Imperfect Financial Expectations: Theory and Evidence

Andrea Vedolin

Boston University

(based on joint work with Pooya Molavi and Alireza Tahbaz-Salehi)

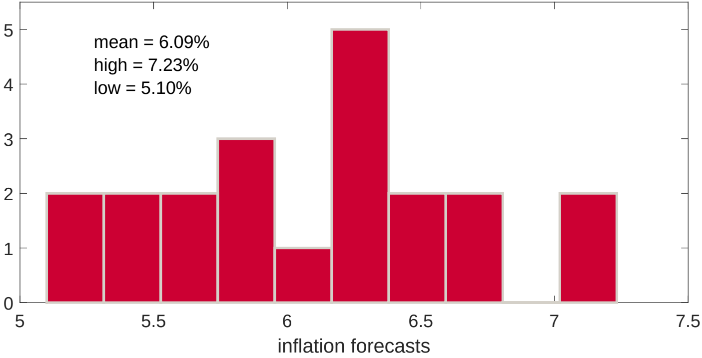
June 2023

Motivation

- Inflation for Italy in June 2023 is 8%
- What do you think is the inflation for 2023?

Motivation

- Inflation for Italy in June 2023 is 8%
- What do you think is the inflation for 2023?



Why Expectations Matter?

- they matter for firms' and consumers' micro decisions
  - how much a particular household consumes or saves
  - how much a firm produces and the price it sets

- they matter for macroeconomic and policy outcomes
  - inflation expectations central to monetary policy

Why Expectations Matter?

- they matter for firms' and consumers' micro decisions
  - how much a particular household consumes or saves
  - how much a firm produces and the price it sets

- they matter for macroeconomic and policy outcomes
  - inflation expectations central to monetary policy

- and obviously for asset prices:

$$p = \mathbb{E}[Md]$$
$$s_t = i_t^\star - i_t + \mathbb{E}_t[s_{t+1}]$$

Why Expectations Matter?

- they matter for firms' and consumers' micro decisions
  - how much a particular household consumes or saves
  - how much a firm produces and the price it sets

- they matter for macroeconomic and policy outcomes
  - inflation expectations central to monetary policy

- and obviously for asset prices:

$$p = \mathbb{E}[Md]$$
$$s_t = i_t^\star - i_t + \mathbb{E}_t[s_{t+1}]$$

- But where do expectations come from?

Rational Expectations (RE) Hypothesis

- The bedrock of modern macro and finance.

- RE maintains that economic agents fully understand the world they live in with all its complexity
  - know the economy's structural equations
  - know the shocks and/or their distributions
  - capable of Bayesian updating

Rational Expectations (RE) Hypothesis

- The bedrock of modern macro and finance.
- RE maintains that economic agents fully understand the world they live in with all its complexity
  - know the economy's structural equations
  - know the shocks and/or their distributions
  - capable of Bayesian updating
- Often combined with another strong assumption: full information
  - all data about the state of the economy are common knowledge

Rational Expectations (RE) Hypothesis

- The bedrock of modern macro and finance.

- RE maintains that economic agents fully understand the world they live in with all its complexity
  - know the economy's structural equations
  - know the shocks and/or their distributions
  - capable of Bayesian updating

- Often combined with another strong assumption: full information
  - all data about the state of the economy are common knowledge

- Full Information Rational Expectations (FIRE): objective and subjective expectations coincide

A Strong Assumption

- **Advantage**: completes the model by fully specifying agents' expectations

- **Disadvantage**: can be a strong assumption. In reality, agents have limited cognitive and computational abilities
  - limited capacity for processing information
  - not understand general equilibrium effects
  - not follow Bayes' rule
  - may not have complete knowledge of information available to others

A Strong Assumption

- **Advantage**: completes the model by fully specifying agents' expectations

- **Disadvantage**: can be a strong assumption. In reality, agents have limited cognitive and computational abilities
  - limited capacity for processing information
  - not understand general equilibrium effects
  - not follow Bayes' rule
  - may not have complete knowledge of information available to others

- Natural to expect **departures from FIRE**

Overview

- Background: Evidence for departures from FIRE
- Wilderness of Non-Rational Expectations
- A Model of Time-Series Misspecification

Evidence against FIRE

Testing FIRE: Kohlhas and Walther (2021)

$$\underbrace{x_{t+h} - \mathbb{E}_t[x_{t+h}]}_{\text{forecast error}} = \alpha_h + \beta_h^{\mathsf{KW}} x_t + \epsilon_{t,t+h}$$

- Under rational expectations, forecast errors should not be predictable

$$\beta_h^{\mathsf{KW}} = 0 \qquad \text{for all } h$$

Testing FIRE: Kohlhas and Walther (2021)

$$\underbrace{x_{t+h} - \mathbb{E}_t[x_{t+h}]}_{\text{forecast error}} = \alpha_h + \beta_h^{\text{KW}} x_t + \epsilon_{t,t+h}$$

- Under rational expectations, forecast errors should not be predictable

$$\beta_h^{\text{KW}} = 0 \qquad \text{for all } h$$

- A consequence of the Law of Iterated Expectations (LIE):

$$\beta_h^{\text{KW}} = \frac{\mathbb{E}^*[x_t(x_{t+h} - \mathbb{E}_t[x_{t+h}])]}{\text{var}(x_t)}$$

Testing FIRE: Kohlhas and Walther (2021)

$$\underbrace{x_{t+h} - \mathbb{E}_t[x_{t+h}]}_{\text{forecast error}} = \alpha_h + \beta_h^{\text{KW}} x_t + \epsilon_{t,t+h}$$

- Under rational expectations, forecast errors should not be predictable

$$\beta_h^{\text{KW}} = 0 \qquad \text{for all } h$$

- A consequence of the Law of Iterated Expectations (LIE):

$$\beta_h = \frac{\mathbb{E}^* \mathbb{E}_t^* [x_t(x_{t+h} - \mathbb{E}_t[x_{t+h}])]}{\text{var}(x_t)}$$

Testing FIRE: Kohlhas and Walther (2021)

$$\underbrace{x_{t+h} - \mathbb{E}_t[x_{t+h}]}_{\text{forecast error}} = \alpha_h + \beta_h^{\text{KW}} x_t + \epsilon_{t,t+h}$$

- Under rational expectations, forecast errors should not be predictable

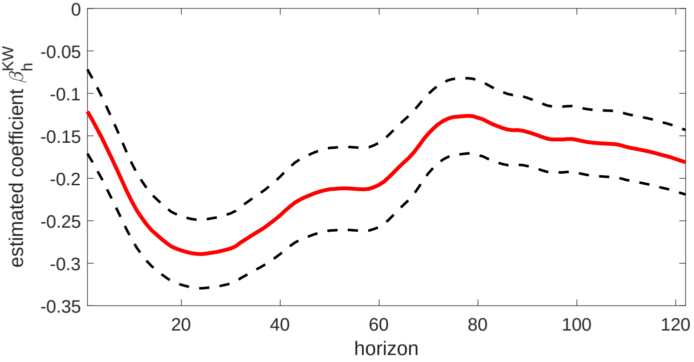
$$\beta_h^{\text{KW}} = 0 \qquad \text{for all } h$$

