Simple Models and Biased Forecasts

Pooya Molavi

July 12, 2022

NBER Summer Institute

Motivation

- People’s optimal decisions often depend on their forecasts about the distant future.
  - Consumption-saving, investment, hiring, price setting,...

Motivation

- People's optimal decisions often depend on their forecasts about the distant future.
  - Consumption-saving, investment, hiring, price setting,...
- Rational expectations: agents forecast as well as someone who...
  - knows the shocks and the their distributions
  - knows the structural relationships in the economy
  - can apply Bayes' rule

Motivation

- People's optimal decisions often depend on their forecasts about the distant future.
  - Consumption-saving, investment, hiring, price setting,...

- Rational expectations: agents forecast as well as someone who...
  - knows the shocks and the their distributions
  - knows the structural relationships in the economy
  - can apply Bayes' rule

- Unrealistic *and* consequential assumption...
  - forecasting is hard even for professional forecasters (CG (2015), BGMS (2020))
  - macro outcomes are sensitive to long-run expectations (FG, FM, eqm determinacy, ...)

Motivation

- People's optimal decisions often depend on their forecasts about the distant future.
    - Consumption-saving, investment, hiring, price setting,...
- Rational expectations: agents forecast as well as someone who...
    - knows the shocks and the their distributions
    - knows the structural relationships in the economy
    - can apply Bayes' rule
- Unrealistic *and* consequential assumption...
    - forecasting is hard even for professional forecasters (CG (2015), BGMS (2020))
    - macro outcomes are sensitive to long-run expectations (FG, FM, eqm determinacy, ...)
- **This paper:** forecasting requires discovering complex time-series relationships.
- **Main idea**: agents can only understand simple time-series relationships.

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.

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

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

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

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

▹ learning foundations

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

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

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. Econometrics 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 d-Factor 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,

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

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 d-factor 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):

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

  - symmetric matrices
  - reduce to the usual autocorrelations when $n = 1$

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):

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

    - symmetric matrices
    - reduce to the usual autocorrelations when $n = 1$
- By linear invariance, can assume without loss that Γ_0 is the identity matrix.
    - I do so for the most part in the rest of the talk.

Main Characterization Result

Theorem

Let λ *denote the eigenvalue of* C_1 *largest in magnitude,*

and let p *denote the corresponding eigenvector (normalized such that* $p'p = 1$*).*

Main Characterization Result

Theorem

Let λ denote the eigenvalue of C_1 largest in magnitude,

and let p denote the corresponding eigenvector (normalized such that $p'p = 1$).

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

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

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

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

▹ ergodicity assumption

Main Characterization Result

Theorem

Let λ denote the eigenvalue of C_1 largest in magnitude,

and let p denote the corresponding eigenvector (normalized such that $p'p = 1$).

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

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

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

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

▹ ergodicity assumption

Main Characterization Result

Theorem

Let λ denote the eigenvalue of C_1 largest in magnitude,

and let p denote the corresponding eigenvector (normalized such that $p'p = 1$).

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

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

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

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

▹ ergodicity assumption

Main Characterization Result

Theorem

Let λ denote the eigenvalue of C_1 largest in magnitude,

and let p denote the corresponding eigenvector (normalized such that $p'p = 1$).

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

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

▹ ergodicity assumption

Second Moments

Theorem

The subjective variance given a pseudo-true 1*-factor model coincides with the true variance:*

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

Second Moments

Theorem

The subjective variance given a pseudo-true 1*-factor model coincides with the true variance:*

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

Principal component analysis...

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

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 eigenvector...

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

Pseudo-True Model in the Diagonal Example

- Eigenvalue and eigenvector...

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

- Forecasts:

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

Pseudo-True Model in the Diagonal Example

- Eigenvalue and eigenvector...

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

- Forecasts:

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

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

Intuition

- The agent perfectly matches the cross-sectional variance-covariance of observables.
  - A consequence of (A, B, Q, R) being unrestricted and KLDR minimization.
- She can do so using white noise or using a *single* persistent factor.
- The least persistent true state are closer to being white noise.
- So, the agent uses the single persistent factor to track the most persistent true factor.
- And explains the other factors by white noise.

Main Shock, Unresponsiveness, and Comovement

Unidimensional Dynamics and a Main Shock

- Recall that with $d = 1$,

$$E_t^*[y_{i,t+s}] = \lambda^s pp' y_t$$

where λ is the top eigenvalue of C_1 and p the corresponding eigenvector.

Unidimensional Dynamics and a Main Shock

- Recall that with $d = 1$,
$$E_t^*[y_{i,t+s}] = \lambda^s pp' y_t$$
where λ is the top eigenvalue of C_1 and p the corresponding eigenvector.

- 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 with $d = 1$,

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

where λ is the top eigenvalue of C_1 and p the corresponding eigenvector.

- 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 with $d = 1$,

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

where λ is the top eigenvalue of C_1 and p the corresponding eigenvector.

