Great Moderation or Great Mistake: Can overconfidence in low macro-risk explain the boom in asset prices?

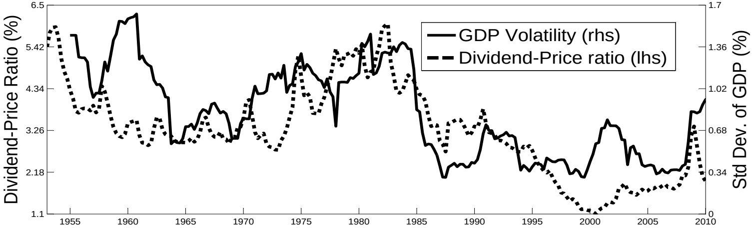
Tobias Broer, IIES Stockholm University and CEPR
Afroditi Kero, University of Cyprus

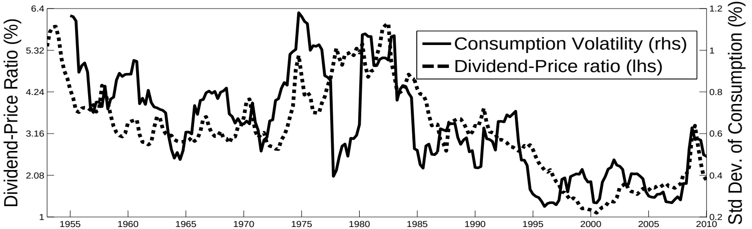
ESSIM 2012

Motivation

1. Macroeconomic Volatility and Asset Prices

Macro-volatility and US PD ratio





2. Uncertainty about origin of GM

2. Uncertainty about origin of GM

- Structural Break in Macro-Volatility noticed by Kim and Nelson (1999), McConnell and Perez-Quiros (1997, 2000)

2. Uncertainty about origin of GM

- Structural Break in Macro-Volatility noticed by Kim and Nelson (1999), McConnell and Perez-Quiros (1997, 2000)
- But: uncertainty about origin and persistence of GM

2. Uncertainty about origin of GM

- Structural Break in Macro-Volatility noticed by Kim and Nelson (1999), McConnell and Perez-Quiros (1997, 2000)
- But: uncertainty about origin and persistence of GM
  1. Academic literature: "Good Luck or Good Policy", better inventory management, globalisation, financial liberalisation or development, etc.

2. Uncertainty about origin of GM

- Structural Break in Macro-Volatility noticed by Kim and Nelson (1999), McConnell and Perez-Quiros (1997, 2000)
- But: uncertainty about origin and persistence of GM
  1. Academic literature: "Good Luck or Good Policy", better inventory management, globalisation, financial liberalisation or development, etc.
  2. Market Participants:
     "*The ongoing deterioration in surprise risk should be seen as one of the arguments behind the declining risk premium. Whether this is due to a more effective central bank policy, a major improvement in the forecast ability of economic observers around the globe, sheer luck or maybe a mix of all three factors can't finally be answered.*"
     Unicredit (2006)

3. "Overconfidence" and the Great Mistake

3. "Overconfidence" and the Great Mistake

"But what matters is how market participants responded to these benign conditions. They are faced with what is, in essence, a complex signal-extraction problem. But whereas many such problems in economics involve learning about first moments of a distribution, this involves making inferences about higher moments. The longer such a period of low volatility lasts, the more reasonable it is to assume that it is permanent. But as tail events are necessarily rarely observed, there is always going to be a danger of underestimating tail risks."

Charles Bean, European Economic Association, 25 August 2009

3. "Overconfidence" and the Great Mistake

"From the Great Moderation to the Great Conflagration: The decline in volatility led the financial institutions to underestimate the amount of risk they faced, thus essentially (though unintentionally) reintroducing a large measure of volatility into the market."
Thomas F. Cooley, Forbes.com, 11 December 2008

The stress-tests required by the authorities over the past few years were too heavily influenced by behavior during the Golden Decade. [...] The sample in question was, with hindsight, most unusual from a macroeconomic perspective. The distribution of outcomes for both macroeconomic and financial variables during the Golden Decade differed very materially from historical distributions."
Andrew Haldane, Bank of England, 13 February 2009

Motivation

- Late 80s / early 90s
  - Abrupt fall in macro-volatility (to $\approx 40 - 45\%$ of pre-GM StDev) followed by strong gradual rise in PD ratio ($75 - 200\%$)
  - GM noticed by academics and market-participants, but uncertainty about its origin and persistence
- Recent crisis
  - Abrupt rise in volatility and fall in PD ratio
  - Policymakers and academics blame *overconfidence* in benign macro-environment for *overvaluation* in asset prices

This Paper

This Paper

- Adds Bayesian learning about the persistence of volatility regimes to standard asset pricing model, to analyse a scenario similar to US post-WW II history.

This Paper

- Adds Bayesian learning about the persistence of volatility regimes to standard asset pricing model, to analyse a scenario similar to US post-WW II history.
- Research Questions
    1. Can learning significantly increase "confidence" in the Great Moderation?

This Paper

- Adds Bayesian learning about the persistence of volatility regimes to standard asset pricing model, to analyse a scenario similar to US post-WW II history.
- Research Questions
  1. Can learning significantly increase "confidence" in the Great Moderation?
  2. How large is the resulting boom in asset prices relative to that observed in the data?

This Paper

- Adds Bayesian learning about the persistence of volatility regimes to standard asset pricing model, to analyse a scenario similar to US post-WW II history.
- Research Questions
  1. Can learning significantly increase "confidence" in the Great Moderation?
  2. How large is the resulting boom in asset prices relative to that observed in the data?
  3. Is there "overvaluation" relative to full information / no learning?

