Estimating Macroeconomic Models: A Likelihood Approach

Jesús Fernández-Villaverde
University of Pennsylvania, NBER, and CEPR

Juan Rubio-Ramírez
Federal Reserve Bank of Atlanta

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 algorithm that tracks the distribution of states conditional on a sequence of observables.

What is the Particle Filter?

- The particle filter is a monte carlo algorithm 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 algorithm 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 Shephard and coauthors.

What is the Particle Filter?

- The particle filter is a monte carlo algorithm 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 Shephard and coauthors.
- Reviewed on Doucet, De Freitas, and Gordon (2001).

What is the Particle Filter?

- The particle filter is a monte carlo algorithm 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 Shephard and coauthors.
- Reviewed on Doucet, De Freitas, and Gordon (2001).
- We modify the filter to be more flexible with shocks.

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.

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.
2. Computational evidence.

Theoretical Arguments

Theoretical Arguments

We develop these arguments in "Convergence Properties of the Likelihood of Computed Dynamic Models" (2006) and in some work-in-progress.

Theoretical Arguments

We develop these arguments in "Convergence Properties of the Likelihood of Computed Dynamic Models" (2006) and in some work-in-progress.

1. Second-order errors in the approximated policy function induce first-order errors in the likelihood function.

Theoretical Arguments

We develop these arguments in "Convergence Properties of the Likelihood of Computed Dynamic Models" (2006) and in some work-in-progress.

1. Second-order errors in the approximated policy function induce first-order errors in the likelihood function.

2. As the sample size grows, the error in the likelihood function also grows and we may have inconsistent point estimates.

Theoretical Arguments

We develop these arguments in "Convergence Properties of the Likelihood of Computed Dynamic Models" (2006) and in some work-in-progress.

1. Second-order errors in the approximated policy function induce first-order errors in the likelihood function.

2. As the sample size grows, the error in the likelihood function also grows and we may have inconsistent point estimates.

3. Linearization complicates the identification of parameters.

Computational Evidence

Computational Evidence

We build the computational evidence in several papers.

Computational Evidence

We build the computational evidence in several papers.

Several researchers (Ann, King, Schorfheide, Winschel) have gathered similar evidence after our paper first circulated.

Computational Evidence

We build the computational evidence in several papers.

Several researchers (Ann, King, Schorfheide, Winschel) have gathered similar evidence after our paper first circulated.

1. Big differences in the level of the likelihood.

Computational Evidence

We build the computational evidence in several papers.

Several researchers (Ann, King, Schorfheide, Winschel) have gathered similar evidence after our paper first circulated.

1. Big differences in the level of the likelihood.

2. Often, important differences in point estimates.

Computational Evidence

We build the computational evidence in several papers.

Several researchers (Ann, King, Schorfheide, Winschel) have gathered similar evidence after our paper first circulated.

1. Big differences in the level of the likelihood.

2. Often, important differences in point estimates.

3. Big differences in smoothed shocks and states.

Computational Evidence

We build the computational evidence in several papers.

Several researchers (Ann, King, Schorfheide, Winschel) have gathered similar evidence after our paper first circulated.

1. Big differences in the level of the likelihood.
2. Often, important differences in point estimates.
3. Big differences in smoothed shocks and states.
4. Better identification of parameters.

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

Filtering problem: forecast and update.

1. Forecast: Chapman-Kolmogorov equation

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

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:

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

Intuition

- Previous equations involve complicated integrals.

Intuition

- Previous equations involve complicated integrals.
- In a linear and normal world: mean and variance are sufficient statistics $\Rightarrow$ translation and spread.

Intuition

- Previous equations involve complicated integrals.
- In a linear and normal world: mean and variance are sufficient statistics $\Rightarrow$ translation and spread.
- This can be done by the Kalman filter.

Intuition

- Previous equations involve complicated integrals.
- In a linear and normal world: mean and variance are sufficient statistics $\Rightarrow$ translation and spread.
- This can be done by the Kalman filter.
- In a nonlinear and/or non-normal world, we need to carry the distribution $\Rightarrow$ deformation.

Intuition

- Previous equations involve complicated integrals.
- In a linear and normal world: mean and variance are sufficient statistics $\Rightarrow$ translation and spread.
- This can be done by the Kalman filter.
- In a nonlinear and/or non-normal world, we need to carry the distribution $\Rightarrow$ deformation.
- This can be done by the Particle filter.

Intuition

- Previous equations involve complicated integrals.
- In a linear and normal world: mean and variance are sufficient statistics $\Rightarrow$ translation and spread.
- This can be done by the Kalman filter.
- In a nonlinear and/or non-normal world, we need to carry the distribution $\Rightarrow$ deformation.
- This can be done by the Particle filter.
- Alternatives? Unscented Kalman filter, Grid Filter,...

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

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

Then:

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

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

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

Introducing Particles

- $\left\{ s^i_{t-1|t-1} \right\}_{i=1}^{N}$ N i.i.d. draws from $p\left(S_{t-1}|y^{t-1};\gamma\right)$.

Introducing Particles

- $\left\{s_{t-1|t-1}^{i}\right\}_{i=1}^{N}$ N i.i.d. draws from $p\left(S_{t-1}|y^{t-1};\gamma\right)$.
- Each $s_{t-1|t-1}^{i}$ is a particle and $\left\{s_{t-1|t-1}^{i}\right\}_{i=1}^{N}$ a swarm of particles.

Introducing Particles

- $\left\{s_{t-1|t-1}^{i}\right\}_{i=1}^{N}$ N i.i.d. draws from $p\left(S_{t-1}|y^{t-1};\gamma\right)$.
- Each $s_{t-1|t-1}^{i}$ is a particle and $\left\{s_{t-1|t-1}^{i}\right\}_{i=1}^{N}$ a swarm of particles.
- $\left\{s_{t|t-1}^{i}\right\}_{i=1}^{N}$ N i.i.d. draws from $p\left(S_{t}|y^{t-1};\gamma\right)$.

Introducing Particles

- $\left\{s^i_{t-1|t-1}\right\}_{i=1}^{N}$ N i.i.d. draws from $p\left(S_{t-1}|y^{t-1};\gamma\right)$.
- Each $s^i_{t-1|t-1}$ is a particle and $\left\{s^i_{t-1|t-1}\right\}_{i=1}^{N}$ a swarm of particles.
- $\left\{s^i_{t|t-1}\right\}_{i=1}^{N}$ N i.i.d. draws from $p\left(S_{t}|y^{t-1};\gamma\right)$.
- Each $s^i_{t|t-1}$ is a proposed particle and $\left\{s^i_{t|t-1}\right\}_{i=1}^{N}$ a swarm of proposed particles.

Introducing Particles

- $\left\{s^i_{t-1|t-1}\right\}_{i=1}^N$ N i.i.d. draws from $p\left(S_{t-1}|y^{t-1};\gamma\right)$.
- Each $s^i_{t-1|t-1}$ is a particle and $\left\{s^i_{t-1|t-1}\right\}_{i=1}^N$ a swarm of particles.
- $\left\{s^i_{t|t-1}\right\}_{i=1}^N$ N i.i.d. draws from $p\left(S_t|y^{t-1};\gamma\right)$.
- Each $s^i_{t|t-1}$ is a proposed particle and $\left\{s^i_{t|t-1}\right\}_{i=1}^N$ a swarm of proposed particles.
- Weight of each proposed particle:

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

A Proposition

- Let $\left\{ s^i_{t|t-1} \right\}_{i=1}^N$ 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.

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.

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

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

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.

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

2. Forecast: We can use a draw $\left\{s_{t|t}^{i}\right\}_{i=1}^{N}$ 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 $\left\{s_{t+1|t}^{i}\right\}_{i=1}^{N}$.

Particle Filtering I

Particle Filtering I

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

Particle Filtering I

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

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

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.

Particle Filtering I

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

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

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 $\{q^i_t\}_{i=1}^N$. 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

An Application: a Business Cycle Model

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.