- 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 $y_t^{\parallel}$.
- Do not update expectations *at all* in response to $y_t^{\perp}$.

Unidimensional Dynamics and a Main Shock

- Recall that with $d = 1$,

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

where λ is the top eigenvalue of C_1 and p the corresponding eigenvector.

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

Aside: The Main Business-Cycle Shock

- Angeletos-Collard-Dellas (2020) identify a "main business-cycle shock."
- The component $y_t^{\|}$ plays a similar role.
- But $y_t^{\|}$ is *not* a new shock.
- It is a composite of existing fundamental shocks.
- It arises endogenously as the agent fits a simple model to the true DGP.

Implications of Unidimensional Dynamics

- Implications of unidimensional dynamics...
  1. Lowered responsiveness of expectations to new information.
  2. Dampened response of forward-looking decisions to shocks.
  3. Comovement of different forward-looking decisions.

Implications of Unidimensional Dynamics

- Implications of unidimensional dynamics...
  1. Lowered responsiveness of expectations to new information.
  2. Dampened response of forward-looking decisions to shocks.
  3. Comovement of different forward-looking decisions.

- I illustrate these assuming the economy is driven by n independent shocks...

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

with $\alpha_1 > \cdots > \alpha_n > 0$.

Lowered Responsiveness of Expectations

- Responsiveness of forecasts to new information:

$$\varepsilon_{ijs} \equiv \frac{\partial E_t[y_{i,t+s}]}{\partial y_{jt}}$$

Lowered Responsiveness of Expectations

- Responsiveness of forecasts to new information:

$$\varepsilon_{ijs} \equiv \frac{\partial E_t\left[y_{i,t+s}\right]}{\partial y_{jt}}$$

- **Result:** for all i, j, and s

$$\left|\varepsilon_{ijs}^{\text{1-factor}}\right| \leq \left|\varepsilon_{ijs}^{\text{RE}}\right|$$

with the inequality strict for some i, j pairs and all s.

Forward-Looking Decisions

- Under rational expectations...

$$x_{jt}^{\text{RE}} = \underbrace{\sum_{i=1}^{n} G_{ji} y_{it}}_{\text{direct effect}} + \underbrace{\sum_{i=1}^{n} \frac{\alpha_i \beta}{1 - \alpha_i \beta} G_{ji} y_{it}}_{\text{changes in expectations}}$$

Forward-Looking Decisions

- Under rational expectations...

$$x_{jt}^{\text{RE}} = \underbrace{\sum_{i=1}^{n} G_{ji} y_{it}}_{\text{direct effect}} + \underbrace{\sum_{i=1}^{n} \frac{\alpha_i \beta}{1 - \alpha_i \beta} G_{ji} y_{it}}_{\text{changes in expectations}}$$

- Given a pseudo-true 1-factor model...

$$x_{jt}^{\text{1-factor}} = \underbrace{\sum_{i=1}^{n} G_{ji} y_{it}}_{\text{direct effect}} + \underbrace{\frac{\alpha_1 \beta}{1 - \alpha_1 \beta} G_{j1} y_{1t}}_{\text{changes in expectations}}$$

Dampening

Result: response of choices to shocks are dampened on impact:

$$\left|\frac{\partial x_{jt}^{\text{1-factor}}}{\partial y_{it}}\right| \leq \left|\frac{\partial x_{jt}^{\text{RE}}}{\partial y_{it}}\right|$$

with the inequality generically strict for $i \neq 1$.

Dampening

Result: response of choices to shocks are dampened on impact:

$$\left|\frac{\partial x_{jt}^{\text{1-factor}}}{\partial y_{it}}\right| \leq \left|\frac{\partial x_{jt}^{\text{RE}}}{\partial y_{it}}\right|$$

with the inequality generically strict for $i \neq 1$.

Intuition:

- Decompose $y_t = y_t^{\parallel} + y_t^{\perp}$ as before.
  - due to linearity, can separately study the responses to $y_t^{\parallel}$ and $y_t^{\perp}$
- Response to $y_t^{\parallel}$ is as in RE.
- Response to $y_t^{\perp}$ is dampened...
  - since expectations do not move in response to $y_t^{\perp}$

Comovement

Result: If β and all of α_i are sufficiently large, then

$$1 \approx \left|\mathrm{Cor}\left(x_{it}^{\text{1-factor}}, x_{jt}^{\text{1-factor}}\right)\right| \geq \left|\mathrm{Cor}\left(x_{it}^{\mathrm{RE}}, x_{jt}^{\mathrm{RE}}\right)\right|$$

for all $i \neq j$ with the inequality generically strict.

Comovement

Result: If β and all of α_i are sufficiently large, then

$$1 \approx \left|\mathrm{Cor}\left(x_{it}^{\text{1-factor}}, x_{jt}^{\text{1-factor}}\right)\right| \geq \left|\mathrm{Cor}\left(x_{it}^{\mathrm{RE}}, x_{jt}^{\mathrm{RE}}\right)\right|$$

for all $i \neq j$ with the inequality generically strict.

