Inference in Structural VARs with External Instruments

José Luis Montiel Olea, Harvard University
James H. Stock, Harvard University
Mark W. Watson, Princeton University

September 2012

VARs, SVARs, and the Identification Problem

Sims (1980):

Structural MAR:

$$Y_t = D_1\varepsilon_{t\text{-}1} + D_2\varepsilon_{t\text{-}2} + \ldots = D(\mathrm{L})\varepsilon_t$$

Reduced form VAR:

$$A(\mathrm{L})Y_t = \eta_t, \quad \text{where } A(\mathrm{L}) = \mathrm{I} - A_1\mathrm{L} - \ldots - A_p\mathrm{L}^p$$

Innovations:

$$\eta_t = Y_t - E_{t-1}Y_t = A(\mathrm{L})Y_t$$

Structural errors ε_t:

$$\eta = H\varepsilon_t \text{ and } \varepsilon_t = H^{-1}\eta_t$$

Structural MAR:

$$Y_t = A(\mathrm{L})^{-1}\eta_t = A(\mathrm{L})^{-1}H\varepsilon_t = C(\mathrm{L})H\varepsilon_t$$

$C(\mathrm{L})H$ is structural impulse response function (dynamic causal effect)

SVAR estimands (focus on shock 1)

Partitioning notation:

$$\eta_t = H\varepsilon_t = \begin{bmatrix} H_1 & \cdots & H_r \end{bmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \vdots \\ \varepsilon_{rt} \end{pmatrix} = \begin{bmatrix} H_1 & H_\bullet \end{bmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \varepsilon_{\bullet t} \end{pmatrix}$$

Structural MAR:

$$Y_t = C(\mathrm{L})H\varepsilon_t = C(\mathrm{L})H_1\varepsilon_{1t} + C(\mathrm{L})H_\bullet\varepsilon_{\bullet t}$$

Structural MAR for j^{th} variable:

$$Y_{jt} = \sum_{k=0}^{\infty} C_{k,j} H_1 \varepsilon_{1t-k} + \sum_{k=0}^{\infty} C_{k,j} H_\bullet \varepsilon_{\bullet t-k}$$

$C_{k,j}$ is a $1\times r$ row vector

SVAR estimands (focus on shock 1), ctd.

(1) Structural IRF of variable j to shock 1 at lag h:

$$IRF = C_{h,j}H_1$$

(2) Historical contribution (decomposition):

$$Y_{jt} = \sum_{k=0}^{\infty} C_{k,j}H_1\varepsilon_{1t-k} + \sum_{k=0}^{\infty} C_{k,j}H_{\bullet}\varepsilon_{\bullet t-k}$$

Historical contribution of shock 1 to variable j over horizon h:

$$HD = \sum_{k=0}^{h} C_{k,j}H_1\varepsilon_{1t-j}$$

SVAR estimands (focus on shock 1), ctd.

(3) Forecast error variance decomposition:

$$Y_{j,t} - Y_{j,t|t\text{-}h} = \sum_{k=0}^{h} C_{k,j} H_1 \varepsilon_{1t-k} + \sum_{k=0}^{h} C_{k,j} H_{\bullet} \varepsilon_{\bullet t-k}$$

Suppose $E\left(\varepsilon_t \varepsilon_t'\right) = \Sigma_{\varepsilon\varepsilon} = D = diag(\sigma_{\varepsilon_1}^2, ..., \sigma_{\varepsilon_r}^2)$ (uncorrelated shocks). Then

$$FEVD = \frac{\text{var}\left(\sum_{k=0}^{h} C_{k,j} H_1 \varepsilon_{1t-k}\right)}{\text{var}\left(\sum_{k=0}^{h} C_{k,j} H \varepsilon_{t-k}\right)} = \frac{\text{var}\left(\sum_{k=0}^{h} C_{k,j} H_1 \varepsilon_{1t-k}\right)}{\text{var}\left(\sum_{k=0}^{h} C_{k,j} \eta_{t-k}\right)}$$

$$= \frac{\sum_{k=0}^{h} C_{k,j} H_1 H_1' C_{k,j}' \sigma_{\varepsilon_1}^2}{\sum_{k=0}^{h} C_{k,j} \Sigma_{\eta\eta} C_{k,j}'}$$

The structural VAR identification problem

r innovations: $\underset{r\times 1}{\eta_t} = \underset{r\times r}{H}\, \underset{r\times 1}{\varepsilon_t} = \begin{bmatrix} H_1 & \cdots & H_r \end{bmatrix} \begin{pmatrix} \varepsilon_{1t} \\ \vdots \\ \varepsilon_{rt} \end{pmatrix}$

System ID: What is H?

Assume $E(\varepsilon_t \varepsilon_t') = Diag = D$:	$r^2 + r$	parameters
$\Sigma_{\eta\eta} = HDH'$:	$- r(r+1)/2$	equations
normalization (e.g. $D = I_r$):	$- r$	normalization restrictions
Need:	$r(r-1)/2$	"theory" restrictions

Single IRF (single shock) ID: What is H_1?

Two approaches:

1. Internal restrictions:

   Short run restrictions (Sims (1980)), long run restrictions, identification by heteroskedasticity, bounds on IRFs)

The structural VAR identification problem, ctd.

2. External information ("method of external instruments")
   Romer and Romer (1989)
   Ramey and Shapiro (1998)
   Selected empirical papers
   - **Monetary shock**: Cochrane and Piazzesi (2002), Faust, Swanson, and Wright (2003. 2004), Romer and Romer (2004), Bernanke and Kuttner (2005), Gürkaynak, Sack, and Swanson (2005)
   - **Fiscal shock**: Romer and Romer (2010), Fisher and Peters (2010), Ramey (2011)
   - **Uncertainty shock**: Bloom (2009), Baker, Bloom, and Davis (2011), Bekaert, Hoerova, and Lo Duca (2010), Bachman, Elstner, and Sims (2010)
   - **Liquidity shocks**: Gilchrist and Zakrajšek's (2011), Bassett, Chosak, Driscoll, and Zakrajšek's (2011)
   - **Oil shock**: Hamilton (1996, 2003), Kilian (2008a), Ramey and Vine (2010)

Outline

1. Introduction
2. Method of external instruments: identification
3. Method of external instruments: estimation
4. Strong instrument asymptotics
5. Weak instrument asymptotics – setup and distributions
6. Inference for IRFs
7. Inference for historical decompositions
8. Extensions
9. Empirical results
10. Conclusions

2. The method of external instruments: Identification

Methods/Literature

- Nearly all empirical papers use OLS & report (only) first stage
- However, these "shocks" are best thought of as instruments (quasi-experiments)
- Treatments of external shocks as instruments:

    Hamilton (2003)

    Kilian (2008 – *JEL*)

    Stock and Watson (2008, 2012)

    Mertens and Ravn (2012) – same setup as here (and as in Stock and Watson (2008)), executed using strong instrument asymptotics

Identification of H_1

$$A(L)Y_t = \eta_t, \quad \eta_t = H\varepsilon_t = \begin{bmatrix} H_1 & \cdots & H_r \end{bmatrix}\begin{pmatrix} \varepsilon_{1t} \\ \vdots \\ \varepsilon_{rt} \end{pmatrix}$$

Suppose you have an instrumental variable Z_t (not in Y_t) such that

(i) $E\left(\varepsilon_{1t}Z_t'\right) = \alpha' \neq 0$ (relevance)

(ii) $E\left(\varepsilon_{jt}Z_t'\right) = 0$, $j = 2,\ldots, r$ (exogeneity)

(iii) $E\left(\varepsilon_t\varepsilon_t'\right) = \Sigma_{\varepsilon\varepsilon} = D = diag(\sigma^2_{\varepsilon_1},...,\sigma^2_{\varepsilon_r})$

Under (i) and (ii), you can identify H_1 up to sign & scale

$$E(\eta_t Z_t') = E(H\varepsilon_t Z_t') = \begin{bmatrix} H_1 & \cdots & H_r \end{bmatrix}\begin{pmatrix} E(\varepsilon_{1t}Z_t') \\ \vdots \\ E(\varepsilon_{rt}Z_t') \end{pmatrix} = \begin{bmatrix} H_1 & \cdots & H_r \end{bmatrix}\begin{pmatrix} \alpha' \\ 0 \\ 0 \end{pmatrix} = H_1\alpha'$$

Identification of H_1, ctd.

$$E(\eta_t Z_t') = E(H\varepsilon_t Z_t') = \begin{bmatrix} H_1 & H_\bullet \end{bmatrix} \begin{pmatrix} E(\varepsilon_{1t} Z_t') \\ E(\varepsilon_{\bullet t} Z_t') \end{pmatrix} = H_1 \alpha'$$

<u>Normalization</u>

- The scale of H_1 and $\sigma^2_{\varepsilon_1}$ is set by a normalization subject to

$$\Sigma_{\eta\eta} = HDH' \quad \text{where } D = diag(\sigma^2_{\varepsilon_1},...,\sigma^2_{\varepsilon_r})$$

- Normalization studied here: a unit positive value of shock 1 is defined to have a unit positive effect on the innovation to variable 1, which is u_{1t}. This corresponds to:

(iv) $H_{11} = 1$ (unit shock normalization)

where H_{11} is the first element of H_1

Identification of H_1, ctd.

Impose normalization (iv):

$$E(\eta_t Z_t') = \begin{pmatrix} E\eta_{1t}Z_t{}' \\ E\eta_{\bullet t}Z_t{}' \end{pmatrix} = H_1\alpha' = \begin{pmatrix} H_{11} \\ H_{1\bullet} \end{pmatrix}\alpha' = \begin{pmatrix} 1 \\ H_{1\bullet} \end{pmatrix}\alpha'$$

So

$$\begin{pmatrix} H_{1\bullet}E\eta_{1t}Z_t{}' \\ E\eta_{\bullet t}Z_t{}' \end{pmatrix} = \begin{pmatrix} H_{1\bullet}\alpha' \\ H_{1\bullet}\alpha' \end{pmatrix}$$

or

$$H_{1\bullet}E\eta_{1t}Z_t{}' = E\eta_{\bullet t}Z_t{}'$$

If Z_t is a scalar (k = 1): $$H_{1\bullet} = \frac{E\eta_{\bullet t}Z_t}{E\eta_{1t}Z_t}$$

Identification of ε_{1t}

$$\varepsilon_t = H^{-1}\eta_t = \begin{bmatrix} H^{1\prime} \\ \vdots \\ H^{r\prime} \end{bmatrix}\eta_t$$

- Identification of first column of H and $\Sigma_{\varepsilon\varepsilon} = D$ identifies first row of H^{-1} up to scale (can show via partitioned matrix inverse formula).
- Alternatively, let Φ be the coefficient matrix of the population regression of Z_t onto η_t:

$$\Phi = E(Z_t\eta_t{}')\Sigma_\eta^{-1} = \alpha H_1{}'(HDH')^{-1} = \alpha H_1{}'H'^{-1}D^{-1}H^{-1} = (\alpha/\sigma_{\varepsilon_1}^2)H^{1\prime}$$

because $H^{-1}H_1 = (1\ 0\ \ldots\ 0)'$. Thus ε_{1t} is identified up to scale by

$$\Phi\eta_t = \frac{\alpha}{\sigma_{\varepsilon_1}^2}H^{1\prime}\eta_t = \frac{\alpha}{\sigma_{\varepsilon_1}^2}\varepsilon_{1t}$$

Identification of ε_{1t}, ctd

$\Phi\eta_t$ is the predicted value from the population projection of Z_t on η_t:

$$\tilde{\varepsilon}_{1t} = \Phi\eta_t = E(Z_t\eta_t{}')\Sigma_\eta^{-1}\eta_t = \frac{\alpha}{\sigma_{\varepsilon_1}^2}\varepsilon_{1t}$$

- Φ has rank 1 (in population), so this is a (population) reduced rank regression
- 2 instruments identify 2 shocks. Suppose they are shocks 1 and 2, identified by Z_{1t} and Z_{2t}. Then

$$E(\tilde{\varepsilon}_{1t}\tilde{\varepsilon}_{2t}) = E(Z_{1t}\eta_t{}')\Sigma_\eta^{-1}E(\eta_t Z_{2t})$$

which = 0 if both instruments satisfy (i) – (iii)

"Reduced form" VARX (single Z case)

VAR: $A(L)Y_t = \eta_t$, $\eta_t = H\varepsilon_t$

Additionally assume:

$$\text{(v) } E\left(Y_{t-k}Z_t'\right) = 0,\ k = 1,\ldots \text{ (}Z\text{ lag dynamics restriction)}$$

Then $\text{Proj}(\eta_t|Z_t,Y_{t-1}) = \text{Proj}(\eta_t|Z_t) = \Gamma Z_t$,

where $\Gamma = E(\eta_t Z_t)/\sigma_Z^2 = (\alpha/\sigma_Z^2)H_1$

Thus under (i) – (iii) and (v), Y_t follows the VARX:

$$A(L)Y_t = \Gamma Z_t + \nu_t, \text{ ("Reduced form" VARX)}$$

where ν_t is the projection residual so $\text{corr}(Z_t, \nu_t) = 0$.

"Reduced form" distributed lag

$$A(L)Y_t = \Gamma Z_t + \nu_t, \text{ ("Reduced form" VARX)}$$

so

$$Y_t = A(L)^{-1}\Gamma Z_t + A(L)^{-1}\nu_t, \text{ ("Reduced form" DL)}$$

where $E(Z_t \nu_t) = 0$.

- $A(L)^{-1}\Gamma$ are the (reduced form) IRFs with respect to the instrument
- Ratios of elements of $A(L)^{-1}\Gamma$ are the structural IRFs.

Empirical practice – *what is done in the literature*?

- Many things: estimation of VARX, of DL, of ADL (single equation)
- In almost cases inference is reported for the IRF with respect to Z_t, not the structural IRF. *Exceptions*: Hamilton (2003), Kilian (2009), Mertens-Ravn (2012)

3. Estimation

Recall notation: $H_1 = \begin{bmatrix} H_{11} \\ H_{1\bullet} \end{bmatrix}$, $\eta_t = \begin{bmatrix} \eta_{1t} \\ \eta_{\bullet t} \end{bmatrix}$

Impose the normalization condition (iv) $H_{11} = 1$, so

$$E(\eta_t Z_t') = H_1 \alpha' = \begin{pmatrix} 1 \\ H_{1\bullet} \end{pmatrix} \alpha \text{ or } E(\eta_t \otimes Z_t) = \begin{pmatrix} 1 \\ H_{1\bullet} \end{pmatrix} \otimes \alpha$$

High level assumption (assume throughout)

$$\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left([\eta_t \otimes Z_t] - [H_1 \otimes \alpha] \right) \xrightarrow{d} \mathrm{N}(0,\Omega) \qquad (*)$$

Estimation of H_1

Efficient GMM objective function:

$S(H_{1\bullet}, \alpha; \hat{\Omega})$

$$= \left(\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left((\hat{\eta}_t \otimes Z_t) - \left(\begin{bmatrix} 1 \\ H_{1\bullet} \end{bmatrix} \otimes \alpha \right) \right) \right)' \hat{\Omega}^{-1} \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \left((\hat{\eta}_t \otimes Z_t) - \left(\begin{bmatrix} 1 \\ H_{1\bullet} \end{bmatrix} \otimes \alpha \right) \right)$$

$k = 1$ (exact identification): $\quad E(\eta_t Z_t') = H_1 \alpha' = \begin{pmatrix} \alpha \\ \alpha H_{1\bullet} \end{pmatrix}$

so GMM estimator solves, $\quad T^{-1} \sum_{t=1}^{T} \hat{\eta}_t Z_t = \begin{pmatrix} \hat{\alpha} \\ \hat{\alpha} \hat{H}_{1\bullet} \end{pmatrix}$

GMM estimator: $\quad \hat{H}_{1\bullet} = \dfrac{T^{-1} \sum_{t=1}^{T} \hat{\eta}_{\bullet t} Z_t}{T^{-1} \sum_{t=1}^{T} \hat{\eta}_{1t} Z_t}$

IV interpretation: $\quad \hat{\eta}_{jt} = H_{1j} \hat{\eta}_{1t} + u_{jt},$

$$\hat{\eta}_{1t} = \Pi_j' Z_t + v_{jt}$$

GMM estimation of $H^{1\prime}$ and ε_{1t}

Recall $\tilde{\varepsilon}_{1t} = E(Z_t\eta_t')\Sigma_\eta^{-1}\eta_t = \Phi\eta_t$

Estimator:

- $k = 1$:

  $\hat{\varepsilon}_{1t}$ is the predicted value (up to scale) in the regression of Z_t on $\hat{\eta}_t$

- $k > 1$(no-HAC):

  Absent serial correlation/no heteroskedasticity, the GMM estimator simplifies to reduced rank regression:

$$Z_t = \Phi\hat{\eta}_t + \nu_t \qquad \text{(RRR)}$$

- If Z_t is available only for a subset of time periods, estimate (RRR) using available data, compute predicted value over full period

4. Strong instrument asymptotics

- $k = 1$ case:

$$\sqrt{T}\left(\hat{H}_{1\bullet} - H_{1\bullet}\right) \xrightarrow{d} N(0, \Gamma'\Omega\Gamma), \text{ where } \Gamma = \begin{bmatrix} -H_{1\bullet}{}' \\ I_{r-1} \end{bmatrix}$$

- Overidentified case ($k > 1$):
  - usual GMM formula
  - J-statistics, etc. are standard textbook GMM

5. Weak instrument asymptotics: k = 1

(a) Distribution of $\hat{H}_{1\bullet}$

$$\hat{H}_{1\bullet} = \frac{T^{-1}\sum_{t=1}^{T}\hat{\eta}_{\bullet t}Z_t}{T^{-1}\sum_{t=1}^{T}\hat{\eta}_{1t}Z_t}$$

Weak IV asymptotic setup – local drift (limit of experiments, etc.):

$$\alpha = \alpha_T = a/\sqrt{T}$$

so

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\Big((\eta_t \otimes Z_t) - (H_1 \otimes \alpha)\Big) \xrightarrow{d} \mathrm{N}(0,\Omega) \qquad (*)$$

becomes

$$\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\left(\eta_t \otimes Z_t\right) \xrightarrow{d} \mathrm{N}(H_1 \otimes a,\ \Omega) \qquad (\text{*-weakIV})$$

Weak instrument asymptotics for H_1, ctd

Estimation of A(L) under (i) – (v) (serially uncorrelated instruments case)

Let $\alpha = [-A_1 \ldots -A_p]$ so $\eta_t = A(\mathrm{L})Y_t = Y_t - \alpha' \underline{Y}_{t-1}$. Then

$$
\begin{aligned}
T^{-1/2}\sum_{t=1}^{T}\hat{\eta}_t Z_t &= T^{-1/2}\sum_{t=1}^{T}\eta_t Z_t + T^{-1/2}\sum_{t=1}^{T}(\hat{\eta}_t - \eta_t)Z_t \\
&= T^{-1/2}\sum_{t=1}^{T}\eta_t Z_t + T^{-1/2}\sum_{t=1}^{T}(\hat{\alpha} - \alpha)\underline{Y}_{t-1}Z_t \\
&= T^{-1/2}\sum_{t=1}^{T}\eta_t Z_t + T^{1/2}(\hat{\alpha} - \alpha)T^{-1}\sum_{t=1}^{T}\underline{Y}_{t-1}Z_t \\
&= T^{-1/2}\sum_{t=1}^{T}\eta_t Z_t + o_p(1)
\end{aligned}
$$

Weak instrument asymptotics for H_1, ctd

Under (iv),

$$\hat{H}_{1\bullet} = \frac{T^{-1}\sum_{t=1}^{T}\hat{\eta}_{\bullet t}Z_t}{T^{-1}\sum_{t=1}^{T}\hat{\eta}_{1t}Z_t} = \frac{T^{-1/2}\sum_{t=1}^{T}\eta_{\bullet t}Z_t}{T^{-1/2}\sum_{t=1}^{T}\eta_{1t}Z_t} + o_p(1)$$

Standardize (*):

$$\sigma_Z^{-1}diag(\Sigma_{\eta\eta})^{-1/2}\frac{1}{\sqrt{T}}\sum_{t=1}^{T}(\eta_t \otimes Z_t) \Rightarrow \lambda + z, \qquad (**)$$

where $\lambda = \sigma_Z^{-1}diag(\Sigma_{\eta\eta})^{-1/2}(H_1 \otimes a)$

and $z = \begin{bmatrix} z_1 \\ z_\bullet \end{bmatrix} \sim \text{N(0,W)}$, $W = \sigma_Z^{-2}diag(\Sigma_{\eta\eta})^{-1/2}\Omega diag(\Sigma_{\eta\eta})^{-1/2\prime}$

Thus, in $k = 1$ case, $\hat{H}_{1\bullet} = \frac{T^{-1}\sum_{t=1}^{T}\eta_{\bullet t}Z_t}{T^{-1}\sum_{t=1}^{T}\eta_{1t}Z_t} \Rightarrow \frac{\lambda_\bullet + z_\bullet}{\lambda_1 + z_1} = H_{1\bullet}^*$

Comments

1. In the no-HAC case, $\Omega = \Sigma_{\eta\eta}\sigma_Z^2$ so $W_{ij} = \text{corr}(\eta_{it}, \eta_{jt})$

Weak instrument asymptotics for H_1, ctd

$$\hat{H}_{1\bullet} = \frac{T^{-1}\sum_{t=1}^{T}\eta_{\bullet t}Z_t}{T^{-1}\sum_{t=1}^{T}\eta_{1t}Z_t} + o_p(1) \Rightarrow \frac{\lambda_\bullet + z_\bullet}{\lambda_1 + z_1} = H_{1\bullet}^*$$

Comments, ctd.