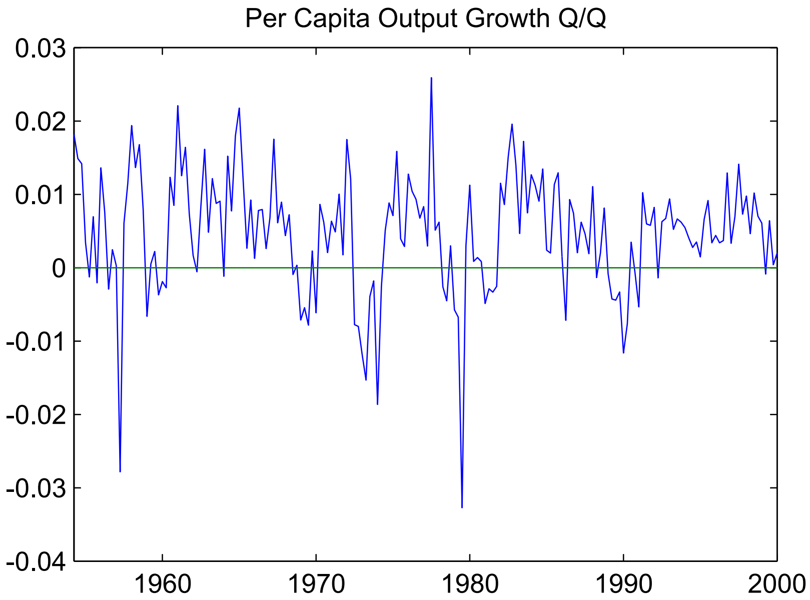
An Application: a Business Cycle Model

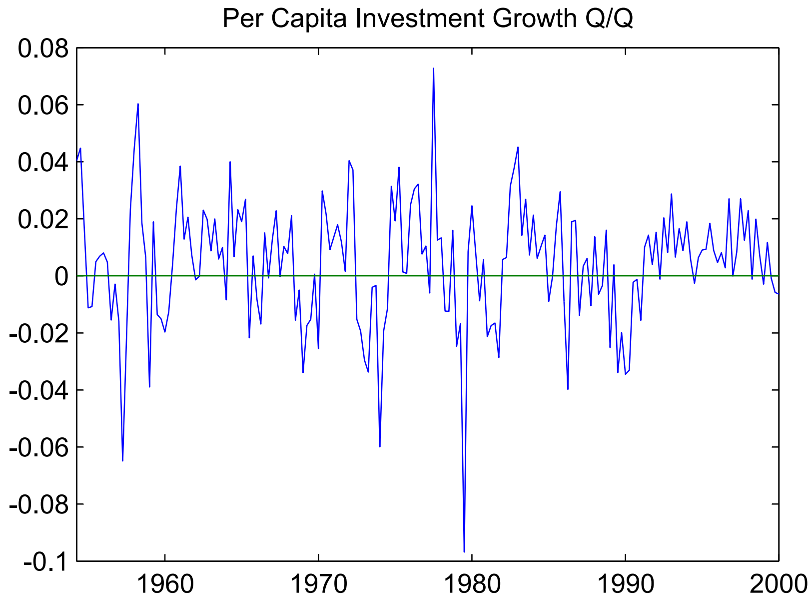
- I want to show the power of the particle filter.
- And learn something about the evolution of the U.S. aggregate fluctuations over the last decades.
- A business cycle model with:

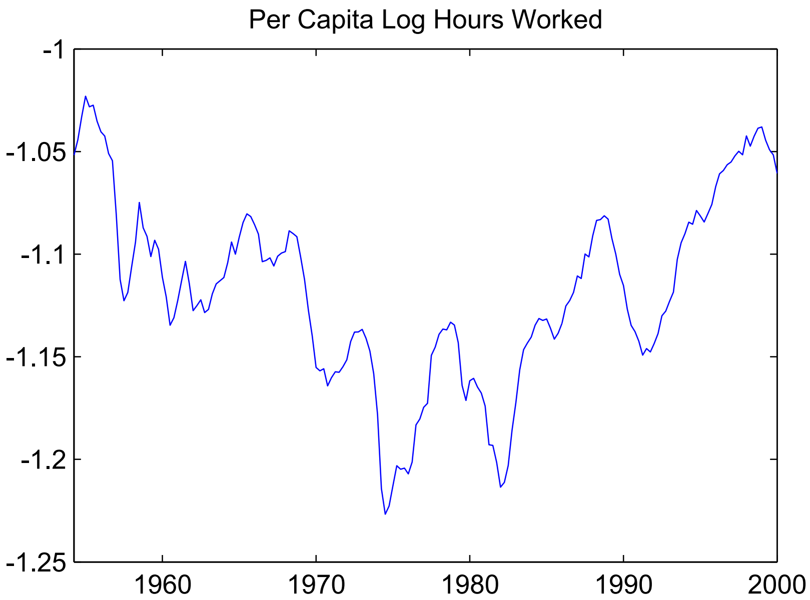
An Application: a Business Cycle Model

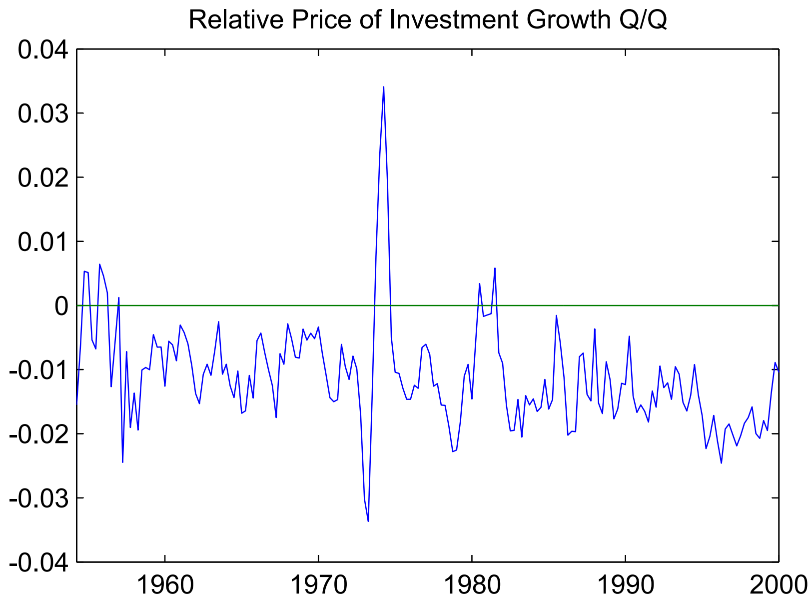
- I want to show the power of the particle filter.
- And learn something about the evolution of the U.S. aggregate fluctuations over the last decades.
- A business cycle model with:

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


Per Capita Output Growth Q/Q
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
1960
1970
1980
1990
2000


Per Capita Investment Growth Q/Q
0.08
0.06
0.04
0.02
0
-0.02
-0.04
-0.06
-0.08
-0.1
1960
1970
1980
1990
2000


Per Capita Log Hours Worked
-1
-1.05
-1.1
-1.15
-1.2
-1.25
1960
1970
1980
1990
2000


Relative Price of Investment Growth Q/Q
0.04
0.03
0.02
0.01
0
-0.01
-0.02
-0.03
-0.04
1960
1970
1980
1990
2000

