Estimating Macroeconomic Models:
A Likelihood Approach

Jesús Fernández-Villaverde
Duke University, NBER, and CEPR

Juan Rubio-Ramírez
Duke University

Outline

- We are interested in performing likelihood-based inference in nonlinear and/or non-normal DSGE models.

Outline

- We are interested in performing likelihood-based inference in nonlinear and/or non-normal DSGE models.
- We apply particle filtering to evaluate the likelihood of the model.

Outline

- We are interested in performing likelihood-based inference in nonlinear and/or non-normal DSGE models.
- We apply particle filtering to evaluate the likelihood of the model.
- We estimate a neoclassical business cycle model with investment-specific technological change and stochastic volatility.

What is the Particle Filter?

- The particle filter is a monte carlo that tracks the distribution of states conditional on a sequence of observables.

What is the Particle Filter?

- The particle filter is a monte carlo that tracks the distribution of states conditional on a sequence of observables.

- Original idea seems to be by Handschin and Mayne (1969).

What is the Particle Filter?

- The particle filter is a monte carlo that tracks the distribution of states conditional on a sequence of observables.

- Original idea seems to be by Handschin and Mayne (1969).

- Applied in financial econometrics by Kim, Shephard, and Chib (1998).

What is the Particle Filter?

- The particle filter is a monte carlo that tracks the distribution of states conditional on a sequence of observables.
- Original idea seems to be by Handschin and Mayne (1969).
- Applied in financial econometrics by Kim, Shephard, and Chib (1998).
- We modify the filter to be more flexible with shocks.

What is the Particle Filter?

- The particle filter is a monte carlo that tracks the distribution of states conditional on a sequence of observables.
- Original idea seems to be by Handschin and Mayne (1969).
- Applied in financial econometrics by Kim, Shephard, and Chib (1998).
- We modify the filter to be more flexible with shocks.
- Alternatives? Kim, Shephard, and Chib (1998), Fiorentini, Sentana, and Shepard (2004)

What is the Particle Filter?

- The particle filter is a monte carlo that tracks the distribution of states conditional on a sequence of observables.
- Original idea seems to be by Handschin and Mayne (1969).
- Applied in financial econometrics by Kim, Shephard, and Chib (1998).
- We modify the filter to be more flexible with shocks.
- Alternatives? Kim, Shephard, and Chib (1998), Fiorentini, Sentana, and Shepard (2004)
- Problems?

Why Are Nonlinearities Important?

Why Are Nonlinearities Important?

- Most DSGE models are nonlinear.

Why Are Nonlinearities Important?

- Most DSGE models are nonlinear.
- Common practice: estimate a linearized version.

Why Are Nonlinearities Important?

- Most DSGE models are nonlinear.
- Common practice: estimate a linearized version.
- Linearization eliminates asymmetries, threshold effects, precautionary behavior, big shocks,....

Why Are Nonlinearities Important?

- Most DSGE models are nonlinear.
- Common practice: estimate a linearized version.
- Linearization eliminates asymmetries, threshold effects, precautionary behavior, big shocks,....
- Moreover, linearization induces an approximation error.

Why Are Nonlinearities Important?

- Most DSGE models are nonlinear.
- Common practice: estimate a linearized version.
- Linearization eliminates asymmetries, threshold effects, precautionary behavior, big shocks,....
- Moreover, linearization induces an approximation error.
- This is worse than you may think:

Why Are Nonlinearities Important?

- Most DSGE models are nonlinear.
- Common practice: estimate a linearized version.
- Linearization eliminates asymmetries, threshold effects, precautionary behavior, big shocks,....
- Moreover, linearization induces an approximation error.
- This is worse than you may think:

1. Theoretical arguments. "Convergence Properties of the Likelihood of Computed Dynamic Models" (2006)

Why Are Nonlinearities Important?

- Most DSGE models are nonlinear.
- Common practice: estimate a linearized version.
- Linearization eliminates asymmetries, threshold effects, precautionary behavior, big shocks,....
- Moreover, linearization induces an approximation error.
- This is worse than you may think:
  1. Theoretical arguments. "Convergence Properties of the Likelihood of Computed Dynamic Models" (2006)
  2. Computational evidence.

Why Are Non-normalities Important?

Why Are Non-normalities Important?

- Evidence of time-varying volatility in time series.

Why Are Non-normalities Important?

- Evidence of time-varying volatility in time series.
- Fundamental issue in financial econometrics.

Why Are Non-normalities Important?

- Evidence of time-varying volatility in time series.
- Fundamental issue in financial econometrics.
- However, macro has lagged despite Engle (1982).

Why Are Non-normalities Important?

- Evidence of time-varying volatility in time series.
- Fundamental issue in financial econometrics.
- However, macro has lagged despite Engle (1982).
- Evidence by Geweke (1993 and 1994).

Why Are Non-normalities Important?

- Evidence of time-varying volatility in time series.
- Fundamental issue in financial econometrics.
- However, macro has lagged despite Engle (1982).
- Evidence by Geweke (1993 and 1994).
- Recent discussion about the Great Moderation:

Why Are Non-normalities Important?

- Evidence of time-varying volatility in time series.
- Fundamental issue in financial econometrics.
- However, macro has lagged despite Engle (1982).
- Evidence by Geweke (1993 and 1994).
- Recent discussion about the Great Moderation:
    1. Kim and Nelson (1999), McConell and Pérez-Quirós (2000), and Stock and Watson (2002).

Why Are Non-normalities Important?

- Evidence of time-varying volatility in time series.
- Fundamental issue in financial econometrics.
- However, macro has lagged despite Engle (1982).
- Evidence by Geweke (1993 and 1994).
- Recent discussion about the Great Moderation:
  1. Kim and Nelson (1999), McConell and Pérez-Quirós (2000), and Stock and Watson (2002).
  2. Good luck? Sims and Zha (2005).

Why Are Non-normalities Important?

- Evidence of time-varying volatility in time series.
- Fundamental issue in financial econometrics.
- However, macro has lagged despite Engle (1982).
- Evidence by Geweke (1993 and 1994).
- Recent discussion about the Great Moderation:

1. Kim and Nelson (1999), McConell and Pérez-Quirós (2000), and Stock and Watson (2002).
2. Good luck? Sims and Zha (2005).
3. Or good policy? Clarida, Galí, and Gertler (2000).

State Space Representation of the Model

- How do we evaluate the likelihood $p\left(y^T;\gamma\right)$ of a DSGE model?

State Space Representation of the Model

- How do we evaluate the likelihood $p\left(y^T;\gamma\right)$ of a DSGE model?

- Transition equation:

$$S_t = f\left(S_{t-1}, W_t; \gamma\right) \Rightarrow p\left(S_t | S_{t-1}; \gamma\right)$$

State Space Representation of the Model

- How do we evaluate the likelihood $p\left(y^T;\gamma\right)$ of a DSGE model?

- Transition equation:

$$S_t = f\left(S_{t-1}, W_t; \gamma\right) \Rightarrow p\left(S_t | S_{t-1}; \gamma\right)$$

- Measurement equation:

$$Y_t = g\left(S_t, V_t; \gamma\right) \Rightarrow p\left(Y_t | S_t; \gamma\right)$$

State Space Representation of the Model

- How do we evaluate the likelihood $p\left(y^T;\gamma\right)$ of a DSGE model?

- Transition equation:

$$S_t = f\left(S_{t-1}, W_t; \gamma\right) \Rightarrow p\left(S_t | S_{t-1}; \gamma\right)$$

- Measurement equation:

$$Y_t = g\left(S_t, V_t; \gamma\right) \Rightarrow p\left(Y_t | S_t; \gamma\right)$$

- We want to track conditional density $p\left(S_t | y^{t-1}; \gamma\right)$.