- A consequence of the Law of Iterated Expectations (LIE):

$$\beta_h = \frac{\mathbb{E}^* \mathbb{E}_t^* [x_t(x_{t+h} - \mathbb{E}_t[x_{t+h}])]}{\text{var}(x_t)}$$
$$= \frac{\mathbb{E}^* [x_t(\mathbb{E}_t^*[x_{t+h}] - \mathbb{E}_t[x_{t+h}])]}{\text{var}(x_t)} = 0$$

Empirical Evidence: Interest Rates

forecast error 3m interest rate $= \alpha + \beta^{\text{KW}}$3m interest rate $+$ error



$\beta_h^{\text{KW}} < 0$: systematic over-reaction to interest rate realizations

Wilderness of Non-Rational Expectations

Wilderness of Non-Rational Expectations

- Without rational expectations, one ends up in the "wilderness" of alternative models of expectation formation (Sims, 1980, Sargent, 2001)
- A whole menu of different alternatives:
  - Rational inattention, sticky info, etc. (Sims, Mankiw & Reis, Mackowiak & Wiederholt)
  - Higher-order uncertainty (Morris & Shin, Woodford, Nimark, Angeletos & Lian)
  - Level-K thinking (Garcia-Schmidt & Woodford, Farhi & Werning, Iovino & Sergeyev)
  - Cognitive discounting (Gabaix)
  - Over-extrapolation (Gennaioli, Ma & Shleifer, Fuster, Laibson & Mendel, Guo & Wachter)
  - Over-confidence (Kohlhas & Broer, Scheinkman & Xiong)
  - Representativeness (Bordalo, Gennaioli & Shleifer)
  - Undue effect of historical experiences (Malmendier & Nagel)
  - ...
- Different models of expectation formation can have very different implications

(Partially) Taming the Wilderness

- **Standard approach**: assume a particular model of expectations and characterize the degree of forecast-error predictability:
- **Alternative**: Characterize forecast-error predictability in terms of the objective and subjective autocorrelation (without imposing a fully specified model)

ξ_h^* : objective autocorrelation of x_t at horizon h

ξ_h : subjective autocorrelation of x_t at horizon h

(Partially) Taming the Wilderness

- **Standard approach**: assume a particular model of expectations and characterize the degree of forecast-error predictability:
- **Alternative**: Characterize forecast-error predictability in terms of the objective and subjective autocorrelation (without imposing a fully specified model)

ξ_h^* : objective autocorrelation of x_t at horizon h

ξ_h : subjective autocorrelation of x_t at horizon h

- Run the following regression

$$x_{t+h} - \mathbb{E}_t[x_{t+h}] = \alpha_h^{\text{KW}} + \beta_h^{\text{KW}} x_t + \epsilon_{t,h}$$

(Partially) Taming the Wilderness

$$x_{t+h} - \mathbb{E}_t[x_{t+h}] = \alpha_h^{\text{KW}} + \beta_h^{\text{KW}} x_t + \epsilon_{t,h}$$

Proposition

Let

$$\xi_h^* : \textit{objective autocorrelation} \quad , \quad \xi_h : \textit{subjective autocorrelation}$$

Then

$$\beta_h^{KW} = (\xi_h^* - \xi_h) - \sum_{\tau=1}^{\infty} \sum_{s=1}^{\infty} \Xi_{\tau s}^{-1} \xi_{s+h-1} (\xi_{\tau-1}^* - \xi_{\tau-1}),$$

where Ξ is an infinite-dimensional matrix such that $\Xi_{ij} = \xi_{i-j}$ for all $i, j \geq 1$.

(Partially) Taming the Wilderness

$$x_{t+h} - \mathbb{E}_t[x_{t+h}] = \alpha_h^{\mathrm{KW}} + \beta_h^{\mathrm{KW}} x_t + \epsilon_{t,h}$$

Proposition

Let

$$\xi_h^* \text{ :objective autocorrelation} \quad , \quad \xi_h : \text{subjective autocorrelation}$$

Then

$$\beta_h^{KW} = (\xi_h^* - \xi_h) - \sum_{\tau=1}^{\infty} \sum_{s=1}^{\infty} \Xi_{\tau s}^{-1} \xi_{s+h-1} (\xi_{\tau-1}^* - \xi_{\tau-1}),$$

where Ξ is an infinite-dimensional matrix such that $\Xi_{ij} = \xi_{i-j}$ for all $i, j \geq 1$.

- The sub. and obj. autocorrelations are summary statistics for predictability
- What matters is the degree of over- or under-extrapolation at different horizons!

(Partially) Taming the Wilderness

$$x_{t+h} - \mathbb{E}_t[x_{t+h}] = \alpha_h^{\mathrm{KW}} + \beta_h^{\mathrm{KW}} x_t + \epsilon_{t,h}$$

Proposition

Let

$$\xi_h^* : \textit{objective autocorrelation} \quad , \quad \xi_h : \textit{subjective autocorrelation}$$

Then

$$\beta_h^{KW} = (\xi_h^* - \xi_h) - \sum_{\tau=1}^{\infty} \sum_{s=1}^{\infty} \Xi_{\tau s}^{-1} \xi_{s+h-1} (\xi_{\tau-1}^* - \xi_{\tau-1}),$$

where Ξ *is an infinite-dimensional matrix such that* $\Xi_{ij} = \xi_{i-j}$ *for all* $i, j \geq 1$.

- The sub. and obj. autocorrelations are summary statistics for predictability
- What matters is the degree of over- or under-extrapolation at different horizons!
- Need a model for how agents may misperceive time-series dependencies

A Model of Time-Series Misspecification

Framework

- A sequence of payoff-relevant variables generated by a latent n-factor model.
- But agents are constrained to thinking through models with at most k factors.

Framework

- A sequence of payoff-relevant variables generated by a latent n-factor model.
- But agents are constrained to thinking through models with at most k factors.
  - agents are otherwise (constrained) rational.
  - they estimate the k-factor model that best fits their observations.

Framework

- A sequence of payoff-relevant variables generated by a latent n-factor model.
- But agents are constrained to thinking through models with at most k factors.
  - agents are otherwise (constrained) rational.
  - they estimate the k-factor model that best fits their observations.

- With **large** k, agents recover the true model (back to RE).
- With **small** k, expectations exhibit deviations from RE. $\rightarrow$ our focus

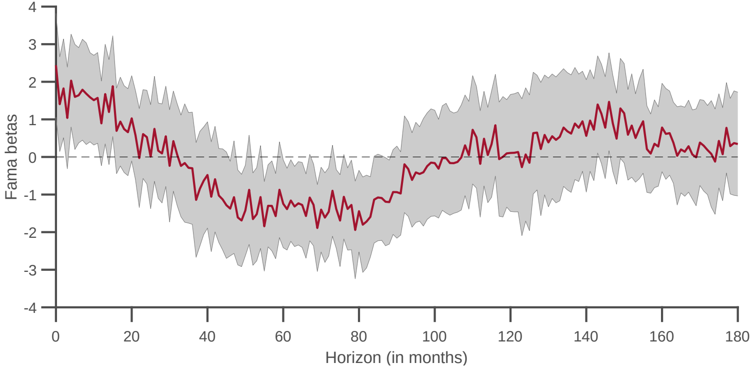
Framework

- A sequence of payoff-relevant variables generated by a latent n-factor model.
- But agents are constrained to thinking through models with at most k factors.
  - agents are otherwise (constrained) rational.
  - they estimate the k-factor model that best fits their observations.