Main Results

Main Results

- Learning about the great moderation can explain a boom and bust in asset prices of 30-80 percent

Main Results

- Learning about the great moderation can explain a boom and bust in asset prices of 30-80 percent
- 50 - 90 percent of this is due to *overconfidence*, as effect of Great Moderation on full information prices is small

Main Results

- Learning about the great moderation can explain a boom and bust in asset prices of 30-80 percent
- 50 - 90 percent of this is due to *overconfidence*, as effect of Great Moderation on full information prices is small
- Non-linearity of prices in transition probabilities gives special role to uncertainty about persistence

Most related literature

1. Asset prices with learning

Most related literature

1. Asset prices with learning
   - ... about mean growth: Adam et al (2010)

Most related literature

1. Asset prices with learning
    - ... about mean growth: Adam et al (2010)
    - ... about volatility: Branch et al (2010)

Most related literature

1. Asset prices with learning
   - ... about mean growth: Adam et al (2010)
   - ... about volatility: Branch et al (2010)
   - ... about different regimes:

Most related literature

1. Asset prices with learning
   - ... about mean growth: Adam et al (2010)
   - ... about volatility: Branch et al (2010)
   - ... about different regimes:
     - Zeira (1999): Unknown length of high prod. regime
     - Cogley et al (2008b): Post-depression pessimism about transition probability to low-growth regime
     - Boz et al (2010): Unknown persistence of pre-crisis high-leverage regime

Most related literature

1. Asset prices with learning
   - ... about mean growth: Adam et al (2010)
   - ... about volatility: Branch et al (2010)
   - ... about different regimes:
     - Zeira (1999): Unknown length of high prod. regime
     - Cogley et al (2008b): Post-depression pessimism about transition probability to low-growth regime
     - Boz et al (2010): Unknown persistence of pre-crisis high-leverage regime
2. Lettau et al (2008 - hf LLW) "The Declining Equity Premium: What Role Does Macroeconomic Risk Play?"
   - Switch to a regime of low volatility as in US leads to asset price boom only if it is (essentially) permanent
   - Uncertainty about current regime (posterior probabilities estimated in statistical model): boom muted, rel to full info

This paper

This paper

- Bayesian learning about persistence of volatility regimes in LLW asset pricing model

Outline

Outline

Outline

1. Model
2. Quantitative results

Outline

1. Model
2. Quantitative results
3. Sensitivity analysis

An asset pricing model with volatility regimes

An asset pricing model with volatility regimes

- Infinitely-lived representative agent

An asset pricing model with volatility regimes

- Infinitely-lived representative agent
- Preferences over $\{C_t\}_{t=0}^{\infty}$:

$$U_t = [(1-\beta)C_t^{\frac{1-\gamma}{\alpha}} + \beta(E_t U_{t+1}^{1-\gamma})^{\frac{1}{\alpha}}]^{\frac{\alpha}{1-\gamma}}$$

with $\alpha = \frac{1-\gamma}{1-\frac{1}{\psi}}$ and ψ IES

HH problem

$$\max_{C_t, S_t} \; U_t$$

$$s.t. \; S_t P_t + C_t = S_{t-1}(P_t + D_t)$$

$$S_0 \; given$$

First order condition

$$P_t = E_t[M_{t+1}(P_{t+1} + D_{t+1})] \tag{1}$$

$$\Longrightarrow \quad p_t \equiv \frac{P_t}{D_t} = E_t[M_{t+1}(p_{t+1} + 1)\frac{D_{t+1}}{D_t}] \tag{2}$$

$$\text{with} \quad M_{t+1} = (\beta(\frac{C_{t+1}}{C_t})^{-\frac{1}{\psi}})^{\alpha} R_{w,t+1}^{\alpha-1}$$

Consumption growth process with volatility regimes

$$\Delta \ln C_t = \bar{g} + \varepsilon_t$$

$$\varepsilon_t \sim N(0, \sigma_t^2)$$

$$\sigma_t^2 \in \{\sigma_l^2, \sigma_h^2\}$$

Consumption growth process with volatility regimes

$$\begin{aligned} \Delta \ln C_t &= \bar{g} + \varepsilon_t \\ &\quad \varepsilon_t \sim N(0, \sigma_t^2) \\ &\quad \sigma_t^2 \in \{\sigma_l^2, \sigma_h^2\} \end{aligned}$$

- $\Pr(\sigma_{t+1}^2 = \sigma_i^2 \mid \sigma_t^2 = \sigma_j^2)$ given by

$$\mathbf{F} = \begin{bmatrix} F_{ll} & 1 - F_{ll} \\ 1 - F_{hh} & F_{hh} \end{bmatrix}$$

"Leveraged" Dividend Process

As in Campbell 1986, Abel 1999, LLW 2008, dividend growth is (potentially) more volatile than consumption

$$\Delta lnD_t = \overline{g} + \lambda\varepsilon_t \quad \lambda \geq 1$$

Full information PD ratios

Full information PD ratios

- Given random walk assumption, p_t depends only on volatility regime

$$
\begin{aligned}
p_i &= p(\sigma_t^2 = \sigma_i^2) \\
&= F_{ii} E_{\varepsilon|\sigma_i^2} \varphi(p_i', \varepsilon') + F_{ij} E_{\varepsilon|\sigma_j^2} \varphi(p_j', \varepsilon') \quad i, j \in \{l, h\} \quad (3)
\end{aligned}
$$

Full information PD ratios

- Given random walk assumption, p_t depends only on volatility regime

$$
\begin{aligned}
p_i &= p(\sigma_t^2 = \sigma_i^2) \\
&= F_{ii} E_{\varepsilon|\sigma_i^2} \varphi(p_i', \varepsilon') + F_{ij} E_{\varepsilon|\sigma_j^2} \varphi(p_j', \varepsilon') \quad i, j \in \{l, h\} \quad (3)
\end{aligned}
$$

- Given log-normality, conditional expectations have closed form

Full information PD ratios

- Given random walk assumption, p_t depends only on volatility regime

$$\begin{aligned} p_i &= p(\sigma_t^2 = \sigma_i^2) \\ &= F_{ii} E_{\varepsilon|\sigma_i^2} \varphi(p_i', \varepsilon') + F_{ij} E_{\varepsilon|\sigma_j^2} \varphi(p_j', \varepsilon') \quad i, j \in \{l, h\} \quad (3) \end{aligned}$$

- Given log-normality, conditional expectations have closed form
- Can calculate p_h, p_l as fixed points to (3)

PD Ratios with Learning

PD Ratios with Learning

- Assume agents know 2-regime nature of environment, and observe realised volatility σ_t^2

PD Ratios with Learning

- Assume agents know 2-regime nature of environment, and observe realised volatility σ_t^2
- 2 Bayesian learning schemes:

PD Ratios with Learning

- Assume agents know 2-regime nature of environment, and observe realised volatility σ_t^2
- 2 Bayesian learning schemes:
  1. ... about transition probabilities $F_{hh}, F_{ll} \in [0, 1]$(Cogley et al 2008a,b)

PD Ratios with Learning

- Assume agents know 2-regime nature of environment, and observe realised volatility σ_t^2
- 2 Bayesian learning schemes:
  1. ... about transition probabilities $F_{hh}, F_{ll} \in [0,1]$(Cogley et al 2008a,b)
     - Beta prior distribution for unknown $\{F_{hh}, F_{ll}\}$
     - Posterior distribution: function of history $\sigma_t^2, t = 1, ..., T$