Intuition:

- When β and α_i are large, expectations matter a lot for choices.
- Decompose $y_t = y_t^{\parallel} + y_t^{\perp}$ as before.
- 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.

Purely-Forward Looking Macro Models

- The analysis is essentially unchanged for linear purely forward looking GE models.
- The problem can be reduced to the diagonal example as long as...
  1. Shocks driving the economy are independent AR(1).
  2. All time-t observables are linear functions of time-t shocks (in equilibrium).
- A direct consequence of the Linear Invariance result.
- The standard three-equation NK model fits this class.
  - get a main shock, unresponsiveness to other shocks, and comovement
- Bigger differences in GE when there are state variables (such as in the RBC model)...

Application: The RBC Model with Simple Factor Models

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* hold away from RE (Preston, 2005)

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* hold away from RE (Preston, 2005)

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

Information Structure

- Need to make an assumption about what agents observe $\rightarrow$ y_t
  - different assumptions lead to different expectations
  - firms vs households

Information Structure

- Need to make an assumption about what agents observe $\rightarrow$ y_t
  - different assumptions lead to different expectations
  - firms vs households
- The benchmark model assumes full information.
- But it isn't always clear what full information means away from RE.
  - replacing y_t with $\tilde{y}_t = f(y_t)$ could change the agents' expectations
  - adding redundant variables could change the agents' expectations

Information Structure

- Need to make an assumption about what agents observe $\rightarrow y_t$
  - different assumptions lead to different expectations
  - firms vs households
- The benchmark model assumes full information.
- But it isn't always clear what full information means away from RE.
  - replacing y_t with $\tilde{y}_t = f(y_t)$ could change the agents' expectations
  - adding redundant variables could change the agents' expectations
- Linear Invariance largely gets around this issue.
  - I set $y_t = (k_t, a_t, w_t, r_t, c_t, i_t)'$.
  - But other "reasonable" choices, e.g., $y_t = (k_t, a_t)'$, lead to the same expectations.

Equilibrium Definition

Definition (1-Factor Equilibrium)

The equilibrium is given by

1. a 1-factor model θ for the agents
2. agents' policy functions
3. and an *endogenous* DGP for y_t

Equilibrium Definition

Definition (1-Factor Equilibrium)

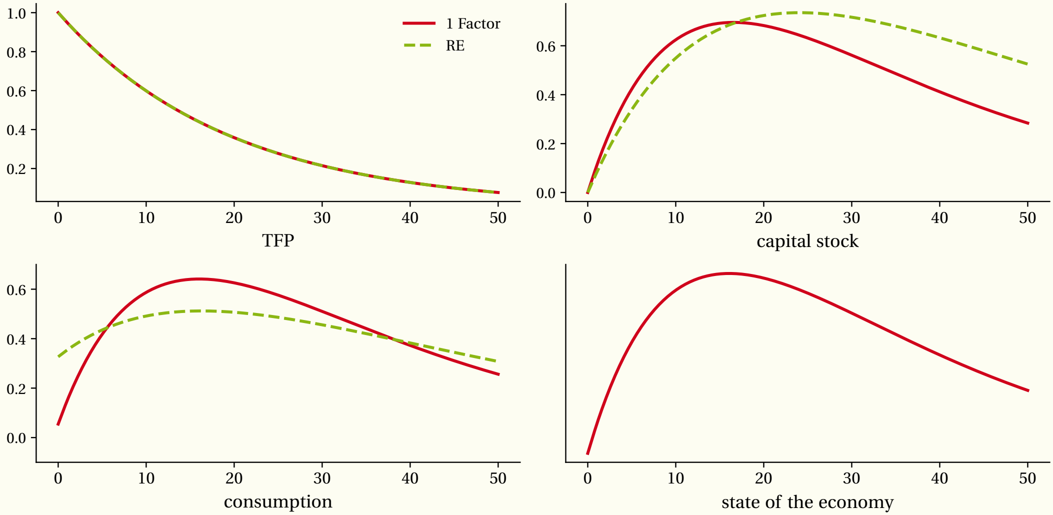
The equilibrium is given by

1. a 1-factor model θ for the agents
2. agents' policy functions
3. and an *endogenous* DGP for y_t

such that

1. agents' model is pseudo true given the DGP
2. agents' choices are optimal given their model
3. the DGP is induced by the agents' choices and the shock process

Impulse Response Functions



Conclusion

- I assume agents forecast using simple factor models that are fit to the DGP.
- This gives rise to persistence bias and stickiness in expectations.
- It also gives rise to dampening of the response of choices to shocks and to comovement in different choices.
- The model endogenously generates a main business-cycle shock.
- The framework can be embedded in workhorse macro models.
- Illustration in the context of the RBC model.

A Complementary Interpretation: Limited Memory

Models with d running statistics

- The summary 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 summary 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.*

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 the application studied in this talk.

▹ back