- With **large** k, agents recover the true model (back to RE).
- With **small** k, expectations exhibit deviations from RE. $\rightarrow$ our focus

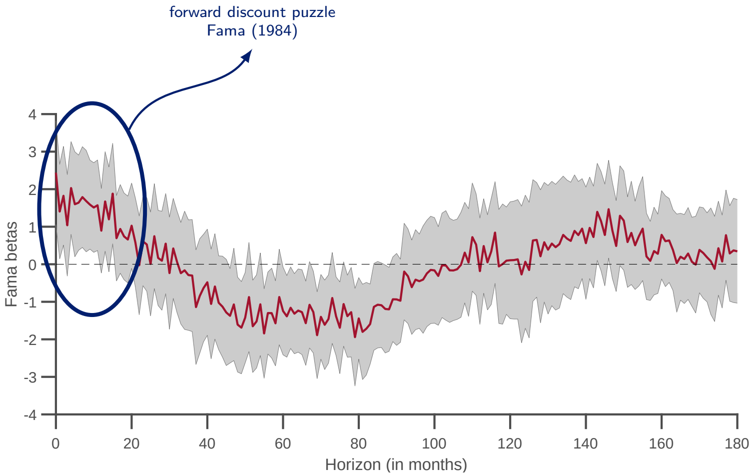
- Asset pricing implications: characterize return predictability as a function of
  (1) true data-generating process
  (2) the complexity/dimension of agents' models

Fact 1: Fama Regressions



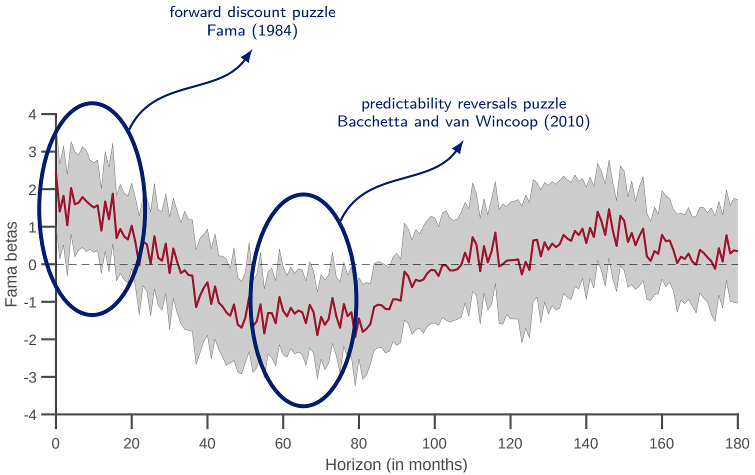
$$\mathrm{rx}_{t+h} = \alpha_h^{\mathrm{rx}} + \beta_h^{\mathrm{rx}}(i_t^\star - i_t) + \epsilon_{t+h}.$$

Fact 1: Fama Regressions



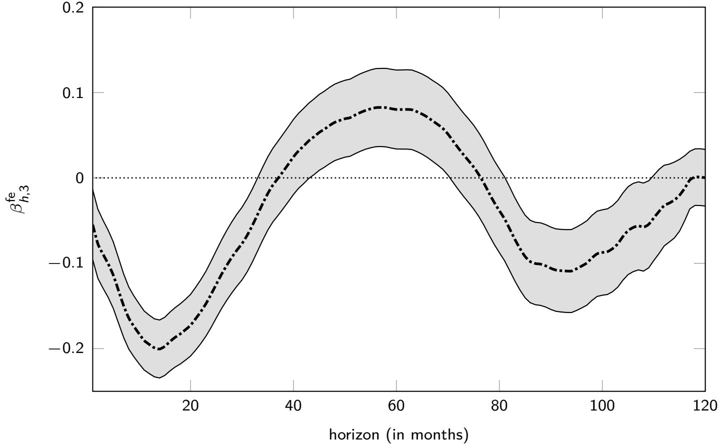
$$\mathsf{rx}_{t+h} = \alpha_h^{\mathsf{rx}} + \beta_h^{\mathsf{rx}}(i_t^\star - i_t) + \epsilon_{t+h}.$$

Fact 1: Fama Regressions



$$rx_{t+h} = \alpha_h^{rx} + \beta_h^{rx}(i_t^\star - i_t) + \epsilon_{t+h}.$$

Fact 2: Forecast-Error Predictability



$$x_{t+h-1} - \mathbb{E}_{t+h-1}[x_{t+h-1}] = \alpha^{\mathrm{fe}}_h + \beta^{\mathrm{fe}}_h x_t + \epsilon_t.$$

Framework

(Reduced-Form) Framework

- Discrete-time economy with unit mass of identical agents.
- Exogenous sequence of **fundamentals** $\{x_t\}_{t=-\infty}^{\infty}$ generated according to some probability distribution $\mathbb{P}^*$.

(Reduced-Form) Framework

- Discrete-time economy with unit mass of identical agents.
- Exogenous sequence of **fundamentals** $\{x_t\}_{t=-\infty}^{\infty}$ generated according to some probability distribution $\mathbb{P}^*$.
- Uncovered Interest Rate Parity (UIP):

$$s_t = i_t^* - i_t + \delta \mathbb{E}_t[s_{t+1}] \tag{1}$$

where $\mathbb{E}[\cdot]$ is agents' subjective expectations and in general $\mathbb{E}[\cdot] \neq \mathbb{E}^*[\cdot]$.

(Reduced-Form) Framework

- Discrete-time economy with unit mass of identical agents.
- Exogenous sequence of **fundamentals** $\{x_t\}_{t=-\infty}^{\infty}$ generated according to some probability distribution $\mathbb{P}^*$.
- Uncovered Interest Rate Parity (UIP):

$$s_t = i_t^* - i_t + \delta \mathbb{E}_t[s_{t+1}] \tag{1}$$

where $\mathbb{E}[\cdot]$ is agents' subjective expectations and in general $\mathbb{E}[\cdot] \neq \mathbb{E}^*[\cdot]$.

- **excess returns**:

$$\mathrm{rx}_{t+1} = \delta s_{t+1} - s_t + i_t^* - i_t \tag{2}$$

Main Assumption (informally)

- Model's structural equation:

$$s_t = i_t^* - i_t + \delta \mathbb{E}_t[s_{t+1}]$$

- The true data-generating process $\mathbb{P}^*$ may not have a simple representation.
- There is a limit to the complexity of statistical models that agents are able to consider.

- They approximate $\mathbb{P}^*$ with a simplified model.

Approximating Complex Models

true model: latent n factors

$$x_t = c^{*\prime} z_t$$

$$z_t = \mathbf{A}^* z_{t-1} + \mathbf{B}^* \epsilon_t$$

$$z_t \in \mathbb{R}^n$$

- agents neither observe nor know the underlying factors $(z_1, \ldots, z_n)$.
- they do not know the collection of parameters $\theta^* = (\mathbf{A}^*, \mathbf{B}^*, c^*)$.