PD Ratios with Learning

- Assume agents know 2-regime nature of environment, and observe realised volatility σ_t^2
- 2 Bayesian learning schemes:
  1. ... about transition probabilities $F_{hh}, F_{ll} \in [0, 1]$(Cogley et al 2008a,b)
     - Beta prior distribution for unknown $\{F_{hh}, F_{ll}\}$
     - Posterior distribution: function of history $\sigma_t^2, t = 1, ..., T$
  2. ... about possible structural break in volatility $F_{ll} \in \{\widehat{F_{ll}}, 1\}$

PD Ratios with Learning

- Assume agents know 2-regime nature of environment, and observe realised volatility σ_t^2
- 2 Bayesian learning schemes:
  1. ... about transition probabilities $F_{hh}, F_{ll} \in [0,1]$(Cogley et al 2008a,b)
     - Beta prior distribution for unknown $\{F_{hh}, F_{ll}\}$
     - Posterior distribution: function of history $\sigma_t^2, t = 1, ..., T$
  2. ... about possible structural break in volatility $F_{ll} \in \{\widehat{F_{ll}}, 1\}$
     - $F_{hh} = \widehat{F_{hh}}, F_{ll} = \widehat{F_{ll}}$ known
     - But small (conditional) prior probability $\widehat{p}$ of structural break to permanent σ_l^2

1. Learning about Transition Probabilities

Beta Prior, Likelihood and Posterior

- Beta distribution on $[0, 1]$ summarises prior information Σ^0

$$f(F_{hh} \mid \Sigma^0) = beta(n_0^{hh}, n_0^{hl}) \propto F_{hh}^{n_0^{hh}}(1 - F_{hh})^{n_0^{hl}}$$
$$f(F_{ll} \mid \Sigma^0) = beta(n_0^{ll}, n_0^{lh}) \propto F_{ll}^{n_0^{ll}}(1 - F_{ll})^{n_0^{lh}}$$

Beta Prior, Likelihood and Posterior

- Beta distribution on $[0,1]$ summarises prior information Σ^0

$$f(F_{hh} \mid \Sigma^0) = beta(n_0^{hh}, n_0^{hl}) \propto F_{hh}^{n_0^{hh}} (1 - F_{hh})^{n_0^{hl}}$$
$$f(F_{ll} \mid \Sigma^0) = beta(n_0^{ll}, n_0^{lh}) \propto F_{ll}^{n_0^{ll}} (1 - F_{ll})^{n_0^{lh}}$$

- Likelihood of observed transitions

$$L(\Sigma^t \quad | \quad F_{hh}) \propto F_{hh}^{n_t^{hh}} (1 - F_{hh})^{n_t^{hl}}$$
$$L(\Sigma^t \quad | \quad F_{ll}) \propto F_{ll}^{n_t^{ll}} (1 - F_{ll})^{n_t^{lh}}$$

Beta Prior, Likelihood and Posterior

- Beta distribution on $[0,1]$ summarises prior information Σ^0

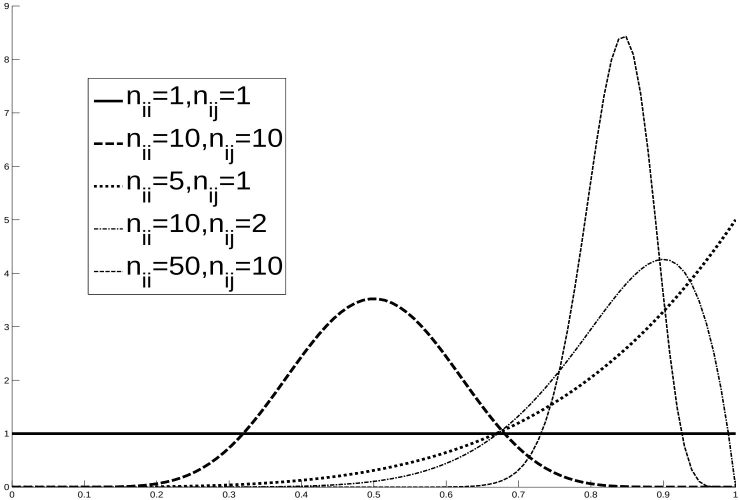
$$f(F_{hh} \mid \Sigma^0) = beta(n_0^{hh}, n_0^{hl}) \propto F_{hh}^{n_0^{hh}} (1 - F_{hh})^{n_0^{hl}}$$
$$f(F_{ll} \mid \Sigma^0) = beta(n_0^{ll}, n_0^{lh}) \propto F_{ll}^{n_0^{ll}} (1 - F_{ll})^{n_0^{lh}}$$

- Likelihood of observed transitions

$$L(\Sigma^t \quad \mid \quad F_{hh}) \propto F_{hh}^{n_t^{hh}} (1 - F_{hh})^{n_t^{hl}}$$
$$L(\Sigma^t \quad \mid \quad F_{ll}) \propto F_{ll}^{n_t^{ll}} (1 - F_{ll})^{n_t^{lh}}$$

- Yields beta posterior with updated "counters" $\widehat{n}_t^{ij} = n_t^{ij} + n_0^{ij}$.

pdf of Beta Distribution



Bayesian Learning vs. Anticipated Utility

Bayesian Learning vs. Anticipated Utility

- $Prob(\sigma^2_{t+1} = \sigma^2_i|\Sigma_t) = \int F_{ij} \, f(F|\Sigma_t) \, dF$, when $\sigma^2_t = \sigma^2_j$ so

$$p_t = \int F \, E_{\varepsilon|\sigma^2_i,F} \, \varphi(p_{t+1}, \varepsilon') \, f(F|\Sigma_t) \, dF \qquad (4)$$

Bayesian Learning vs. Anticipated Utility

- $Prob(\sigma^2_{t+1} = \sigma^2_i | \Sigma_t) = \int \; F_{ij} \; f(F|\Sigma_t) \; dF$, when $\sigma^2_t = \sigma^2_j$ so

$$p_t \quad = \quad \int F \; E_{\varepsilon|\sigma^2_i, F} \; \varphi(p_{t+1}, \varepsilon') \; f(F|\Sigma_t) \; dF \qquad (4)$$

- Bayesian Learning:

Bayesian Learning vs. Anticipated Utility

- $Prob(\sigma^2_{t+1} = \sigma^2_i | \Sigma_t) = \int \; F_{ij} \; f(F|\Sigma_t) \; dF$, when $\sigma^2_t = \sigma^2_j$ so

$$p_t \quad = \quad \int F \; E_{\varepsilon|\sigma^2_i, F} \; \varphi(p_{t+1}, \varepsilon') \; f(F|\Sigma_t) \; dF \qquad (4)$$