Factorization of the Likelihood

- Why?

$$
\begin{aligned}
p\left(y^{T};\gamma\right) &= \prod_{t=1}^{T} p\left(y_t|y^{t-1};\gamma\right) \\
&= \prod_{t=1}^{T} \int p\left(y_t|S_t;\gamma\right) p\left(S_t|y^{t-1};\gamma\right) dS_t
\end{aligned}
$$

Factorization of the Likelihood

- Why?

$$
\begin{aligned}
p\left(y^T;\gamma\right) &= \prod_{t=1}^{T} p\left(y_t|y^{t-1};\gamma\right) \\
&= \prod_{t=1}^{T} \int p\left(y_t|S_t;\gamma\right) p\left(S_t|y^{t-1};\gamma\right) dS_t
\end{aligned}
$$

- Knowledge of $\left\{p\left(S_t|y^{t-1};\gamma\right)\right\}_{t=1}^{T}$ allows the evaluation of the likelihood of the model.

Tracking the Conditional Distribution of States S_t

Filtering problem: forecast and update.

Tracking the Conditional Distribution of States S_t

Filtering problem: forecast and update.

1. Forecast: Chapman-Kolmogorov equation

$$p\left(S_t|y^{t-1};\gamma\right) = \int p\left(S_t|S_{t-1};\gamma\right) p\left(S_{t-1}|y^{t-1};\gamma\right) dS_{t-1}$$

Tracking the Conditional Distribution of States S_t

Filing problem: forecast and update.

1. Forecast: Chapman-Kolmogorov equation

$$p\left(S_t|y^{t-1};\gamma\right) = \int p\left(S_t|S_{t-1};\gamma\right) p\left(S_{t-1}|y^{t-1};\gamma\right) dS_{t-1}$$

2. Update: Bayes' theorem

$$p\left(S_t|y^{t};\gamma\right) = \frac{p\left(y_t|S_t;\gamma\right) p\left(S_t|y^{t-1};\gamma\right)}{p\left(y_t|y^{t-1};\gamma\right)}$$

where:

$$p\left(y_t|y^{t-1};\gamma\right) = \int p\left(y_t|S_t;\gamma\right) p\left(S_t|y^{t-1};\gamma\right) dS_t$$

A Law of Large Numbers

$$\prod_{t=1}^{T} \int p\left(y_t|S_t;\gamma\right) p\left(S_t|y^{t-1};\gamma\right) dS_t$$

A Law of Large Numbers

$$\prod_{t=1}^{T} \int p\left(y_t | S_t; \gamma\right) p\left(S_t | y^{t-1}; \gamma\right) dS_t$$

Suppose we have $\left\{ \left\{ s_{t|t-1}^{i} \right\}_{i=1}^{N} \right\}_{t=1}^{T} \sim \left\{ p\left(S_t | y^{t-1}; \gamma\right) \right\}_{t=1}^{T}$.

A Law of Large Numbers

$$\prod_{t=1}^{T} \int p\left(y_t | S_t; \gamma\right) p\left(S_t | y^{t-1}; \gamma\right) dS_t$$

Suppose we have $\left\{ \left\{ s_{t|t-1}^{i} \right\}_{i=1}^{N} \right\}_{t=1}^{T} \sim \left\{ p\left(S_t | y^{t-1}; \gamma\right) \right\}_{t=1}^{T}$.

Then:

$$p\left(y^T; \gamma\right) \simeq \prod_{t=1}^{T} \frac{1}{N} \sum_{i=1}^{N} p\left(y_t | s_{t|t-1}^{i}; \gamma\right)$$

A Law of Large Numbers

$$\prod_{t=1}^{T} \int p\left(y_t | S_t; \gamma\right) p\left(S_t | y^{t-1}; \gamma\right) dS_t$$

Suppose we have $\left\{ \left\{ s_{t|t-1}^{i} \right\}_{i=1}^{N} \right\}_{t=1}^{T} \sim \left\{ p\left(S_t | y^{t-1}; \gamma\right) \right\}_{t=1}^{T}$.

Then:

$$p\left(y^T; \gamma\right) \simeq \prod_{t=1}^{T} \frac{1}{N} \sum_{i=1}^{N} p\left(y_t | s_{t|t-1}^{i}; \gamma\right)$$

Evaluating the likelihood function $\Leftrightarrow$ Drawing from density:

$$\left\{ p\left(S_t | y^{t-1}; \gamma\right) \right\}_{t=1}^{T}$$

A Proposition

- Let $\left\{ s^{i}_{t|t-1} \right\}_{i=1}^{N}$ be a draw from $p\left(S_t|y^{t-1};\gamma\right)$.

A Proposition

- Let $\left\{s^i_{t|t-1}\right\}_{i=1}^N$ be a draw from $p\left(S_t|y^{t-1};\gamma\right)$.
- Let $\{\widetilde{s}^i_t\}_{i=1}^N$ be a draw with replacement from $\left\{s^i_{t|t-1}\right\}_{i=1}^N$ and probabilities:

$$q^i_t = \frac{p\left(y_t|s^i_{t|t-1};\gamma\right)}{\sum_{i=1}^N p\left(y_t|s^i_{t|t-1};\gamma\right)}$$

A Proposition

- Let $\left\{s_{t|t-1}^{i}\right\}_{i=1}^{N}$ be a draw from $p\left(S_t|y^{t-1};\gamma\right)$.
- Let $\{\widetilde{s}_t^i\}_{i=1}^{N}$ be a draw with replacement from $\left\{s_{t|t-1}^{i}\right\}_{i=1}^{N}$ and probabilities:

$$q_t^i = \frac{p\left(y_t|s_{t|t-1}^i;\gamma\right)}{\sum_{i=1}^{N} p\left(y_t|s_{t|t-1}^i;\gamma\right)}$$

- Then, $\{\widetilde{s}_t^i\}_{i=1}^{N}$ is a draw from $p\left(S_t|y^t;\gamma\right)$:

$$\left\{s_{t|t}^{i}\right\}_{i=1}^{N} = \{\widetilde{s}_t^i\}_{i=1}^{N}$$

A Proposition

- Let $\left\{s^i_{t|t-1}\right\}_{i=1}^N$ be a draw from $p\left(S_t|y^{t-1};\gamma\right)$.
- Let $\{\widetilde{s}^i_t\}_{i=1}^N$ be a draw with replacement from $\left\{s^i_{t|t-1}\right\}_{i=1}^N$ and probabilities:

$$q^i_t = \frac{p\left(y_t|s^i_{t|t-1};\gamma\right)}{\sum_{i=1}^N p\left(y_t|s^i_{t|t-1};\gamma\right)}$$

- Then, $\{\widetilde{s}^i_t\}_{i=1}^N$ is a draw from $p\left(S_t|y^t;\gamma\right)$:

$$\left\{s^i_{t|t}\right\}_{i=1}^N = \{\widetilde{s}^i_t\}_{i=1}^N$$

- Proof: Importance sampling and Bayes' theorem.

Importance of the Proposition

Importance of the Proposition

1. Update: We can use a draw $\left\{s_{t|t-1}^{i}\right\}_{i=1}^{N}$ from $p\left(S_t|y^{t-1};\gamma\right)$ to get a draw $\left\{s_{t|t}^{i}\right\}_{i=1}^{N}$ from $p\left(S_t|y^{t};\gamma\right)$.

Importance of the Proposition

1. Update: We can use a draw $\left\{s^i_{t|t-1}\right\}_{i=1}^N$ from $p\left(S_t|y^{t-1};\gamma\right)$ to get a draw $\left\{s^i_{t|t}\right\}_{i=1}^N$ from $p\left(S_t|y^t;\gamma\right)$.

2. Forecast: We can use a draw $\left\{s^i_{t|t}\right\}_{i=1}^N$ from $p\left(S_t|y^t;\gamma\right)$, a draw from $p\left(W_{t+1};\gamma\right)$, and $S_{t+1}=f\left(S_t,W_{t+1};\gamma\right)$ to get a draw $\left\{s^i_{t+1|t}\right\}_{i=1}^N$.

Particle Filtering I

Step 0, Initialization: Sample N values $\left\{s^i_{1|0}\right\}_{i=1}^N$ from $p\left(S_1;\gamma\right)$. Go to step 2.

Step 1, Forecast: Sample N values $\left\{s^i_{t|t-1}\right\}_{i=1}^N$ with $p\left(W_t;\gamma\right)$ and $S_t = f\left(S_{t-1}, W_t;\gamma\right)$.