Approximating Complex Models

true model: latent n factors

$$x_t = c^{*\prime} z_t$$
$$z_t = \mathbf{A}^* z_{t-1} + \mathbf{B}^* \epsilon_t$$
$$z_t \in \mathbb{R}^n$$

- agents neither observe nor know the underlying factors $(z_1, \dots, z_n)$.
- they do not know the collection of parameters $\theta^* = (\mathbf{A}^*, \mathbf{B}^*, c^*)$.

subjective model: up to k factors

$$x_t = c' \omega_t$$
$$\omega_t = \mathbf{A}\, \omega_{t-1} + \mathbf{B} \varepsilon_t$$
$$\omega_t \in \mathbb{R}^k$$

- if $k < n$ set of models entertained by agents does not contain the true model $\rightarrow$ **misspecification**
- Agent's model, $\theta = (\mathbf{A}, \mathbf{B}, c)$, is endogenous outcome of learning

Approximating Complex Models

true model: latent n factors

$$x_t = c^{*\prime} z_t$$
$$z_t = \mathbf{A}^* z_{t-1} + \mathbf{B}^* \epsilon_t$$
$$z_t \in \mathbb{R}^n$$

- agents neither observe nor know the underlying factors $(z_1, \ldots, z_n)$.
- they do not know the collection of parameters $\theta^* = (\mathbf{A}^*, \mathbf{B}^*, c^*)$.

subjective model: up to k factors

$$x_t = c' \omega_t$$
$$\omega_t = \mathbf{A}\, \omega_{t-1} + \mathbf{B} \varepsilon_t$$
$$\omega_t \in \mathbb{R}^k$$

- if $k < n$ set of models entertained by agents does not contain the true model $\rightarrow$ **misspecification**
- Agent's model, $\theta = (\mathbf{A}, \mathbf{B}, c)$, is endogenous outcome of learning

- only exogenously-specified feature of the agents: **k**
- "low rank approximation" of the true data-generating process
- similar to model-order reduction in control theory

minimal-order representation

limited memory formulation

example: ARMA(2,1)

Learning

- Agents start with a common prior belief at initial period t_0 with full support over the set of *k-factor models*

$$\mu_{t_0} \in \Delta\Theta_k$$

Learning

- Agents start with a common prior belief at initial period t_0 with full support over the set of k-factor models

$$\mu_{t_0} \in \Delta\Theta_k$$

- Form Bayesian posteriors after observing $(x_{t_0}, \dots, x_{t-1}, x_t)$:

$$\mu_t \in \Delta\Theta_k$$

Learning

- Agents start with a common prior belief at initial period t_0 with full support over the set of k-factor models

$$\mu_{t_0} \in \Delta\Theta_k$$

- Form Bayesian posteriors after observing $(x_{t_0}, \ldots, x_{t-1}, x_t)$:

$$\mu_t \in \Delta\Theta_k$$

- Abstract from finite-sample issues by taking the limit as $t_0 \to -\infty$
  - any model misspecification is due to mismatch between k and n

Learning

- Agents are fully **Bayesian** and their expectations are internally consistent (satisfy law of iterated expectations) but assign **zero prior beliefs** on models with more than k factors.

- Agents' model is
  - endogenous and depends on the true data-generating process and **k**
  - independent of other characteristics of the environment (preferences, etc.)

Learning

- Agents are fully **Bayesian** and their expectations are internally consistent (satisfy law of iterated expectations) but assign **zero prior beliefs** on models with more than k factors.

- Agents' model is
  - endogenous and depends on the true data-generating process and **k**
  - independent of other characteristics of the environment (preferences, etc.)

- If $\mathbf{k} \geq \mathbf{n}$,
  - agents learn the true model (up to an observational equivalence)
  - back to rational expectations

Learning

- Agents are fully **Bayesian** and their expectations are internally consistent (satisfy law of iterated expectations) but assign **zero prior beliefs** on models with more than k factors.

- Agents' model is
  - endogenous and depends on the true data-generating process and $\mathbf{k}$
  - independent of other characteristics of the environment (preferences, etc.)

- If $\mathbf{k} \geq \mathbf{n}$,
  - agents learn the true model (up to an observational equivalence)
  - back to rational expectations

- If $\mathbf{k} < \mathbf{n}$,
  - agents end up with a misspecified model of the world
  - but one that is the "best" k-dimensional approximation to the true model
  - no data would make them revise k

Subjective Expectations

Subjective Expectations: Characterization

Proposition

Let $\widetilde{\Theta} \subseteq \Theta_k$ *have positive prior measure, i.e.,* $\bar{\mu}(\widetilde{\Theta}) > 0$. *If*

$$\operatorname*{ess\,inf}_{\theta \in \widetilde{\Theta}} \ \mathrm{KL}(\theta^* \| \theta) \ > \ \operatorname*{ess\,inf}_{\theta \in \Theta_k} \ \mathrm{KL}(\theta^* \| \theta),$$

then

$$\lim_{t_0 \to -\infty} \mu_t(\widetilde{\Theta}) = 0 \qquad \mathbb{P}^*\text{-almost surely.}$$

- Kullback-Leibler divergence between the agents' and the true model

$$\mathrm{KL}(\theta^* \| \theta) = \ \underbrace{\mathbb{E}^*[-\log f^{\theta}(x_{t+1} | x_t, \dots)]}_{\text{uncertainty about one-step-ahead predictions under agents' model}} \ - \ \underbrace{\mathbb{E}^*[-\log f^{*}(x_{t+1} | x_t, \dots)]}_{\text{uncertainty under the true model}}$$

Non-Bayesian estimation

Subjective Expectations: Characterization

Proposition

Let $\widetilde{\Theta} \subseteq \Theta_k$ *have positive prior measure, i.e.,* $\bar{\mu}(\widetilde{\Theta}) > 0$. *If*

$$\operatorname*{ess\,inf}_{\theta \in \widetilde{\Theta}} \ \mathsf{KL}(\theta^* \| \theta) \ > \ \operatorname*{ess\,inf}_{\theta \in \Theta_k} \ \mathsf{KL}(\theta^* \| \theta),$$

then

$$\lim_{t_0 \to -\infty} \mu_t(\widetilde{\Theta}) = 0 \qquad \mathbb{P}^*\text{-almost surely.}$$

- Kullback-Leibler divergence between the agents' and the true model

$$\mathsf{KL}(\theta^* \| \theta) = \ \mathbb{E}^*[-\log f^{\theta}(x_{t+1}|x_t, \dots)] \ - \ \mathbb{E}^*[-\log f^*(x_{t+1}|x_t, \dots)]$$

- Agents' posteriors concentrate on the set of models with minimum KL divergence to the true model.
- When $k \geq n$, agents learn the true model.