- Bayesian Learning:
  - $f(F|\Sigma_s), s = t, t+1, ...$ is not time-invariant

Bayesian Learning vs. Anticipated Utility

- $Prob(\sigma^2_{t+1} = \sigma^2_i | \Sigma_t) = \int F_{ij} \; f(F|\Sigma_t) \; dF$, when $\sigma^2_t = \sigma^2_j$ so

$$p_t = \int F \; E_{\varepsilon|\sigma^2_i, F} \; \varphi(p_{t+1}, \varepsilon') \; f(F|\Sigma_t) \; dF \qquad (4)$$

- Bayesian Learning:
  - $f(F|\Sigma_s), s = t, t+1, ...$ is not time-invariant
  - $Prob(\sigma^2_s|\Sigma_{s-1})$ is not a Markov Process

Bayesian Learning vs. Anticipated Utility

- $Prob(\sigma^2_{t+1} = \sigma^2_i | \Sigma_t) = \int \; F_{ij} \; f(F|\Sigma_t) \; dF$, when $\sigma^2_t = \sigma^2_j$ so

$$p_t \quad = \quad \int F \; E_{\varepsilon|\sigma^2_i, F} \; \varphi(p_{t+1}, \varepsilon') \; f(F|\Sigma_t) \; dF \qquad (4)$$

- Bayesian Learning:
  - $f(F|\Sigma_s), s = t, t+1, ...$ is not time-invariant
  - $Prob(\sigma^2_s | \Sigma_{s-1})$ is not a Markov Process
  - But $Prob(\sigma^2_s, n^{hh}_s, n^{hl}_s, n^{ll}_s, n^{lh}_s | \Sigma_s)$ is a (non-homog.) MP

Bayesian Learning vs. Anticipated Utility

- $Prob(\sigma^2_{t+1} = \sigma^2_i | \Sigma_t) = \int \; F_{ij} \; f(F|\Sigma_t) \; dF$, when $\sigma^2_t = \sigma^2_j$ so

$$p_t \quad = \quad \int F \; E_{\varepsilon|\sigma^2_i, F} \; \varphi(p_{t+1}, \varepsilon') \; f(F|\Sigma_t) \; dF \qquad (4)$$

- Bayesian Learning:
  - $f(F|\Sigma_s), s = t, t+1, \ldots$ is not time-invariant
  - $Prob(\sigma^2_s|\Sigma_{s-1})$ is not a Markov Process
  - But $Prob(\sigma^2_s, n^{hh}_s, n^{hl}_s, n^{ll}_s, n^{lh}_s|\Sigma_s)$ is a (non-homog.) MP
  - State space increases with s

Bayesian Learning vs. Anticipated Utility

- $Prob(\sigma^2_{t+1} = \sigma^2_i | \Sigma_t) = \int \; F_{ij} \; f(F|\Sigma_t) \; dF$, when $\sigma^2_t = \sigma^2_j$ so

$$p_t \quad = \quad \int F \; E_{\varepsilon|\sigma^2_i, F} \; \varphi(p_{t+1}, \varepsilon') \; f(F|\Sigma_t) \; dF \qquad (4)$$

- Bayesian Learning:
  - $f(F|\Sigma_s), s = t, t+1, \ldots$ is not time-invariant
  - $Prob(\sigma^2_s | \Sigma_{s-1})$ is not a Markov Process
  - But $Prob(\sigma^2_s, n^{hh}_s, n^{hl}_s, n^{ll}_s, n^{lh}_s | \Sigma_s)$ is a (non-homog.) MP
  - State space increases with s
- Here: "Anticipated Utility" approach (Cogley and Sargent 2008a)

Bayesian Learning vs. Anticipated Utility

- $Prob(\sigma^2_{t+1} = \sigma^2_i | \Sigma_t) = \int F_{ij} \ f(F|\Sigma_t) \ dF$, when $\sigma^2_t = \sigma^2_j$ so

$$p_t = \int F \ E_{\varepsilon|\sigma^2_i,F} \ \varphi(p_{t+1}, \varepsilon') \ f(F|\Sigma_t) \ dF \qquad (4)$$

- Bayesian Learning:
  - $f(F|\Sigma_s), s = t, t+1, ...$ is not time-invariant
  - $Prob(\sigma^2_s|\Sigma_{s-1})$ is not a Markov Process
  - But $Prob(\sigma^2_s, n^{hh}_s, n^{hl}_s, n^{ll}_s, n^{lh}_s|\Sigma_s)$ is a (non-homog.) MP
  - State space increases with s
- Here: "Anticipated Utility" approach (Cogley and Sargent 2008a)
  - Agents update $f(F|\Sigma_t)$ in line with history

Bayesian Learning vs. Anticipated Utility

- $Prob(\sigma^2_{t+1} = \sigma^2_i|\Sigma_t) = \int \; F_{ij} \; f(F|\Sigma_t) \; dF$, when $\sigma^2_t = \sigma^2_j$ so

$$p_t \quad = \quad \int F \; E_{\varepsilon|\sigma^2_i, F} \; \varphi(p_{t+1}, \varepsilon') \; f(F|\Sigma_t) \; dF \qquad (4)$$

- Bayesian Learning:
    - $f(F|\Sigma_s), s = t, t+1, ...$ is not time-invariant
    - $Prob(\sigma^2_s|\Sigma_{s-1})$ is not a Markov Process
    - But $Prob(\sigma^2_s, n^{hh}_s, n^{hl}_s, n^{ll}_s, n^{lh}_s|\Sigma_s)$ is a (non-homog.) MP
    - State space increases with s
- Here: "Anticipated Utility" approach (Cogley and Sargent 2008a)
    - Agents update $f(F|\Sigma_t)$ in line with history
    - Take $f(F|\Sigma_s) = f(F|\Sigma_t), s = t+1, t+2, ...$

Bayesian Learning vs. Anticipated Utility

- $Prob(\sigma_{t+1}^2 = \sigma_i^2 | \Sigma_t) = \int \; F_{ij} \; f(F|\Sigma_t) \; dF$, when $\sigma_t^2 = \sigma_j^2$ so

$$p_t \quad = \quad \int F \; E_{\varepsilon|\sigma_i^2, F} \; \varphi(p_{t+1}, \varepsilon') \; f(F|\Sigma_t) \; dF \qquad (4)$$

- Bayesian Learning:
  - $f(F|\Sigma_s), s = t, t+1, \ldots$ is not time-invariant
  - $Prob(\sigma_s^2 | \Sigma_{s-1})$ is not a Markov Process
  - But $Prob(\sigma_s^2, n_s^{hh}, n_s^{hl}, n_s^{ll}, n_s^{lh} | \Sigma_s)$ is a (non-homog.) MP
  - State space increases with s
- Here: "Anticipated Utility" approach (Cogley and Sargent 2008a)
  - Agents update $f(F|\Sigma_t)$ in line with history
  - Take $f(F|\Sigma_s) = f(F|\Sigma_t), s = t+1, t+2, \ldots$
  - Can calculate p_l, p_h as fixed point to (4)