Step 2, Weighting: Assign to each draw $s^i_{t|t-1}$ the weight q^i_t.

Step 3, Update: Draw $\left\{s^i_{t|t}\right\}_{i=1}^N$ with replacement from $\left\{s^i_{t|t-1}\right\}_{i=1}^N$ with probabilities $\left\{q^i_t\right\}_{i=1}^N$. If $t < T$ set $t \rightsquigarrow t+1$ and go to step 1. Otherwise stop.

Particle Filtering II

Use $\left\{\left\{s_{t|t-1}^{i}\right\}_{i=1}^{N}\right\}_{t=1}^{T}$ to compute:

$$p\left(y^{T};\gamma\right)\simeq\prod_{t=1}^{T}\frac{1}{N}\sum_{i=1}^{N}p\left(y_{t}|s_{t|t-1}^{i};\gamma\right)$$

We can filter, forecast, and smooth

Numerical accuracy: effective sample size, density ratio test, etc.

An Application: a Business Cycle Model

- I want to show the power of the particle filter.

An Application: a Business Cycle Model

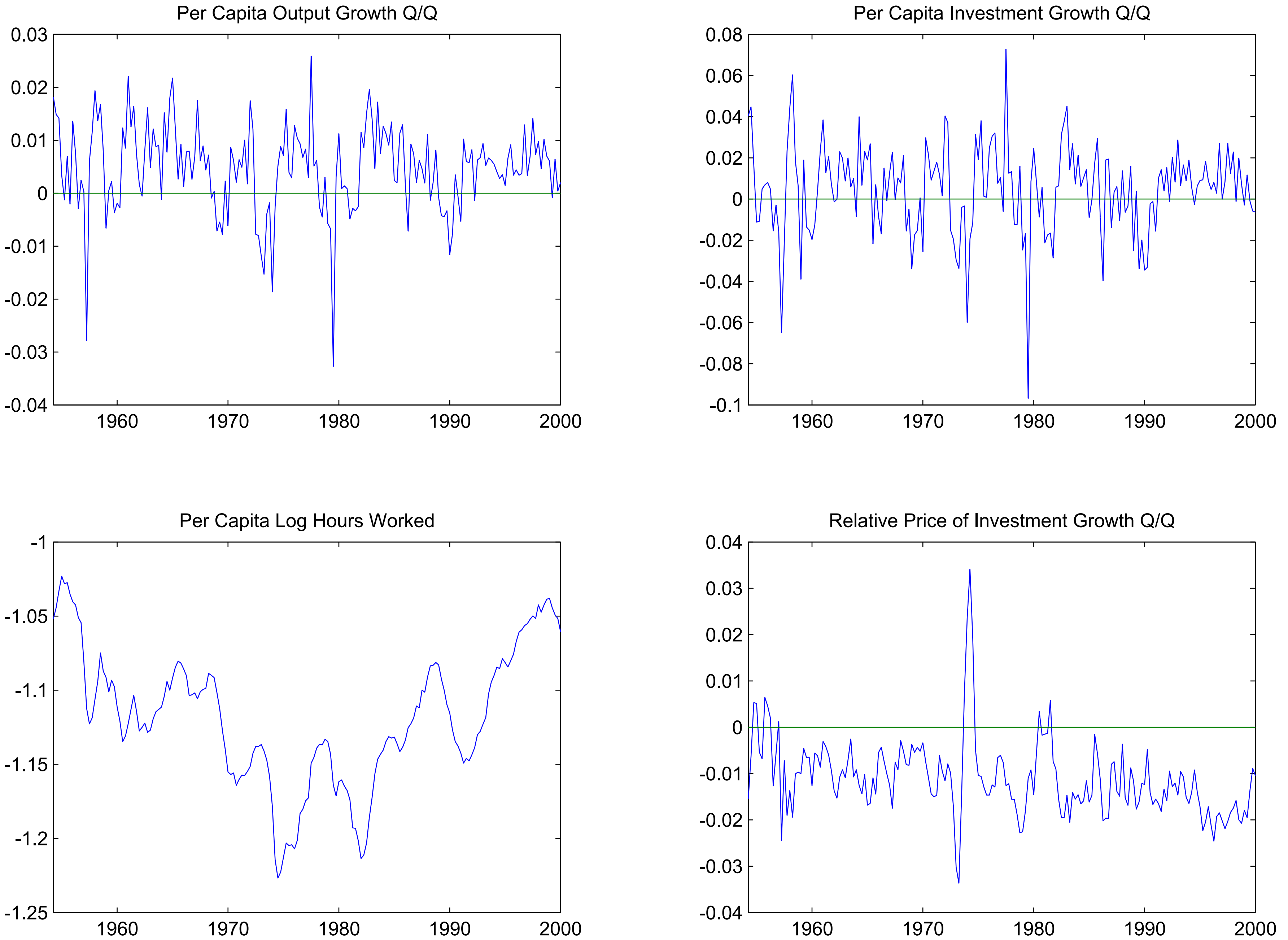
- I want to show the power of the particle filter.

- And learn something about the evolution of the U.S. aggregate fluctuations over the last decades.

An Application: a Business Cycle Model

- I want to show the power of the particle filter.
- And learn something about the evolution of the U.S. aggregate fluctuations over the last decades.
- A business cycle model with:

1. Investment-specific technological change. Greenwood, Herkowitz, and Krusell (1997 and 2000)
2. Stochastic volatility.


Per Capita Output Growth Q/Q
Per Capita Investment Growth Q/Q
Per Capita Log Hours Worked
Relative Price of Investment Growth Q/Q
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
0.08
0.06
0.04
0.02
-0.02
-0.04
-0.06
-0.08
-0.1
-1
-1.05
-1.1
-1.15
-1.2
-1.25
1960
1970
1980
1990
2000

Environment

- Representative household with utility function:

$$E_0 \sum_{t=0}^{\infty} \beta^t \left(e^{d_t} \log C_t + \psi \log (1 - L_t) \right)$$

- Final Good: $C_t + X_t = A_t K_t^{\alpha} L_t^{1-\alpha}$
- Law of motion of capital: $K_{t+1} = (1 - \delta) K_t + V_t X_t$
- Shocks:

$$\begin{aligned} d_t &= \rho d_{t-1} + \sigma_{dt} \varepsilon_{dt}, \ \varepsilon_{dt} \sim \mathcal{N}(0,1) \\ \log A_t &= \zeta + \log A_{t-1} + \sigma_{at} \varepsilon_{at}, \ \gamma \geq 0 \text{ and } \varepsilon_{at} \sim \mathcal{N}(0,1) \\ \log V_t &= \upsilon + \log V_{t-1} + \sigma_{\upsilon t} \varepsilon_{\upsilon t}, \ \upsilon \geq 0 \text{ and } \varepsilon_{\upsilon t} \sim \mathcal{N}(0,1) \end{aligned}$$

Stochastic Volatility

We follow a standard specification:

$$\log \sigma_{dt} = (1 - \lambda_d) \log \overline{\sigma}_d + \lambda_d \log \sigma_{dt-1} + \tau_d \eta_{dt} \text{ and } \eta_{dt} \sim \mathcal{N}(0,1)$$

$$\log \sigma_{at} = (1 - \lambda_a) \log \overline{\sigma}_a + \lambda_a \log \sigma_{at-1} + \tau_a \eta_{at} \text{ and } \eta_{at} \sim \mathcal{N}(0,1)$$

$$\log \sigma_{\upsilon t} = (1 - \lambda_\upsilon) \log \overline{\sigma}_\upsilon + \lambda_\upsilon \log \sigma_{\upsilon t-1} + \tau_\upsilon \eta_{\upsilon t} \text{ and } \eta_{\upsilon t} \sim \mathcal{N}(0,1)$$

Performing Likelihood-Based Inference

- We compute the model using a perturbation method.
- Time series:

  1. Relative price of capital, output, investment, and hours.
  2. Sample: 1955:Q1 to 2000:Q4.

- Vector of parameters γ is:

  $$(\rho, \beta, \psi, \alpha, \delta, \upsilon, \zeta, \tau_d, \tau_a, \tau_\upsilon, \overline{\sigma}_d, \overline{\sigma}_a, \overline{\sigma}_\upsilon, \lambda_a, \lambda_\upsilon, \lambda_d, \sigma_1^\epsilon, \sigma_2^\epsilon, \sigma_3^\epsilon)$$

- Use a Random-walk Metropolis-Hastings to explore the likelihood: Classical and Bayesian.