An Application: a Business Cycle Model

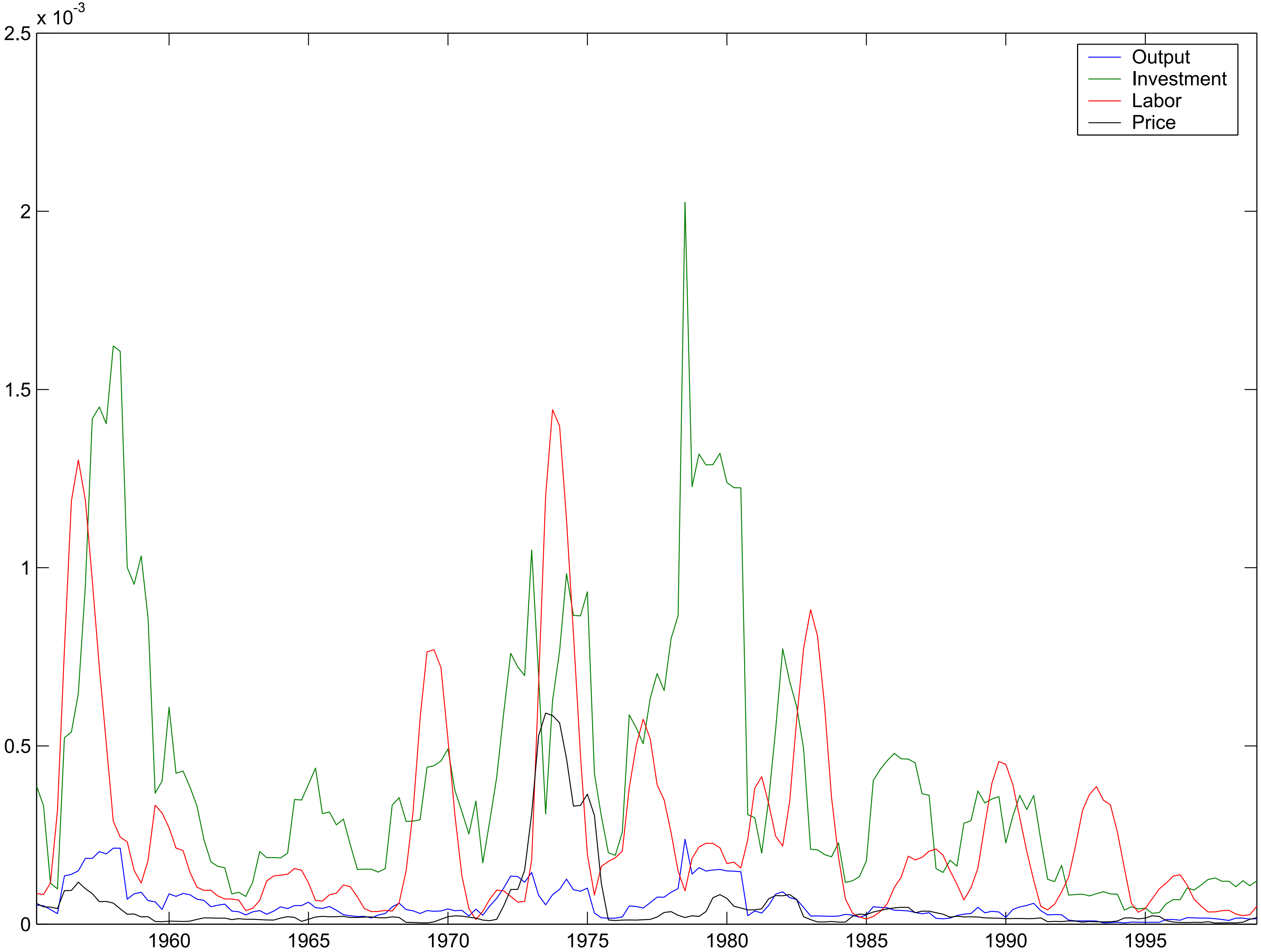
- I want to show the power of the particle filter.

- 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.


x 10^-3
2.5
2
1.5
1
0.5
0
1960
1965
1970
1975
1980
1985
1990
1995
Output
Investment
Labor
Price

Households

- Representative household with utility function:

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

- Law of motion of d_t, the preference shock:

$$d_t = \rho d_{t-1} + \sigma_{dt} \varepsilon_{dt}, \; \varepsilon_{dt} \sim \mathcal{N}(0,1)$$

- We will explain later the law of motion of σ_{dt}.

Technology

- 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:

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

Stochastic Volatility

We follow a standard specification:

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

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

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

Analysis of the Model

- It can be shown that along a balanced growth path the following variables are stationary:

$$Y_t/Z_t,\ C_t/Z_t,\ X_t/Z_t,\ K_{t+1}/(Z_t V_t),\ (Y_t/L_t)/Z_t, \text{ and } L_t$$

  where $Z_t = A_t^{1/(1-\alpha)} V_t^{\alpha/(1-\alpha)}$.

Analysis of the Model

- It can be shown that along a balanced growth path the following variables are stationary:

  $$Y_t/Z_t,\ C_t/Z_t,\ X_t/Z_t,\ K_{t+1}/(Z_tV_t),\ (Y_t/L_t)/Z_t, \text{ and } L_t$$

  where $Z_t = A_t^{1/(1-\alpha)} V_t^{\alpha/(1-\alpha)}$.
- The growth rates for the exogenous shocks are:

$$\begin{aligned} \log A_t - \log A_{t-1} &= \zeta + \sigma_{at}\varepsilon_{at} \\ \log V_t - \log V_{t-1} &= \upsilon + \sigma_{\upsilon t}\varepsilon_{\upsilon t} \end{aligned}$$

Analysis of the Model

- It can be shown that along a balanced growth path the following variables are stationary:

$$Y_t/Z_t,\ C_t/Z_t,\ X_t/Z_t,\ K_{t+1}/(Z_tV_t),\ (Y_t/L_t)/Z_t, \text{ and } L_t$$

  where $Z_t = A_t^{1/(1-\alpha)} V_t^{\alpha/(1-\alpha)}$.

- The growth rates for the exogenous shocks are:

$$\begin{aligned}\log A_t - \log A_{t-1} &= \zeta + \sigma_{at}\varepsilon_{at} \\ \log V_t - \log V_{t-1} &= \upsilon + \sigma_{vt}\varepsilon_{vt}\end{aligned}$$

- Therefore, Y_t, C_t, X_t, and Y_t/L_t grow at rate $(\zeta + \alpha\upsilon)/(1-\alpha)$, K_t grows at rate $(\zeta+\upsilon)/(1-\alpha)$, and L_t is stationary.

Transforming the Model

- The model is nonstationary because of the presence of two unit roots, one in each technological process.
- We need to transform the model into a stationary problem.
- We scale variables as $\widetilde{C}_t = \frac{C_t}{Z_t}$, $\widetilde{X}_t = \frac{X_t}{Z_t}$, and $\widetilde{K}_t = \frac{K_t}{Z_t V_{t-1}}$.
- Then:

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

$$\widetilde{C}_t + e^{\frac{\gamma + \alpha \upsilon + \sigma_{at}\varepsilon_{at} + \alpha \sigma_{\upsilon t}\varepsilon_{\upsilon t}}{1-\alpha}} \widetilde{K}_{t+1} =$$

$$e^{\gamma + \sigma_{at}\varepsilon_{at}} \widetilde{K}_t^{\alpha} L_t^{1-\alpha} + (1 - \delta)\, e^{-\upsilon - \sigma_{\upsilon t}\varepsilon_{\upsilon t}} \widetilde{K}_t$$

Equilibrium Conditions

1. Euler equation:

$$\frac{e^{d_t} e^{\frac{\gamma+\alpha\upsilon+\varepsilon_{at}+\alpha\sigma_{\upsilon t}\varepsilon_{\upsilon t}}{1-\alpha}}}{\widetilde{C}_t} =$$

$$\beta E_t \frac{e^{d_{t+1}}}{\widetilde{C}_{t+1}} \left(\alpha e^{\gamma+\sigma_{at+1}\varepsilon_{at+1}} \widetilde{K}_{t+1}^{\alpha} L_{t+1}^{1-\alpha} + (1-\delta)\, e^{-\upsilon-\sigma_{\upsilon t+1}\varepsilon_{\upsilon t+1}} \right)$$

2. A labor supply condition:

$$\psi \frac{e^{d_t} \widetilde{C}_t}{1-L_t} = (1-\alpha)\, e^{\gamma+\sigma_{at}\varepsilon_{at}} \widetilde{K}_t^{\alpha} L_t^{-\alpha}$$

3. The resource constraint:

$$\widetilde{C}_t + e^{\frac{\gamma+\alpha\upsilon+\sigma_{at}\varepsilon_{at}+\alpha\sigma_{\upsilon t}\varepsilon_{\upsilon t}}{1-\alpha}} \widetilde{K}_{t+1} = e^{\gamma+\sigma_{at}\varepsilon_{at}} \widetilde{K}_t^{\alpha} L_t^{1-\alpha} + (1-\delta)\, e^{-\upsilon-\sigma_{\upsilon t}\varepsilon_{\upsilon t}} \widetilde{K}_t$$

The Steady State

- We can find the steady state of the transformed model.