Non-Bayesian estimation

Return and Forecast-Error Predictability

Excess Returns

$$\text{exchange rates:} \qquad s_t = i_t^* - i_t + \delta \mathbb{E}_t[s_{t+1}]$$

$$\text{currency returns:} \qquad \mathrm{rx}_{t+1} = \delta s_{t+1} - s_t + i_t^* - i_t$$

- With rational expectations:

$$\mathbb{E}_t[\mathrm{rx}_{t+h}] = \mathbb{E}_t^*[\mathrm{rx}_{t+h}] = 0$$

Excess Returns

$$\text{exchange rates:} \qquad s_t = i_t^* - i_t + \delta \mathbb{E}_t[s_{t+1}]$$

$$\text{currency returns:} \qquad \mathrm{rx}_{t+1} = \delta s_{t+1} - s_t + i_t^* - i_t$$

- With rational expectations:

$$\mathbb{E}_t[\mathrm{rx}_{t+h}] = \mathbb{E}_t^*[\mathrm{rx}_{t+h}] = 0$$

- Under complexity constraint:

$$\mathbb{E}_t[\mathrm{rx}_{t+h}] = 0$$

$$\mathbb{E}_t^*[\mathrm{rx}_{t+h}] \neq 0 \qquad \text{in general}$$

Predictability Regressions

- Quantify deviations from the rational expectations benchmark using two families of predictability regressions:

  (1) (excess) return predictability:

  $$\mathrm{rx}_{t+h} = \alpha_h^{\mathrm{rx}} + \beta_h^{\mathrm{rx}} x_t + \epsilon_{t,h}$$

  (2) forecast-error predictability:

  $$x_{t+h+m-1} - \mathbb{E}_{t+h-1}[x_{t+h+m-1}] = \alpha_{h,m}^{\mathrm{fe}} + \beta_{h,m}^{\mathrm{fe}} x_t + \epsilon_{t,h}$$

- Rational expectations: returns and forecast errors are unpredictable: $\beta_h^{\mathrm{rx}} = \beta_{h,m}^{\mathrm{fe}} = 0$ for all $h, m \geq 1$

Predictability Regressions

- Quantify deviations from the rational expectations benchmark using two families of predictability regressions:

  (1) (excess) return predictability:

  $$\mathrm{rx}_{t+h} = \alpha_h^{\mathrm{rx}} + \beta_h^{\mathrm{rx}} x_t + \epsilon_{t,h}$$

  (2) forecast-error predictability:

  $$x_{t+h+m-1} - \mathbb{E}_{t+h-1}[x_{t+h+m-1}] = \alpha_{h,m}^{\mathrm{fe}} + \beta_{h,m}^{\mathrm{fe}} x_t + \epsilon_{t,h}$$

- Rational expectations: returns and forecast errors are unpredictable: $\beta_h^{\mathrm{rx}} = \beta_{h,m}^{\mathrm{fe}} = 0$ for all $h, m \geq 1$

- The **term structure of coefficients** $\{\beta_h^{\mathrm{rx}}\}_{h=1}^{\infty}$ and $\{\beta_{h,m}^{\mathrm{fe}}\}_{h=1}^{\infty}$ inform us about deviations from the rational-expectations benchmark at each horizon h.

Return and Forecast-Error Predictability: Main Result

Theorem

Suppose agents are constrained to k-factor models. Then,

(a) *subjective autocorrelation:*

$$\xi_h = \frac{\sum_{s=0}^{\infty} (u' \mathbf{M}^{h+s} u)(u' \mathbf{M}^s u)}{\sum_{s=0}^{\infty} (u' \mathbf{M}^s u)^2} \qquad \textit{for all } h \geq 0$$

where

$$(\boldsymbol{M}, u) = \underset{\boldsymbol{M} \in \mathbb{S}^k, u \in \mathbb{R}^k}{\operatorname{argmin}} \quad 1 - 2 \sum_{s=1}^{\infty} \phi_s^{(1)} \xi_s^* + \sum_{s=1}^{\infty} \sum_{\tau=1}^{\infty} \phi_s^{(1)} \phi_\tau^{(1)} \xi_{\tau-s}^*$$

$$\textit{s.t.} \qquad \phi_s^{(m)} = u' M^{m-1} [\boldsymbol{M}(I - uu')]^{s-1} \boldsymbol{M} u.$$

Return and Forecast-Error Predictability: Main Result

Theorem

Suppose agents are constrained to k-factor models. Then,

(a) *subjective autocorrelation:*

$$\xi_h = \frac{\sum_{s=0}^{\infty} (u' \mathbf{M}^{h+s} u)(u' \mathbf{M}^{s} u)}{\sum_{s=0}^{\infty} (u' \mathbf{M}^{s} u)^2} \qquad \textit{for all } h \geq 0$$

where

$$(\boldsymbol{M}, u) = \underset{\boldsymbol{M} \in \mathbb{S}^k, u \in \mathbb{R}^k}{\operatorname{argmin}} \quad 1 - 2 \sum_{s=1}^{\infty} \phi_s^{(1)} \xi_s^* + \sum_{s=1}^{\infty} \sum_{\tau=1}^{\infty} \phi_s^{(1)} \phi_\tau^{(1)} \xi_{\tau-s}^*$$

$$\text{s.t.} \qquad \phi_s^{(m)} = u' M^{m-1} [\boldsymbol{M}(I - uu')]^{s-1} \boldsymbol{M} u.$$

(b) *coefficients of the predictability regressions:*

$$\beta_h^{rx} = \delta \frac{\xi_h^* - \sum_{\tau=1}^{\infty} \phi_\tau^{(1)} \xi_{h-\tau}^*}{1 - \sum_{\tau=1}^{\infty} \delta^\tau \phi_\tau^{(1)}} \quad , \quad \beta_{h,m}^{fe} = \xi_{h+m-1}^* - \sum_{\tau=1}^{\infty} \phi_\tau^{(m)} \xi_{h-\tau}^*.$$

Return and Forecast-Error Predictability: Main Result

$$\left.\begin{array}{l}\text{ACF of the fundamental under the true model: } \xi_h^* = \frac{\mathbb{E}^*[x_t x_{t+h}]}{\mathbb{E}^*[x_t^2]} \\ \\ + \\ \\ \text{number of factors in agents' subjective model: } k\end{array}\right\} \longrightarrow \beta_h^{\mathsf{rx}}, \beta_{h,m}^{\mathsf{fe}}$$

- the way the true data-generating process matters is via its ACF.
- only free parameter is the number of factors in the agents' model.

Single-Factor Subjective Model ($k = 1$)

Proposition

If agents are constrained to single-factor models ($k = 1$), then

$$\beta_h^{rx} = \frac{\delta}{1 - \delta \xi_1^*} (\xi_h^* - \xi_{h-1}^* \xi_1^*)$$
$$\beta_{h,m}^{fe} = \xi_{h+m-1}^* - {\xi_1^*}^m \xi_{h-1}^*$$

respectively, where ξ_h^* is the autocorrelation of the fundamental at lag h.