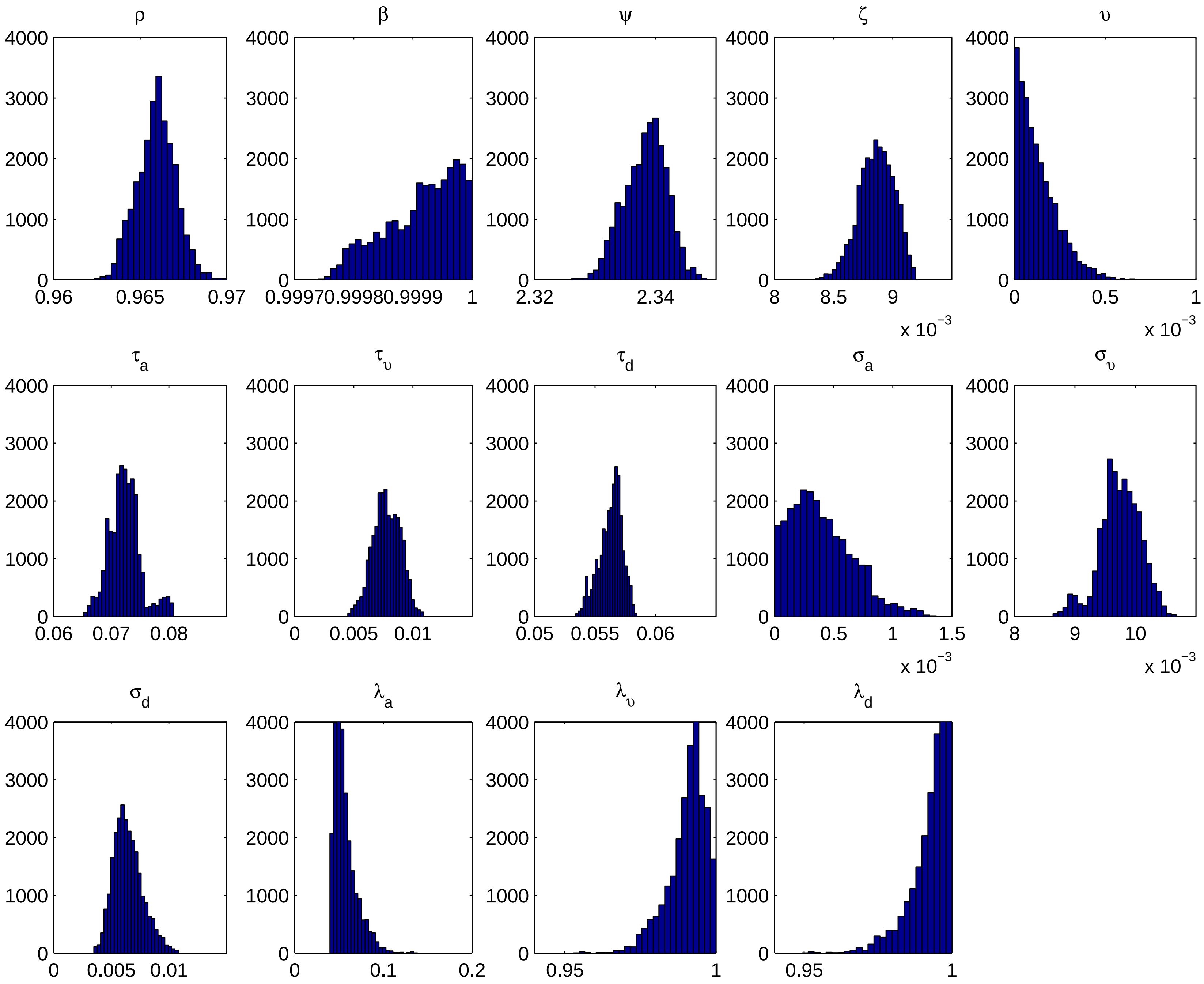

ρ
β
ψ
ζ
υ
τa
τυ
τd
σa
συ
σd
λa
λυ
λd
x 10^−3

Figure 6.1: Model versus Data

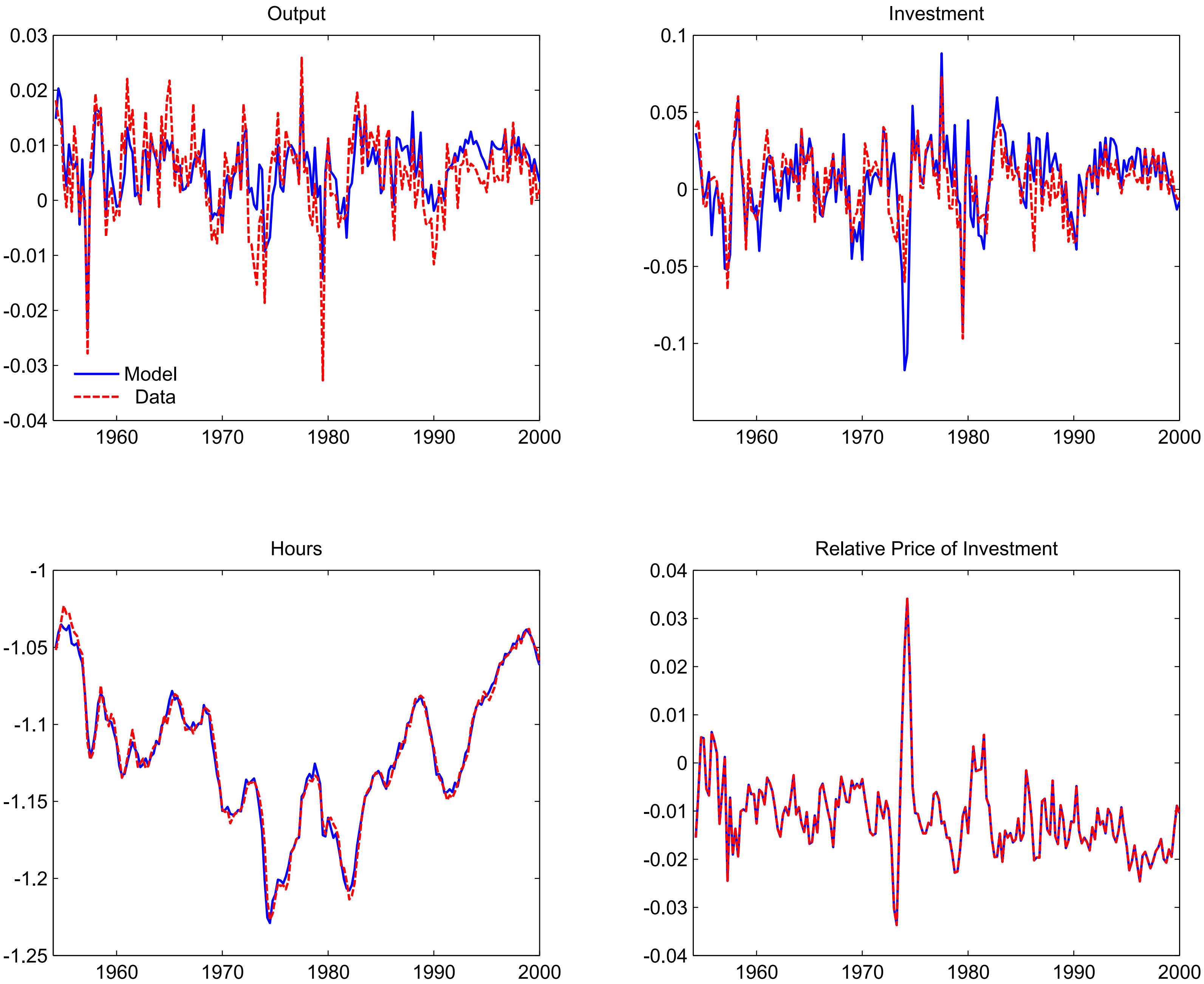


Figure 6.2: Smoothed Capital and Shocks

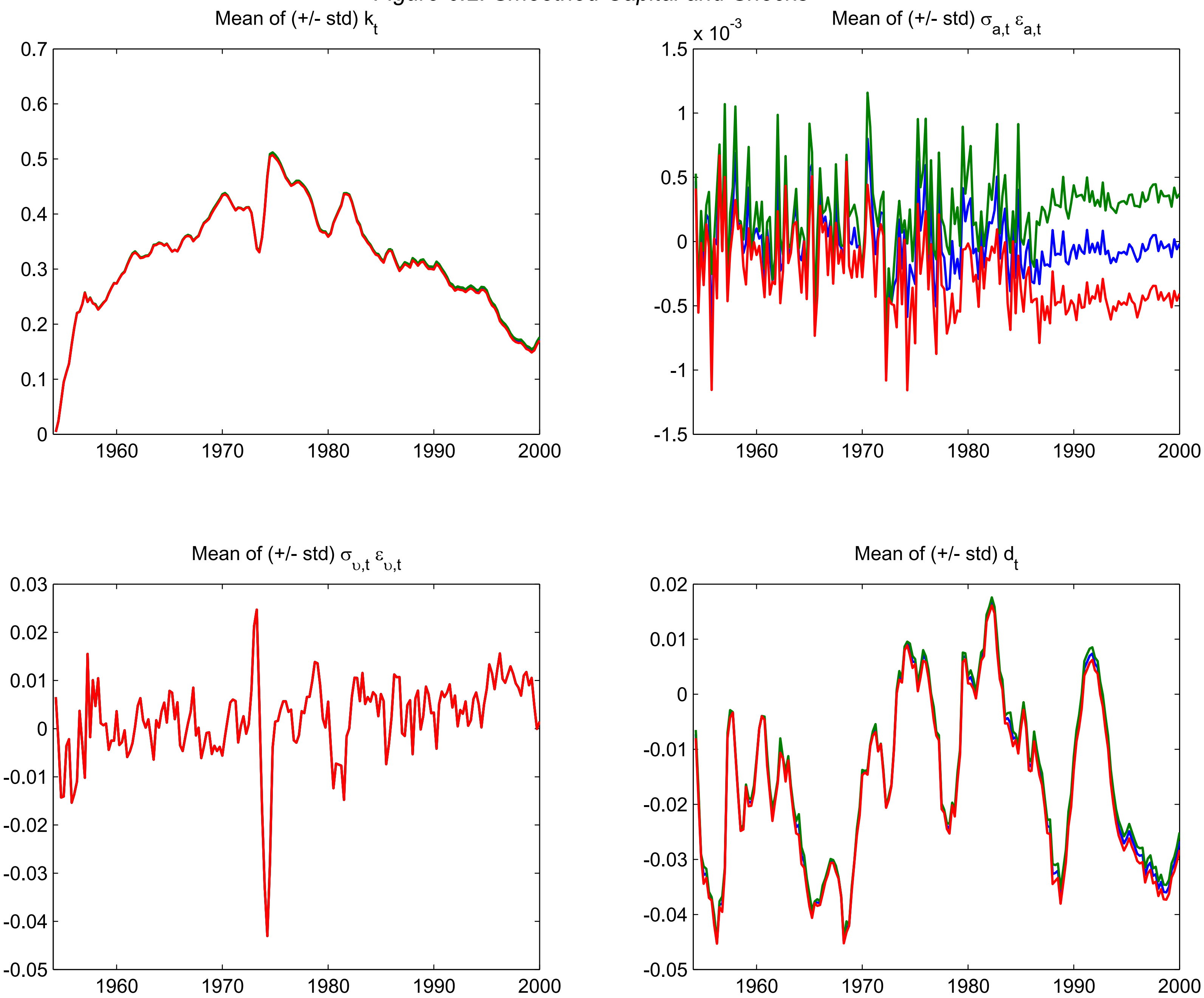