- We have a cointegration relation between output and investment in nominal terms:

$$\frac{\widetilde{X}_{ss}}{\widetilde{Y}_{ss}} = \alpha \frac{\left(e^{\frac{\zeta+\alpha\upsilon}{1-\alpha}} - (1-\delta)\, e^{-\upsilon}\right)}{\frac{\exp\left(\frac{\zeta+\alpha\upsilon}{1-\alpha}\right)}{\beta} - (1-\delta)\exp\left(-\upsilon\right)}$$

- Let

$$\log \widehat{K}_{t+1} = \log \frac{\widetilde{K}_{t+1}}{\widetilde{K}_{ss}} \text{ and } \log \widehat{X}_t = \log \frac{\widetilde{X}_t}{\widetilde{X}_{ss}}$$

Solution

- We solve the model using 2nd order perturbation.
- Other non-linear solution methods are possible. Aruoba, Fernández-Villaverde, and Rubio-Ramírez (2005).
- Structure of a 2nd order approximation:

$$\begin{aligned} \log \widehat{K}_{t+1} &= \Psi_{k1} s_t + \frac{1}{2} s_t' \Psi_{k2} s_t \\ \log \widehat{X}_t &= \Psi_{x1} s_t + \frac{1}{2} s_t' \Psi_{x2} s_t \\ \log L_t &= \Psi_{l1} s_t + \frac{1}{2} s_t' \Psi_{l2} s_t \end{aligned}$$

where:

$$s_t = \left(1, \log \widehat{K}_t, d_t, \sigma_{at}\varepsilon_{at}, \sigma_{\upsilon t}\varepsilon_{\upsilon t}, \log \sigma_{dt}, \log \sigma_{at}, \log \sigma_{\upsilon t}\right)'$$

State Space Representation I: Transition Equation

$$
\begin{aligned}
f_1(\mathcal{S}_t, W_t) &= 1 \\
f_2(\mathcal{S}_t, W_t) &= \Psi_{k1} s_t + \frac{1}{2} s_t' \Psi_{k2} s_t \\
f_3(\mathcal{S}_t, W_t) &= \left((1-\lambda_a)\log \overline{\sigma}_a + \lambda_a \log \sigma_{at} + \tau_a \eta_{at+1}\right) \varepsilon_{at+1} \\
f_4(\mathcal{S}_t, W_t) &= \left((1-\lambda_\upsilon)\log \overline{\sigma}_\upsilon + \lambda_\upsilon \log \sigma_{\upsilon t} + \tau_\upsilon \eta_{\upsilon t+1}\right) \varepsilon_{\upsilon t+1} \\
f_5(\mathcal{S}_t, W_t) &= \rho d_t + e^{(1-\lambda_d)\log \overline{\sigma}_d + \lambda_d \log \sigma_{dt} + \tau_d \eta_{dt+1}} \varepsilon_{dt+1} \\
f_6(\mathcal{S}_t, W_t) &= (1-\lambda_a)\log \overline{\sigma}_a + \lambda_a \log \sigma_{at} + \tau_a \eta_{at+1} \\
f_7(\mathcal{S}_t, W_t) &= (1-\lambda_\upsilon)\log \overline{\sigma}_\upsilon + \lambda_\upsilon \log \sigma_{\upsilon t} + \tau_\upsilon \eta_{\upsilon t+1} \\
f_8(\mathcal{S}_t, W_t) &= (1-\lambda_d)\log \overline{\sigma}_d + \lambda_d \log \sigma_{dt} + \tau_d \eta_{dt+1} \\
f_{9-16}(\mathcal{S}_t, W_t) &= s_t
\end{aligned}
$$

where $\mathcal{S}_t = (s_t, s_{t-1})$.

State Space Representation II: Measurement Equation

$$\begin{pmatrix} \Delta \log P_t \\ \Delta \log Y_t \\ \Delta \log X_t \\ \log L_t \end{pmatrix} = \begin{pmatrix} -\upsilon \\ \frac{\gamma+\alpha\upsilon}{1-\alpha} \\ \frac{\gamma+\alpha\upsilon}{1-\alpha} \\ \log L_{ss} + \Psi_{l3} \end{pmatrix} +$$

$$\begin{pmatrix} -\varepsilon_{\upsilon t} \\ \Psi_{y1}\left(s_t - s_{t-1}\right) + \frac{1}{2}\left(s_t'\Psi_{y2}s_t - s_{t-1}'\Psi_{y2}s_{t-1}\right) + \frac{\sigma_{at-1}\varepsilon_{at-1}+\alpha\sigma_{\upsilon t-1}\varepsilon_{\upsilon t-1}}{1-\alpha} \\ \Psi_{x1}\left(s_t - s_{t-1}\right) + \frac{1}{2}\left(s_t'\Psi_{x2}s_t - \frac{1}{2}s_{t-1}'\Psi_{x2}s_{t-1}\right) + \frac{\sigma_{at-1}\varepsilon_{at-1}+\alpha\sigma_{\upsilon t-1}\varepsilon_{\upsilon t-1}}{1-\alpha} \\ \Psi_{l1}s_t + \frac{1}{2}s_t'\Psi_{l2}s_t \end{pmatrix}$$

$$+ \begin{pmatrix} 0 \\ \epsilon_{1t} \\ \epsilon_{2t} \\ \epsilon_{3t} \end{pmatrix}$$

Performing Likelihood-Based Inference

- 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.

Table 5.1: Maximum Likelihood Estimates

Parameter	Point Estimate	Standard Error ($x10^{-3}$)
ρ	0.967	3.743
β	0.999	0.460
ψ	2.343	6.825
υ	8.960E-003	0.828
ζ	3.594E-005	2.254
τ_a	7.120E-002	1.589
τ_υ	7.772E-003	2.940
τ_d	5.653E-002	2.034
$\overline{\sigma}_a$	4.008E-004	0.692
$\overline{\sigma}_\upsilon$	8.523E-003	0.101
$\overline{\sigma}_d$	5.016E-003	2.344
λ_a	4.460E-002	6.788
λ_υ	0.998	8.248
λ_d	0.998	2.302
$\sigma_{1\epsilon}$	1.031E-005	0.424
$\sigma_{2\epsilon}$	1.024E-004	0.495
$\sigma_{3\epsilon}$	1.110E-005	0.082