Application: Violations of Uncovered Interest Rate Parity

Background and Setting

- Fundamental: (log) interest rate differential of foreign and U.S. deposit rates:

$$x_t = i_t^\star - i_t$$

- Price: (log) exchange rate (expressed as the U.S. dollar price of the foreign currency) satisfying the uncovered interest rate parity (UIP) condition:

$$s_t = i_t^\star - i_t + \mathbb{E}_t[s_{t+1}]$$

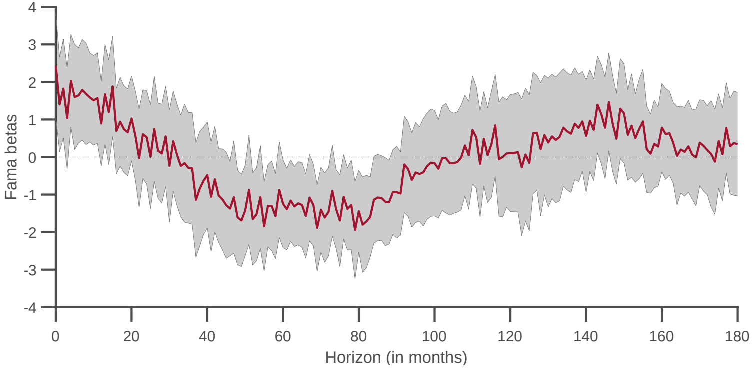
- Currency excess return:

$$\mathrm{rx}_{t+1} = s_{t+1} - s_t + (i_t^\star - i_t)$$

- Return Predictability Regression (Fama, 1984):

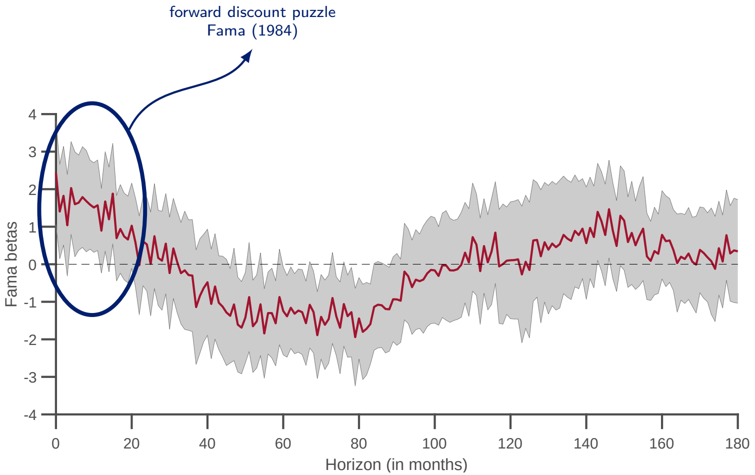
$$\mathrm{rx}_{t+h} = \alpha_h^{\mathrm{rx}} + \beta_h^{\mathrm{rx}}(i_t^\star - i_t) + \epsilon_{t,h}.$$

Fama Regressions



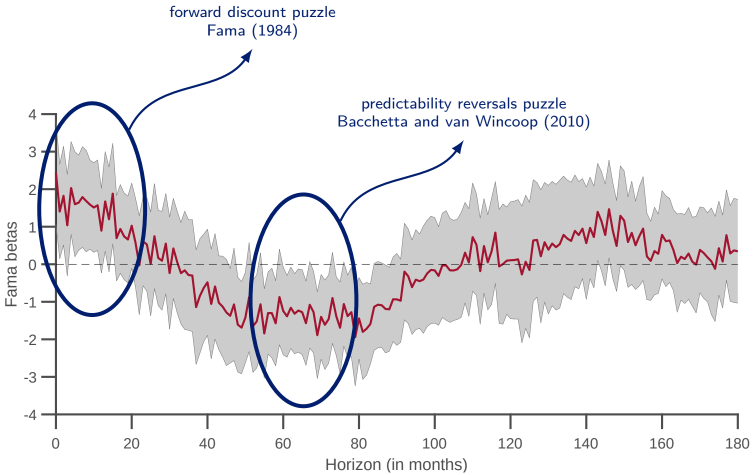
$$\mathrm{rx}_{t+h} = \alpha_h^{\mathrm{rx}} + \beta_h^{\mathrm{rx}}(i_t^\star - i_t) + \epsilon_{t+h}.$$

Fama Regressions



$$\mathrm{rx}_{t+h} = \alpha_h^{\mathrm{rx}} + \beta_h^{\mathrm{rx}}(i_t^{\star} - i_t) + \epsilon_{t+h}.$$

Fama Regressions



$$\mathrm{rx}_{t+h} = \alpha_h^{\mathrm{rx}} + \beta_h^{\mathrm{rx}}(i_t^\star - i_t) + \epsilon_{t+h}.$$

Term Structure of UIP Violations in the Model

- Fundamental: log-interest rate differential

$$x_t = i_t^\star - i_t$$

- Construct the autocorrelation of the interest rate differential:

$$\xi_h^* = \frac{\mathbb{E}^*[x_t x_{t+h}]}{\mathbb{E}^*[x_t^2]}.$$

- Plug into the formula:

$$\beta_h^{\text{rx}} = \frac{\xi_h^* - \xi_1^* \xi_{h-1}^*}{1 - \xi_1^*}$$

(optimize numerically for $k > 1$)

Term Structure of UIP Violations in the Model

- Fundamental: log-interest rate differential

$$x_t = i_t^\star - i_t$$

- Construct the autocorrelation of the interest rate differential:

$$\xi_h^* = \frac{\mathbb{E}^*[x_t x_{t+h}]}{\mathbb{E}^*[x_t^2]}.$$

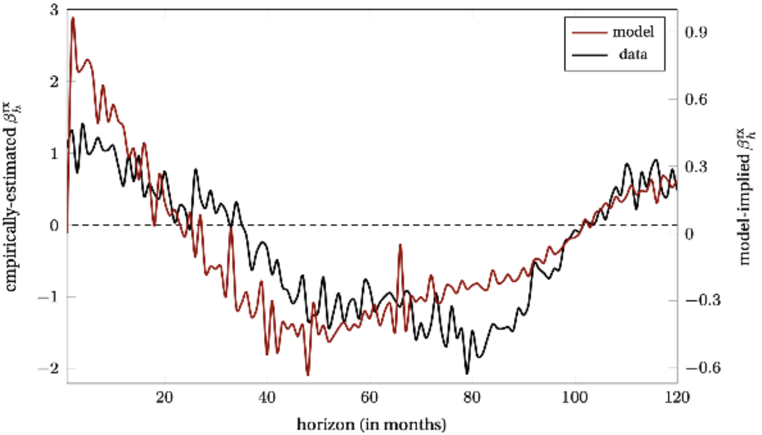
- Plug into the formula:

$$\beta_h^{\text{rx}} = \frac{\xi_h^* - \xi_1^* \xi_{h-1}^*}{1 - \xi_1^*}$$

  (optimize numerically for $k > 1$)

- β_h^{rx} only depend on the shape of the ACF of interest rate differential.
  we do not use information on exchange rate or currency excess returns.

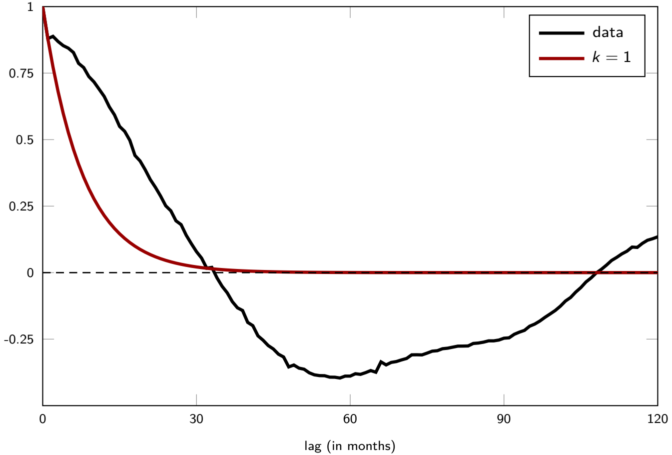
UIP Violations



- generates patterns simultaneously consistent with forward premium and return predictability puzzles **only using information from interest rate differentials**.