1. Learning about Transition Probabilities - Discussion

1. Learning about Transition Probabilities - Discussion

- Posterior can be summarised by
  1. Mean persistence: $\widehat{F_{ll}} = n_t^{ll}/(n_t^{ll} + n_t^{hl})$
  2. Amount of information: $\widehat{n} = (n_t^{ll} + n_t^{hl})$

1. Learning about Transition Probabilities - Discussion

- Posterior can be summarised by
  1. Mean persistence: $\widehat{F_{ll}} = n_t^{ll}/(n_t^{ll} + n_t^{hl})$
  2. Amount of information: $\widehat{n} = (n_t^{ll} + n_t^{hl})$
- True transition probabilities play no role

1. Learning about Transition Probabilities - Discussion

- Posterior can be summarised by
  1. Mean persistence: $\widehat{F_{ll}} = n_t^{ll}/(n_t^{ll} + n_t^{hl})$
  2. Amount of information: $\widehat{n} = (n_t^{ll} + n_t^{hl})$
- True transition probabilities play no role
- Expectations over beta distribution give role to non-linearity of full-info asset prices in F_{ll}, F_{hh}

2. Learning about a structural break to permanent GM

2. Learning about a structural break to permanent GM

- Agents know "normal" transition probabilities $\widehat{F_{ll}}, \widehat{F_{hh}}$

2. *Learning about a structural break to permanent GM*

- Agents know "normal" transition probabilities $\widehat{F_{ll}}, \widehat{F_{hh}}$
- When $\sigma_h^2 \longrightarrow \sigma_l^2$, prior probability of permanent low volatility $P(F_{ll} = 1|\Sigma_0) = \widehat{p}$

2. Learning about a structural break to permanent GM

- Agents know "normal" transition probabilities $\widehat{F_{ll}}, \widehat{F_{hh}}$
- When $\sigma_h^2 \longrightarrow \sigma_l^2$, prior probability of permanent low volatility $P(F_{ll} = 1|\Sigma_0) = \widehat{p}$
- Likelihood of low-volatility sequence in "normal" times

$$L(\{\sigma_s^2 = \sigma_l^2, s = t, t+1, ..., t+N | \sigma_t^2 = \sigma_l^2, F_{ll} = \widehat{F_{ll}}\}) = \widehat{F_{ll}}^N$$

2. Learning about a structural break to permanent GM

- Agents know "normal" transition probabilities $\widehat{F_{ll}}, \widehat{F_{hh}}$
- When $\sigma_h^2 \longrightarrow \sigma_l^2$, prior probability of permanent low volatility $P(F_{ll} = 1|\Sigma_0) = \widehat{p}$
- Likelihood of low-volatility sequence in "normal" times

$$L(\{\sigma_s^2 = \sigma_l^2, s = t, t+1, ..., t+N|\sigma_t^2 = \sigma_l^2, F_{ll} = \widehat{F_{ll}}\}) = \widehat{F_{ll}}^N$$

- Posterior

$$\begin{aligned} P(F_{ll} = 1|\sigma_l^N) &= \frac{P(F_{ll} = 1 \wedge \sigma_l^N)}{P(F_{ll} = 1 \wedge \sigma_l^N) + P(F_{ll} = \widehat{F_{ll}} \wedge \sigma_l^N)} \\ &= \frac{\widehat{p}}{\widehat{p} + \widehat{F_{ll}}^N (1 - \widehat{p})} \qquad (5) \end{aligned}$$

2. Learning about a structural break to permanent GM

- Agents know "normal" transition probabilities $\widehat{F_{ll}}, \widehat{F_{hh}}$
- When $\sigma_h^2 \longrightarrow \sigma_l^2$, prior probability of permanent low volatility $P(F_{ll} = 1|\Sigma_0) = \widehat{p}$
- Likelihood of low-volatility sequence in "normal" times

$$L(\{\sigma_s^2 = \sigma_l^2, s = t, t+1, ..., t+N | \sigma_t^2 = \sigma_l^2, F_{ll} = \widehat{F_{ll}}\}) = \widehat{F_{ll}}^N$$

- Posterior

$$\begin{aligned} P(F_{ll} = 1|\sigma_l^N) &= \frac{P(F_{ll} = 1 \wedge \sigma_l^N)}{P(F_{ll} = 1 \wedge \sigma_l^N) + P(F_{ll} = \widehat{F_{ll}} \wedge \sigma_l^N)} \\ &= \frac{\widehat{p}}{\widehat{p} + \widehat{F_{ll}}^N (1 - \widehat{p})} \end{aligned} \quad (5)$$

- Bayesian Approach: p_t function of $\{p_{ls}, p_{hs}, P(F_{ll} = 1|\Sigma_s)\}_{s=t}^{\infty}$.

2. Learning about a structural break to permanent GM - Discussion

2. Learning about a structural break to permanent GM - Discussion

- Likelihood $\widehat{F_{ll}}^{t}$ falls geometrically

2. Learning about a structural break to permanent GM - Discussion

- Likelihood $\widehat{F_{ll}}^{t}$ falls geometrically
- True transition probability $\widehat{F_{ll}}$ is important

2. Learning about a structural break to permanent GM - Discussion

- Likelihood $\widehat{F_{ll}}^{t}$ falls geometrically
- True transition probability $\widehat{F_{ll}}$ is important
- First switch to high volatility reveals "no structural break"

Quantitative results

The exercise

The exercise

- Analyse scenario of volatility regimes similar to post-WW II US experience

The exercise

- Analyse scenario of volatility regimes similar to post-WW II US experience
    1. High volatility regime 1952Q2 - 1984Q4

The exercise

- Analyse scenario of volatility regimes similar to post-WW II US experience
  1. High volatility regime 1952Q2 - 1984Q4
  2. Great Moderation 1985Q1 - 2006Q4