Figure 6.1: Model versus Data

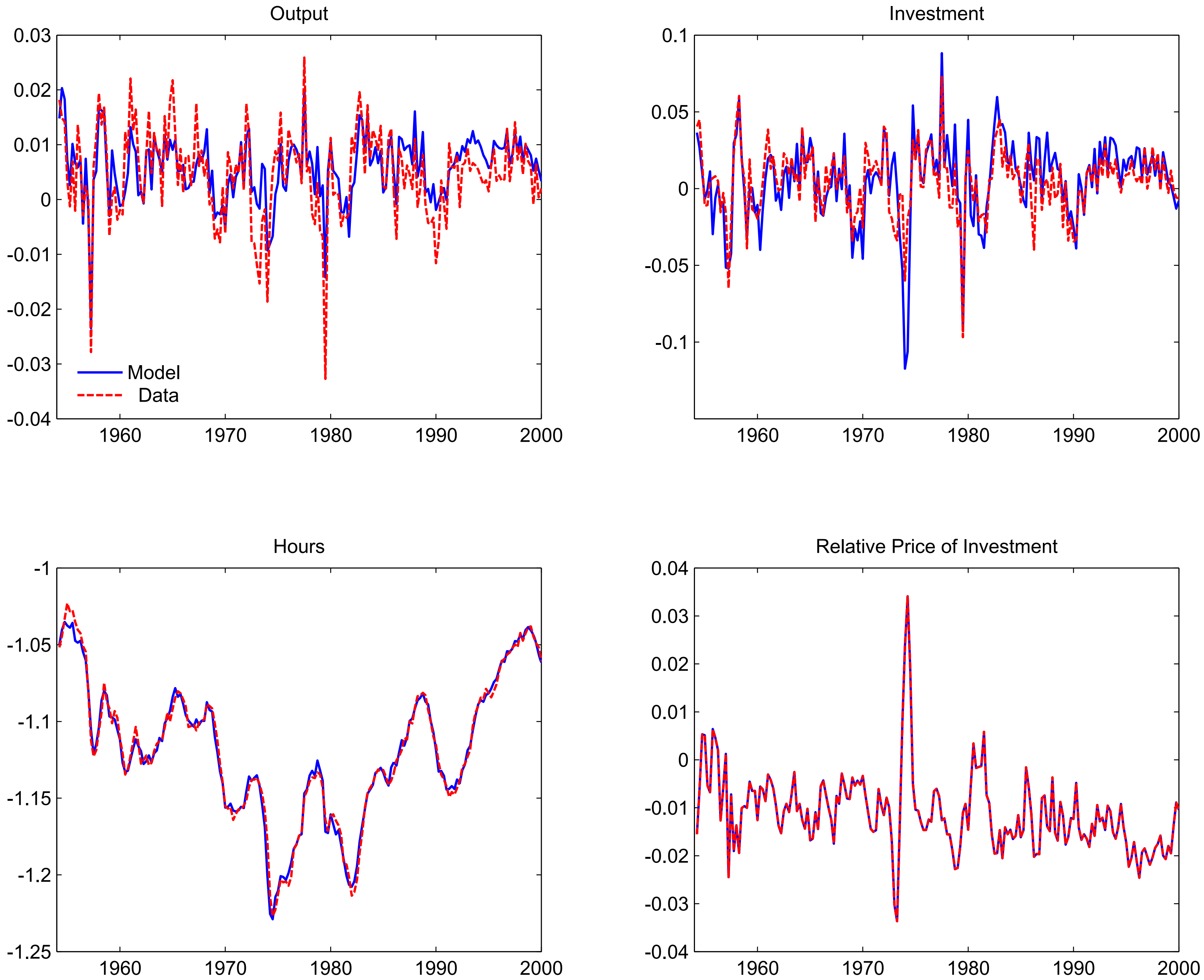


Figure 6.2: Smoothed Capital and Shocks

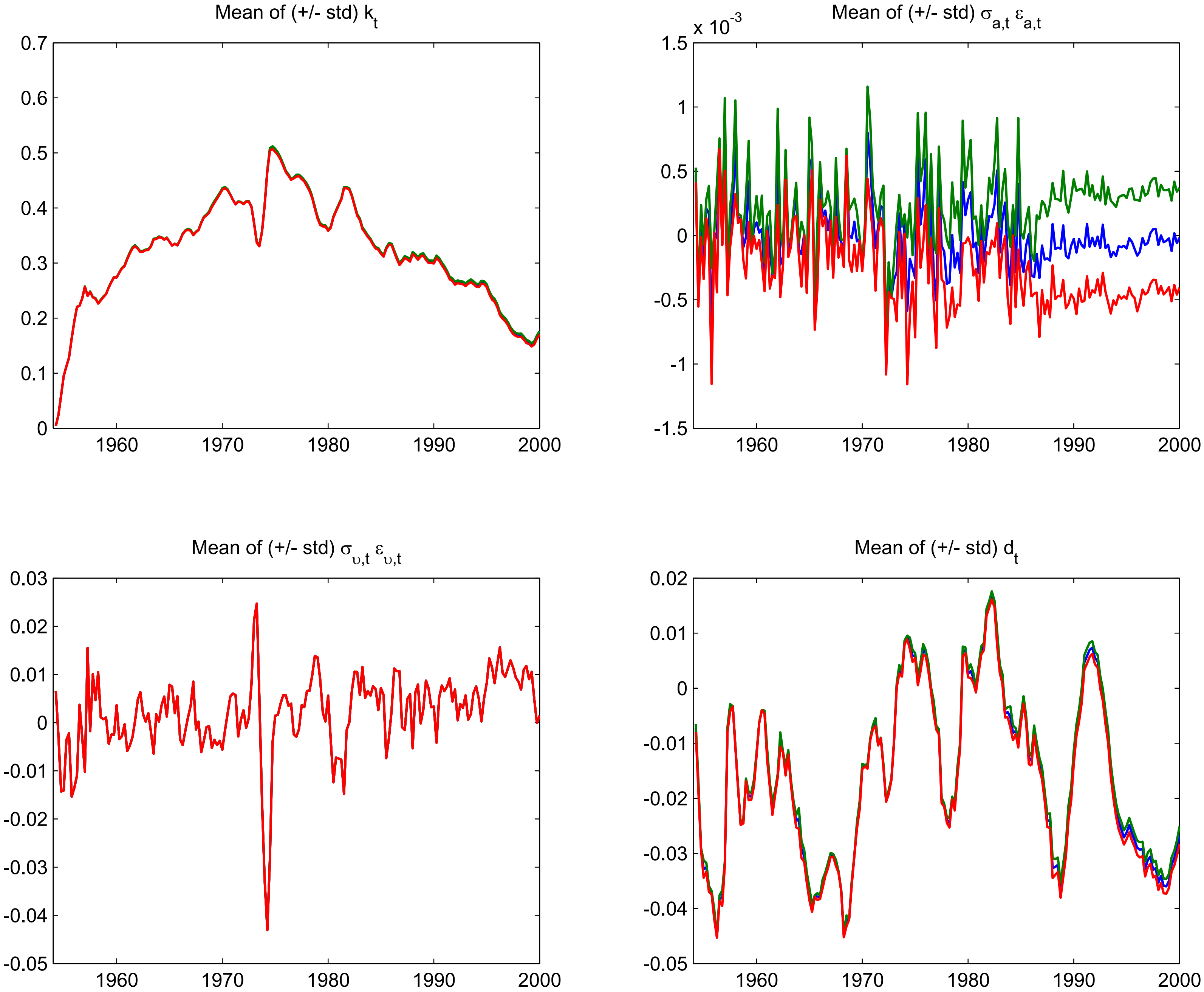


Figure 6.3: Smoothed Volatilities

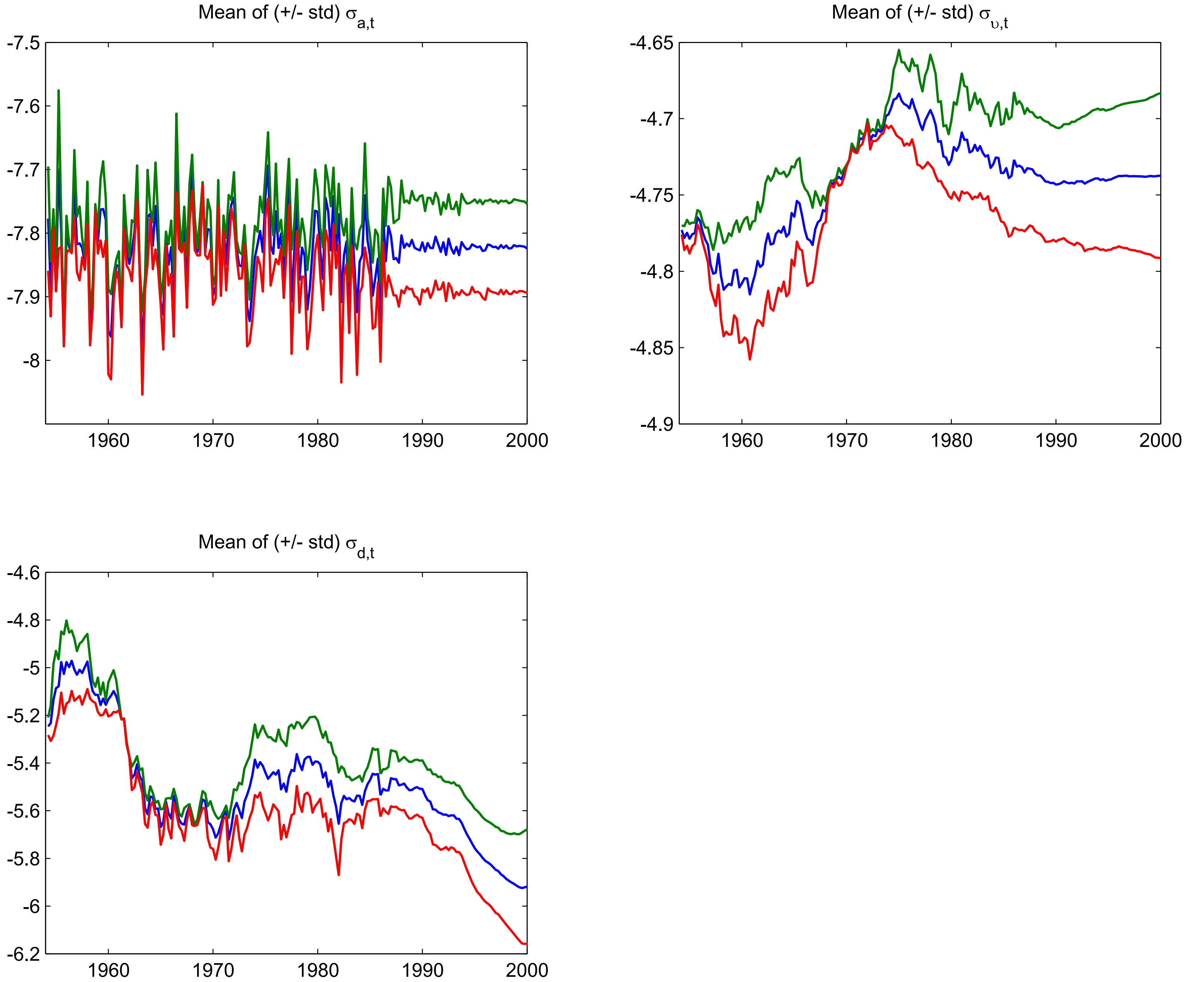