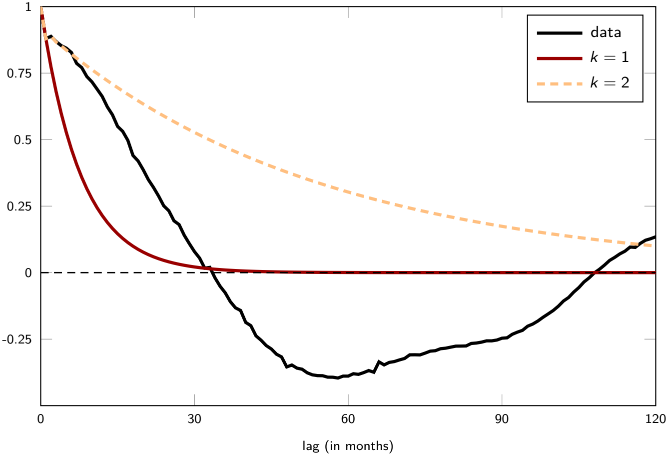
Varying the Complexity of Agents' Model: Autocorrelation

- Compare ACF of agents' implied model to that of the true process.



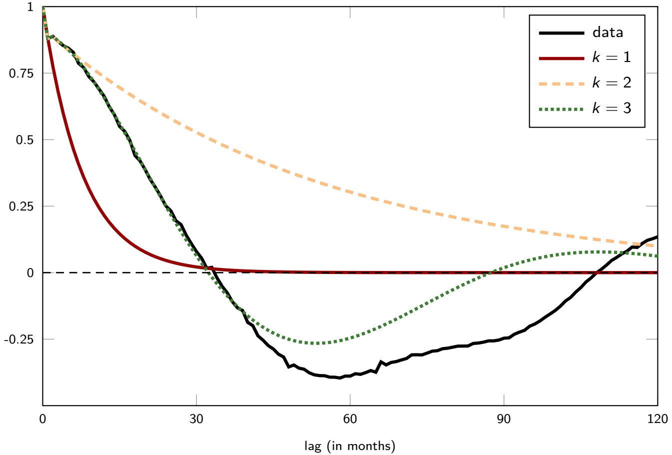
Varying the Complexity of Agents' Model: Autocorrelation

- Compare ACF of agents' implied model to that of the true process.



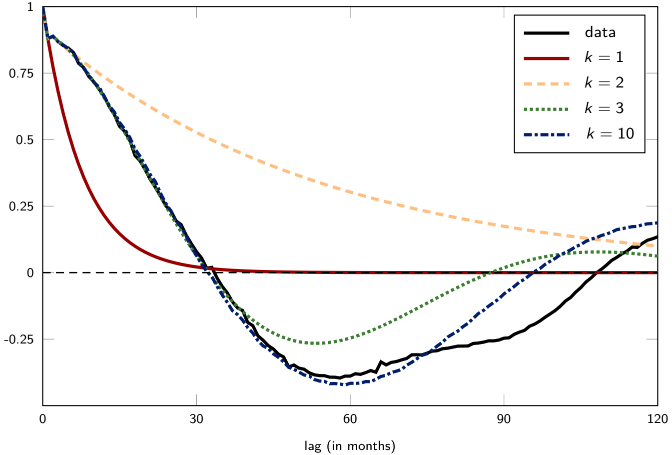
Varying the Complexity of Agents' Model: Autocorrelation

- Compare ACF of agents' implied model to that of the true process.

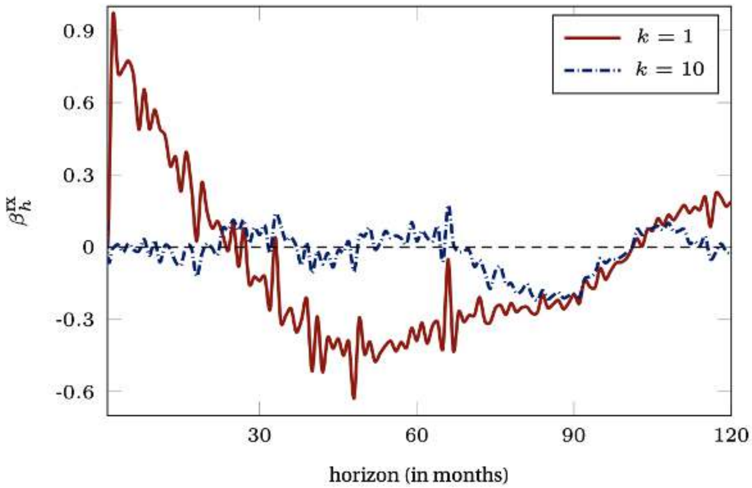


Varying the Complexity of Agents' Model: Autocorrelation

- Compare ACF of agents' implied model to that of the true process.



Return Predictability: Varying the Complexity of Agents' Model



Forecast-Error Predictability

- Forecasts: survey data from *Consensus Economics*:

$$\mathbb{E}_t[x_{t+3}]$$

- Forecast-Error Predictability regression:

$$x_{t+h+m-1} - \mathbb{E}_{t+h-1}[x_{t+h+m-1}] = \alpha^{\text{fe}}_{h,m} + \beta^{\text{fe}}_{h,m} x_t + \epsilon_{t,h}$$

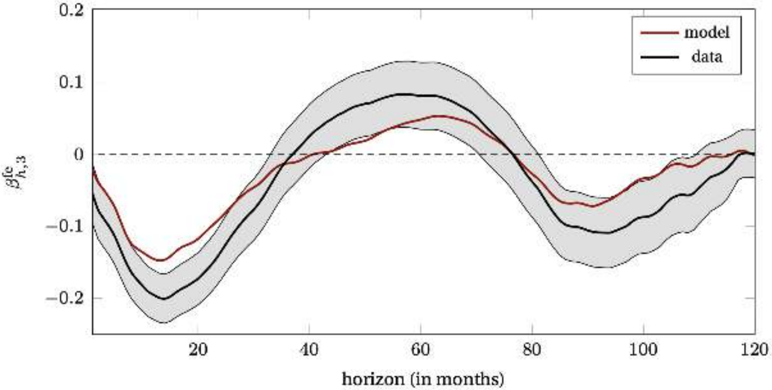
- Model-implied coefficient:

$$\beta^{\text{fe}}_{h,3} = \xi^*_{h+2} - \xi^{*\,3}_1 \xi^*_{h-1}$$

(optimize numerically for $k > 1$)

- $\beta^{\text{fe}}_{h,3}$ only depend on the shape of the ACF of interest rate differential and does not depend on survey data, exchange rates, or realized returns.