2. In the no-HAC case, convergence to strong instrument normal is governed by

$$\lambda_1^2 = a^2 / \sigma_{\eta_1}^2 \sigma_Z^2 \text{ = noncentrality parameter of first-stage } F$$

For the HAC case, see Montiel Olea and Pflueger (2012)

3. Consider unidentified case: $a = 0$ so $\lambda = 0$ so

$$\hat{H}_{1j} = \frac{T^{-1}\sum_{t=1}^{T}\eta_{jt}Z_t}{T^{-1}\sum_{t=1}^{T}\eta_{1t}Z_t} \Rightarrow \frac{z_j}{z_1} \sim \int N(\delta_j, \frac{\tau_j^2}{z_1^2})dF_{z_1^2}$$

where δ_j = plim of OLS estimator in the regression, $\eta_{jt} = \delta_j\eta_{1t} + \nu_{jt}$

- $\hat{H}_1$ is median-biased towards $\delta = E(\eta_t\eta_{1t})/\sigma_{\eta_1}^2$ = the first column of the Cholesky decomposition whit η_{1t} ordered first

Weak instrument asymptotics for structural IRFs

Structural IRF: $C(\mathrm{L})H_1$

where $C(\mathrm{L}) = A(\mathrm{L})^{-1} = C_0 + C_1\mathrm{L} + C_2\mathrm{L}^2 + \ldots$

Effect on variable j of shock 1 after h periods: $C_{h,j}H_1$

<u>Weak instrument asymptotic distribution of IRF</u>

$\hat{A}(L)$ is identified from the reduced form:

$$\sqrt{T}(\hat{A}(L) - A(L)) = O_p(1) \text{ (asymptotically normal)}$$

so

$$\hat{C}(L)\hat{H}_1 \Rightarrow C(\mathrm{L})H_1^*$$

Estimator of h-step IRF on variable j: $\hat{C}_{h,j}\hat{H}_1 \Rightarrow C_{h,j}H_1^*$

- This won't be a good approximation in practice – need to incorporate $O_p(T^{-1/2})$ term!

Numerical results for IRFs – asymptotic distributions

DGP calibration: $r = 2$

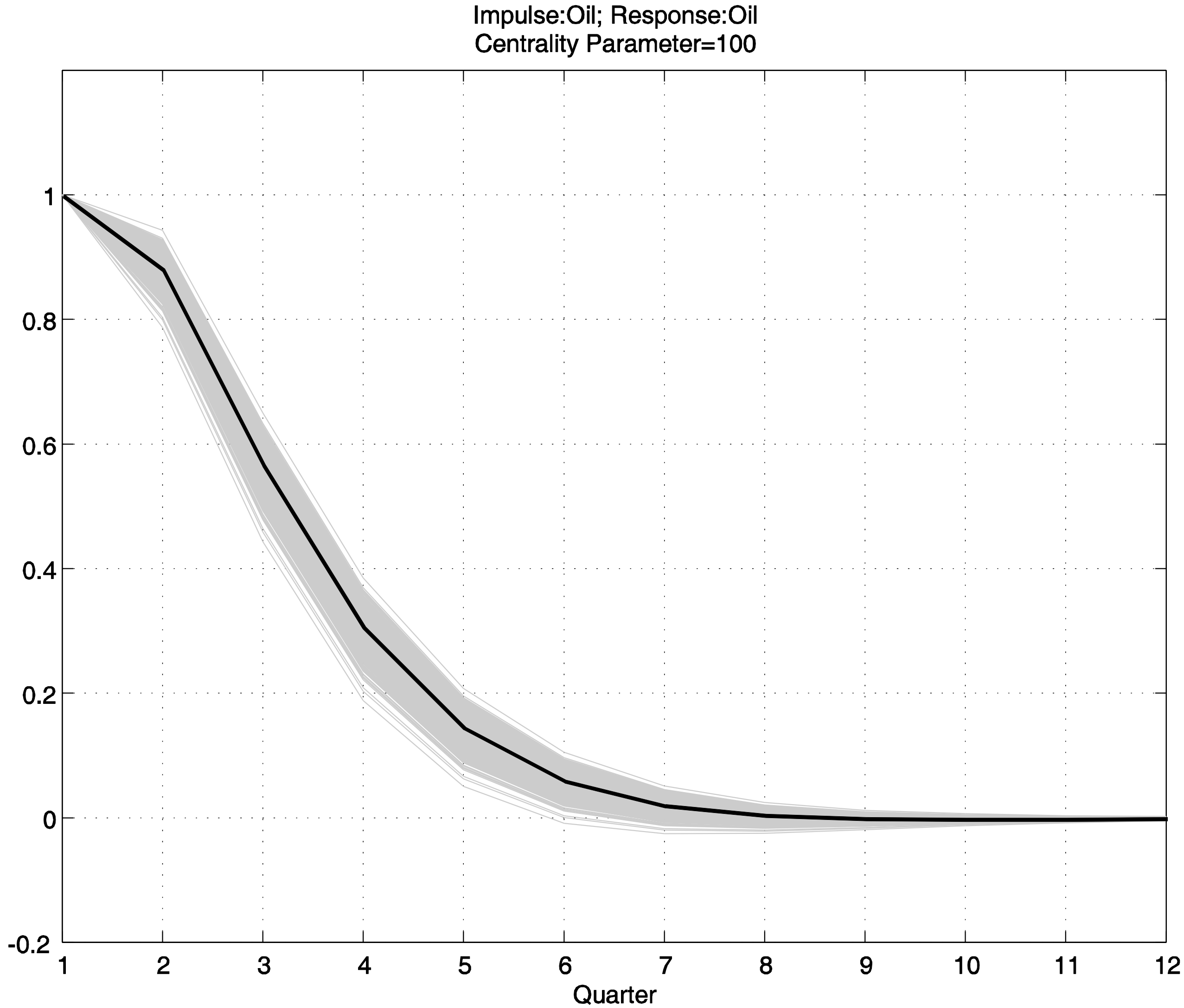
- $Y = (\Delta\ln POIL_t, \Delta\ln GDP_t)$, US, 1959Q1-2011Q2
- Estimate $A(L)$, $\Sigma_{\eta\eta}$, and H_1, then fix throughout
  - $A(L)$, $\Sigma_{\eta\eta}$: VAR(2)
  - H_1: estimated using Z_t = Kilian (2008 – *REStat*) OPEC supply shortfall (available 1971Q1-2004Q3)

Weak instrument asymptotic distribution:

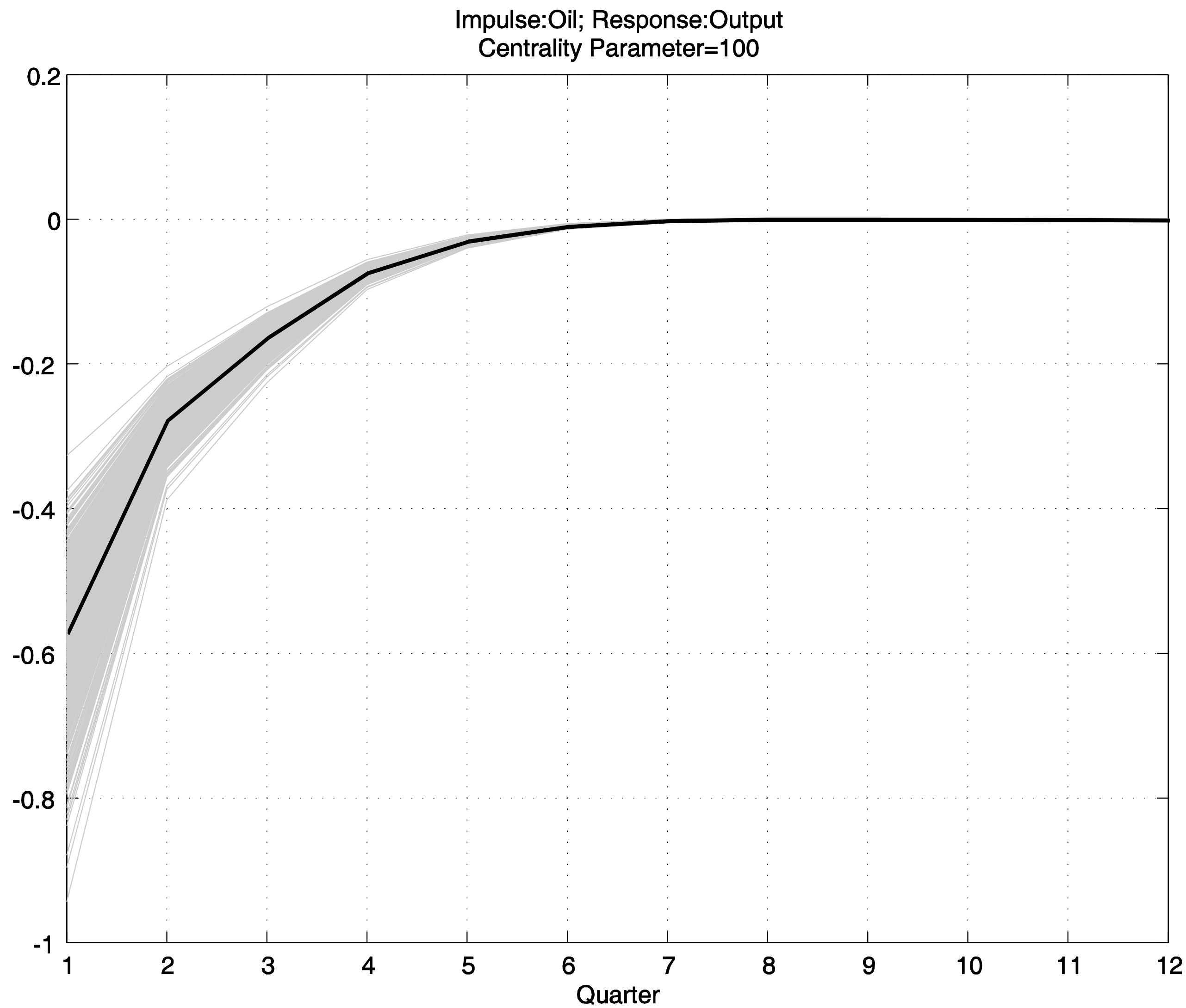
h-period IR, shock 1 on variable j: because $r = 2$,

$$C_{h,j}H_1^* = C_{h,j1} + C_{h,j2}H_{12}^*,\ H_{12}^* = \frac{\lambda_2 + z_2}{\lambda_1 + z_1}$$

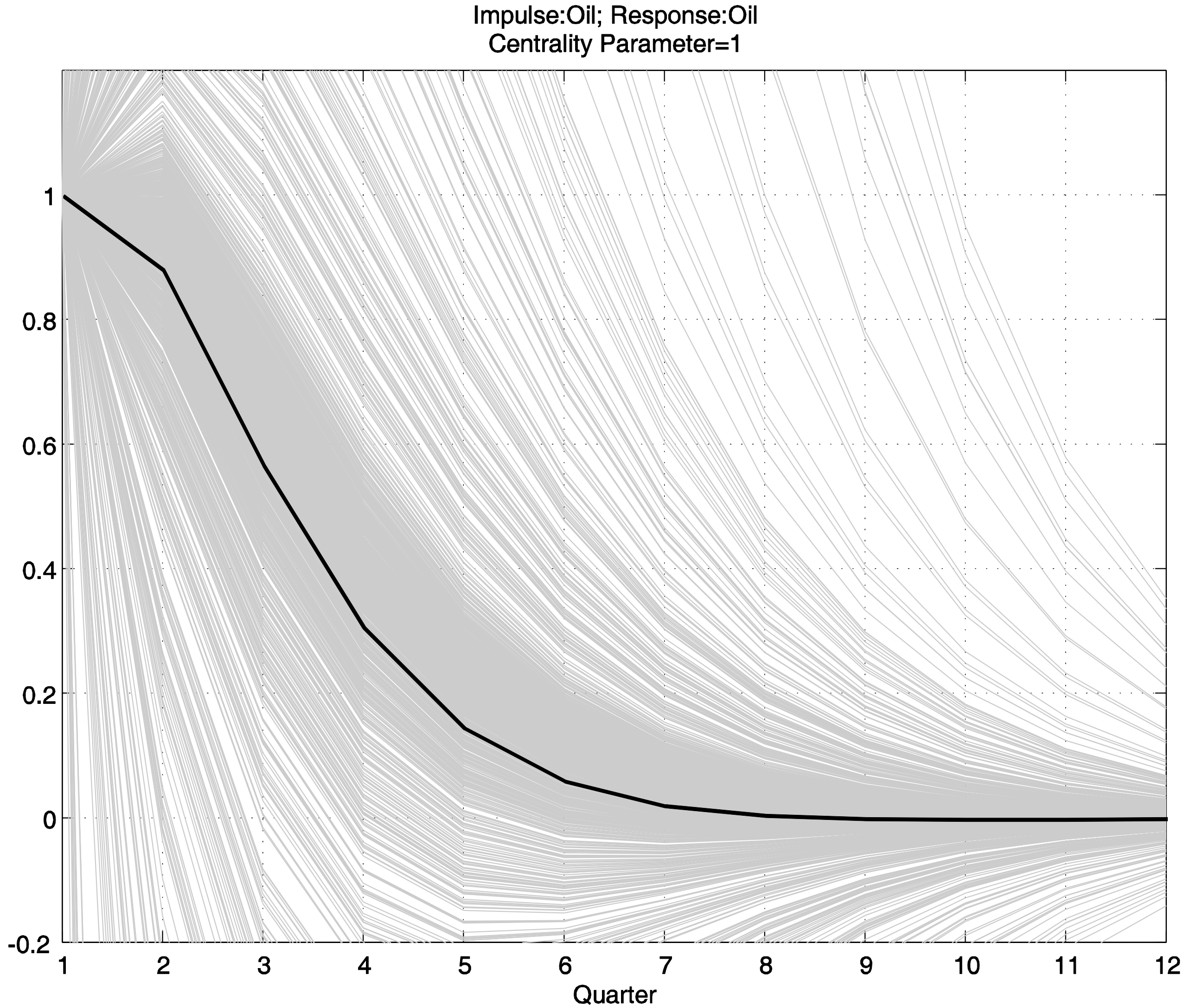
where $\begin{pmatrix} \lambda_1 \\ \lambda_2 \end{pmatrix} = a\begin{pmatrix} 1/\sigma_Z\sigma_{\eta_1} \\ H_{12}/\sigma_Z\sigma_{\eta_2} \end{pmatrix}$ and $\begin{pmatrix} z_1 \\ z_2 \end{pmatrix} \sim N\left(0, \begin{bmatrix} 1 & corr(\eta_1,\eta_2) \\ . & 1 \end{bmatrix}\right)$



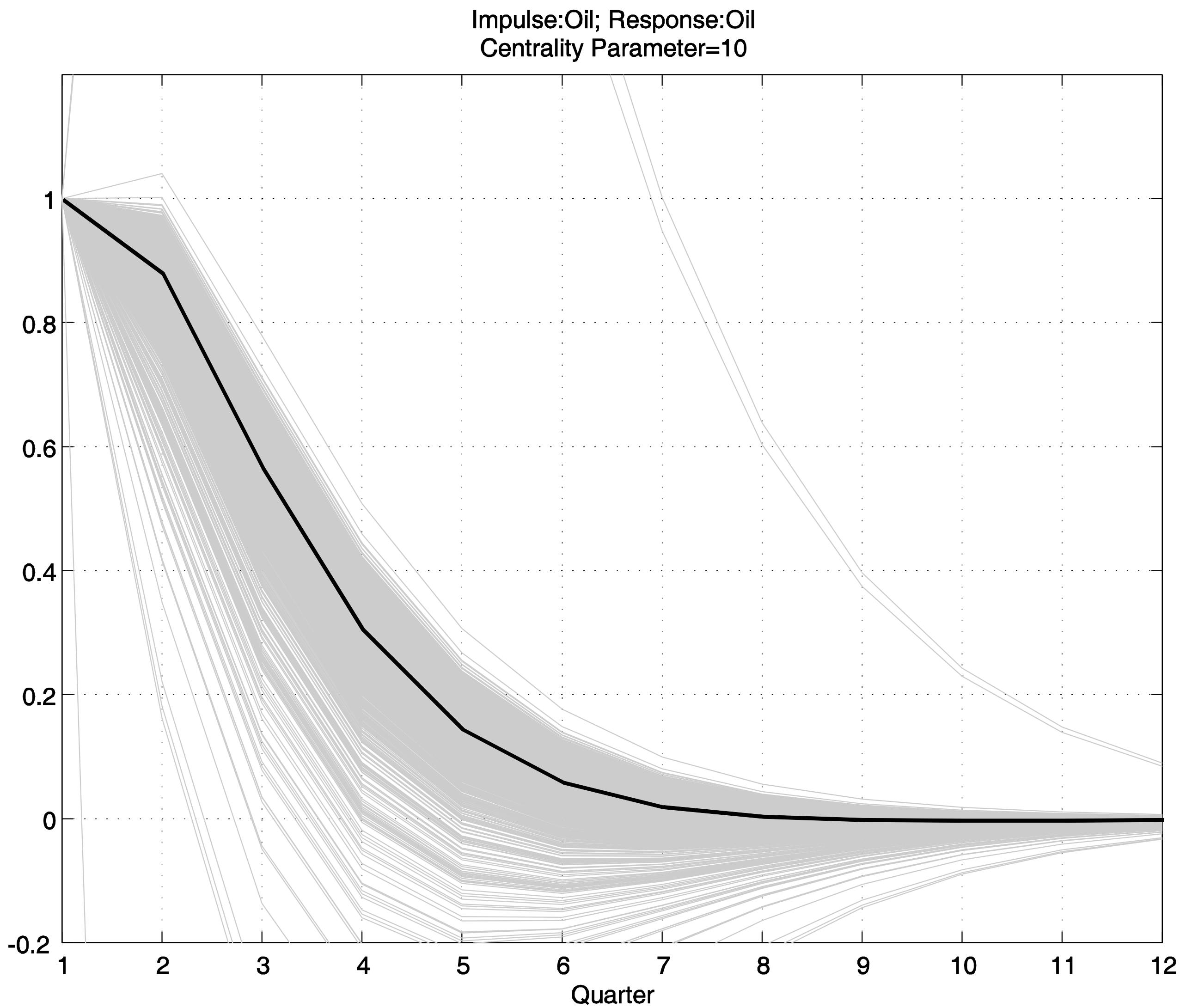
Effect of oil on oil growth: $\lambda_1^2 = 100$



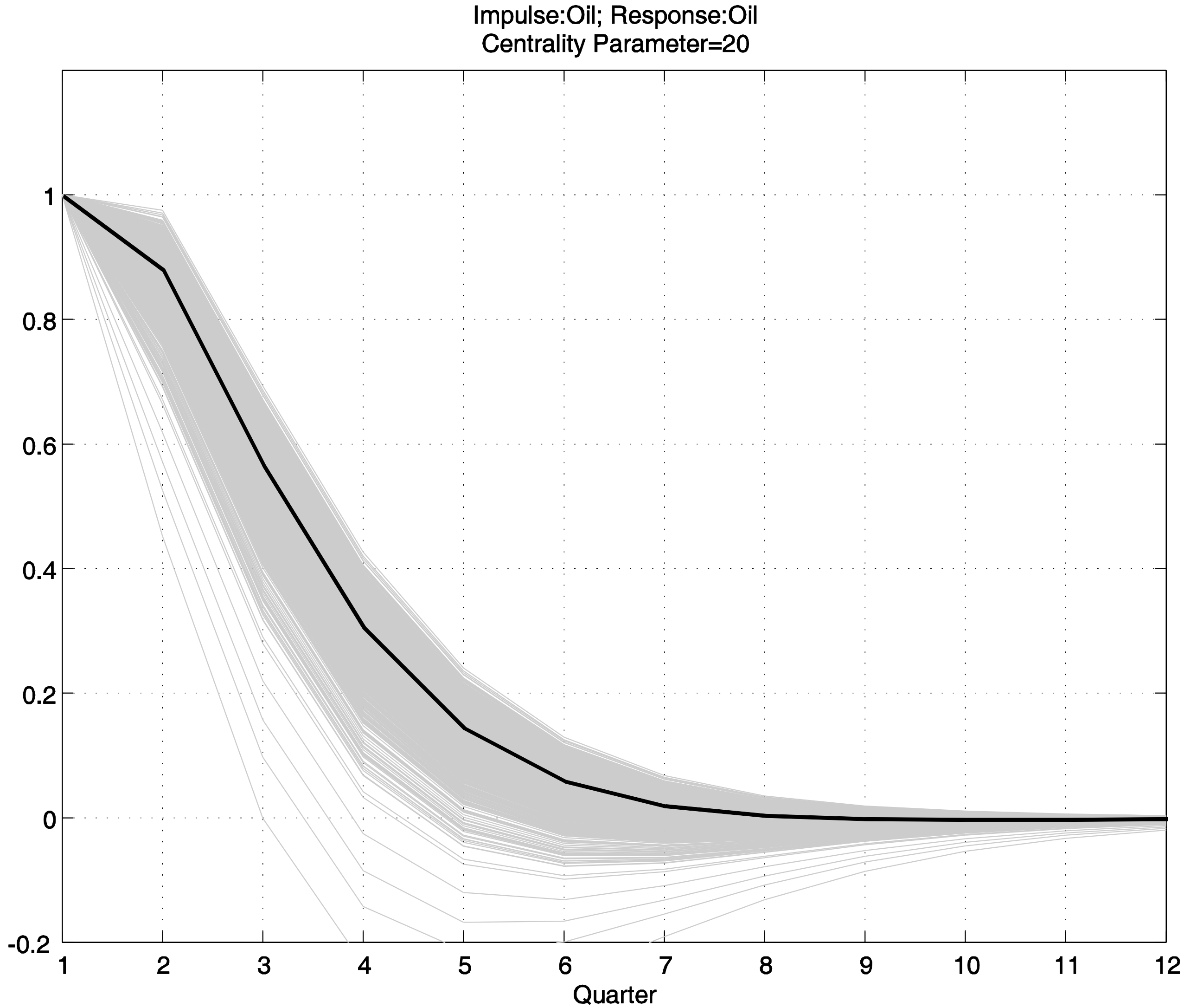
Effect of oil on GDP growth: $\lambda_1^2 = 100$