Figure 6.4: Instantaneous Standard Deviation

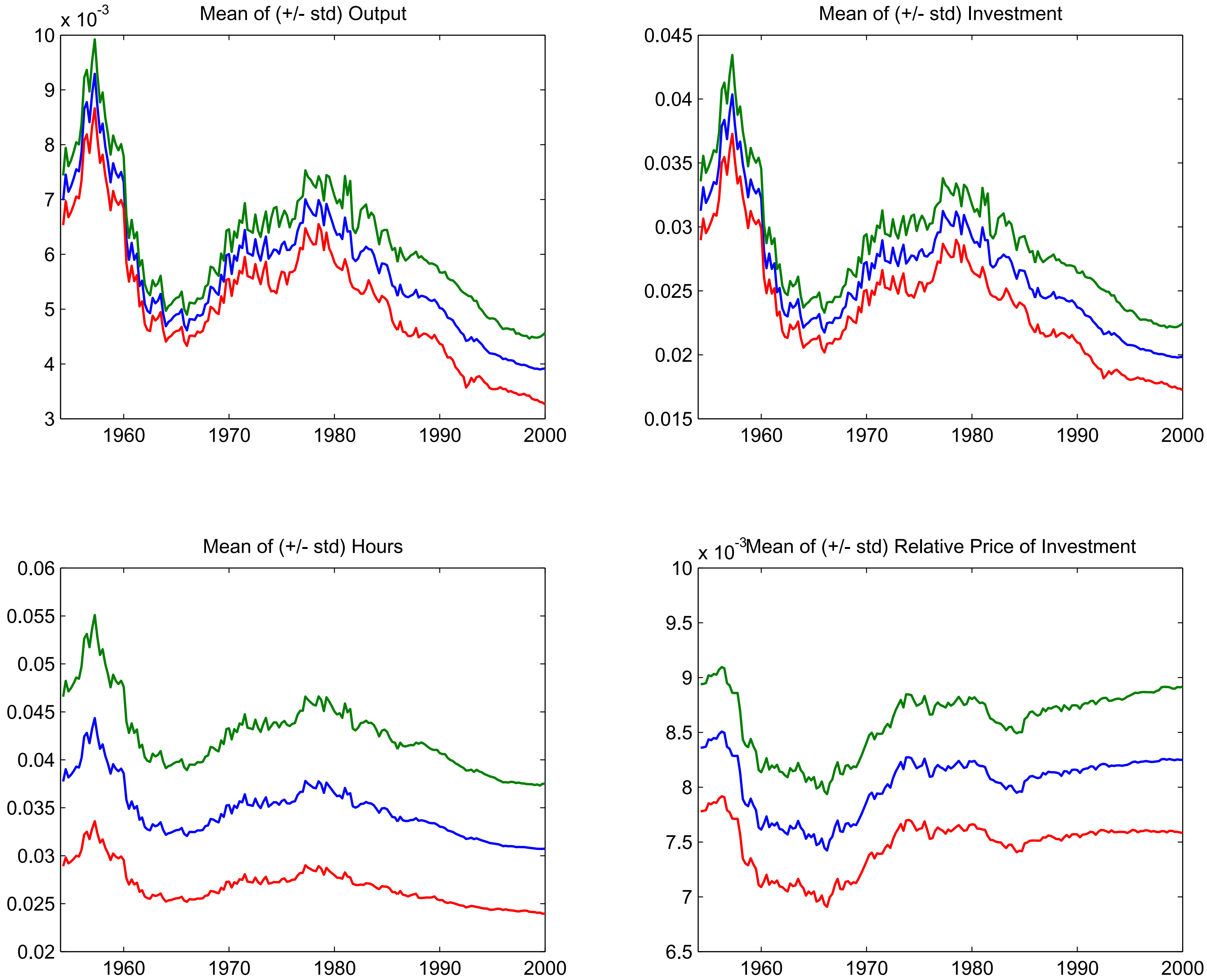


Figure 6.5: Counterfactual Exercise 1

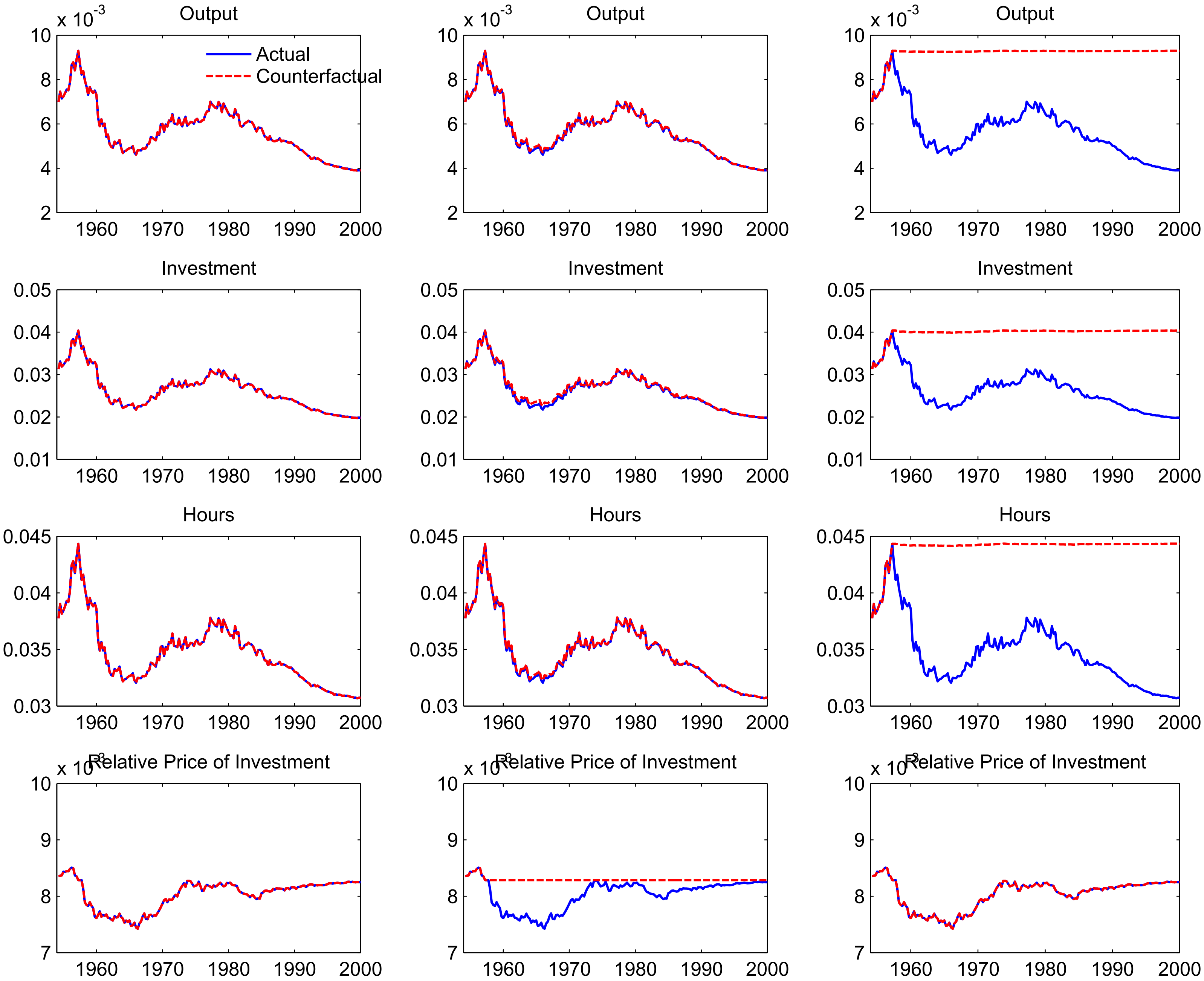


Figure 6.6: Counterfactual Exercise 2

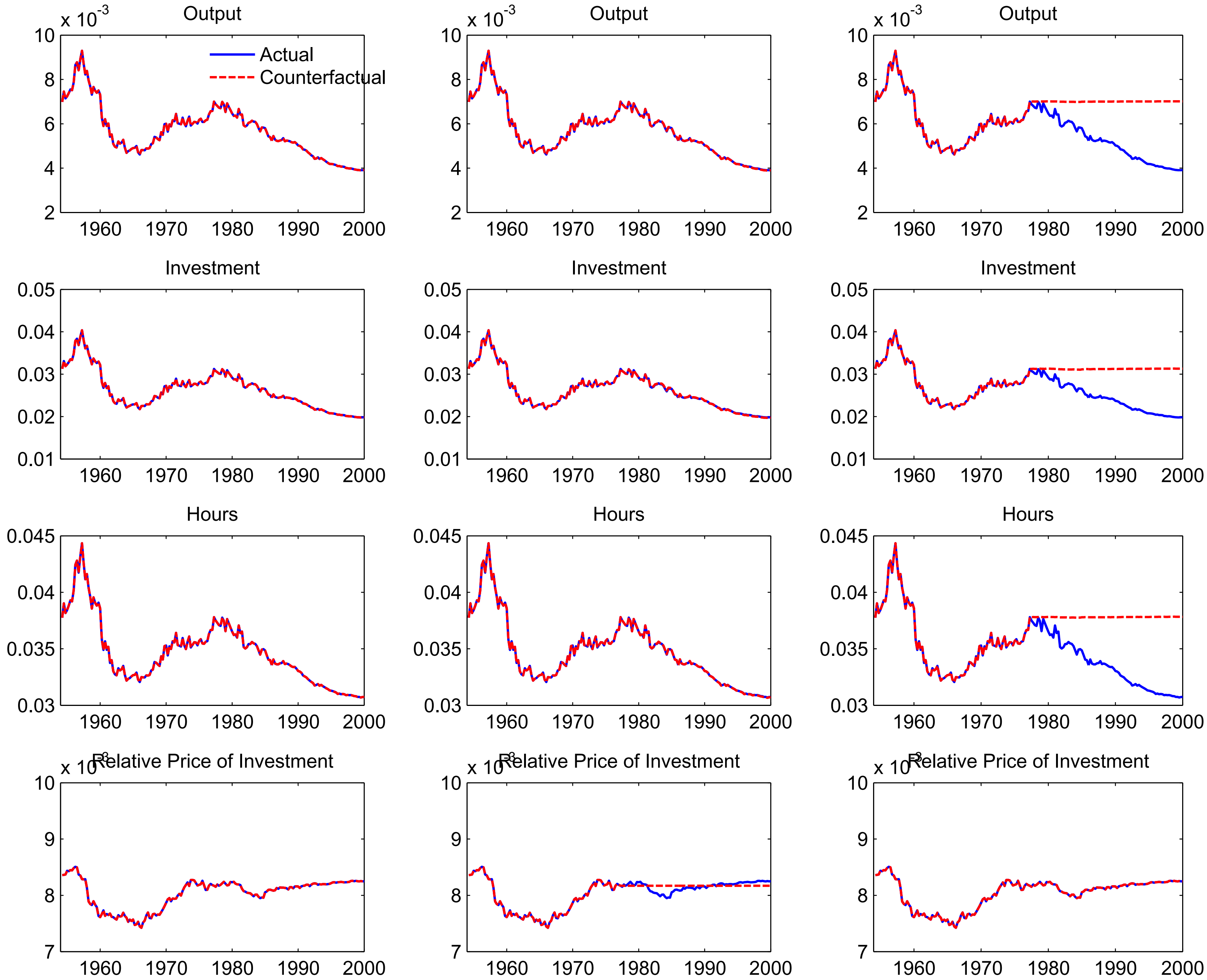