Figure 6.3: Smoothed Volatilities

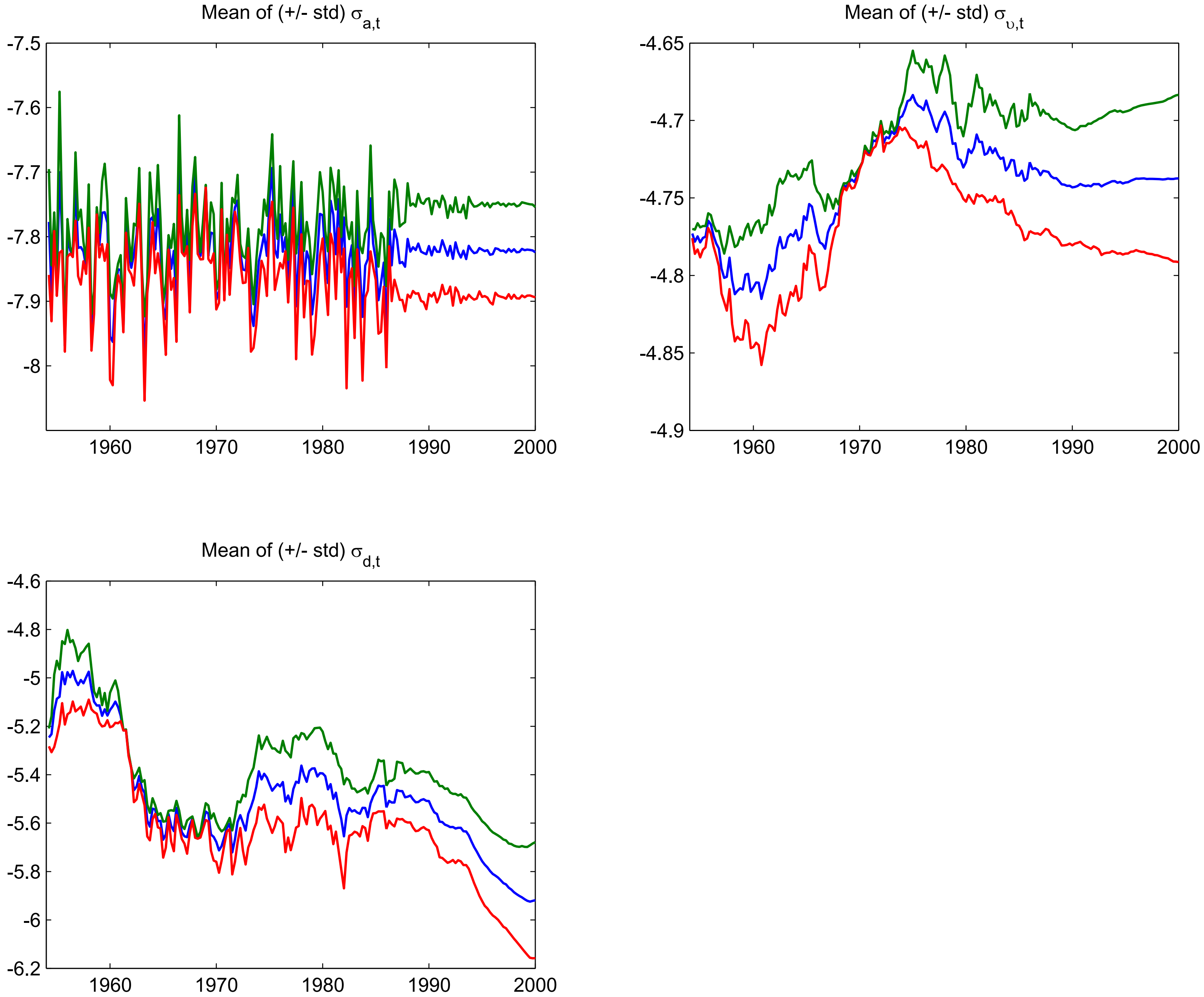


Figure 6.4: Instantaneous Standard Deviation

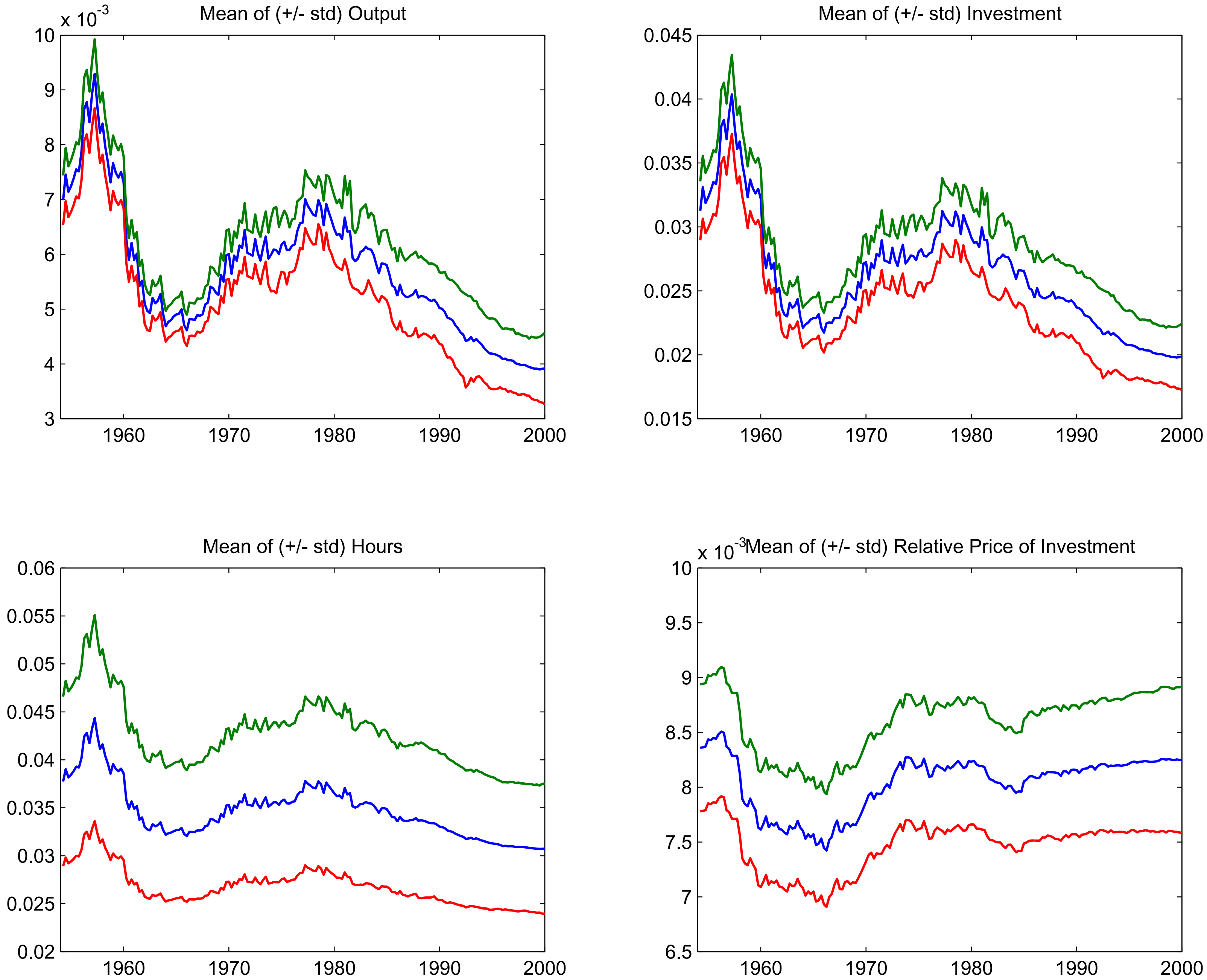


Figure 6.5: Counterfactual Exercise 1

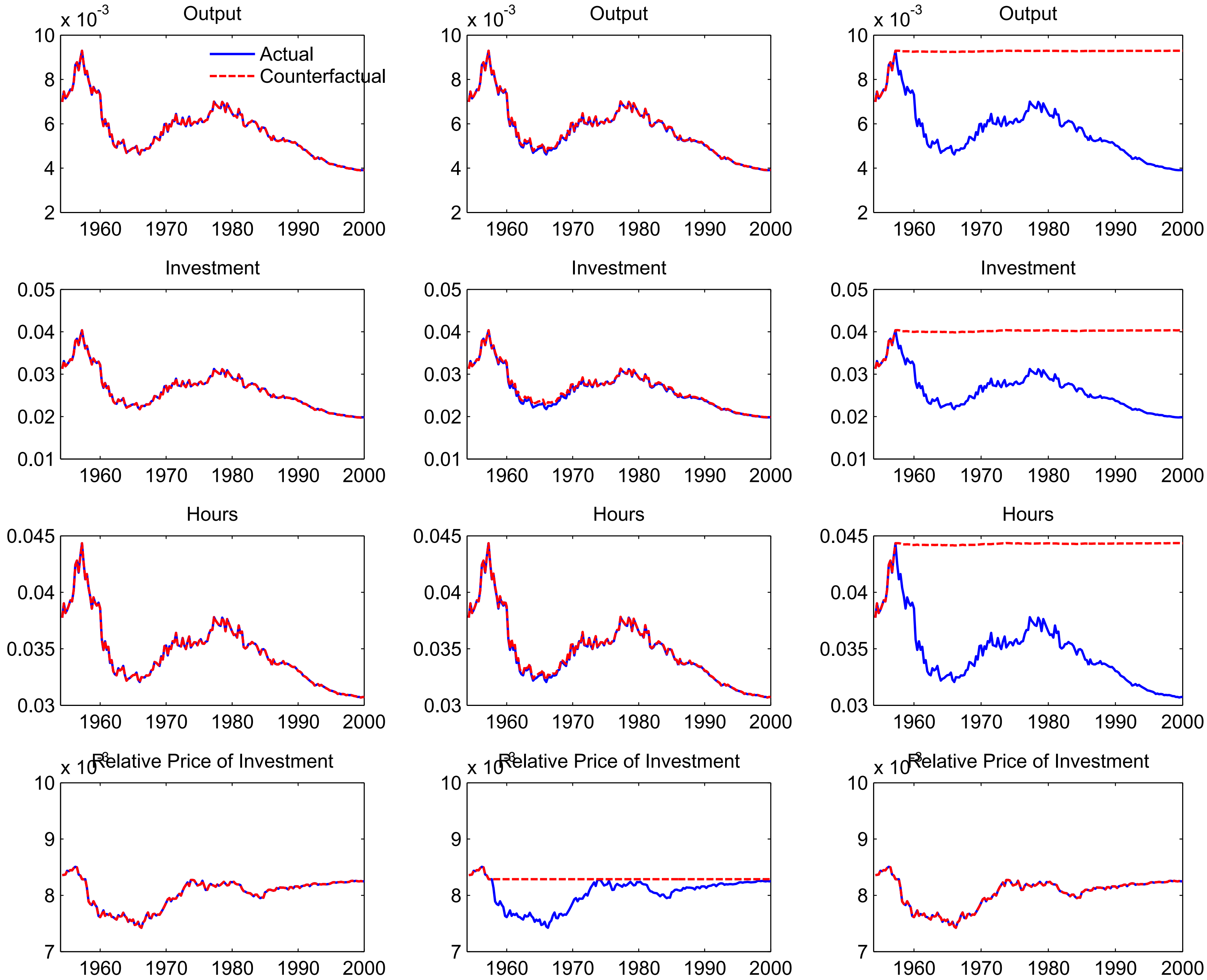