Forecast-Error Predictability



Conclusions

Conclusions

- A framework in which investors are constrained in model complexity.
- Implies rich dynamics for the term structure of deviations from the benchmark rational expectations.
- Implications for violations of uncovered interest rate parity

Conclusions

- A framework in which investors are constrained in model complexity.
- Implies rich dynamics for the term structure of deviations from the benchmark rational expectations.
- Implications for violations of uncovered interest rate parity

Grazie

Minimal-Order Representation

- Data-generating process: $z_t \in \mathbb{R}^n$

$$z_t = \mathbf{A}^* z_{t-1} + \mathbf{B}^* \epsilon_t$$
$$x_t = c^{*\prime} z_t$$

- Model order: dimension of the model's minimal representation

$$n = \mathrm{rank}(\mathbf{Q}^*),$$

where

$$\mathbf{Q}^* = \begin{bmatrix} q_0^* & q_1^* & q_2^* & \cdots \\ q_1^* & q_2^* & q_3^* & \cdots \\ \vdots & \vdots & & \ddots \end{bmatrix} \qquad \text{and} \qquad q_h^* = \mathrm{cov}^*[x_t x_{t+h}]$$

back

Representation as Limited Memory

- An agent who constructs k moving averages $s_t = (s_{1t}, s_{2t}, \ldots, s_{kt})$ of the past realizations of the fundamentals:

$$s_{it} = w_i x_t + \sum_{j=1}^{k} q_{ij} s_{jt-1},$$

and treats these moving averages as summary statistics for making predictions about all future realizations of the fundamental:

$$\mathbb{E}_t[x_{t+\tau}] = \sum_{i=1}^{k} v_{\tau i} s_{it} \qquad \text{for all } \tau \geq 1,$$

back

Representation as Limited Memory

- An agent who constructs k moving averages $s_t = (s_{1t}, s_{2t}, \ldots, s_{kt})$ of the past realizations of the fundamentals:

$$s_{it} = w_i x_t + \sum_{j=1}^{k} q_{ij} s_{jt-1},$$

and treats these moving averages as summary statistics for making predictions about all future realizations of the fundamental:

$$\mathbb{E}_t[x_{t+\tau}] = \sum_{i=1}^{k} v_{\tau i} s_{it} \qquad \text{for all } \tau \geq 1,$$

- Parameterize such an agent: $\psi = (w, \mathbf{Q}, \{v_\tau\}_{\tau=1}^{\infty}) \in \Psi_k$.

back

Representation as Limited Memory

- An agent who constructs k moving averages $s_t = (s_{1t}, s_{2t}, \ldots, s_{kt})$ of the past realizations of the fundamentals:

$$s_{it} = w_i x_t + \sum_{j=1}^{k} q_{ij} s_{jt-1},$$

and treats these moving averages as summary statistics for making predictions about all future realizations of the fundamental:

$$\mathbb{E}_t[x_{t+\tau}] = \sum_{i=1}^{k} v_{\tau i} s_{it} \qquad \text{for all } \tau \geq 1,$$

- Parameterize such an agent: $\psi = (w, \mathbf{Q}, \{v_\tau\}_{\tau=1}^{\infty}) \in \Psi_k$.

Proposition

There is a one-to-one correspondence between Ψ_k and Θ_k.

back

Example

- ARMA(2,1) stochastic process:

$$x_t - \phi_1 x_{t-1} - \phi_2 x_{t-2} = \epsilon_t + \psi_1 \epsilon_{t-1}$$

- Hidden-factor model representation:

$$z_{t+1} = \begin{bmatrix} \phi_1 & 1 \\ \phi_2 & 0 \end{bmatrix} z_t + \begin{bmatrix} 1 \\ \psi_1 \end{bmatrix} \epsilon_t$$
$$x_t = \begin{bmatrix} 1 & 0 \end{bmatrix} z_t \, .$$

- In this case,

$$z_t = \begin{bmatrix} x_t \\ \phi_2 x_{t-1} + \psi_1 \epsilon_t \end{bmatrix}$$

back

Alternative (non-Bayesian) Estimation

Theorem

Let $\widehat{\Theta}_k$ denote an arbitrary compact subset of Θ_k.

(a) *If $\hat{\theta}_t^{\mathrm{ML}} \in \arg\max_{\theta \in \widehat{\Theta}_k} f^{\theta}(x_t, \ldots, x_0)$ is the maximum likelihood estimator,*

$$\lim_{t \to \infty} \mathrm{KL}(\theta^* \| \hat{\theta}_t^{\mathrm{ML}}) = \min_{\theta \in \widehat{\Theta}_k} \mathrm{KL}(\theta^* \| \theta) \qquad \mathbb{P}^*\text{-almost surely.}$$

- Same expectations whether agents...
  - (a) use maximum likelihood to estimate their model

back

Alternative (non-Bayesian) Estimation

Theorem

Let $\widehat{\Theta}_k$ denote an arbitrary compact subset of Θ_k.

(a) *If* $\hat{\theta}_t^{\mathrm{ML}} \in \arg\max_{\theta \in \widehat{\Theta}_k} f^\theta(x_t, \ldots, x_0)$ *is the maximum likelihood estimator,*

$$\lim_{t \to \infty} \mathrm{KL}(\theta^* \| \hat{\theta}_t^{\mathrm{ML}}) = \min_{\theta \in \widehat{\Theta}_k} \mathrm{KL}(\theta^* \| \theta) \qquad \mathbb{P}^*\text{-almost surely.}$$

(b) *If* $\hat{\theta}_t^{\mathrm{MSE}} \in \arg\min_{\theta \in \widehat{\Theta}_k} \frac{1}{t} \sum_{\tau=0}^{t} (x_{\tau+1} - \mathbb{E}_\tau^\theta[x_{\tau+1}])^2$, *then*

$$\lim_{t \to \infty} \mathrm{KL}(\theta^* \| \hat{\theta}_t^{\mathrm{MSE}}) = \min_{\theta \in \widehat{\Theta}_k} \mathrm{KL}(\theta^* \| \theta) \qquad \mathbb{P}^*\text{-almost surely.}$$

- Same expectations whether agents...
  - (a) use maximum likelihood to estimate their model
  - (b) pick their model to minimize the MSE of their one-step-ahead predictions

back

Micro-Founded Model (Allen, Morris, and Shin, 2006)

- A single asset in zero net supply with dividend stream $\{x_t\}_{t=-\infty}^{\infty}$
- Overlapping generations of traders who each live for two periods
- At each date,
  - young traders build up a position in the asset but do not consume
  - old traders unwind their position and acquire the consumption good

- Utility of acquiring q_{it} units of the asset

$$u_i(q_{it}) = (y_{t+1} - y_t + x_t)q_{it} - \frac{1}{2}\gamma q_{it}^2$$

Micro-Founded Model

- Utility of acquiring q_{it} units of the asset

$$u_i(q_{it}) = (y_{t+1} - y_t + x_t)q_{it} - \frac{1}{2}\gamma q_{it}^2$$

- First-order conditions:

$$q_{it} = \frac{1}{\gamma}(x_t - y_t + \mathbb{E}_{it}[y_{t+1}])$$

- Market clearing:

$$x_t - y_t + \int_0^1 \mathbb{E}_{it}[y_{t+1}] = 0.$$

back