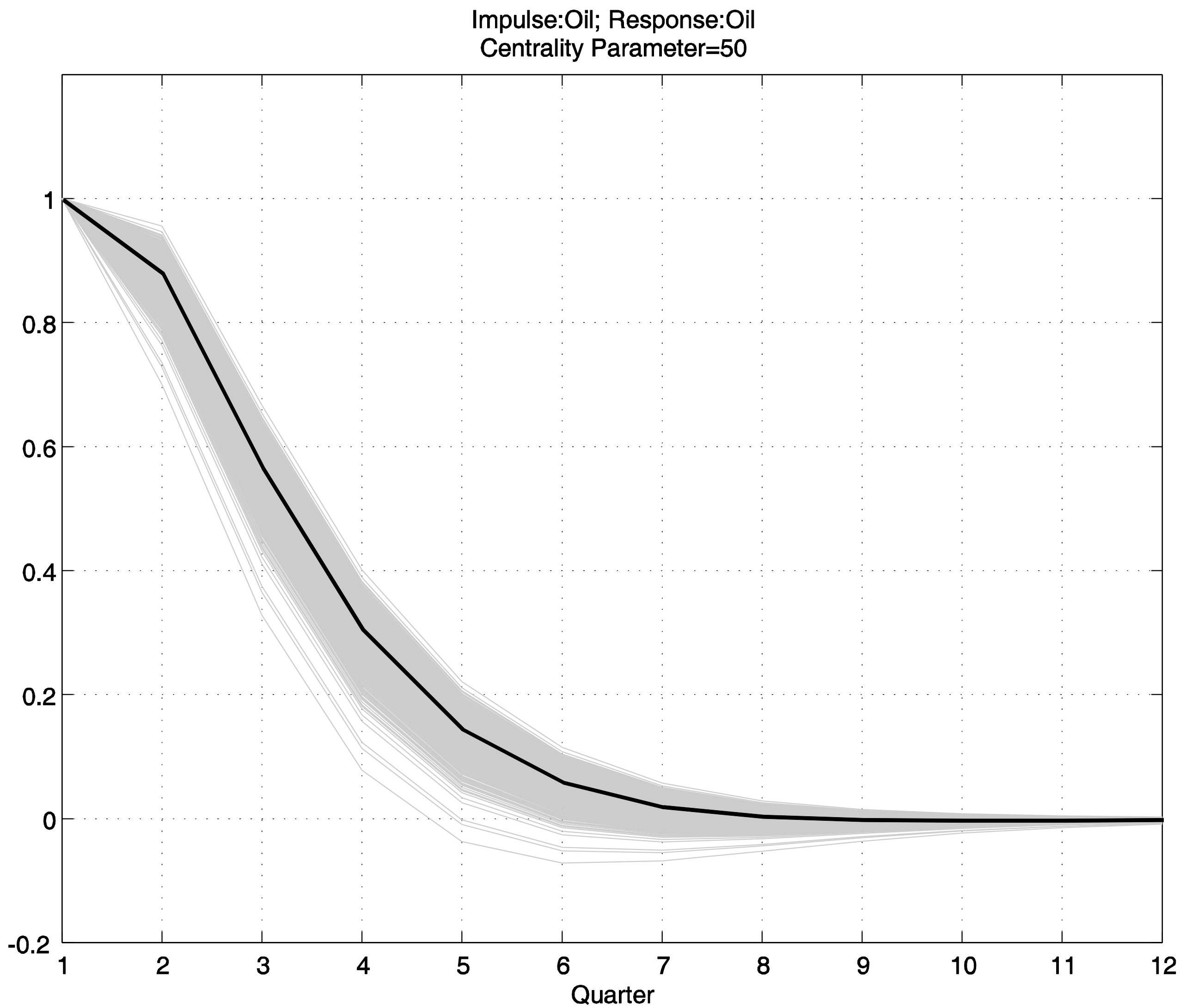
Effect of oil on oil growth: $\lambda_1^2 = 1$



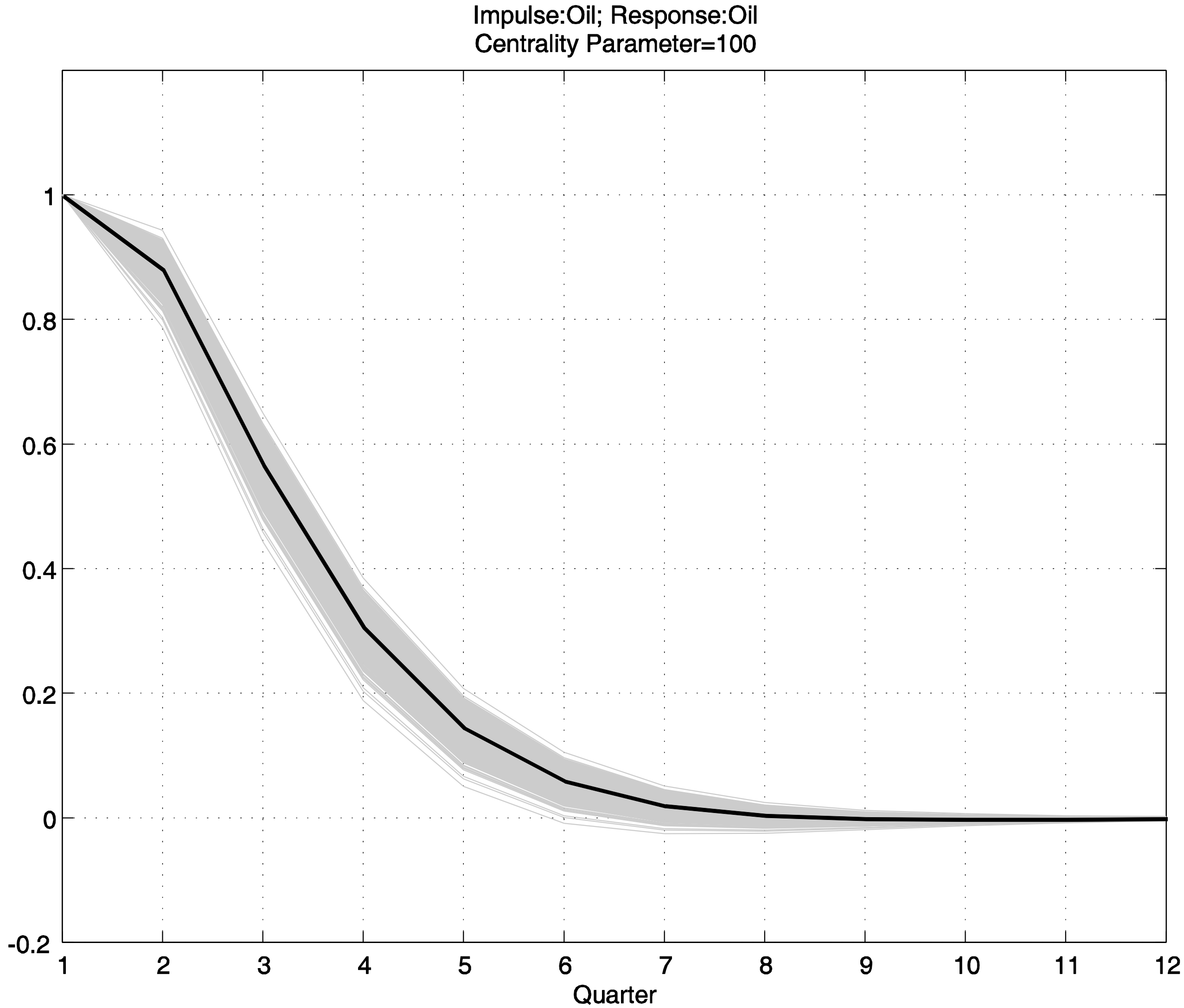
Effect of oil on oil growth: $\lambda_1^2 = 10$



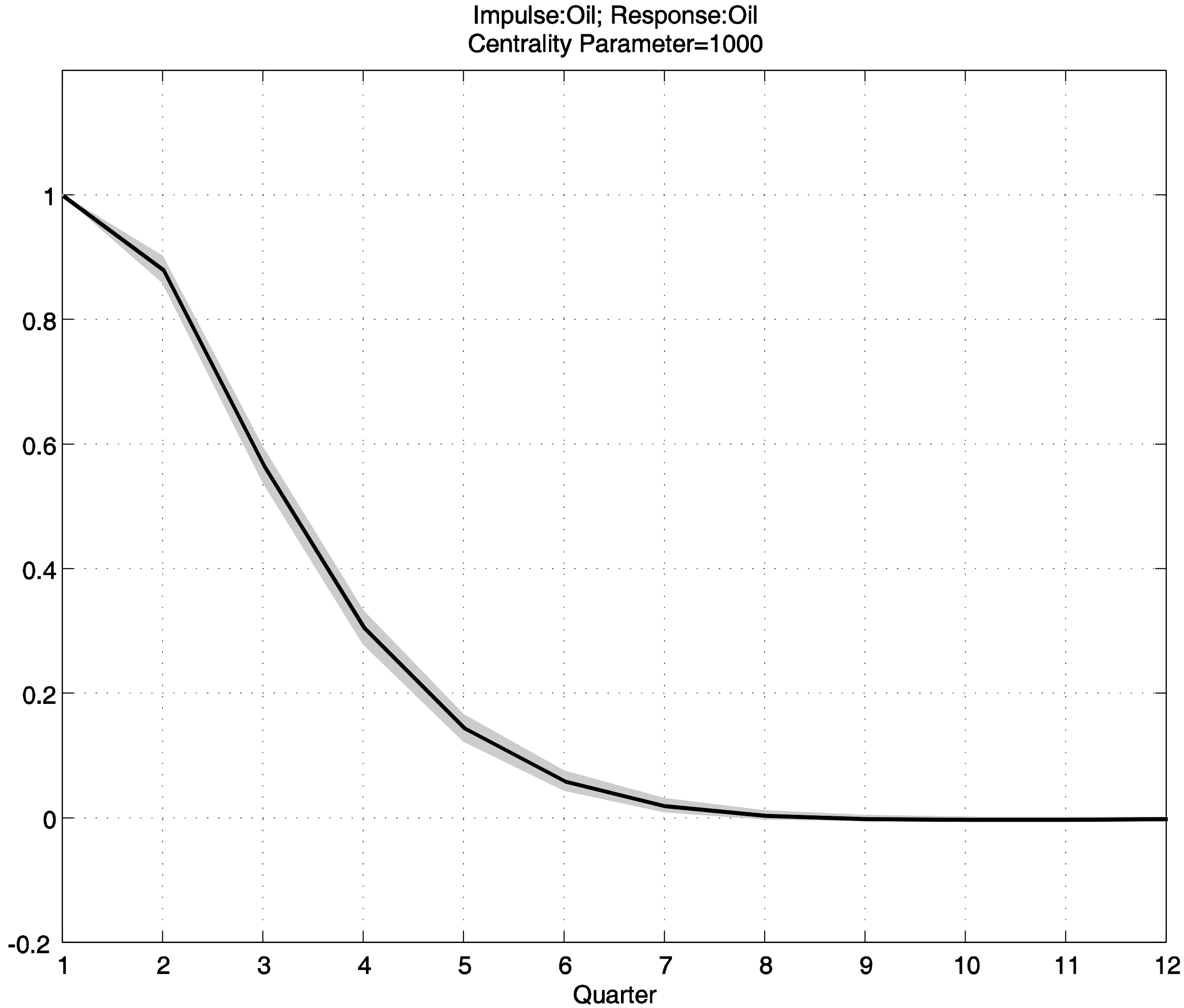
Effect of oil on oil growth: $\lambda_1^2 = 20$



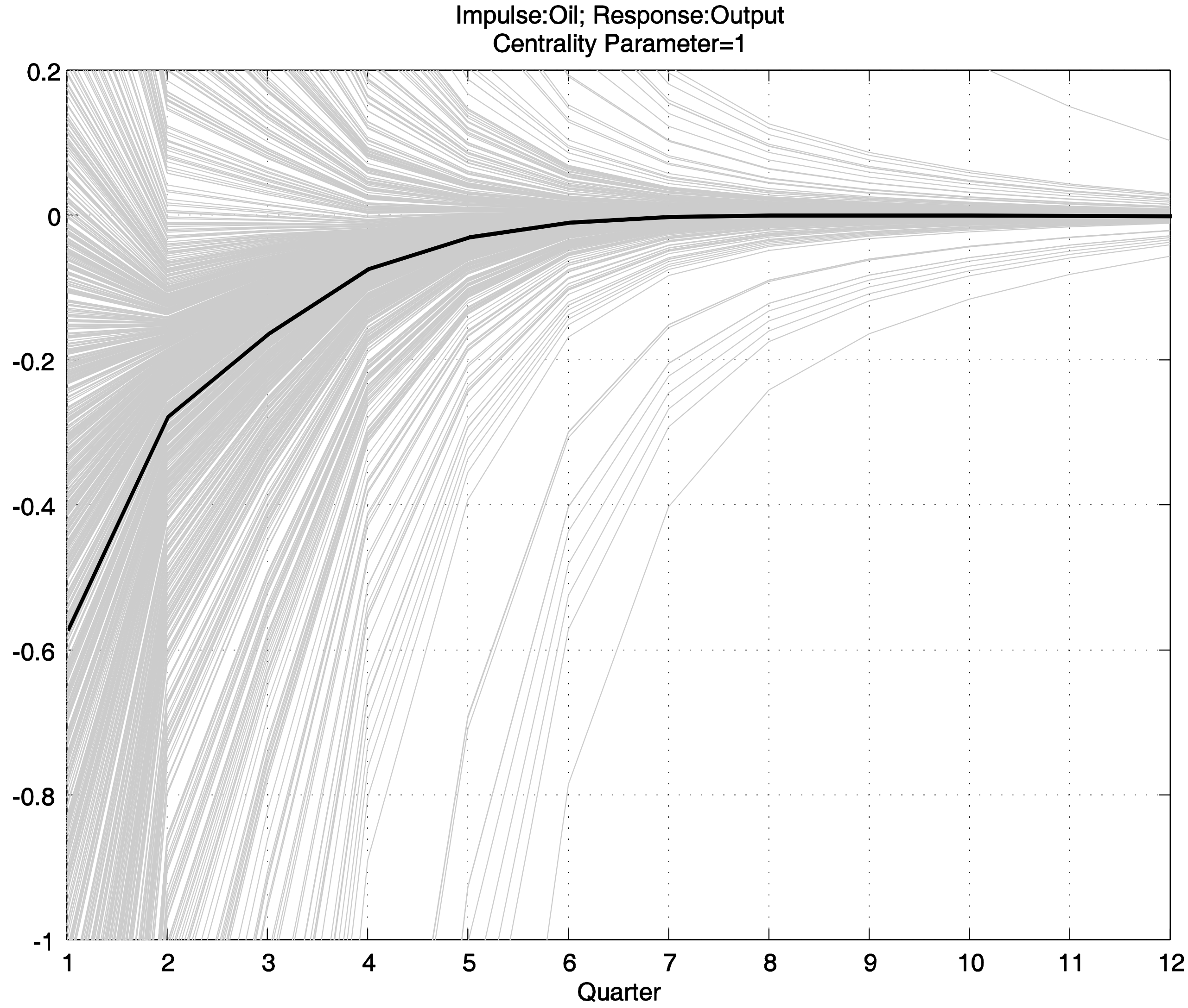
Effect of oil on oil growth: $\lambda_1^2 = 50$



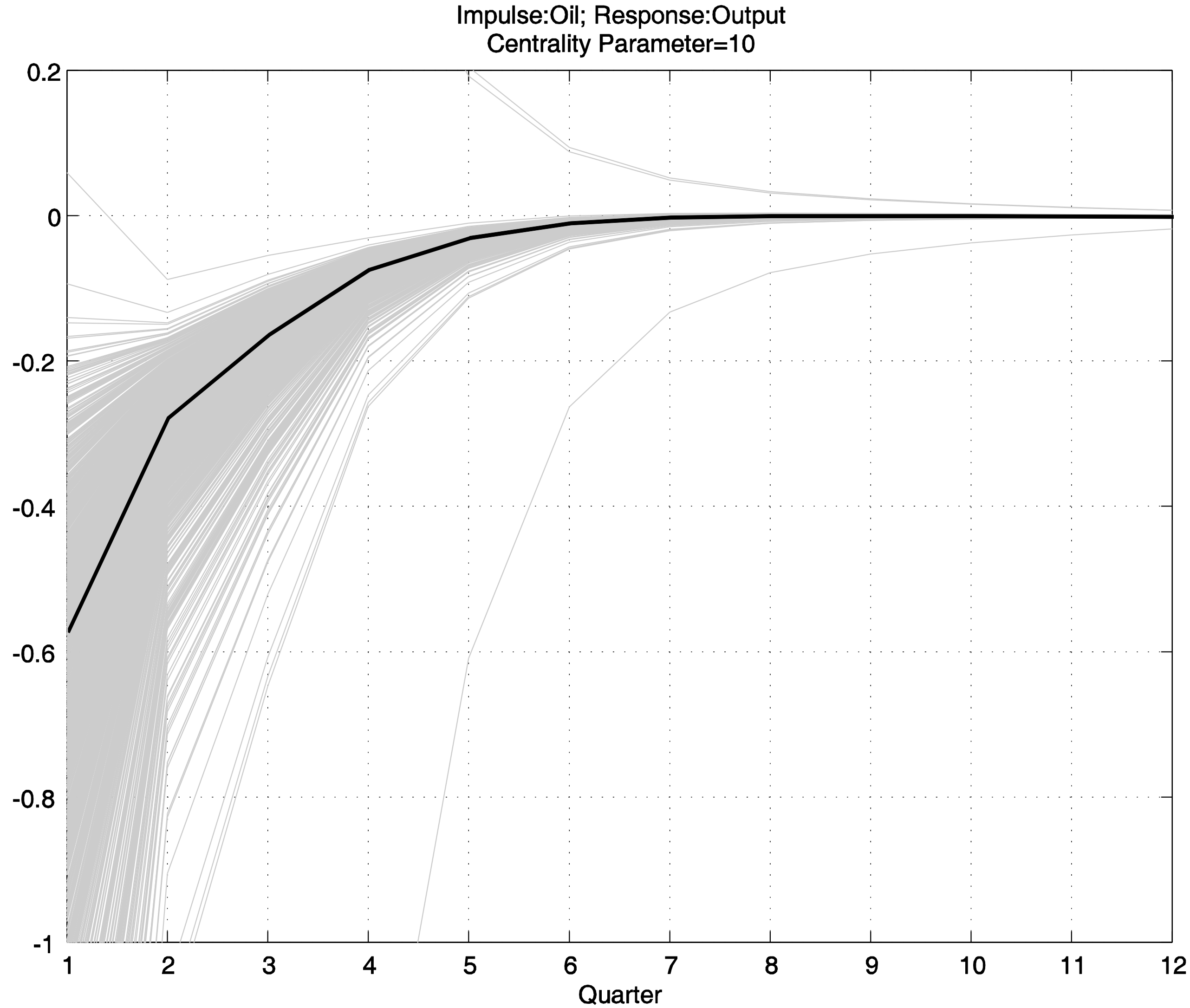
Effect of oil on oil growth: $\lambda_1^2 = 100$



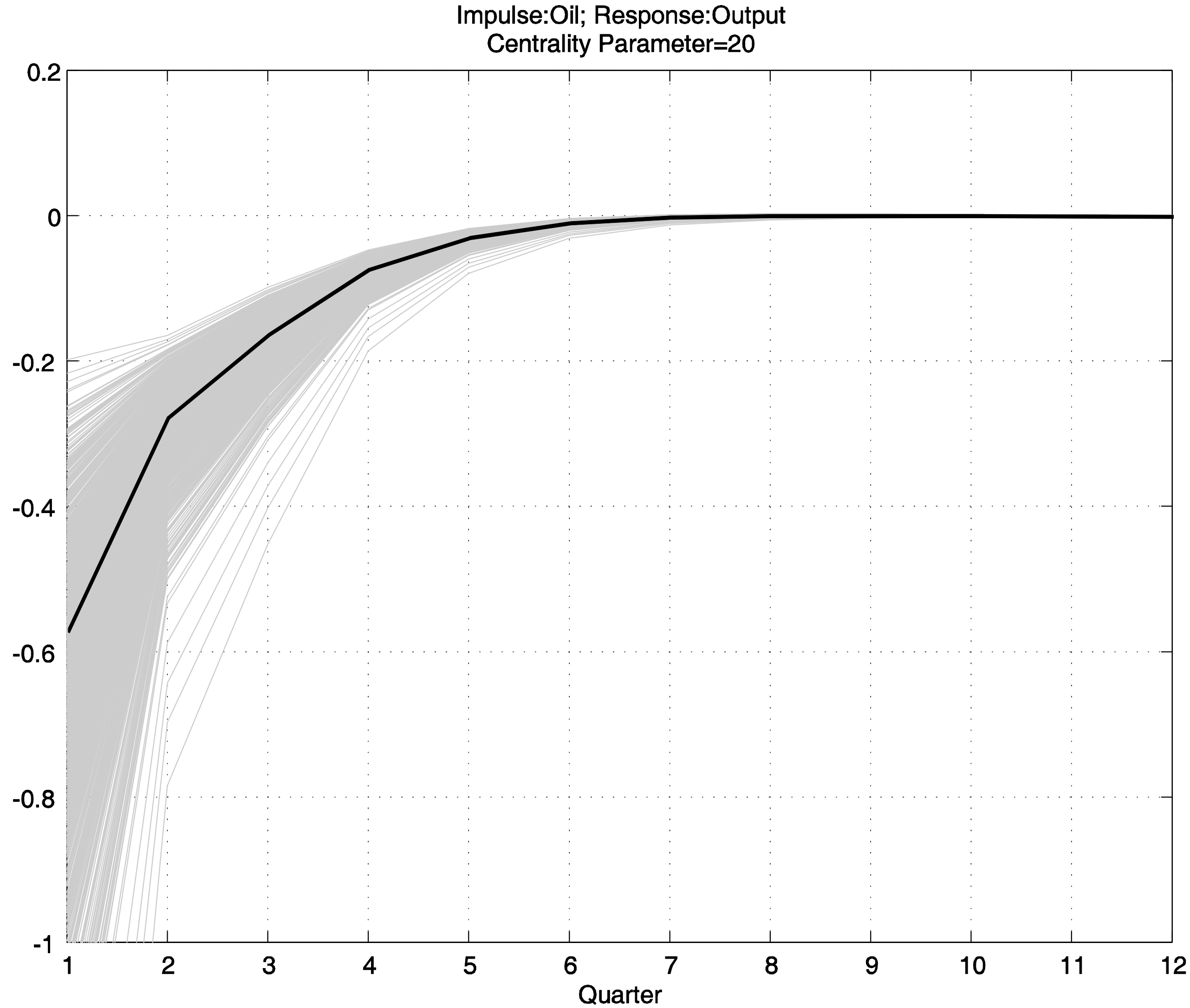
Effect of oil on oil growth: $\lambda_1^2 = 1000$