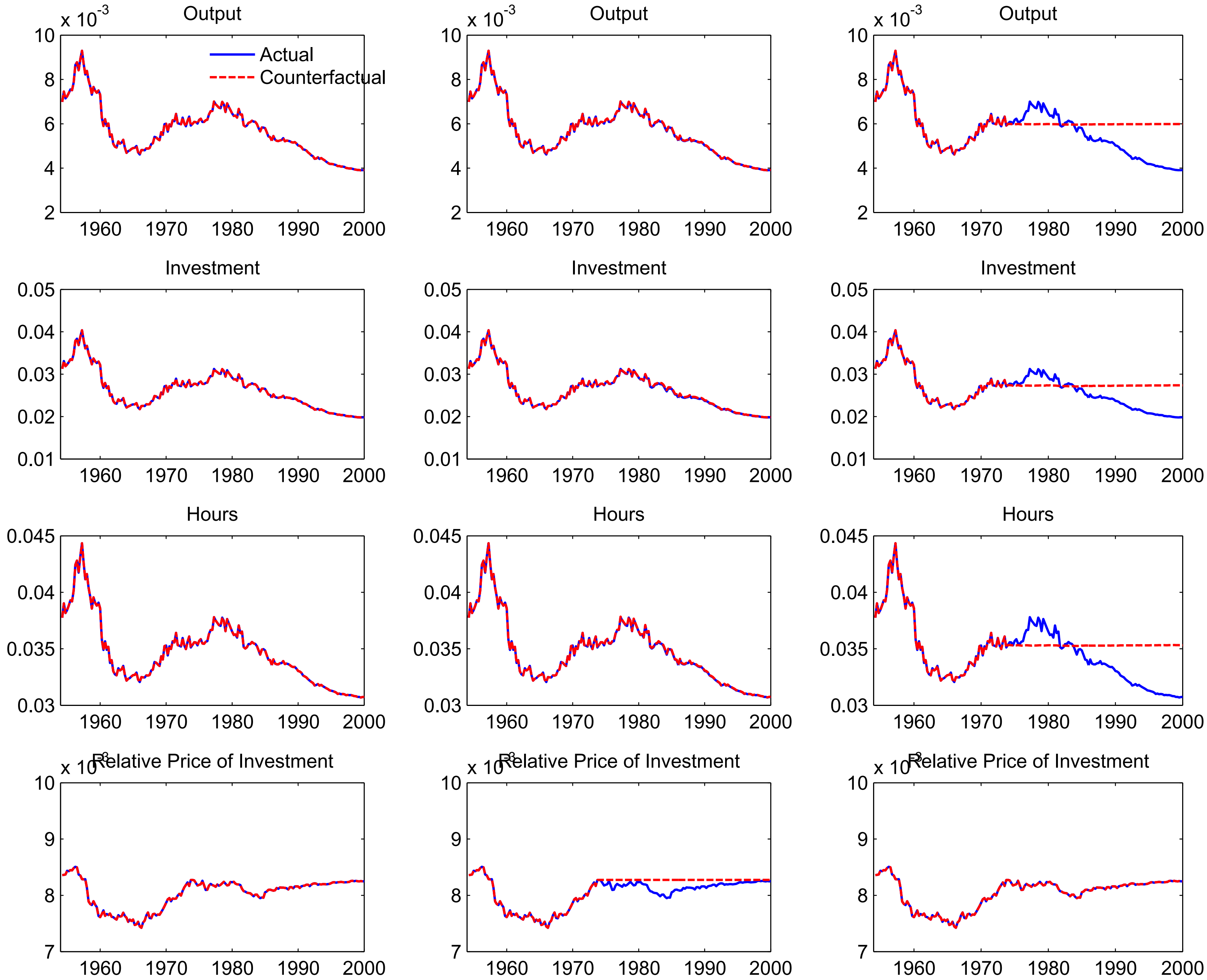


Figure 6.7: Counterfactual Exercise 3



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

- We use Likelihood Ratio tests to compare models, Rivers and Vuong (2002).