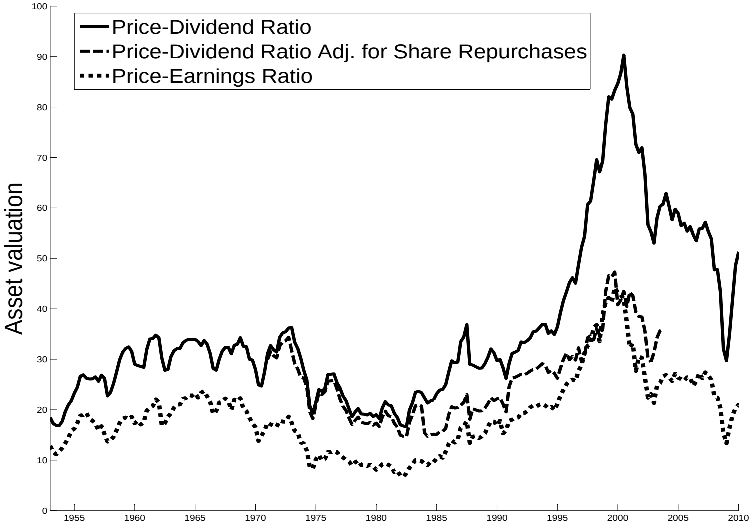
The exercise

- Analyse scenario of volatility regimes similar to post-WW II US experience
  1. High volatility regime 1952Q2 - 1984Q4
  2. Great Moderation 1985Q1 - 2006Q4
  3. Assume: Crisis marks return to high volatility regime

The exercise

- Analyse scenario of volatility regimes similar to post-WW II US experience
  1. High volatility regime 1952Q2 - 1984Q4
  2. Great Moderation 1985Q1 - 2006Q4
  3. Assume: Crisis marks return to high volatility regime
- Compare price-dividend ratios across
  - Data
  - 2 learning schemes
  - Full information price based on ex-post estimate of F

Different Valuation Measures



Calibration of preferences

β	0.9935	Discount Factor	$R = 1.95\%$ pre-GM
γ	30	Risk Aversion	$\approx$ PD-ratio pre-GM
ψ	1.5	IES	> 1 to get $\frac{dp}{d\sigma^2} < 0$

Calibration of consumption and dividend process

$\bar{g}$	0.0059		Mean ΔC	US data 1952 - 2010
σ_l	0.0037		Low Vol	US data 1985 - 2006
σ_h	0.0082		High Vol	US data 1952 - 1984
λ	4.5		Leverage	LLW
Ex-post **F**	0.989 $1-0.992$	$1-0.989$ 0.992	Trans Prob	$E[T_{\sigma_h}] = 1984 - 1952$ $E[T_{\sigma_l}] = 2006 - 1985$
	Comparison			
$\mathbf{F}^{LLW}$	0.991 $1-0.994$	$1-0.991$ 0.994	LLW (2008)	Est Reg Switch model

Results I

Learning about Transition Probabilities

Calibration of learning parameters n_0^{ij}

Calibration of learning parameters n_0^{ij}

1. Benchmark: Uninformative uniform prior $n_0^{ij} = 1,\ i, j \in \{h, l\}$

Calibration of learning parameters n_0^{ij}

1. Benchmark: Uninformative uniform prior $n_0^{ij} = 1, \; i, j \in \{h, l\}$
2. Alternative Assumption: Observed persistence of high-volatility regime increases prior persistence of low volatility

Calibration of learning parameters n_0^{ij}

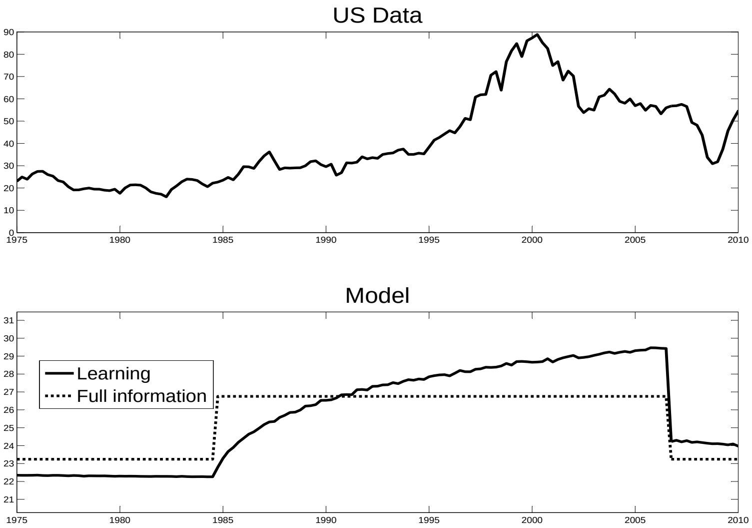
1. Benchmark: Uninformative uniform prior $n_0^{ij} = 1,\ i, j \in \{h, l\}$
2. Alternative Assumption: Observed persistence of high-volatility regime increases prior persistence of low volatility
   - Uninformative prior after WWII: $n_0^{hh} = n_0^{hl} = 1$

Calibration of learning parameters n_0^{ij}

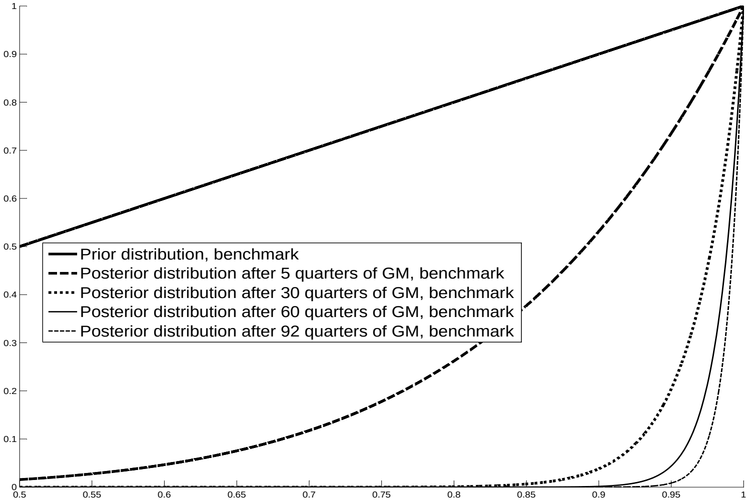
1. Benchmark: Uninformative uniform prior $n_0^{ij} = 1,\ i, j \in \{h, l\}$
2. Alternative Assumption: Observed persistence of high-volatility regime increases prior persistence of low volatility
   - Uninformative prior after WWII: $n_0^{hh} = n_0^{hl} = 1$
   - But persistent prior for GM with mean $n_0^{ll}/(n_0^{ll} + n_0^{lh}) = 0.9$