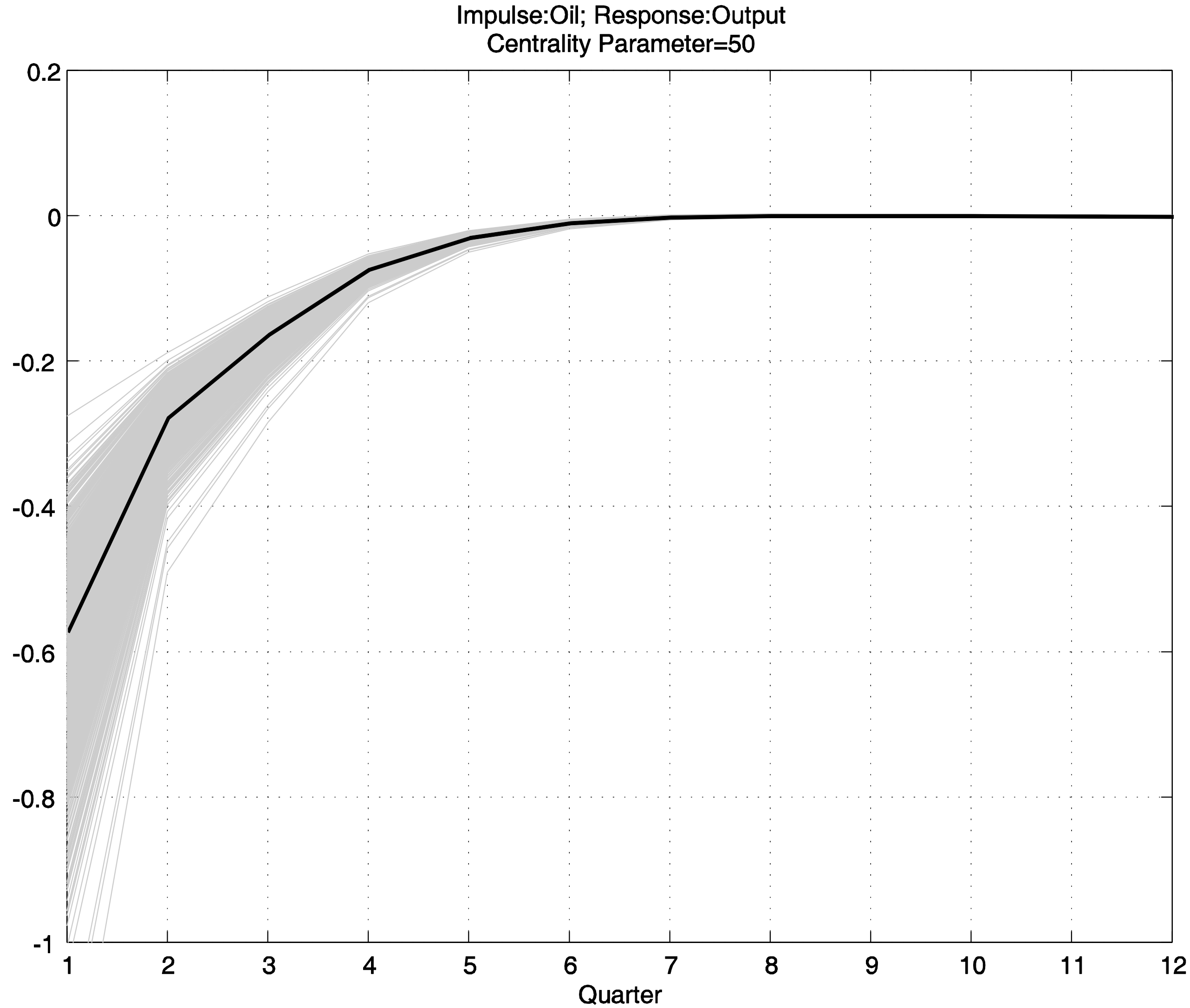
Effect of oil on GDP growth: $\lambda_1^2 = 1$



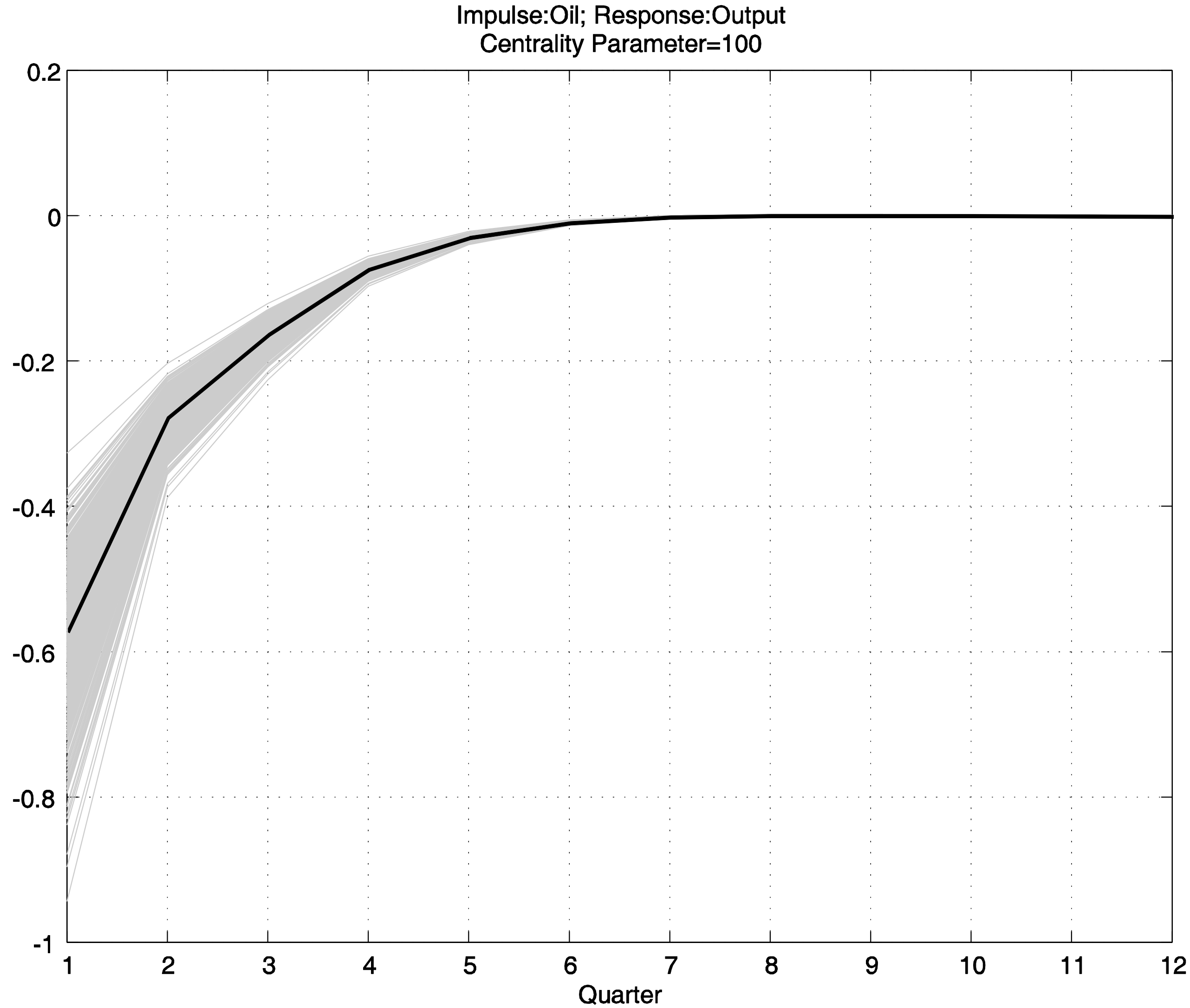
Effect of oil on GDP growth: $\lambda_1^2 = 10$



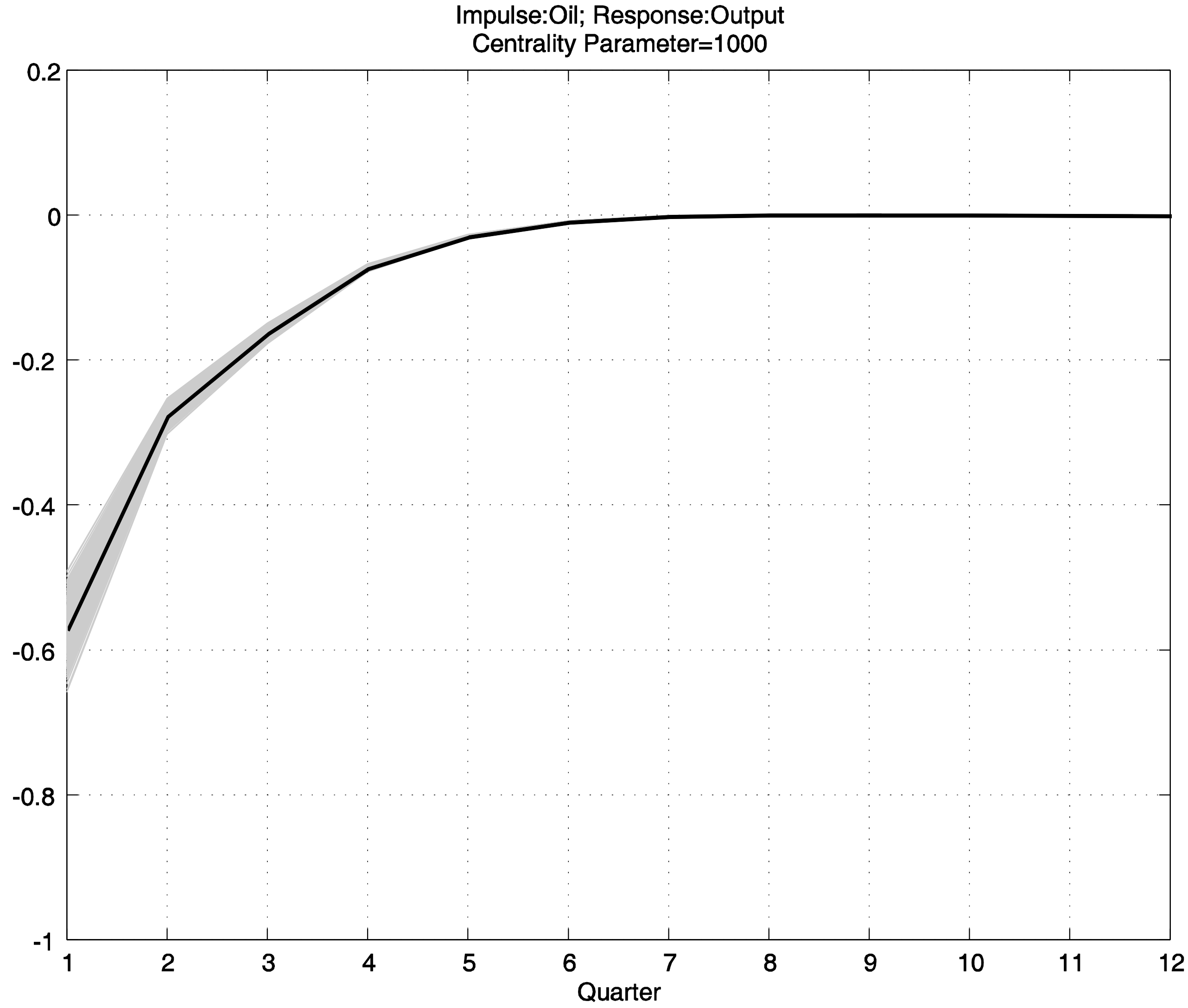
Effect of oil on GDP growth: $\lambda_1^2 = 20$



Effect of oil on GDP growth: $\lambda_1^2 = 50$



Effect of oil on GDP growth: $\lambda_1^2 = 100$



Effect of oil on GDP growth: $\lambda_1^2 = 1000$

Weak instrument asymptotics for cross-shock correlation

Correlation between two identified shocks:

Let Z_{1t} and Z_{2t} be scalar instruments that identify ε_{1t} and ε_{2t}:

$$\hat{\varepsilon}_{1t} = \left(T^{-1} \sum_{t=1}^{T} Z_{1t} \hat{\eta}_t \right) \hat{\Sigma}_{\eta\eta}^{-1} \eta_t$$

$$\hat{\varepsilon}_{2t} = \left(T^{-1} \sum_{t=1}^{T} Z_{2t} \hat{\eta}_t \right) \hat{\Sigma}_{\eta\eta}^{-1} \eta_t$$

$$r_{12} = \frac{T^{-1} \sum \hat{\varepsilon}_{1t} \hat{\varepsilon}_{2t}}{\sqrt{T^{-1} \sum \hat{\varepsilon}_{1t}^2} \sqrt{T^{-1} \sum \hat{\varepsilon}_{2t}^2}}$$

What is the null distribution (when (i)-(ii) hold for both instruments and $\Sigma_{\varepsilon\varepsilon} = I$)?

Weak instrument asymptotics for cross-shock correlation, ctd.

Expression for no-HAC case: $\Omega = \sigma_Z^2\Sigma_{\eta\eta}$, so

$$T^{-1/2}\sum Z_{1t}\hat{\eta}_t = T^{-1/2}\sum Z_{1t}\eta_t + o_p(1) \xrightarrow{d} \text{N}(0, \sigma_Z^2\Sigma_{\eta\eta})$$

so $$r_{12} = \frac{T^{-1}\sum \hat{\varepsilon}_{1t}\hat{\varepsilon}_{2t}}{\sqrt{T^{-1}\sum \hat{\varepsilon}_{1t}^2}\sqrt{T^{-1}\sum \hat{\varepsilon}_{2t}^2}}$$

$$= \frac{\left(T^{-1/2}\sum Z_{1t}\hat{\eta}_t\right)'\hat{\Sigma}_{\eta\eta}^{-1}\left(T^{-1/2}\sum \hat{\eta}_t Z_{2t}\right)}{\sqrt{\left(T^{-1/2}\sum Z_{1t}\hat{\eta}_t\right)'\hat{\Sigma}_{\eta\eta}^{-1}\left(T^{-1/2}\sum \hat{\eta}_t Z_{1t}\right)}\sqrt{\left(T^{-1/2}\sum Z_{2t}\hat{\eta}_t\right)'\hat{\Sigma}_{\eta\eta}^{-1}\left(T^{-1/2}\sum \hat{\eta}_t Z_{2t}\right)}}$$

$$\Rightarrow \frac{(\gamma_1+\zeta_1)'(\gamma_2+\zeta_2)}{\sqrt{(\gamma_1+\zeta_1)'(\gamma_1+\zeta_1)}\sqrt{(\gamma_2+\zeta_2)'(\gamma_2+\zeta_2)}}$$

Function of noncentral Wishart r.v.s (Anderson & Girshick (1944))

Weak instrument asymptotics for cross-shock correlation, ctd.

$$r_{12} \Rightarrow \frac{(\gamma_1+\zeta_1)'(\gamma_2+\zeta_2)}{\sqrt{(\gamma_1+\zeta_1)'(\gamma_1+\zeta_1)}\sqrt{(\gamma_2+\zeta_2)'(\gamma_2+\zeta_2)}}$$

where $$\varsigma = \begin{pmatrix}\zeta_1 \\ \zeta_2\end{pmatrix} \sim \mathrm{N}(0,\bar{\Sigma}\otimes \mathrm{I}),\ \ \bar{\Sigma} = \begin{bmatrix} 1 & corr(Z_1,Z_2) \\ corr(Z_1,Z_2) & 1 \end{bmatrix}$$

$$\gamma_1'\gamma_1 = a_1^2/\sigma_{\varepsilon_1}^2\sigma_{Z_1}^2,\ \gamma_2'\gamma_2 = a_2^2/\sigma_{\varepsilon_2}^2\sigma_{Z_2}^2$$

$$\gamma_1'\gamma_2 = 0 \ \text{ under (i) – (iii)}$$

<u>Comments</u>

1. Nonstandard distribution – function of noncentral Wishart rvs
2. Normal under null as $\gamma_1'\gamma_1$ and $\gamma_2'\gamma_2 \to \infty$
3. Strong instruments under alternative: $r_{12} \xrightarrow{p} \dfrac{\gamma_1'\gamma_2}{\sqrt{\gamma_1'\gamma_1}\sqrt{\gamma_2'\gamma_2}}$

Weak instrument asymptotics for cross-shock correlation, ctd.

Numerical results

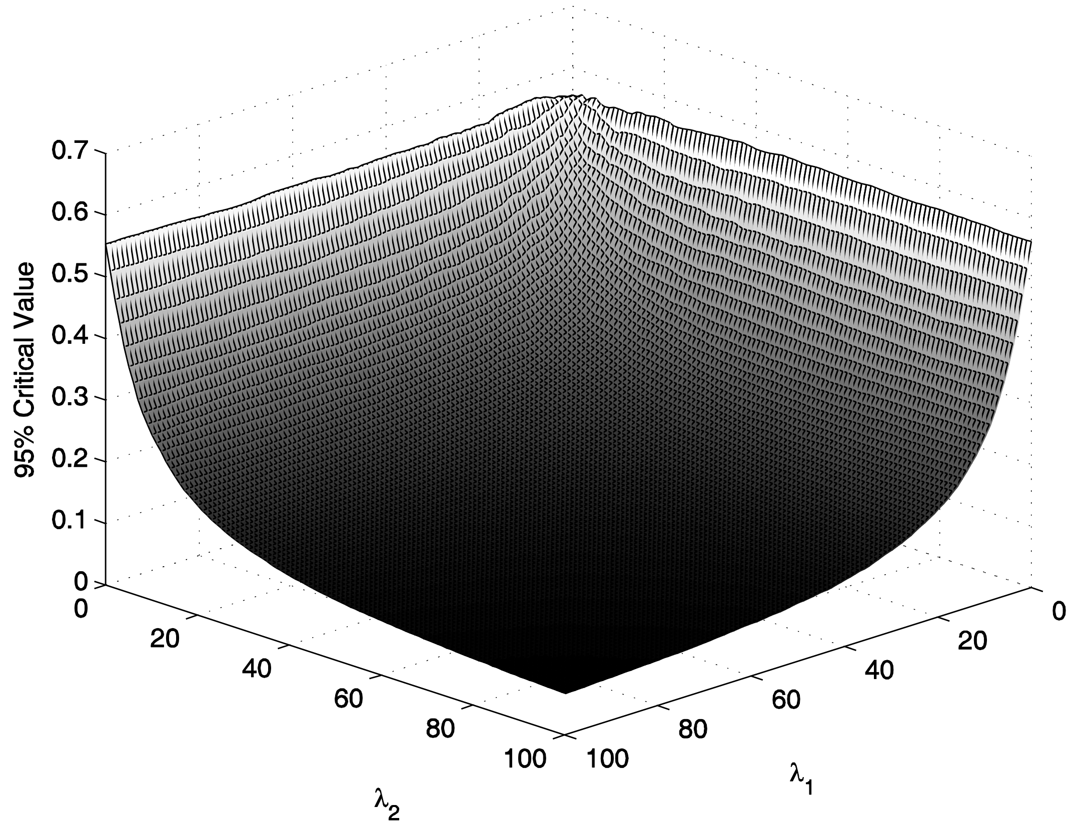
Asymptotic null distribution is a function of

$$\gamma_1'\gamma_1 = a_1^2 / \sigma_{\varepsilon_1}^2 \sigma_{Z_1}^2,$$

$$\gamma_2'\gamma_2 = a_2^2 / \sigma_{\varepsilon_2}^2 \sigma_{Z_2}^2$$

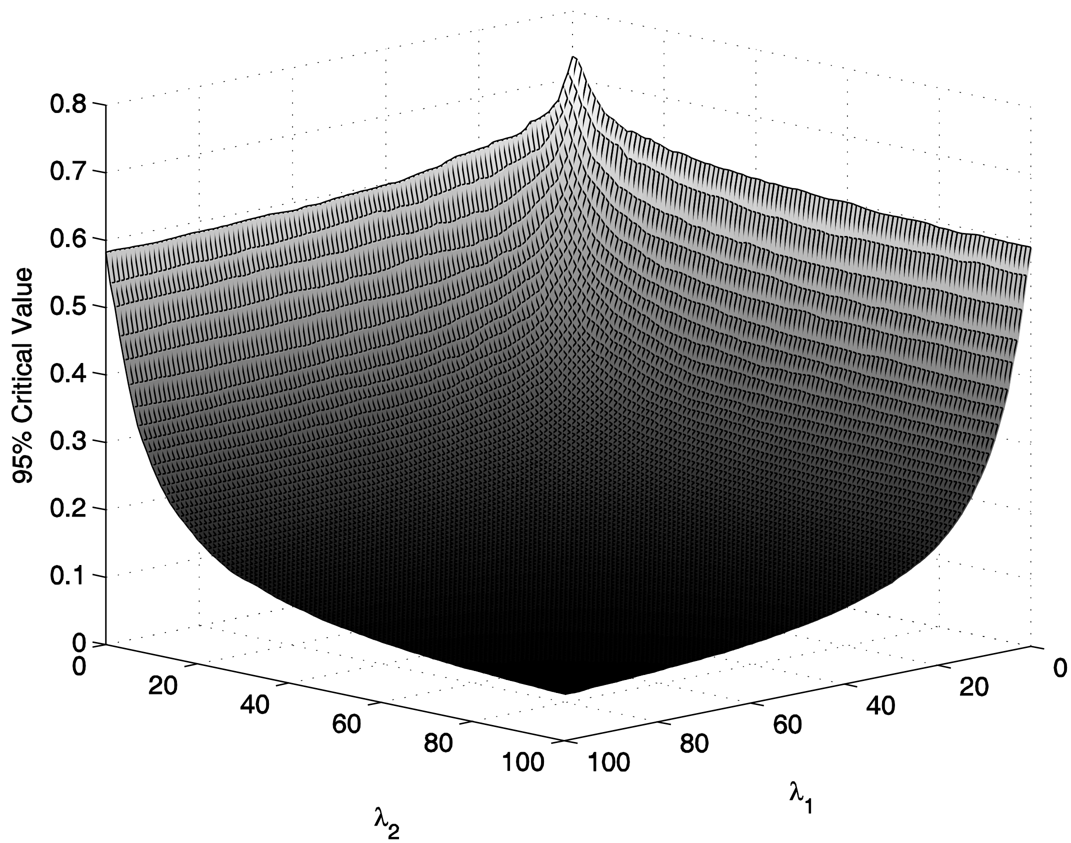
$$\text{corr}(Z_1, Z_2)$$

Weak instrument asymptotics for cross-shock correlation, ctd.