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

- We use Likelihood Ratio tests to compare models, Rivers and Vuong (2002).
- Loglike benchmark: 2350.6, loglike version 2: 2230.4

Figure 7.1: Comparison of Smoothed Capital and Shocks

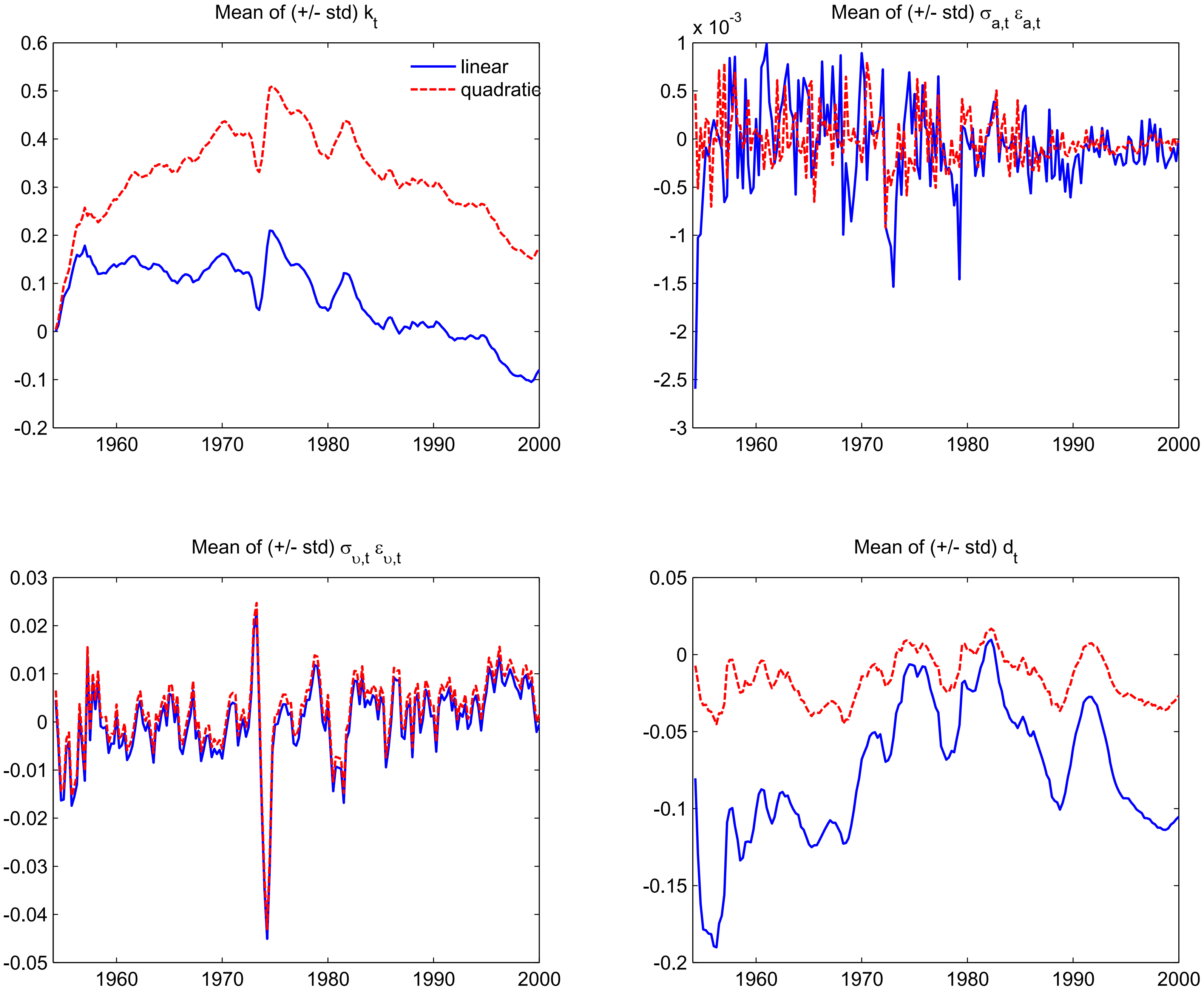
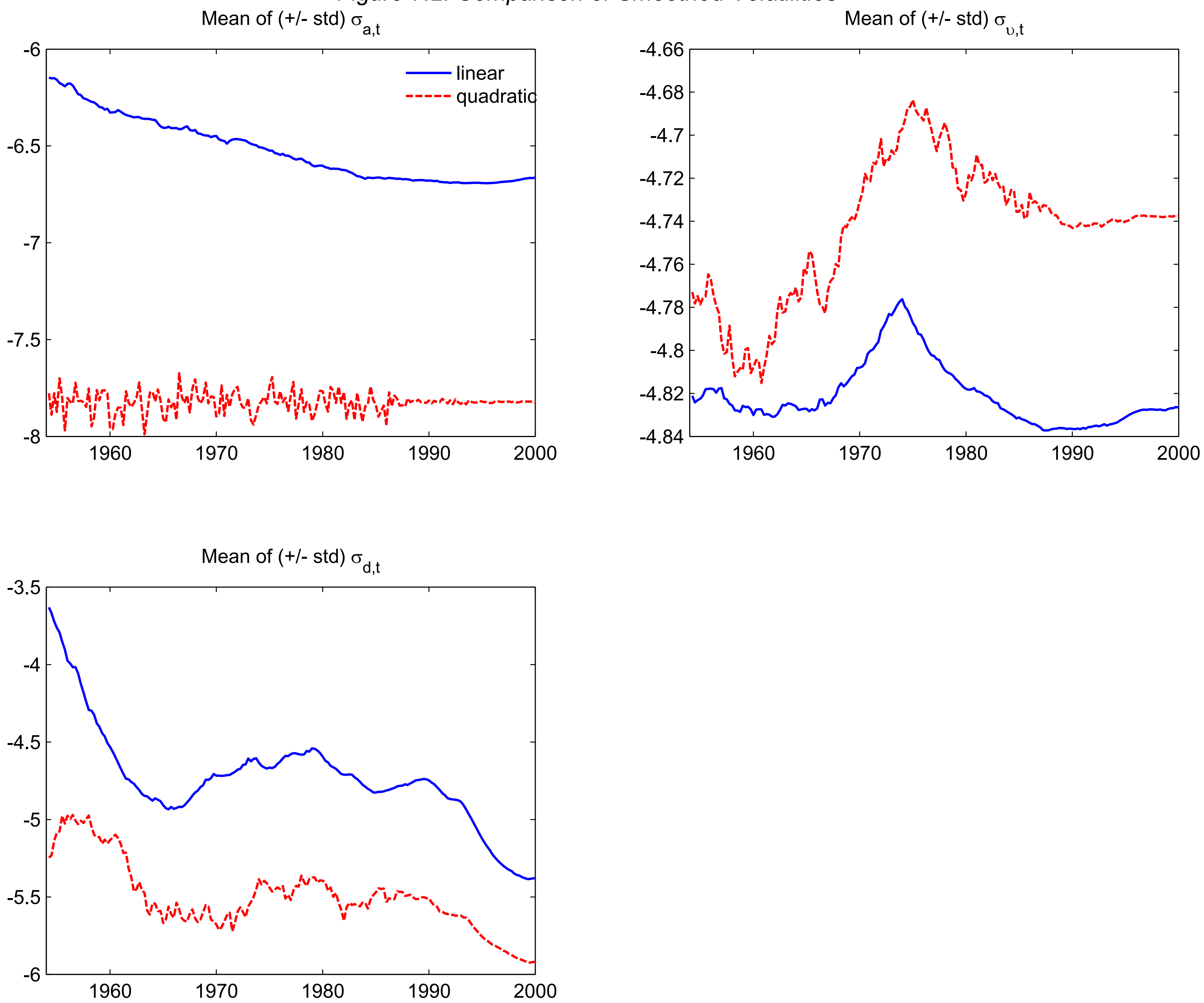


Figure 7.2: Comparison of Smoothed Volatilities



What are We Doing Now?

- We are estimating a richer DSGE model with:

  1. Nominal and real rigidities.
  2. Monetary and fiscal policy.
  3. Stochastic volatility.
  4. Parameter drifting.

What are We Doing Now?

- We are estimating a richer DSGE model with:

  1. Nominal and real rigidities.
  2. Monetary and fiscal policy.
  3. Stochastic volatility.
  4. Parameter drifting.

- We are working on a model with micro heterogeneity.

What are We Doing Now?

- We are estimating a richer DSGE model with:
  1. Nominal and real rigidities.
  2. Monetary and fiscal policy.
  3. Stochastic volatility.
  4. Parameter drifting.

- We are working on a model with micro heterogeneity.

- We are exploring the semi-nonparametric estimation of DSGE models.

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!