Results

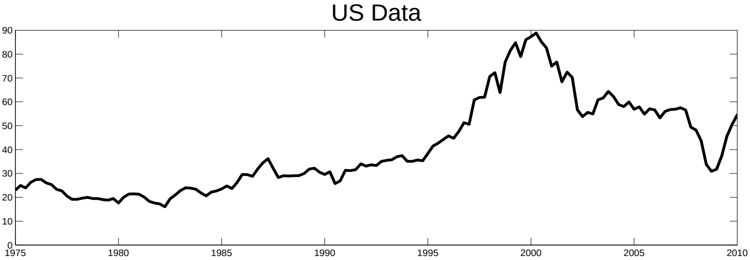
Dividend Ratios: Benchmark Model

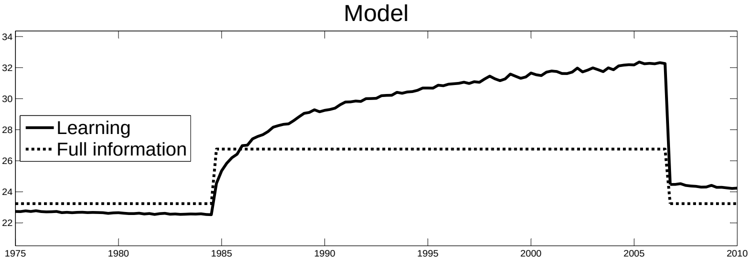


Posterior CDFs: Benchmark Model



Dividend Ratios: Persistent prior for GM





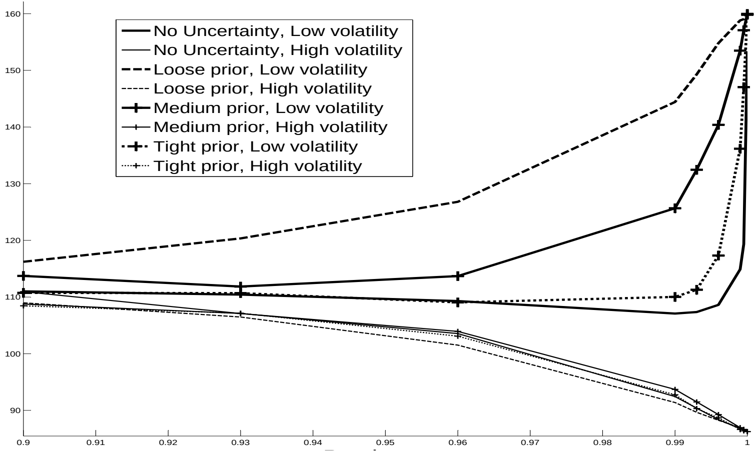
Results

Asset Prices - Learning about Transition Probabilities

	Boom	Overvaluation	Bust
Full Info	15%	0	15%
Uninformed prior	32%	11%	22%
Moderately persistent prior	42%	20%	32%

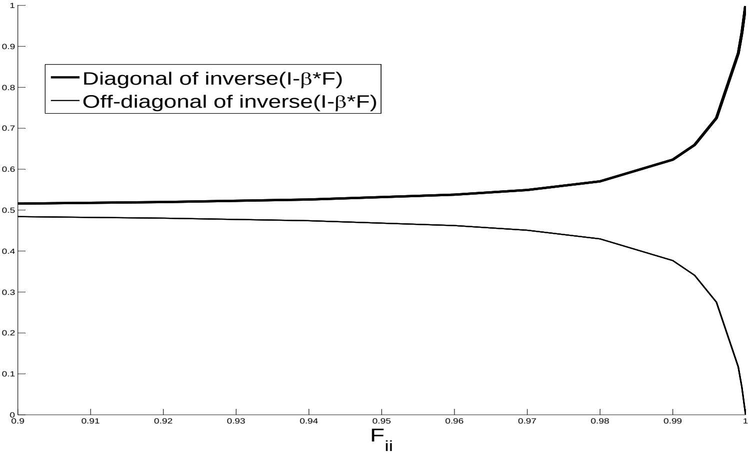
Discussion: Mean and Variance Effects of Learning about Transition Probabilities

Dividend Ratios as a Function of (symmetric) mean Persistence and Prior Tightness



Persistence F_{ll}, F_{hh}, Posterior means

Entries of $(I - \beta\mathbb{F})^{-1}$ *for symmetric* $\mathbb{F}$



Results II

Learning about a structural break to permanent GM

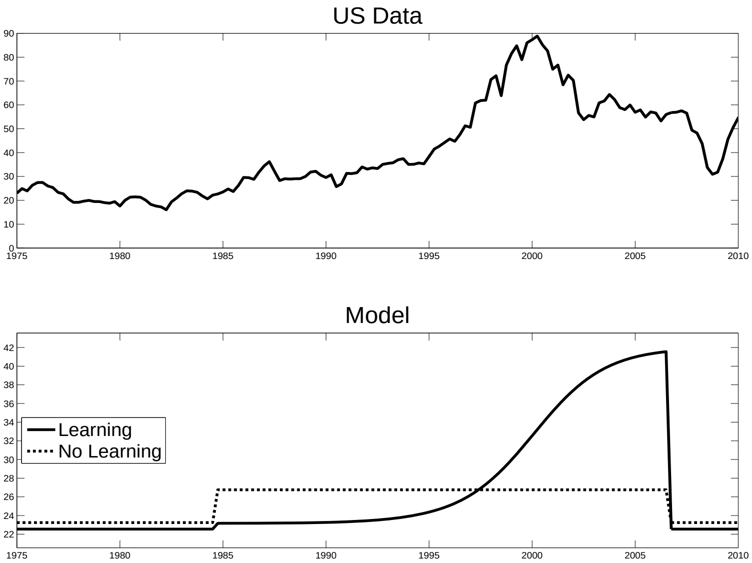
Calibration of learning parameters

- Need:
  1. Prior probability of structural break $\widehat{p}$
  2. Trans probabilities $\widehat{F_{ll}}, \widehat{F_{hh}}$ in "normal times"

Calibration of learning parameters

- Conditional prior probability of structural break: 1 percent
- Transition probablities in "normal" times:
  1. Take ex-post estimate of F_{hh}.
  2. Choose $\widehat{F_{ll}}$ to imply "suspicion" about structural break in mid-1990s:
     $Prob(\sigma_t^2 = \sigma_l^2, t = 1, ..., 48 | \widehat{F_{ll}}) = 0.1$
  3. Yields $\widehat{F_{ll}} = 0.87$