Figure 6.6: Counterfactual Exercise 2

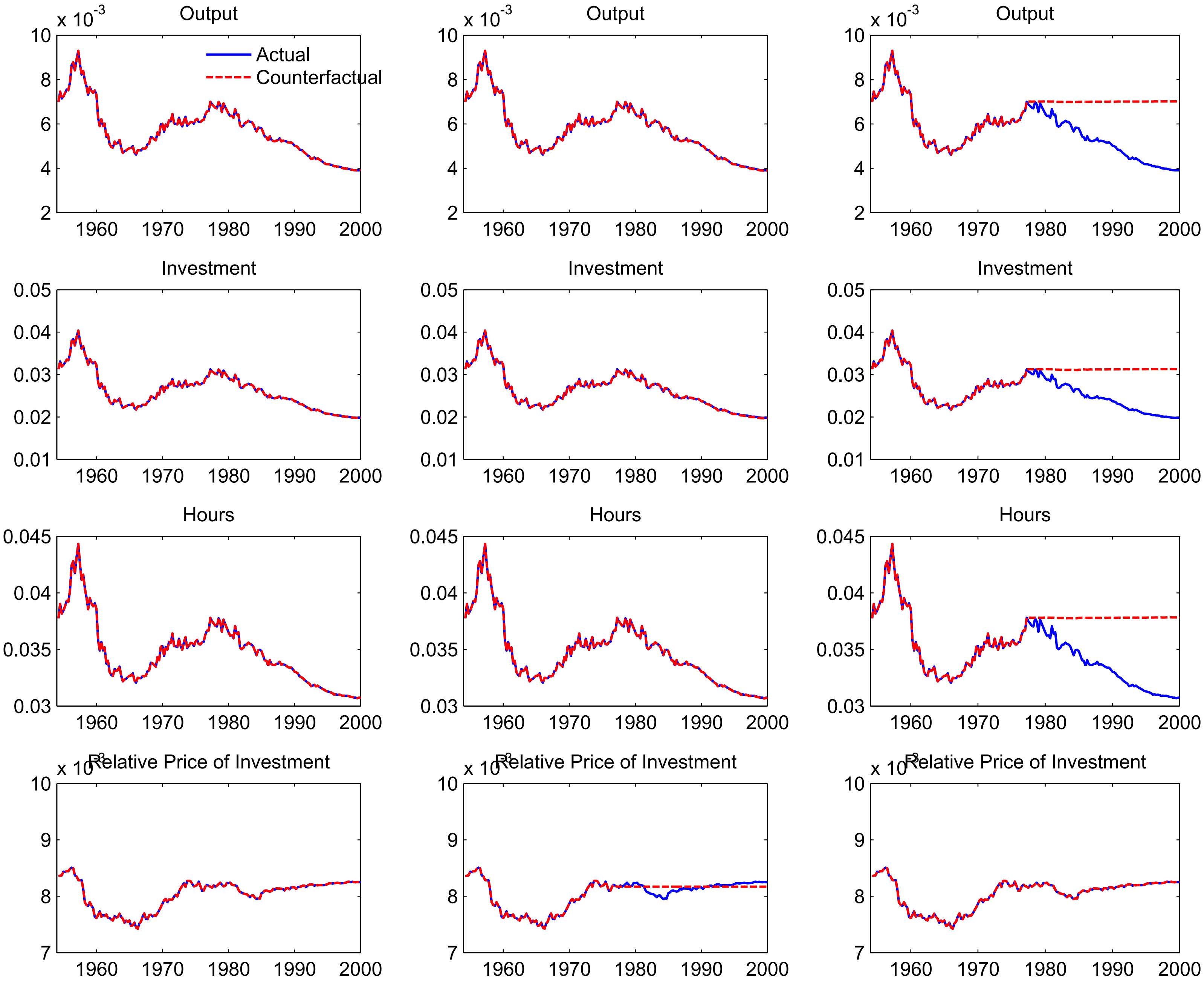
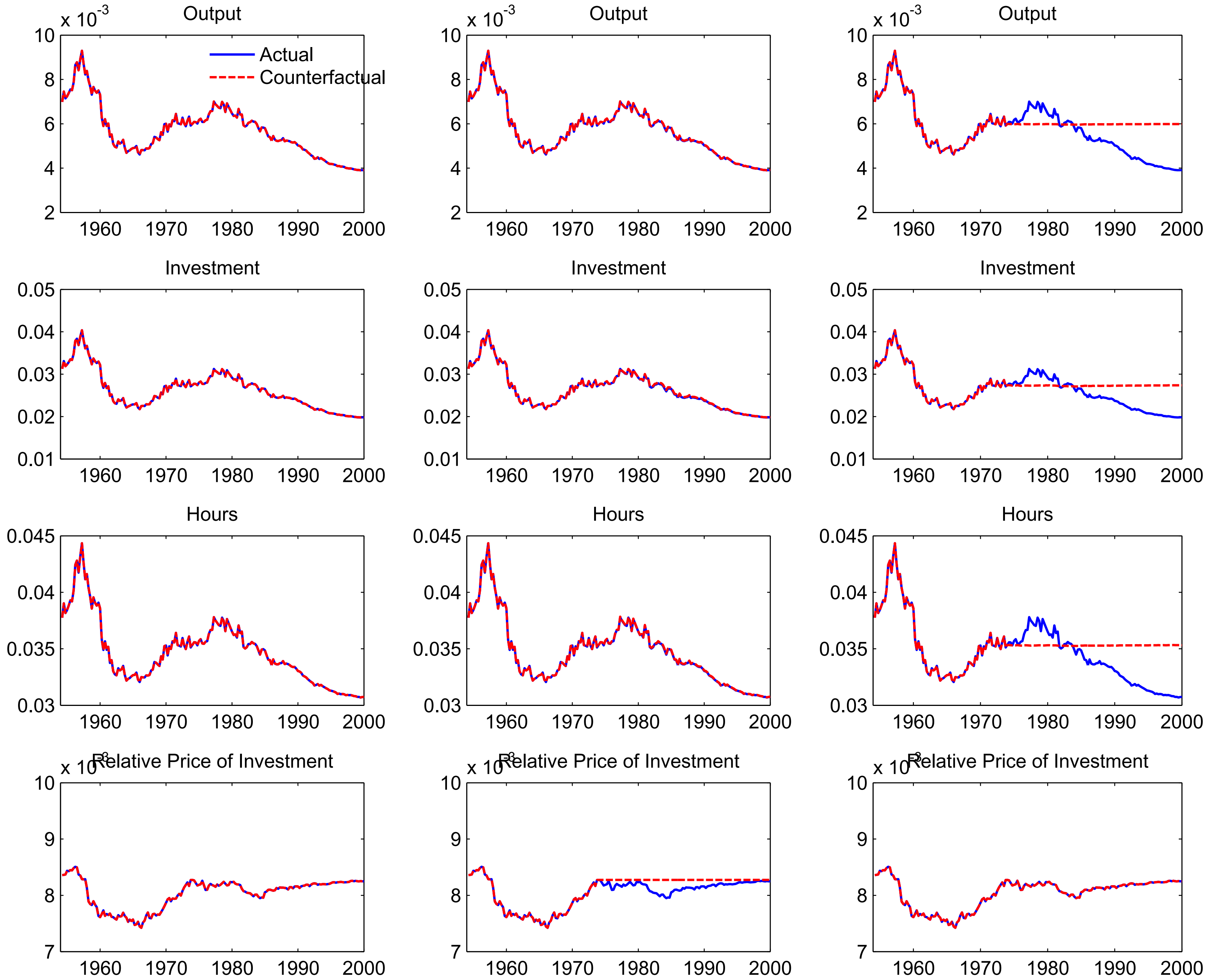


Figure 6.7: Counterfactual Exercise 3



Are Nonlinearities and Non-normalities Important?

Are Nonlinearities and Non-normalities Important?

- We estimate four version of the model:

Table 7.1: Versions of the Model

Solution	No Stochastic Volatility	Stochastic Volatility
Linear	Version 1	Version 2
Quadratic	Version 3	Benchmark

Are Nonlinearities and Non-normalities Important?