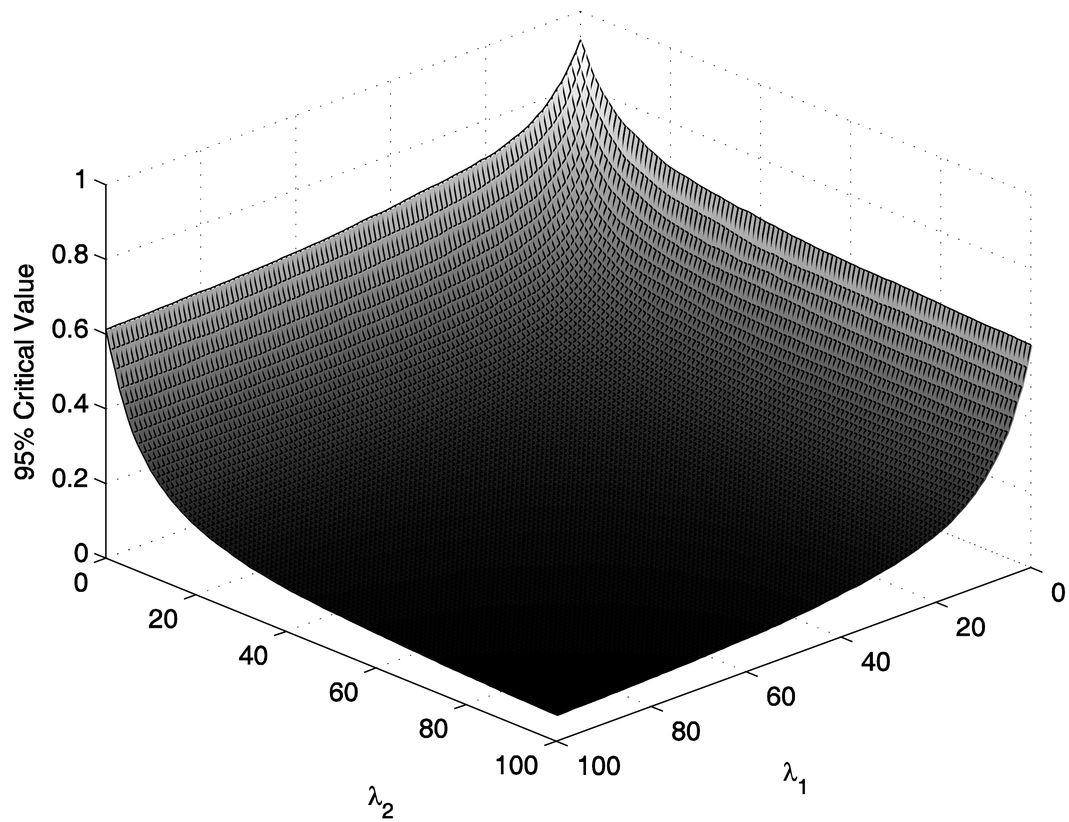
Weak instrument asymptotic null distribution of r_{12}: $|\text{Corr}(Z_1,Z_2)| = 0$

Weak instrument asymptotics for cross-shock correlation, ctd.



Weak instrument asymptotic null distribution of r_{12}: $|\mathrm{Corr}(Z_1,Z_2)| = 0.4$

Weak instrument asymptotics for cross-shock correlation, ctd.



Weak instrument asymptotic null distribution of r_{12}: $|\text{Corr}(Z_1,Z_2)| = 0.8$

Weak instrument asymptotics for cross-shock correlation, ctd.

Sup critical values (worst case over $\gamma_1'\gamma_1$ and $\gamma_2'\gamma_2$):

$\|\text{corr}(Z_1, Z_2)\|$	**95 % critical value**
0	.5705
.2	.6253
.4	.7327
.6	.8406
.8	.9231

6. Weak-instrument robust inference for structural IRFs

$$IRF = C_{h,j}H_1$$

Consider null hypothesis $C_{h,j}H_1 = \kappa_0$ and a single Z.

Use (iv) to write the null as,

$$C_{h,j}H_1 = (C_{h,j1} \quad C_{h,j\bullet})H_1 = C_{h,j1} + C_{h,j\bullet}H_{1\bullet} = \kappa_0$$

or

$$C_{h,j\bullet}H_{1\bullet} = \kappa_0 - C_{h,j1}$$

Recall moment restriction:

$$H_{1\bullet}E(\eta_{1t}Z_t) - E(\eta_{\bullet t}Z_t) = 0$$

so

$$C_{h,j\bullet}H_{1\bullet}E(\eta_{1t}Z_t) - C_{h,j\bullet}E(\eta_{\bullet t}Z_t) = 0$$

Thus under the null,

$$(\kappa_0 - C_{h,j1})E(\eta_{1t}Z_t) - C_{h,j\bullet}E(\eta_{\bullet t}Z_t) = 0$$

Weak-instrument robust inference for IRFs, ctd

Under null that $IRF = C_{h,j}H_1 = \kappa_0$,

$$(\kappa_0 - C_{h,j1})E(\eta_{1t}Z_t) - C_{h,j\bullet}E(\eta_{\bullet t}Z_t) = 0$$

or

$$E\gamma_0{}'\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\eta_t Z_t = 0 \text{ where } \gamma_0 = \begin{pmatrix} \kappa_0 - C_{h,j1} \\ -C_{h,j\bullet}{}' \end{pmatrix}$$

Test: reject κ_0 if $\left(\gamma_0{}'\frac{1}{\sqrt{T}}\sum_{t=1}^{T}\eta_t Z_t\right)^2 \Big/ \gamma_0{}'\Omega\gamma_0 > \chi^2_{1;.95}$

Note: Under weak instrument nesting, C(L) is known

Weak-instrument robust inference for IRFs, ctd

$$\left(\gamma_0' \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \eta_t Z_t\right)^2 \Big/ \gamma_0' \Omega \gamma_0 > \chi^2_{1;.95}$$

Comments

- This is one degree of freedom test (not r-1 d.f. AR set for $H_{1\bullet}$)
- Conf. int. inversion can be done analytically (ratio of quadratics)
- Strong-instrument efficient (asy equivalent to standard GMM test)
- Scalar Z: this test is UMPU in limit experiment using the sufficient statistic $\frac{1}{\sqrt{T}} \sum_{t=1}^{T} \eta_t Z_t$ in sense of Müller (2011) (proof: rotate so that you are testing mean of first element of independent normal), so confidence intervals are (limit experiment) UMAU

Weak-instrument robust inference for IRFs, ctd

- Multiple Z: The testing problem of H_0: $\kappa = \kappa_0$ can be rewritten as H_0: $\beta = \beta_0$ in the standard IV regression form,

$$C_{h,j\bullet}\eta_{\bullet t} - (\kappa_0 - C_{h,j1})\eta_{1t} = \beta_0\, \eta_{1t} + u_t$$
$$\eta_{1t} = \pi Z_t + v_t$$

so for multiple Z_t the CLR confidence interval can be used.
(Working on efficiency improvements)

7. Inference for Historical Decompositions

$$HD = \sum_{k=0}^{h} C_{k,j} H_1 \varepsilon_{1t-j} = \left(\sum_{k=0}^{h} C_{k,j1} \varepsilon_{1t-k} \right) + \left(\sum_{k=0}^{h} C_{k,j\bullet} \varepsilon_{1t-k} \right) H_{1\bullet}$$

Treat $\varepsilon_{1t},\ldots,\varepsilon_{t\text{-}h}$ as nonrandom, and C(L) as known. Then this is also testing a linear combination of $H_{1\bullet}$ so the approach for IRFs applies directly.

Test: reject $\tilde{\kappa}_0$ if $\left(\tilde{\gamma}_0{}' \frac{1}{\sqrt{T}} \sum_{t=1}^{T} \eta_t Z_t \right)^2 \Big/ \tilde{\gamma}_0{}' \Omega \tilde{\gamma}_0 > \chi^2_{1;.95}$

where

$$\tilde{\gamma}_0 = \begin{pmatrix} \tilde{\kappa}_0 - \sum_{k=0}^{h} C_{k,j1} \varepsilon_{1t-k} \\ -\sum_{k=0}^{h} C_{k,j\bullet} \varepsilon_{1t-k} \end{pmatrix}$$

8. Extensions

8.1 When Z_t is serially correlated

Let $\hat{Z}_t$ = residual from regression of Z_t onto $\underline{Y}_{t-1}$

and $\zeta_t = Z_t - \text{Proj}(Z_t|\underline{Y}_{t-1})$

$$
\begin{aligned}
T^{-1/2}\sum_{t=1}^{T}\hat{\eta}_t Z_t &= T^{-1/2}\sum_{t=1}^{T}\eta_t \hat{Z}_t \\
&= T^{-1/2}\sum_{t=1}^{T}\eta_t (Z_t - \underline{Y}_{t-1}\hat{\Sigma}_{Y_{-1}Y_{-1}}^{-1}\hat{\Sigma}_{Y_{-1}Z}^{-1}) \\
&= T^{-1/2}\sum_{t=1}^{T}\eta_t (Z_t - \underline{Y}_{t-1}\Sigma_{Y_{-1}Y_{-1}}^{-1}\Sigma_{Y_{-1}Z}^{-1}) + o_p(1) \\
&= T^{-1/2}\sum_{t=1}^{T}\eta_t \zeta_t \\
&\xrightarrow{d} \text{N}(\text{H}_1\alpha', \Omega),
\end{aligned}
$$

where $\Omega = 2\pi S_{\eta\zeta}(0)$. Under the no-HAC assumption, $\Omega = \Sigma_{\eta\eta}\sigma_{\zeta}^2$ so all goes through as above with ζ_t replacing Z_t

8.2 When Z_t is a generated instrument

- For example, Z_t is the residual from a preliminary regression
- Additional adjustment to the variance formula

8.3 Dynamic Factor Models

Dynamic factor model (Geweke (1977), Sargent & Sims (1977)):

$X_t = \Lambda F_t + e_t$ (F_t = 6 factors, e_t = idiosyncratic disturbance)

$A(\mathrm{L})F_t = \eta_t$ (factors follow a reduced form VAR)

$\eta_t = H\varepsilon_t$, H invertible (same as in SVAR setup)

Moving average representations:

$X_t = \Lambda A(\mathrm{L})^{-1}\eta_t + e_t$ (reduced form)

$X_t = \Lambda A(\mathrm{L})^{-1}H\varepsilon_t + e_t$ (S-DFM, MA form)

Extension to DFMs, ctd.

$$X_t = \Lambda A(L)^{-1}H\varepsilon_t + e_t \qquad \text{(S-DFM, MA form)}$$

IRF of variable j with respect to shock 1: $\Lambda_j{}'C(L)H_1$

Extension of foregoing results to S-DFM requires:

- Estimation of F_t's (e.g. principal components);
- no "generated regressor" problem under Bai-Ng (2006) conditions
- Modification for normalization condition (iv): ε_{1t} has positive unit impact effect on X_{jt}: because C_0= I,

$$\text{(iv}'\text{)} \qquad \Lambda_j{}'H_1 = 1$$

- If you renormalize F_t so that Λ is lower triangular on r variables with "variable 1 first" then the foregoing formulas apply directly (no modifications)

9. Empirical Results

Empirical framework

Dynamic factor model:

$X_t = \Lambda F_t + e_t$ (F_t = 6 factors, e_t = idiosyncratic disturbance)

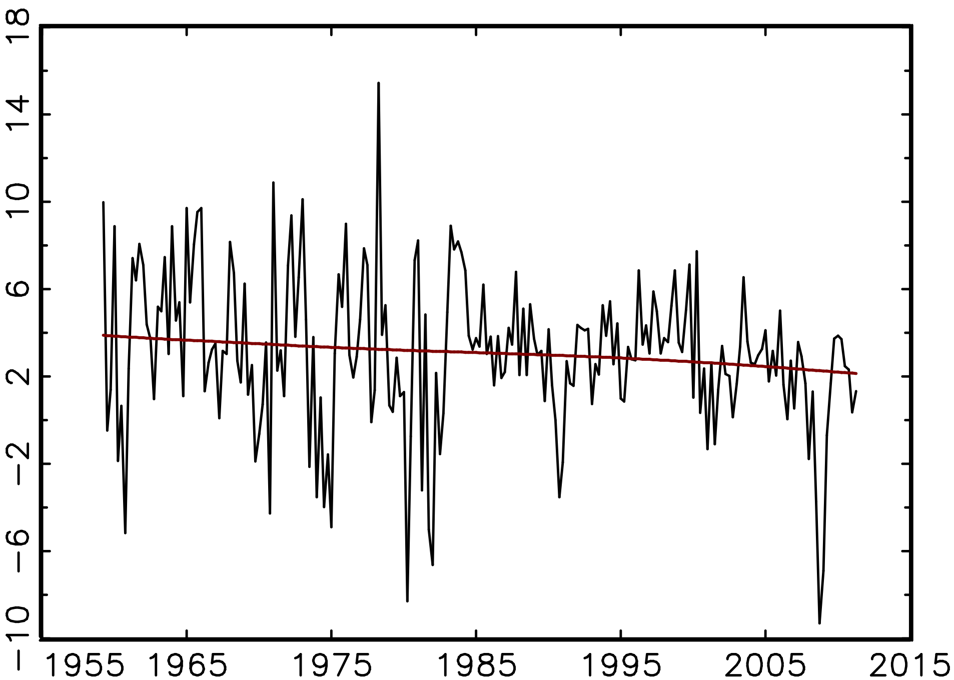
$\Phi(\mathrm{L})F_t = \eta_t$ (factors follow a VAR)

Notes:

- Large *n* beneficial for estimation of factor space
- Only 132 series are used to estimate factors (disaggregates only)
- Estimate F_t by principal components, then treat F_t as data
- Factor space is identified, factors aren't: $\Lambda F_t = \Lambda \mathrm{H}\mathrm{H}^{-1} F_t$

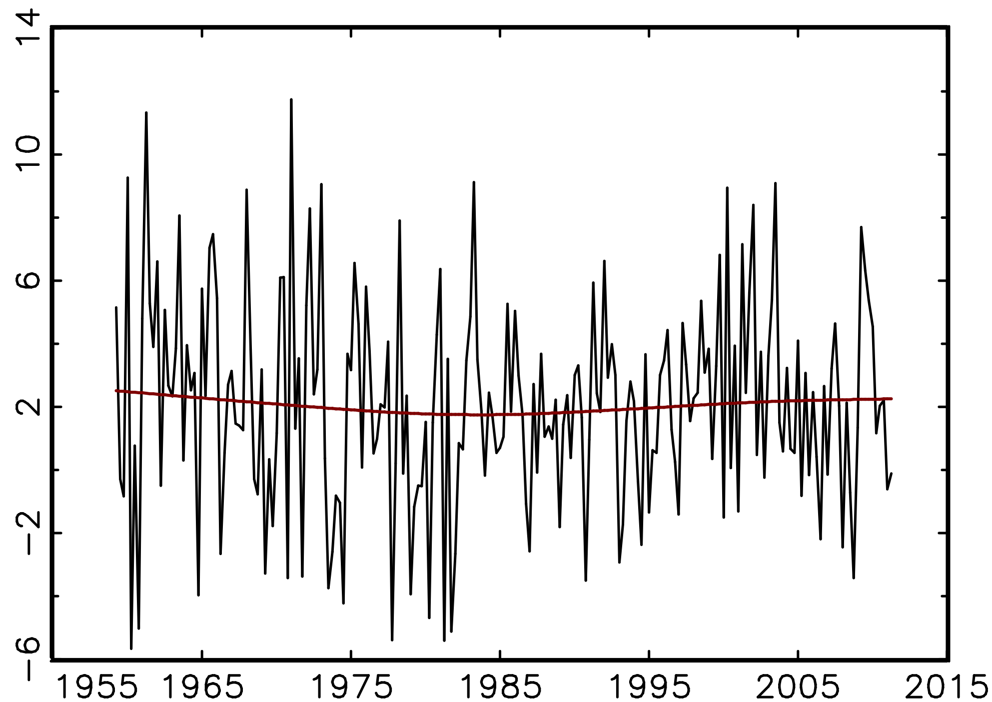
Data

- U.S., quarterly, 1959-2011Q2, 200 time series
- Almost all series analyzed in changes or growth rates
- All series detrended by local demeaning – approximately 15 year centered moving average:



Quarterly GDP growth (a.r.)

Trend: 3.7% → 2.5%



Quarterly productivity growth

2.3% → 1.8% → 2.2%

Instruments

1. Oil Shocks
   - a. Hamilton (2003) net oil price increases
   - b. Killian (2008) OPEC supply shortfalls
   - c. Ramey-Vine (2010) innovations in adjusted gasoline prices
2. Monetary Policy
   - a. Romer and Romer (2004) policy
   - b. Smets-Wouters (2007) monetary policy shock
   - c. Sims-Zha (2007) MS-VAR-based shock
   - d. Gürkaynak, Sack, and Swanson (2005), FF futures market
3. Productivity
   - a. Fernald (2009) adjusted productivity
   - b. Gali (200x) long-run shock to labor productivity
   - c. Smets-Wouters (2007) productivity shock

Instruments, ctd.

4. Uncertainty
   - a. VIX/Bloom (2009)
   - b. Baker, Bloom, and Davis (2009) Policy Uncertainty

5. Liquidity/risk
   - a. Spread: Gilchrist-Zakrajšek (2011) excess bond premium
   - b. Bank loan supply: Bassett, Chosak, Driscoll, Zakrajšek (2011)
   - c. TED Spread

6. Fiscal Policy
   - a. Ramey (2011) spending news
   - b. Fisher-Peters (2010) excess returns gov. defense contractors
   - c. Romer and Romer (2010) "all exogenous" tax changes.