Dividend Ratios

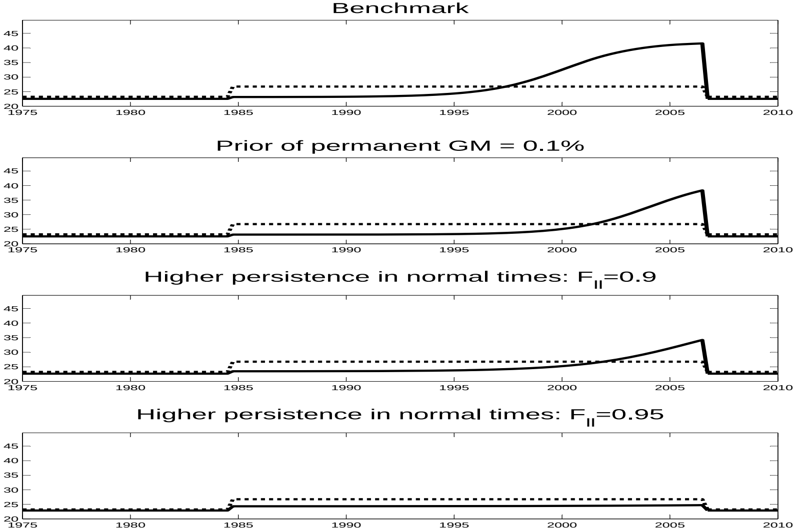


Results

Asset prices - learning about structural break

	Boom	Overvaluation	Bust
Full Info	3%	0	3%
Struct Break	77%	79%	84%

Struct-break learning: Alternative priors



Sensitivity Analysis

Sensitivity Analysis

1. Non-Bayesian, ad-hoc Learning

Sensitivity Analysis

1. Non-Bayesian, ad-hoc Learning
2. Risk Aversion and Dividend Volatility

Ad-hoc / recursive learning

Ad-hoc / recursive learning

- Agents ignore 2-regime nature of dividend process, use simple rules to infer "average" volatility of dividends

Ad-hoc / recursive learning

- Agents ignore 2-regime nature of dividend process, use simple rules to infer "average" volatility of dividends
    1. OLS estimate on full sample

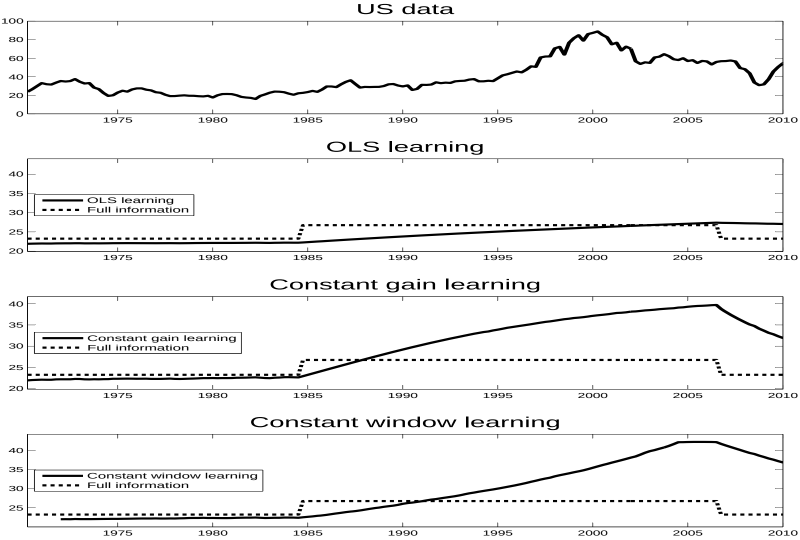
Ad-hoc / recursive learning

- Agents ignore 2-regime nature of dividend process, use simple rules to infer "average" volatility of dividends
  1. OLS estimate on full sample
  2. OLS estimate on history of constant length

Ad-hoc / recursive learning

- Agents ignore 2-regime nature of dividend process, use simple rules to infer "average" volatility of dividends
  1. OLS estimate on full sample
  2. OLS estimate on history of constant length
  3. Constant gain learning: higher weight on recent observations

Ad-Hoc Learning: Results



Sensitivity: lower leverage and risk aversion

Sensitivity: lower leverage and risk aversion

Asset prices - Lower Risk Aversion and Dividend Volatility

	Boom	Overvaluation	Bust
Full Info (benchmark)	15%	0	15%
Full Info ($\gamma = 20$)	12%	0	12%
Full Info ($\lambda = 2.5$)	8%	0	7%
Trans Probabilities (benchmark)	32%	11%	22%
Trans Probabilities ($\gamma = 20$)	30%	9%	18%
Trans Probabilities ($\lambda = 2.5$)	17%	6%	11%
Perm vs Trans (benchmark)	77%	79%	84%
Perm vs Trans ($\gamma = 20$)	59%	56%	59%
Perm vs Trans ($\lambda = 2.5$)	45%	43%	45%

Conclusion

- Standard asset pricing model with Bayesian learning responds with strong boom and bust to GM
- Much of this is due to *overconfidence*, as effect of GM on full information prices is small
- Non-linearity of prices in regime-persistence gives special role to uncertainty about transition probabilities
- Bayesian assumption crucial for strong bust

Great Moderation or Great Mistake: Can overconfidence in low macro-risk explain the boom in asset prices?

Tobias Broer, IIES Stockholm University and CEPR
Afroditi Kero, University of Cyprus

ESSIM 2012

Moments of Consumption Growth

Date	*Mean*	*StDev*
$1952Q2 : 1991Q4$	0.57%	0.82%
$1992Q1 : 2006Q4$	0.61%	0.36%
$2007Q1 : 2010Q2$	−0.19%	0.50%

Moments of GDP growth

Date	*Mean*	*StDev*
$1952Q2:1983Q4$	0.53%	1.1%
$1984Q1:2006Q4$	0.51%	0.51%
$2007Q1:2010Q2$	−0.16%	0.90%

Full information asset prices

$$p_{\sigma_t^2=\sigma_i^2} = \rho_i^{1-a}\beta^a e^{\left(-\frac{a}{\psi}+a\right)\bar{g}}$$
$$\{F_{ii} e^{\frac{\left(-\frac{a}{\psi}+a-1+\lambda\right)^2}{2}\sigma_i^2}(1+\rho_i)^{a-1}(1+p_i)$$
$$+F_{ij} e^{\frac{\left(-\frac{a}{\psi}+a-1+\lambda\right)^2}{2}\sigma_j^2}(1+\rho_j)^{a-1}(1+p_j)\}$$

with

$$\rho^a_{\sigma_t^2=\sigma_i^2} = (\frac{P^C_{it}}{C_{it}})^a = \rho_i^{1-a}\beta^a e^{\left(-\frac{a}{\psi}+a\right)\bar{g}}$$
$$\left(F_{ii} e^{\frac{\left(-\frac{a}{\psi}+a\right)^2}{2}\sigma_i^2}(1+\rho_i)^a + F_{ij} e^{\frac{\left(-\frac{a}{\psi}+a\right)^2}{2}\sigma_j^2}(1+\rho_j)^a\right)$$

Ad-hoc learning: Lower gain and longer windows