where

$$|\alpha_1| > |\alpha_2| > \dots > |\alpha_n|$$

A Diagonal Example

- True data-generating process:

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

where

$$|\alpha_1| > |\alpha_2| > \dots > |\alpha_n|$$

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

$$\left(\frac{\Gamma_1 + \Gamma_1'}{2} \right) \Gamma_0^\dagger = \begin{pmatrix} \alpha_1 & 0 & \dots & 0 \\ 0 & \alpha_2 & \dots & 0 \\ \vdots & \vdots & \ddots & \vdots \\ 0 & 0 & \dots & \alpha_n \end{pmatrix}$$

Pseudo-True One-State Model in the Diagonal Example

- Eigenvalue and eigenvector...

$$a = \alpha_1, \quad p = q = (1, 0, \dots, 0)'$$

Pseudo-True One-State Model in the Diagonal Example

- Eigenvalue and eigenvector...

$$a = \alpha_1, \quad p = q = (1, 0, \dots, 0)'$$

- Forecasts:

$$E_t^\theta[y_{1,t+s}] = \alpha_1^s y_{1t}$$

$$E_t^\theta[y_{j,t+s}] = 0, \quad \forall j \neq 1$$

Pseudo-True One-State Model in the Diagonal Example

- Eigenvalue and eigenvector...

$$a = \alpha_1, \quad p = q = (1, 0, \dots, 0)'$$

- Forecasts:

$$E_t^\theta[y_{1,t+s}] = \alpha_1^s y_{1t}$$

$$E_t^\theta[y_{j,t+s}] = 0, \quad \forall j \neq 1$$

- Persistence bias:
 - Forecasts of y_1 coincide with the RE.
 - Forecast y_j for $j \neq 1$ as if i.i.d.

Agents behave *as if* only the most persistent component of y_t is relevant.

- Ignore the other components *even though they're observed perfectly*.

Agents behave *as if* only the most persistent component of y_t is relevant.

- Ignore the other components *even though they're observed perfectly*.

Why don't volatilities matter?

- By linear invariance, can scale different components to have the same volatilities!
- Furthermore, the agent can always match the volatilities.

Agents behave *as if* only the most persistent component of y_t is relevant.

- Ignore the other components *even though they're observed perfectly*.

Why don't volatilities matter?

- By linear invariance, can scale different components to have the same volatilities!
- Furthermore, the agent can always match the volatilities.

Why focus on persistent component?

- Any component that is ignored is approximated as i.i.d.
- Less persistent components are closer to being i.i.d.

- Consider J agents, with agent j using a pseudo-true d_j -state model θ_j .
- Suppose each agent takes an action of the form

$$x_{jt} = E_{jt}^{\theta_j} \left[\sum_{s=1}^{\infty} c'_{js} y_{t+s} \right]$$

Increased Comovement

Theorem

*If $d_j = 1$ for all j , the agents' actions are **perfectly correlated**.*

Theorem

*If $d_j = 1$ for all j , the agents' actions are **perfectly correlated**.*

Intuition:

- Agents all decompose the observable in the same way.
- They all *agree* on the persistence ranking of components of y_t .
- Their actions only respond to changes in the most persistent component of y_t .

Theorem

*If $d_j = 1$ for all j , the agents' actions are **perfectly correlated**.*

Intuition:

- Agents all decompose the observable in the same way.
- They all *agree* on the persistence ranking of components of y_t .
- Their actions only respond to changes in the most persistent component of y_t .

A main shock:

- The economy looks *as if* it is driven by a single “**main shock**.”
- Angeletos, Collard, and Dellas (2020) find a “main business cycle shock.”

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 .

Primitives:

- Each period is a quarter.
- Textbook calibration of primitives:

$$\beta = 0.99, \quad \sigma = \phi = 1, \quad \delta = 0.012, \quad \alpha = 0.3, \quad \rho = 0.95$$

Expectations:

- Agents have full information: perfectly observe everything.

Primitives:

- Each period is a quarter.
- Textbook calibration of primitives:

$$\beta = 0.99, \quad \sigma = \phi = 1, \quad \delta = 0.012, \quad \alpha = 0.3, \quad \rho = 0.95$$

Expectations:

- Agents have full information: perfectly observe everything.
- d is the *only* additional free parameter.
- $d \geq 2 \implies$ REE
- So, we only need to consider $d = 1$.

The (endogenously-determined) main shock:

$$0.947k_t + 0.053a_t$$

- Almost perfectly correlated with the capital stock.
- In equilibrium the capital stock is more persistent than TFP (**persistence bias**).
- Agents almost ignore innovations to the TFP.

The (endogenously-determined) main shock:

$$0.947k_t + 0.053a_t$$

- Almost perfectly correlated with the capital stock.
- In equilibrium the capital stock is more persistent than TFP (**persistence bias**).
- Agents almost ignore innovations to the TFP.

This does *not* require imperfect information:

- Agents perfectly observe TFP.
- But find it optimal (in the sense of maximum likelihood) to almost ignore it.

The (endogenously-determined) main shock:

$$0.947k_t + 0.053a_t$$

- Almost perfectly correlated with the capital stock.
- In equilibrium the capital stock is more persistent than TFP (**persistence bias**).
- Agents almost ignore innovations to the TFP.