"First stage": F_1: regression of Z_t on η_t, F_2: regression of η_{1t} on Z_t

Structural Shock	F_1	F_2
1. Oil		
Hamilton	2.9	**15.7**
Killian	1.1	1.6
Ramey-Vine	1.8	0.6
2. Monetary policy		
Romer and Romer	4.5	**21.4**
Smets-Wouters	9.0	5.3
Sims-Zha	6.5	**32.5**
GSS	0.6	0.1
3. Productivity		
Fernald TFP	**14.5**	**59.6**
Smets-Wouters	7.0	**32.3**

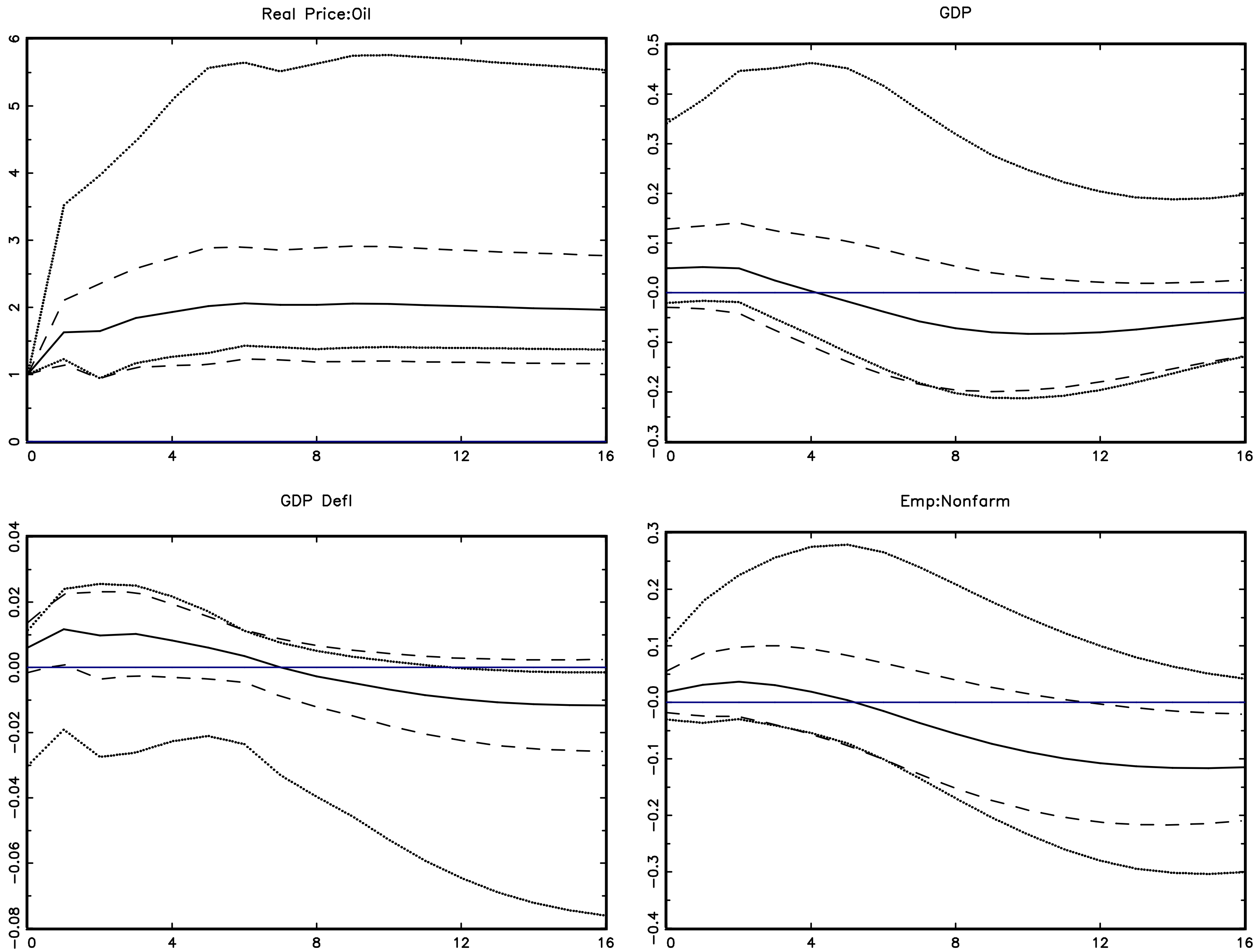
Structural Shock	F_1	F_2
4. Uncertainty		
Fin Unc (VIX)	**43.2**	**239.6**
Pol Unc (BBD)	**12.5**	73.1
5. Liquidity/risk		
GZ EBP Spread	4.5	**23.8**
TED Spread	**12.3**	**61.1**
BCDZ Bank Loan	4.4	4.2
6. Fiscal policy		
Ramey Spending	0.5	1.0
Fisher-Peters Spending	1.3	0.1
Romer-Romer Taxes	0.5	2.1

Correlations among selected structural shocks

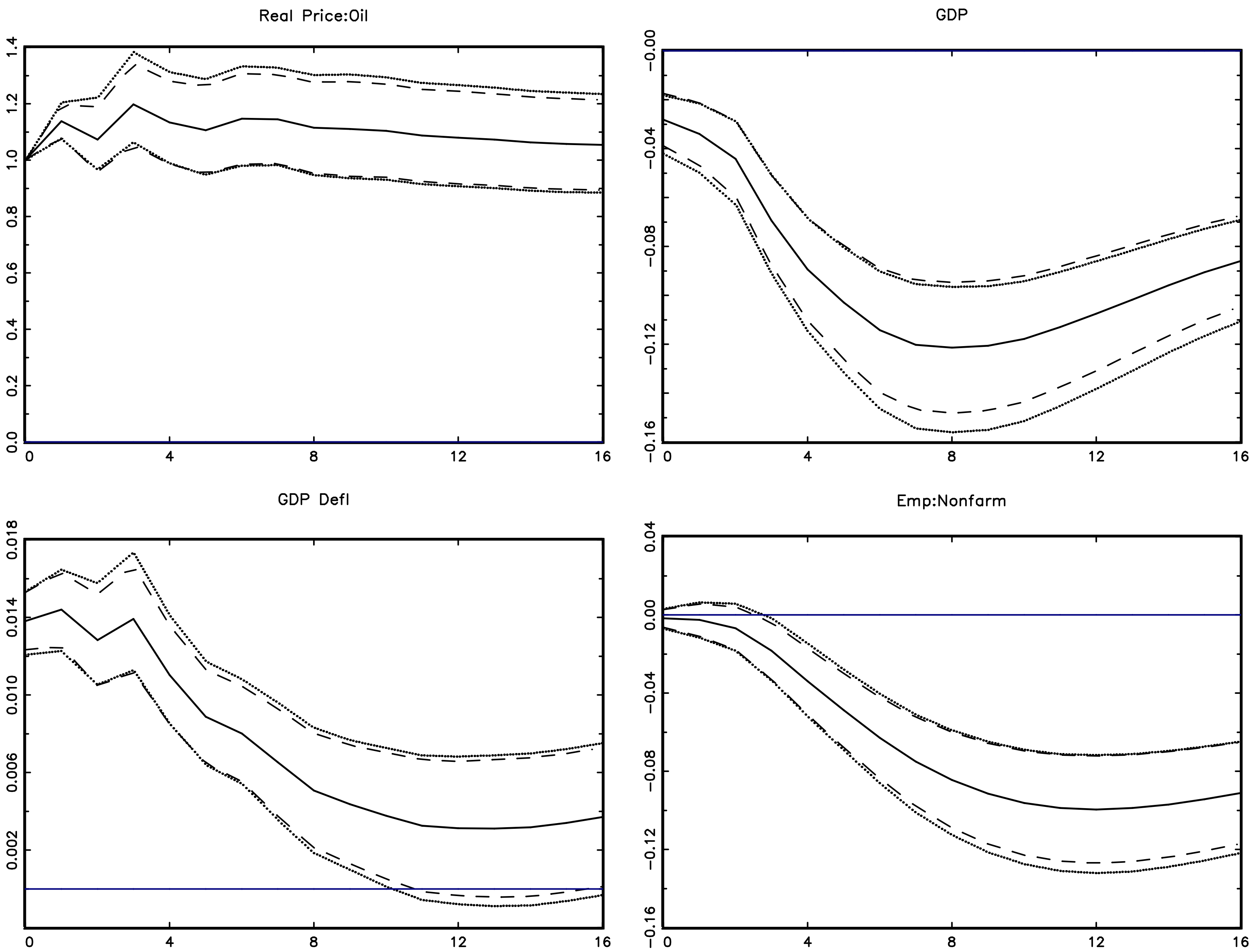
	O_K	M_{RR}	M_{SZ}	P_F	U_B	U_{BBD}	S_{GZ}	B_{BCDZ}	F_R	F_{RR}
O_K	1.00									
M_{RR}	**0.65**	1.00								
M_{SZ}	0.35	0.93	1.00							
P_F	0.30	0.20	0.06	1.00						
U_B	-0.37	-0.39	-0.29	0.19	1.00					
U_{BBD}	0.11	-0.17	-0.22	-0.06	0.78	1.00				
L_{GZ}	-0.42	-0.41	-0.24	0.07	**0.92**	**0.66**	1.00			
L_{BCDZ}	0.22	0.56	0.55	-0.09	**-0.69**	-0.54	-0.73	1.00		
F_R	**-0.64**	**-0.84**	**-0.72**	-0.17	0.26	-0.08	0.40	-0.13	1.00	
F_{RR}	0.15	**0.77**	**0.88**	0.18	0.01	-0.10	0.02	0.19	-0.45	1.00

Oil_{Kilian} oil – Kilian (2009)
M_{RR} monetary policy – Romer and Romer (2004)
M_{SZ} monetary policy – Sims-Zha (2006)
P_F productivity – Fernald (2009)
U_B Uncertainty – VIX/Bloom (2009)
U_{BBD} uncertainty (policy) – Baker, Bloom, and Davis (2012)
L_{GZ} liquidity/risk – Gilchrist-Zakrajšek (2011) excess bond premium
L_{BCDZ} liquidity/risk – BCDZ (2011) SLOOS shock
F_R fiscal policy – Ramey (2011) federal spending
F_{RR} fiscal policy – Romer-Romer (2010) federal tax

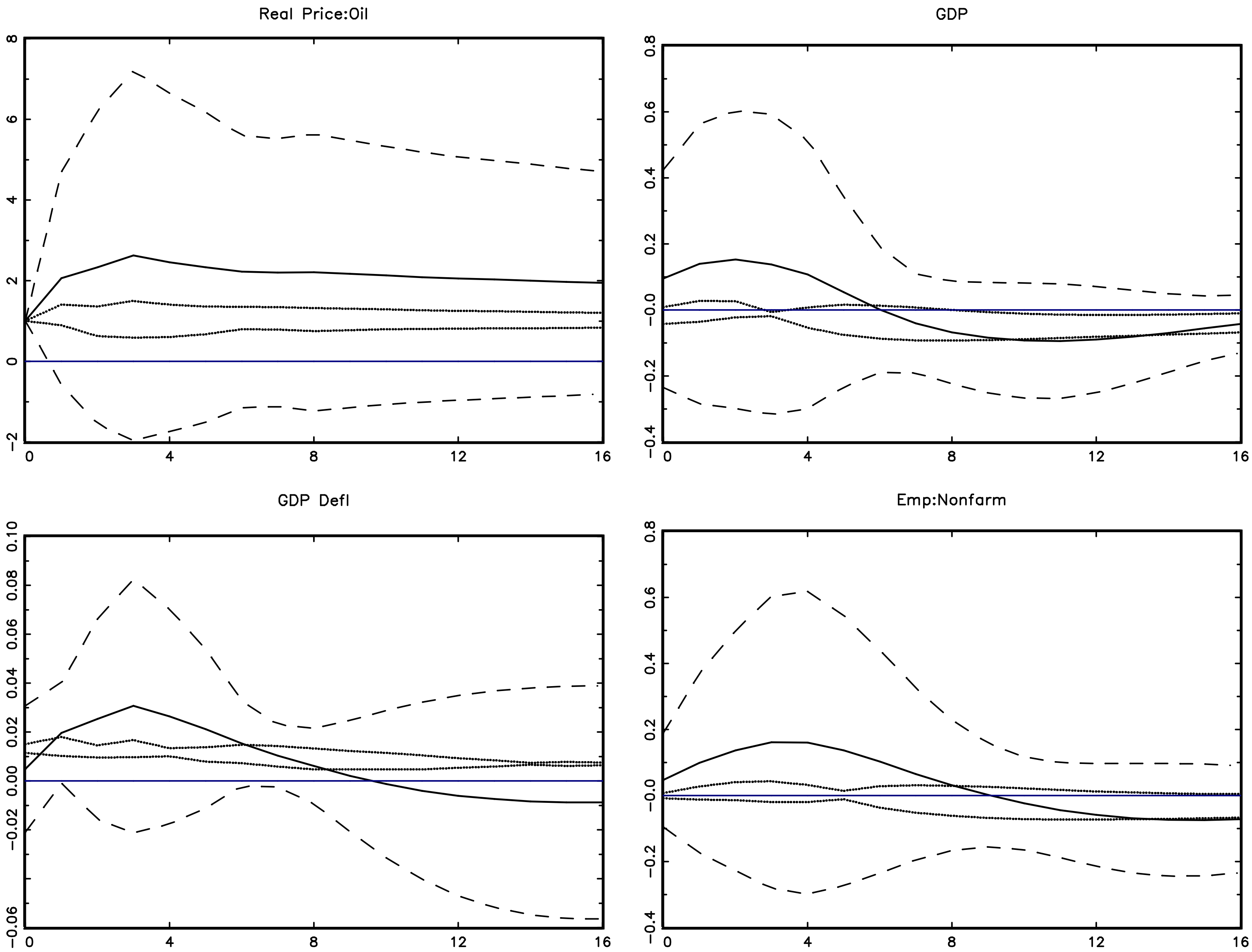
IRFs: strong-IV (dashed) and weak-IV robust (solid) pointwise bands



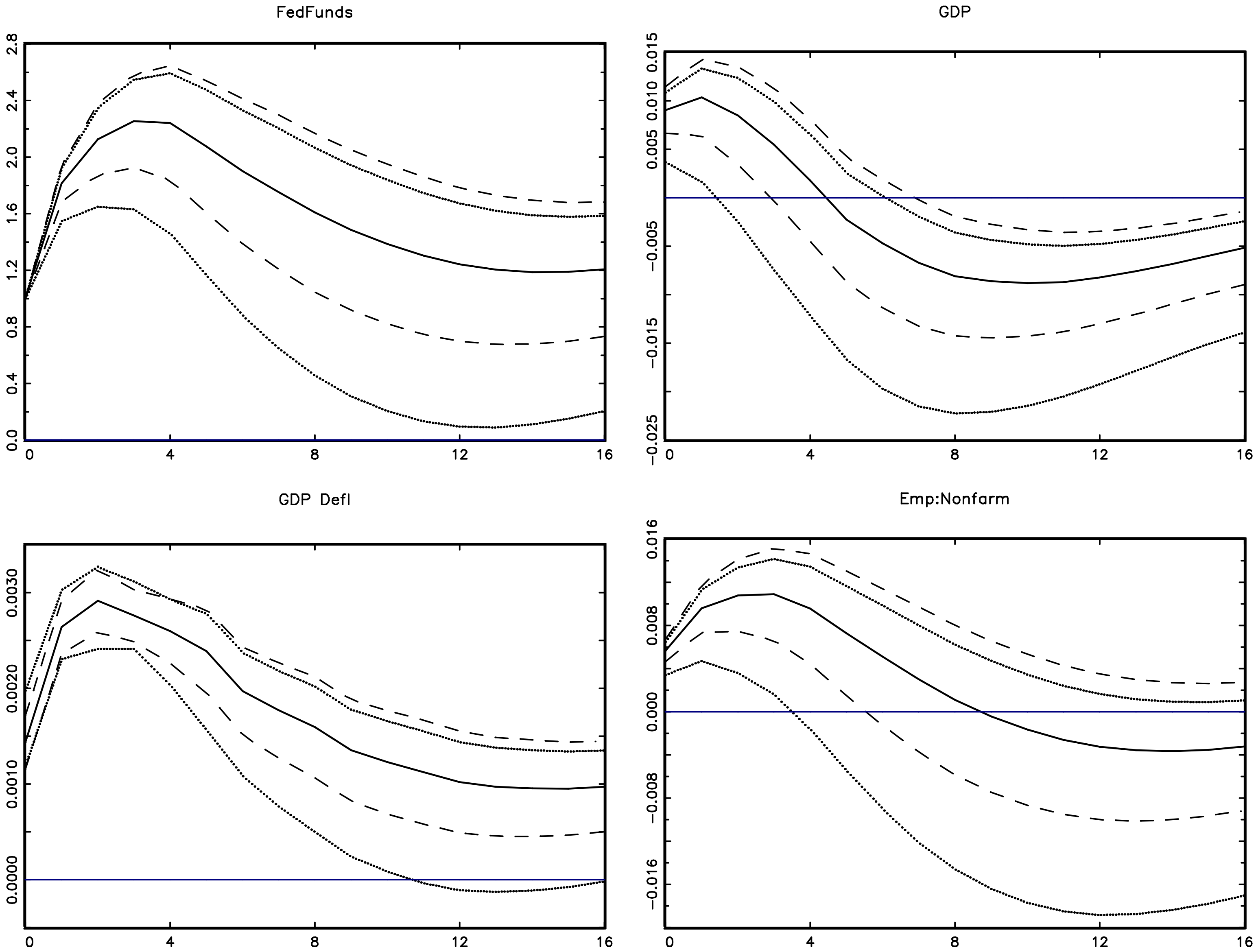
Kilian (2008) oil shock ($F_2 = 1.6$)



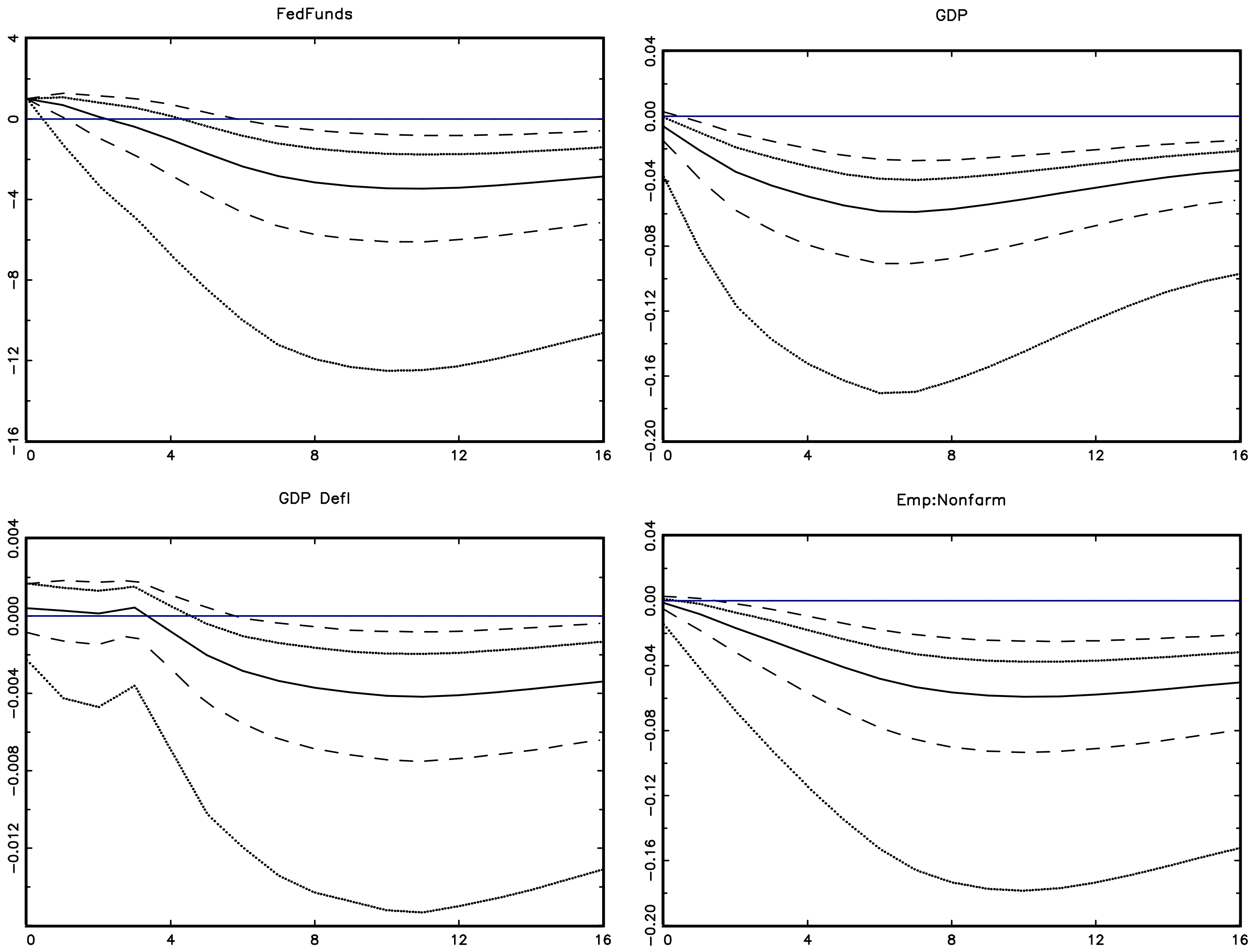
Hamilton (1996, 2003) oil shock (F_2 = 15.7)



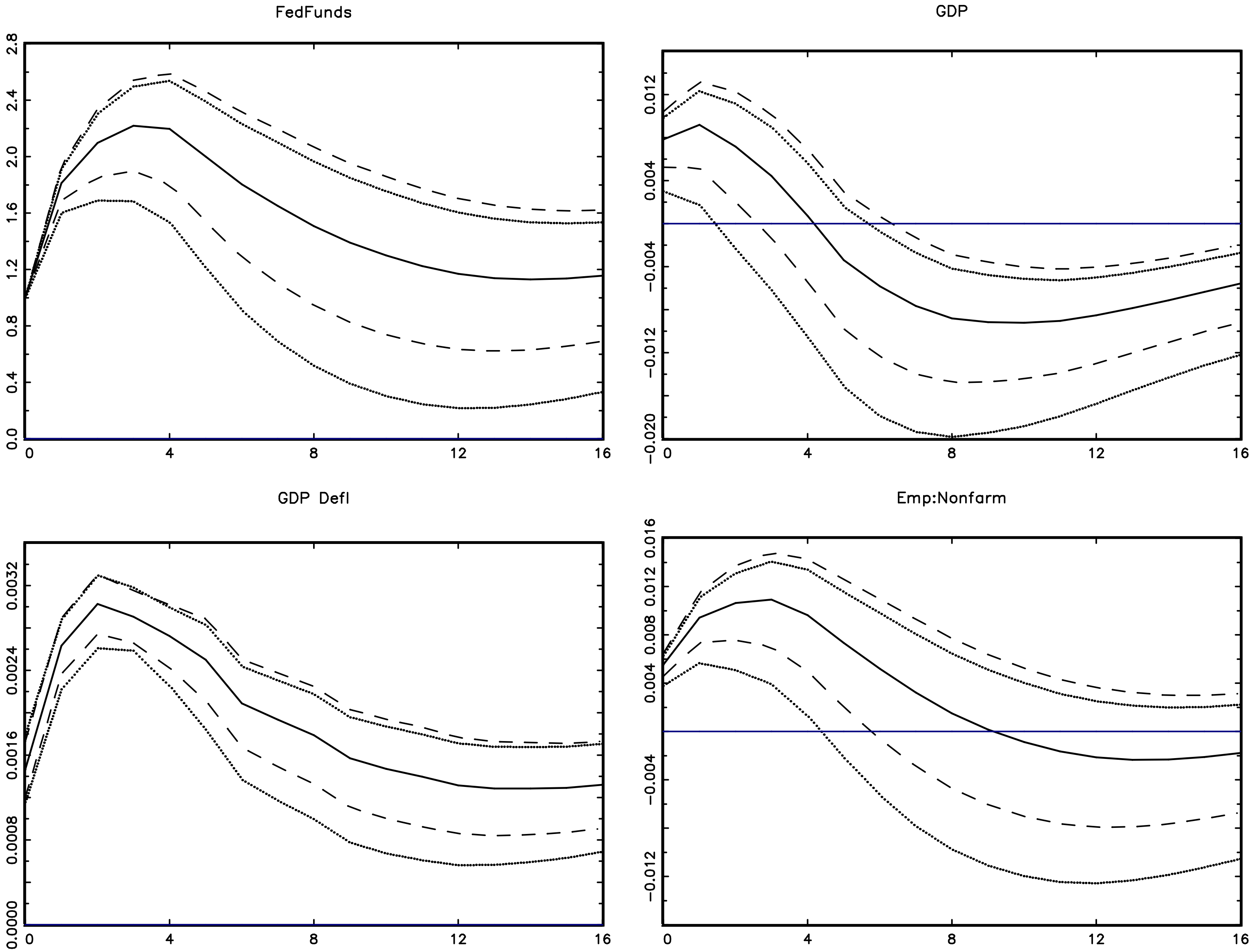
Ramey-Vine (2010) oil shock (F_2 = 0.6)