- We estimate four version of the model:

Table 7.1: Versions of the Model

Solution	No Stochastic Volatility	Stochastic Volatility
Linear	Version 1	Version 2
Quadratic	Version 3	Benchmark

- Loglike benchmark: 2350.6, loglike version 2: 2230.4

Figure 7.1: Comparison of Smoothed Capital and Shocks

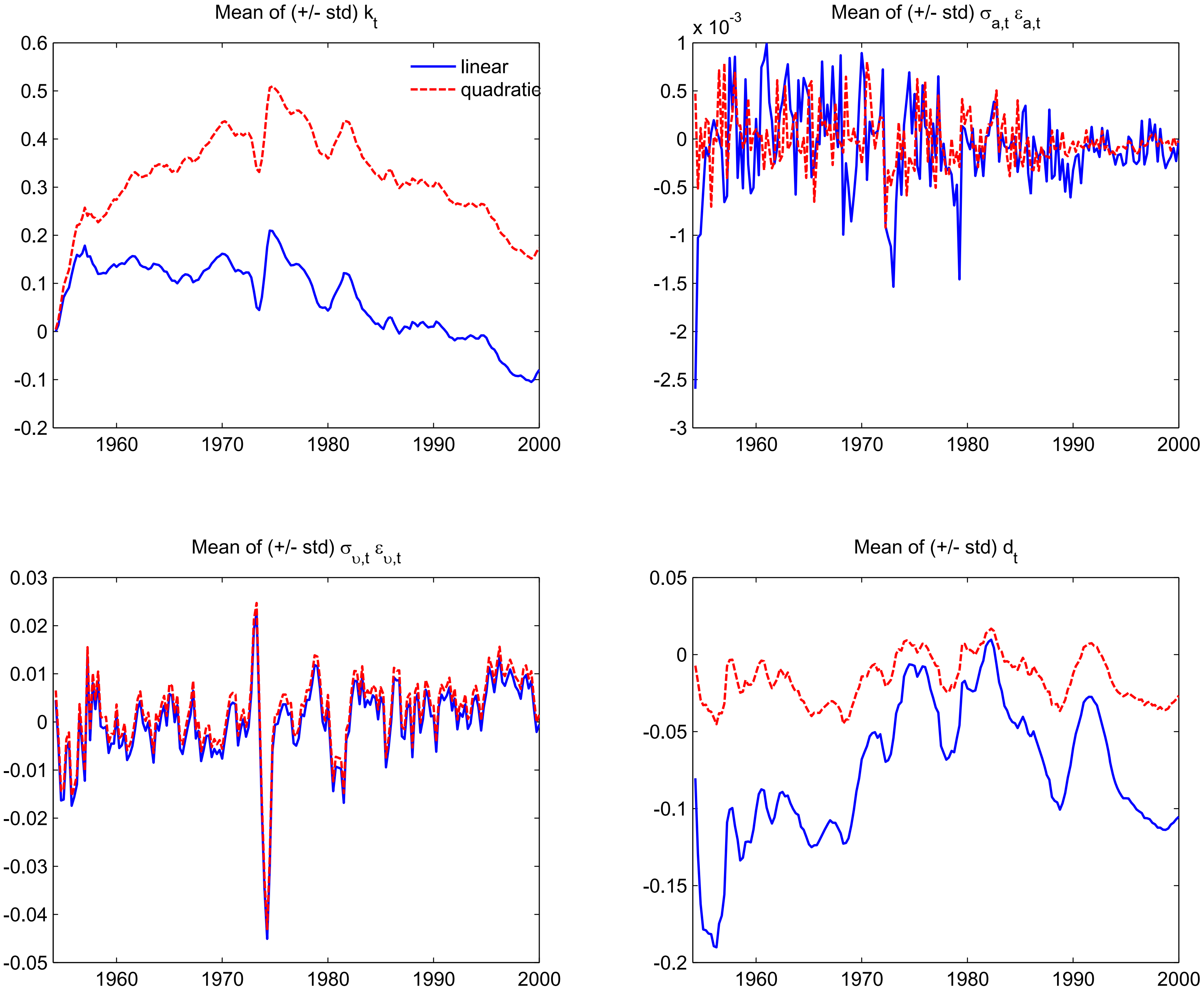
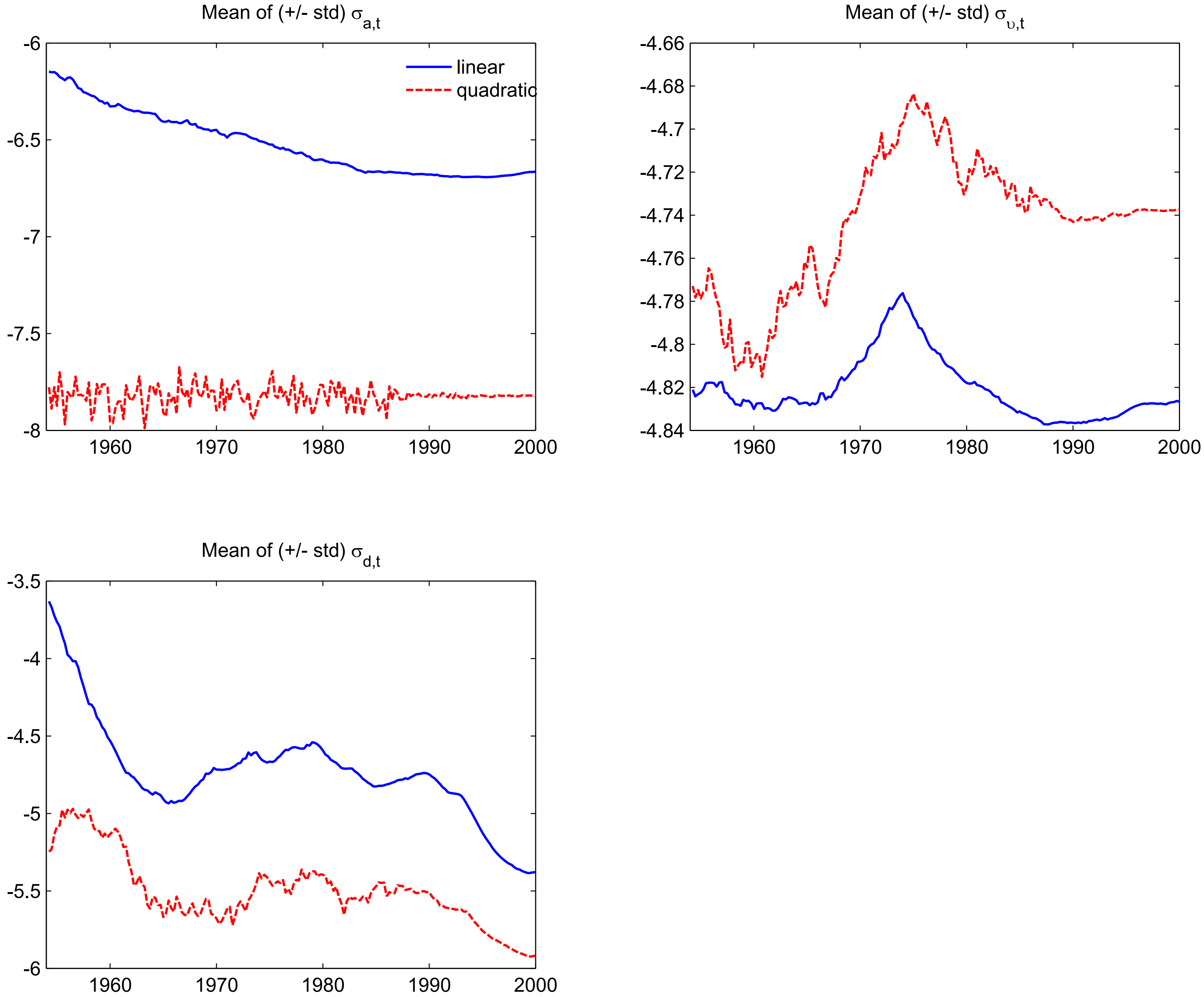


Figure 7.2: Comparison of Smoothed Volatilities



Conclusions

1. Particle filtering is a general purpose and efficient method to estimate DSGE models.

2. We learned about the importance of stochastic volatility to account for U.S. Business Cycle.

3. Much exciting work to do in the next few years!