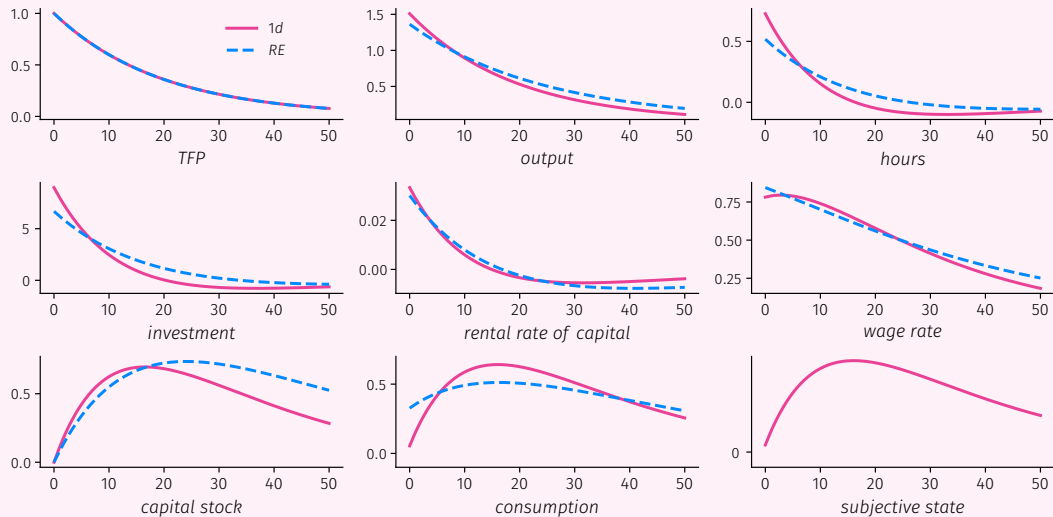
This does *not* require imperfect information:

- Agents perfectly observe TFP.
- But find it optimal (in the sense of maximum likelihood) to almost ignore it.

Sluggishness:

- Expectations move almost one-for-one with capital.
- Capital is sluggish, so expectatations are sluggish.
- Consumption is forward-looking, so it is also sluggish.

Impulse Response Functions



Productivity and Separation Shocks in the DMP Model

Primitives:

- Standard undirected labor search model: discrete-time version of Shimer (2005).
- The only interesting decision: firms' vacancy-posting decision.

Primitives:

- Standard undirected labor search model: discrete-time version of Shimer (2005).
- The only interesting decision: firms' vacancy-posting decision.

Calibration:

- Almost everything is calibrated as in Shimer (2005).
- Parameters that don't matter under RE:
 - correlation of the two shocks: -0.4
 - SDs of the shocks: ten times larger for productivity

Primitives:

- Standard undirected labor search model: discrete-time version of Shimer (2005).
- The only interesting decision: firms' vacancy-posting decision.

Calibration:

- Almost everything is calibrated as in Shimer (2005).
- Parameters that don't matter under RE:
 - correlation of the two shocks: -0.4
 - SDs of the shocks: ten times larger for productivity

Expectations:

- Agents have full information.
- Set $d = 1$ (only other free parameter).

The main shock:

$$- 0.812u_t + 0.010a_t - 0.177s_t$$

- Productivity and separations move the firms' expectations in **opposite** directions.

The main shock:

$$- 0.812u_t + 0.010a_t - 0.177s_t$$

- Productivity and separations move the firms' expectations in **opposite** directions.
- Unlike in RE:
 - high productivity \rightarrow more profitable to fill vacancies
 - separations \rightarrow easier to fill vacancies

The main shock:

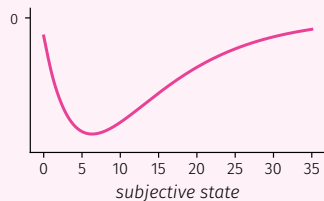
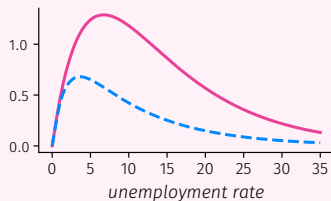
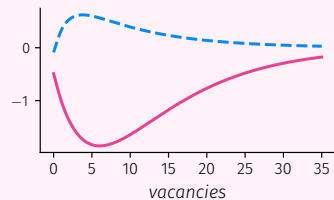
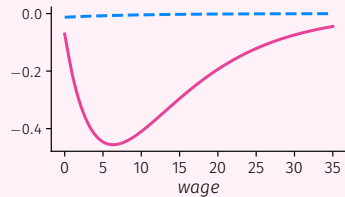
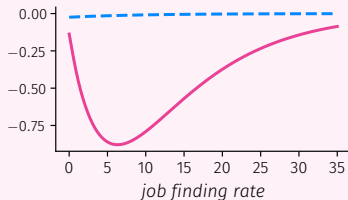
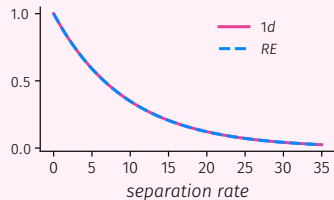
$$- 0.812u_t + 0.010a_t - 0.177s_t$$

- Productivity and separations move the firms' expectations in **opposite** directions.
- Unlike in RE:
 - high productivity \rightarrow more profitable to fill vacancies
 - separations \rightarrow easier to fill vacancies

Intuition: “Keynesian” complementarity between the firms' decisions and models...

- Suppose $u \uparrow + a \downarrow + s \uparrow$ make firms “pessimistic.”
- Pessimistic firms post few vacancies. \implies Persistent recession after $u \uparrow + a \downarrow + s \uparrow$.
- In equilibrium, the most persistent combination has $u \uparrow + a \downarrow + s \uparrow$.

Impulse Response Functions to Separation Shock



Concluding Remarks

- I assume agents forecast using simple models that are fit to the DGP.
- Illustration in the context of the RBC and DMP 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 DMP 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.