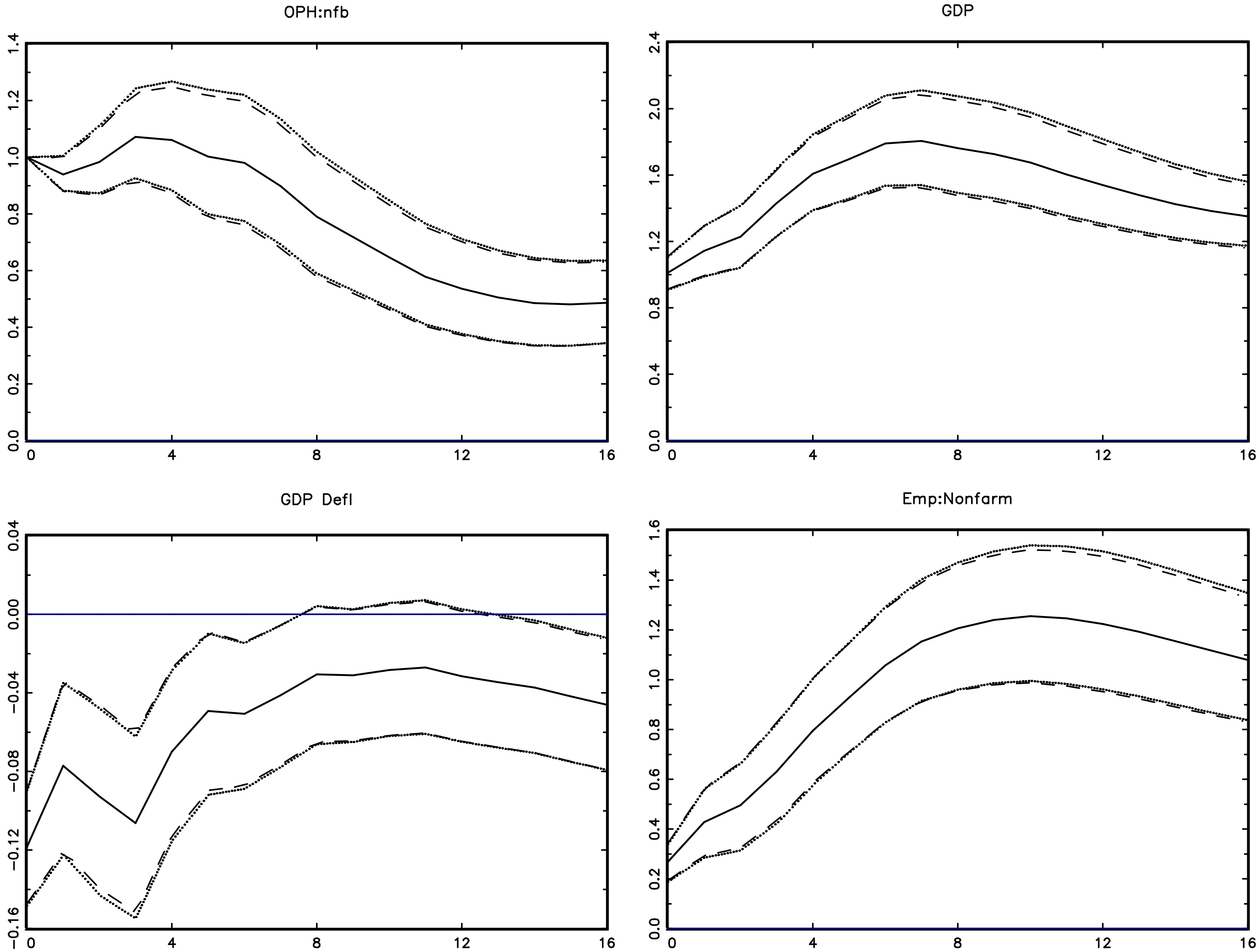
Romer and Romer (2004) monetary policy shock (F_2 = 21.4)



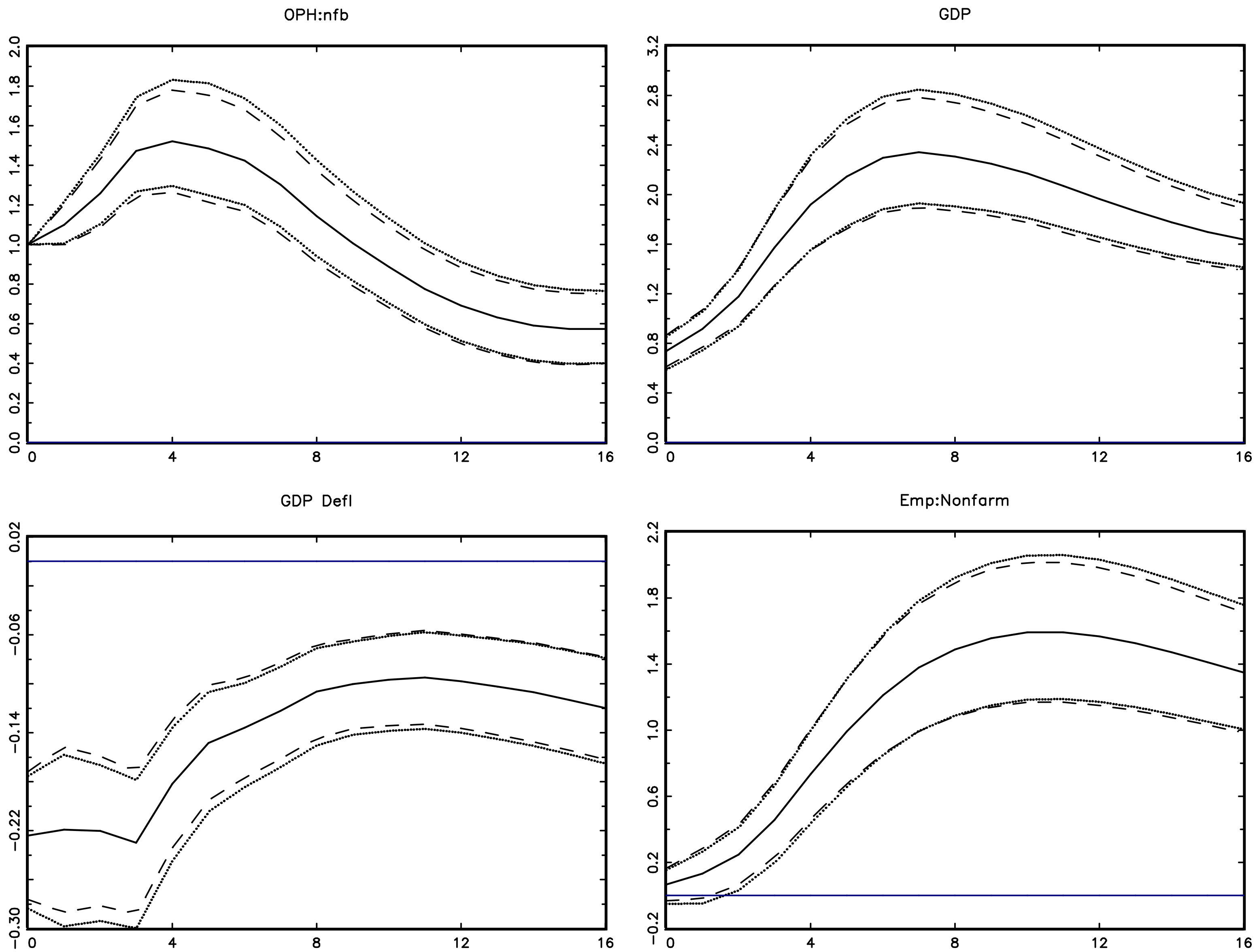
Smets-Wouters (2007) monetary policy shock ($F_2 = 5.3$)



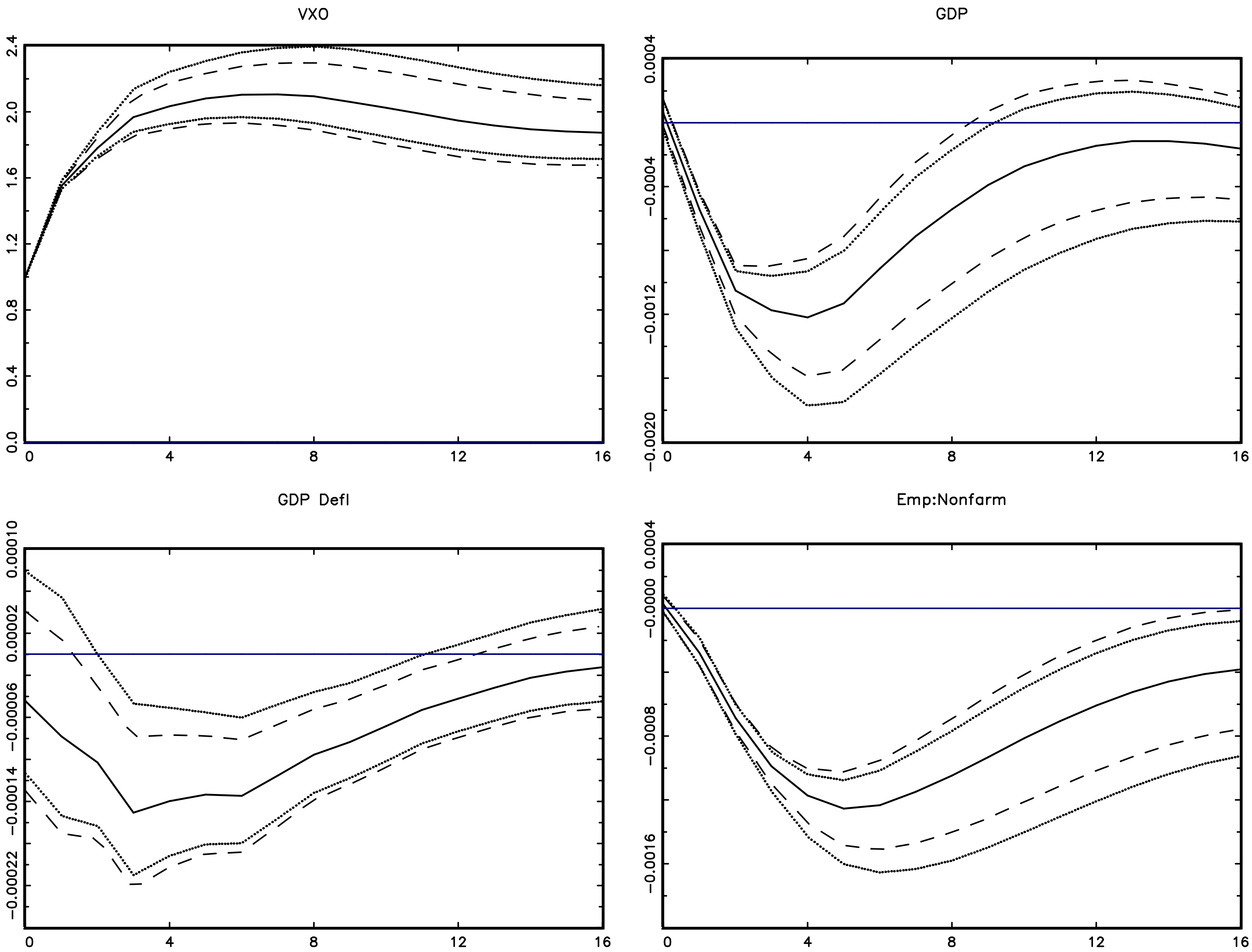
Sims-Zha (2006) monetary policy shock (F_2 = 32.5)



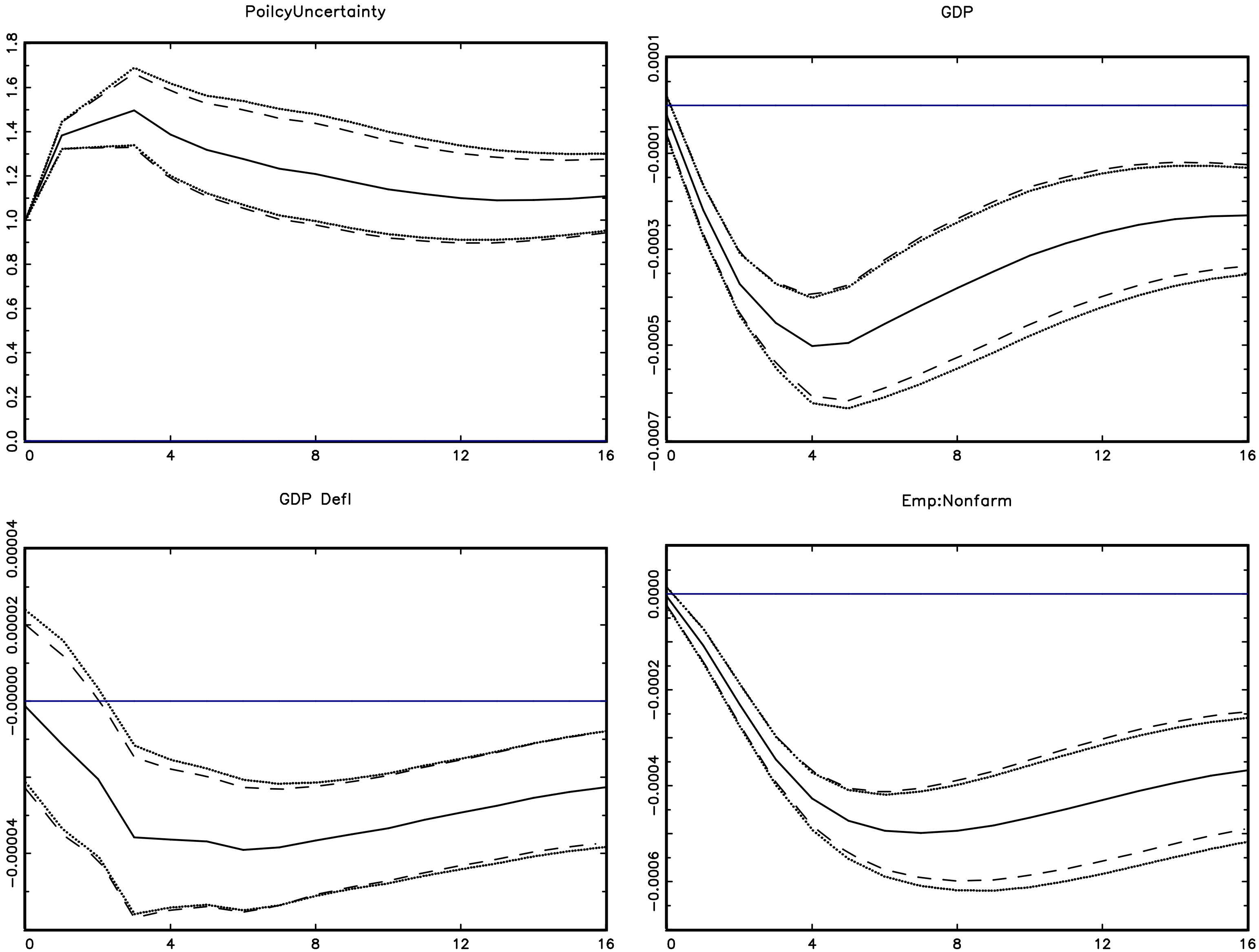
Fernald (2009) productivity shock (F_2 = 59.6)



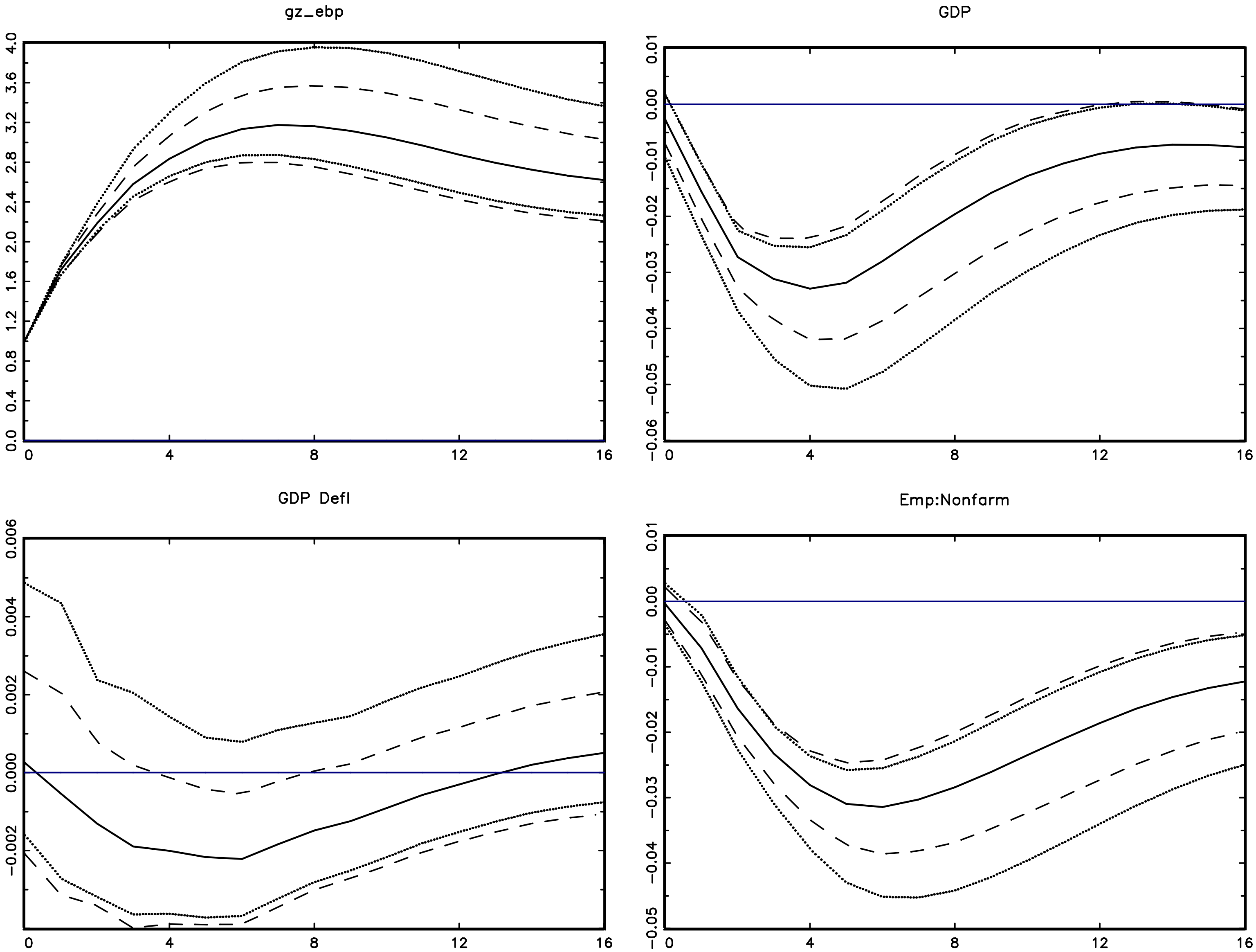
Smets-Wouters (2007) productivity shock (F_2 = 32.3)



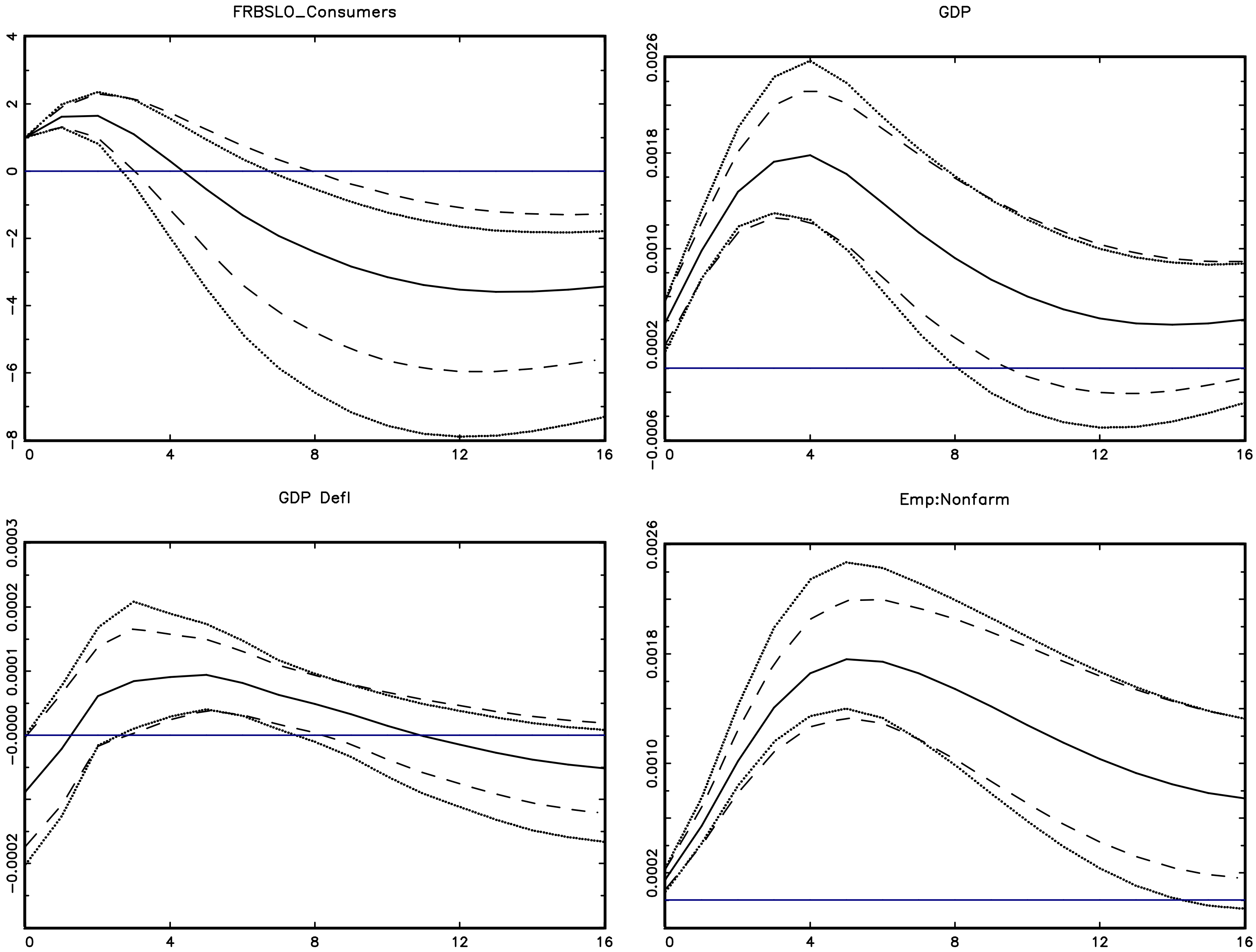
Bloom (2009) (VIX) uncertainty shock (F_2 = 239.6)



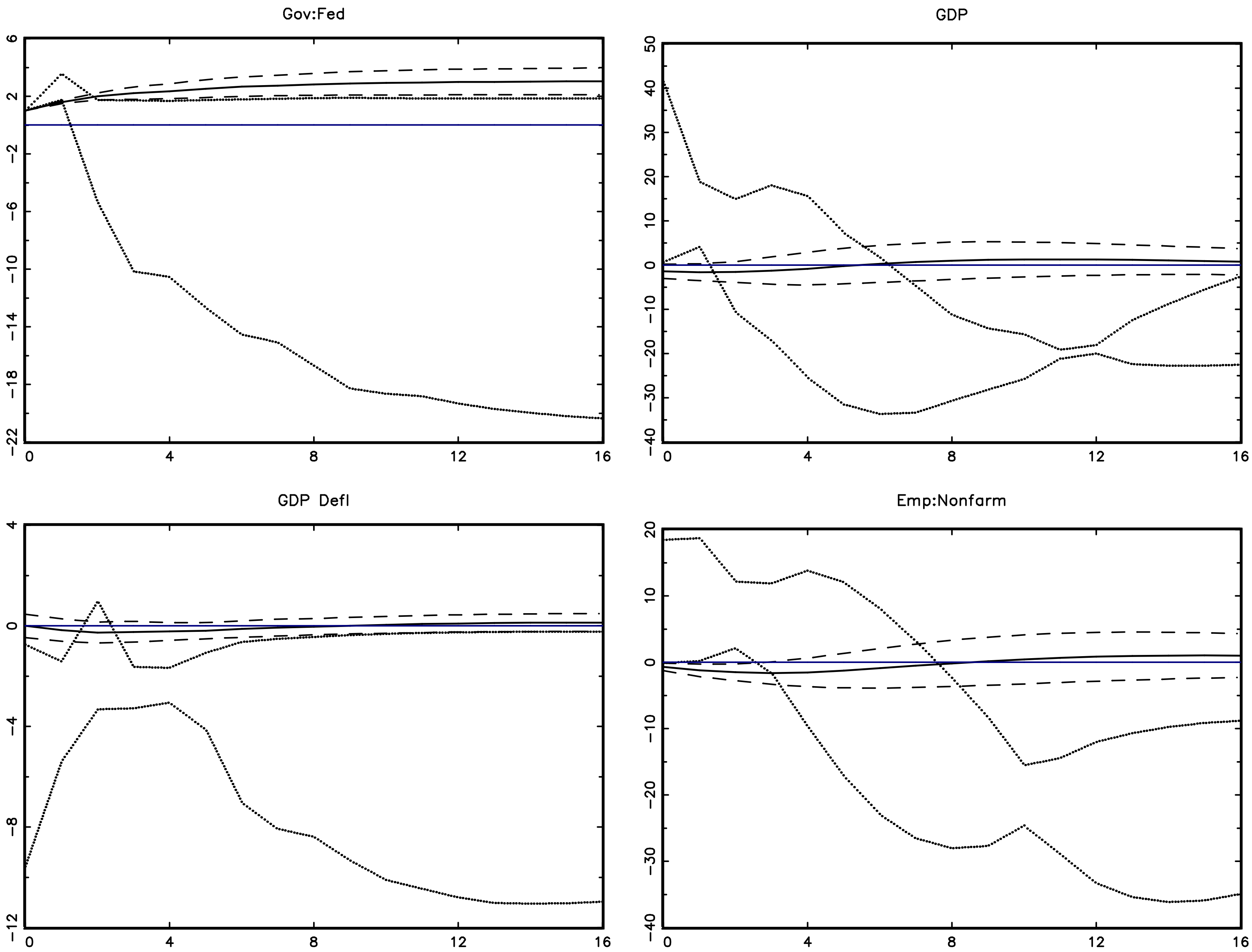
Baker, Bloom, Davis (2012) policy uncertainty shock (F_2 = 73.1)



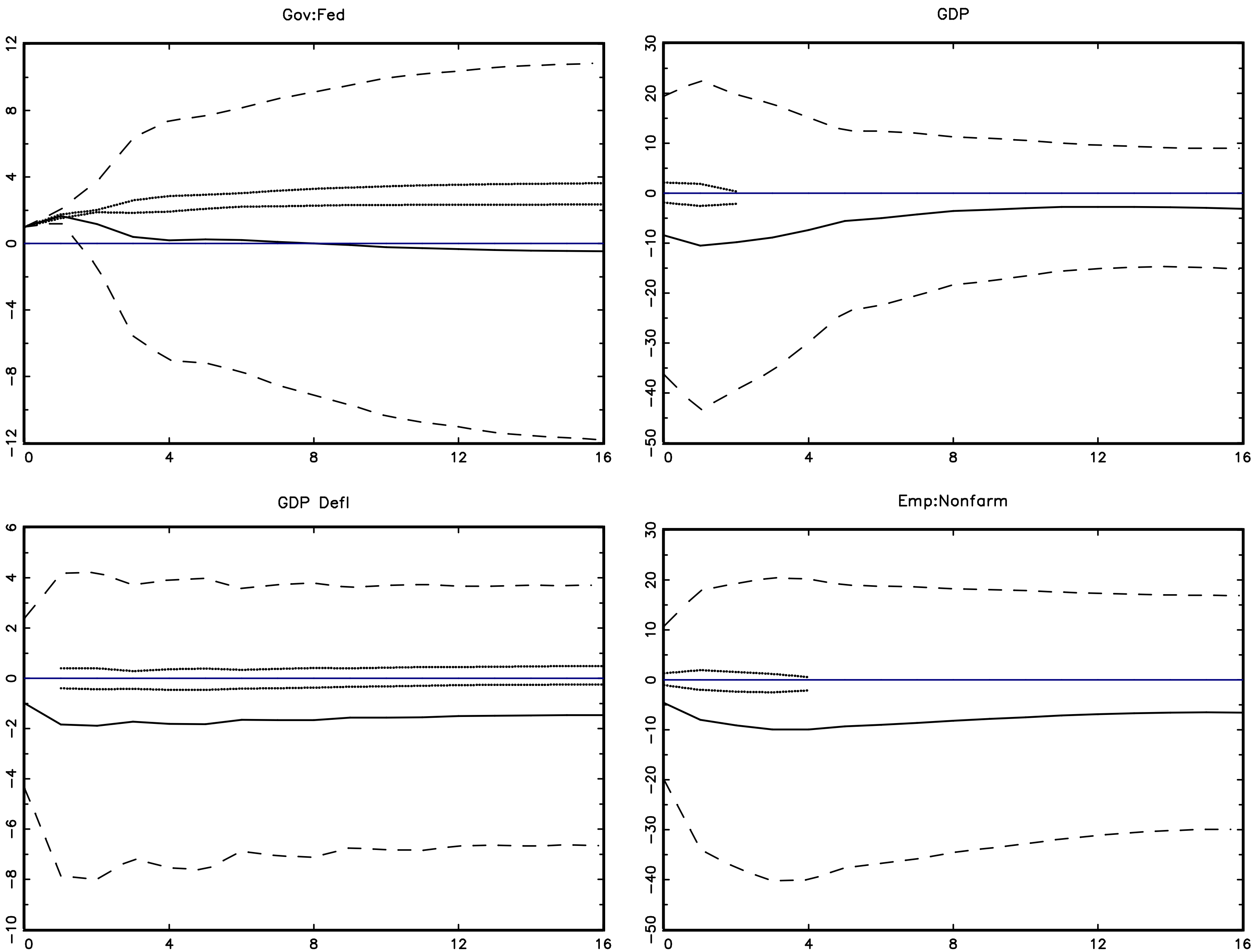
Gilchrist and Zakrajšek (2011) excess bond premium liquidity/risk shock (F_2 = 23.8)



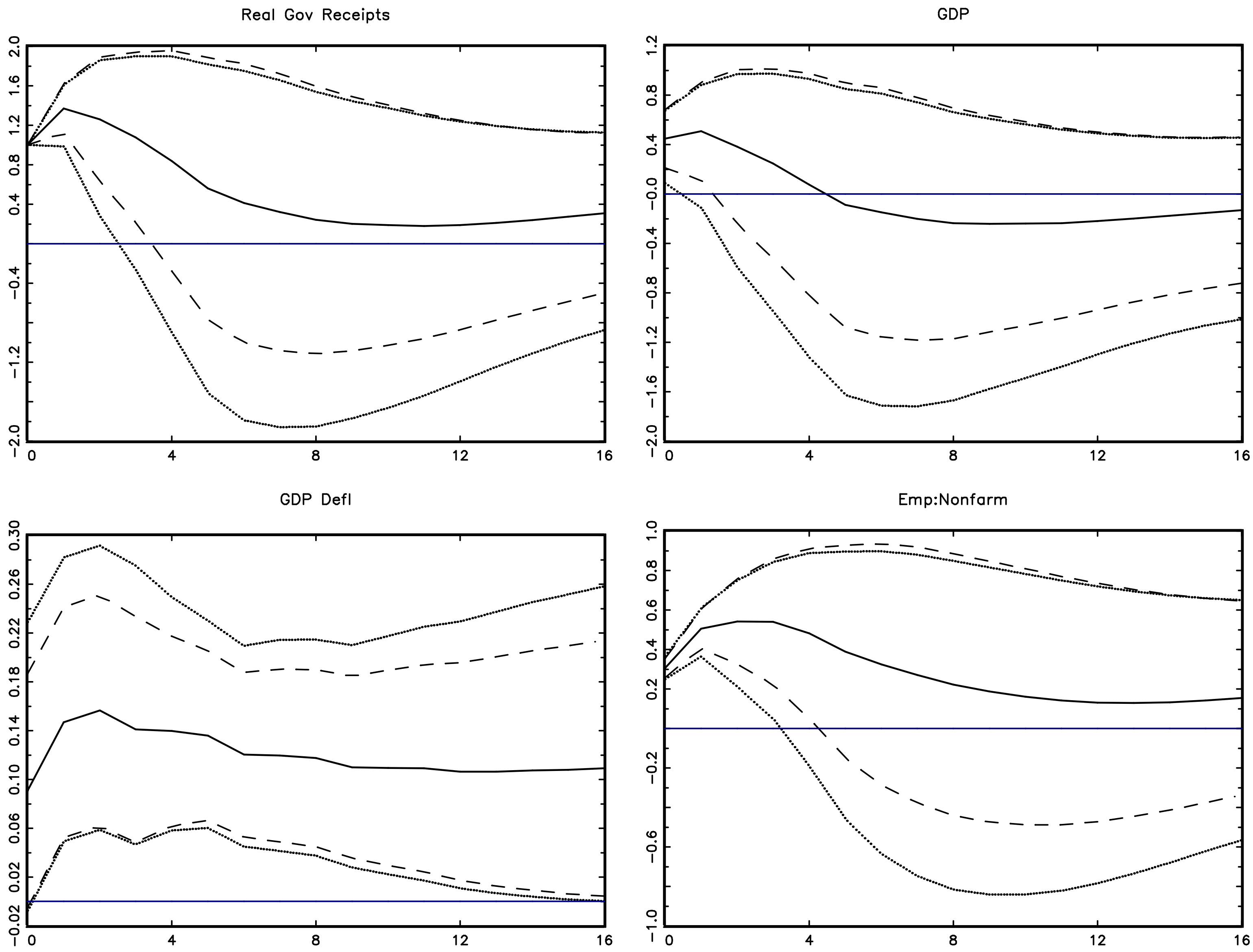
Bassett, Chosak, Driscoll, and Zakrajšek (2011) bank loan supply liquidity/risk shock (F_2 = 4.2)



Ramey (2011) fiscal (spending) shock ($F_2 = 1.0$)



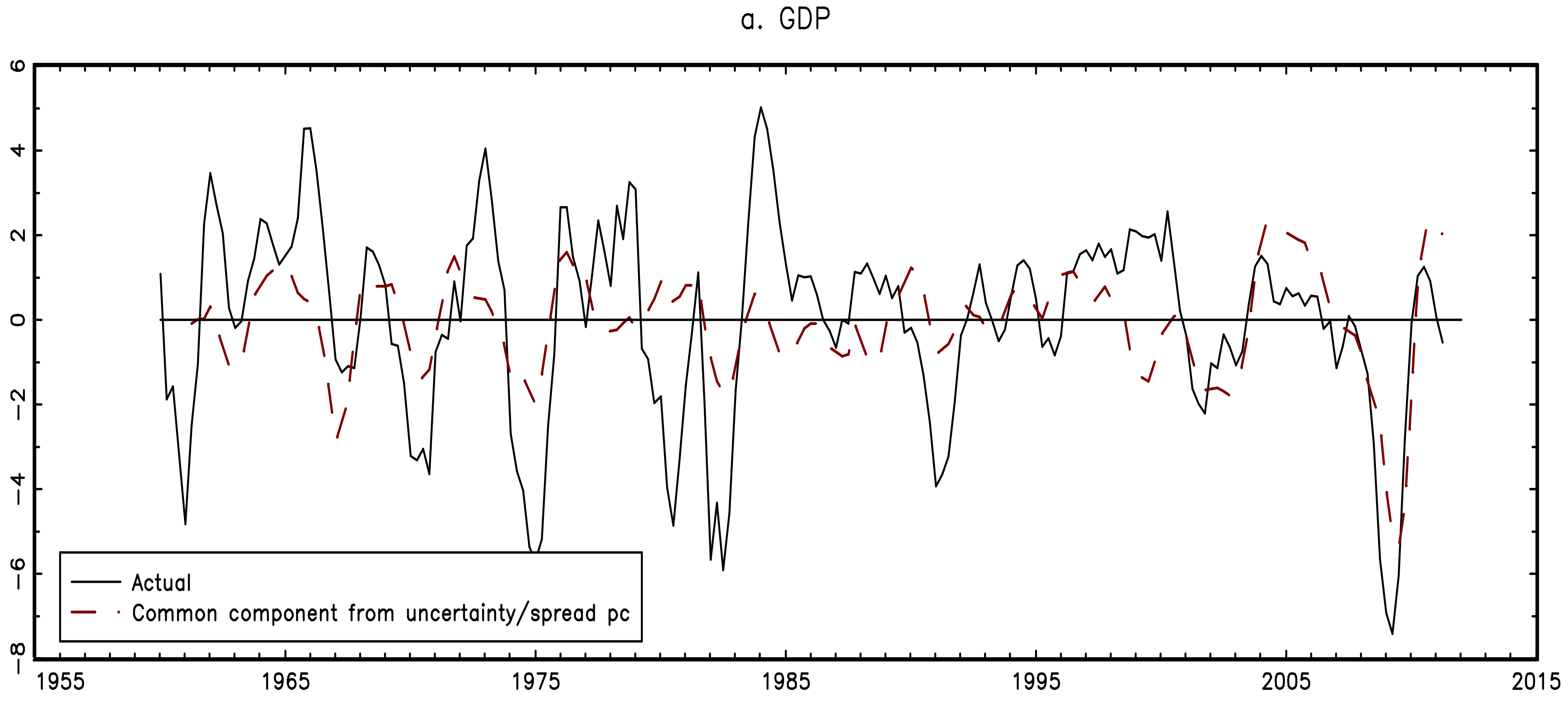
Fisher and Peters (2010) fiscal (spending) shock (F_2 = 0.1)



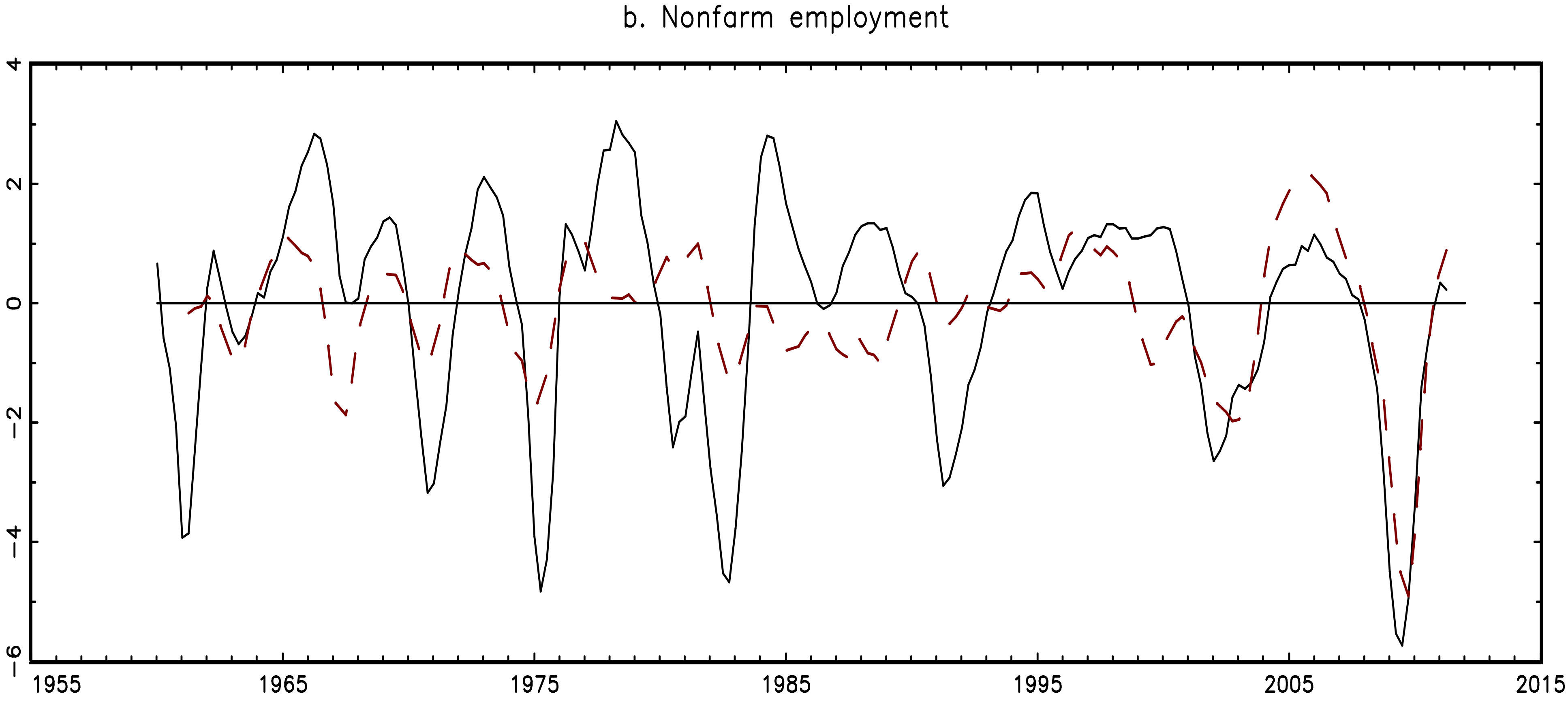
Romer and Romer (2010) fiscal (tax) schock (F_2 = 2.1)

Decomposition (estimated common component) for composite uncertainty/liquidity shock

Contribution to 4-Q **GDP** growth (1959-2011Q2) of first principal component of two term spread shocks & two uncertainty shocks



Contribution to 4-Q **Employment** growth (1959-2011Q2) of first principal component of two term spread shocks & two uncertainty shocks



10. Conclusions

<u>Work to do includes</u>

- Inference on correlations and on tests of overID restrictions in general
- Efficient inference for $k > 1$ (beyond CLR confidence sets) – exploit equivariance restriction to left-rotations (respecify SVAR in terms of linear combination of Y's – this should reduce the dimension of the sufficient statistics in the limit experiment)
- Inference on variance decomps – via the reduced form MARX?
- Inference in systems imposing uncorrelated shocks
- Formally taking into account "higher order" ($O_p(T^{-1/2})$) sampling uncertainty of reduced-form VAR parameters (conjecture: work via the (asymptotically normal) reduced form VARX but continue to use the "Fieller" trick)
- HAC (non-Kronecker) case: (a) robustify; (b) efficient inference?