A Critique of Structural VARs Using Business Cycle Theory

V.V. Chari, Patrick Kehoe, and Ellen McGrattan

Federal Reserve Bank of Minneapolis and University of Minnesota

(Materials available at www.minneapolisfed.org, Staff Report 364)

Introduction

- Main claim of structural VAR literature:
  - Given minimal set of identifying assumptions,

    can accurately estimate impulse responses of economic shocks

    regardless of the other details of the economy

- We provide class of **counterexamples** to main claim

Motivation

- RBC theory used to study business cycles, stock market, labor market...
- Want to investigate evidence ruling out RBC theory

> *The original technology-driven real business cycle hypothesis does appear to be dead.* — Francis and Ramey

> *Overall, the [SVAR] evidence seems to be clearly at odds with the predictions of standard RBC models...* — Gali

- Want to investigate evidence ruling in RBC theory

> *We find that a permanent shock to technology has qualitative consequences that a student of real business cycles would anticipate.* — Christiano et al.

Test of SVARs in the Laboratory

Test of SVARs in the Laboratory

- Use standard RBC model satisfying key identifying assumptions of SVARs
- Focus on response of hours to technology shock
- Derive **theoretical** impulse response from model
- Generate data from the model
- Calculate **empirical** impulse response identified by SVAR procedure
- Test: Are they close?

Main Findings

- SVAR procedure **does not** robustly uncover impulse responses

- Works better:

  - the lower is the capital share

  - the larger is contribution of technology shocks to fluctuations

- Works poorly for large class of parameters, including estimated ones

Main Findings

- SVAR procedure **does not** robustly uncover impulse responses
- Works better:
    - the lower is the capital share
    - the larger is contribution of technology shocks to fluctuations
- Works poorly for large class of parameters, including estimated ones
- Conjectured solution:
    - adding variables helps
    - we find it does not help in general

What is the Root of the Problem?

- An auxiliary assumption: small number of lags in VAR is sufficient
- How do we prove this?
  - Begin by abstracting from small sample issues
  - Derive population moments when number of lags in VAR small
- Small sample issues
  - Small sample bias is small
  - Confidence bands often so wide that SVAR is uninformative

The Deep Root of the Problem

- Not computing same statistics in model and data

- Three candidates to compare:

  1. Theoretical RBC Model

  2. SVAR applied to RBC Model

  3. SVAR applied to US data

- SVAR "compares" 1 and 3 (apples to oranges)

- Should compare 2 and 3 (apples to apples)

Related Literature

- Economic model's MA noninvertible (Hansen and Sargent)
- Shock processes often misspecified (Cooley and Dwyer)
- Specification-mining drive results (Uhlig)
- Inference in ∞-dimensional space hard (Sims, Faust, Leeper)
- Over-differencing problems (Christiano, Eichenbaum, and Vigfusson)
- Small samples a practical problem (Erceg, Guerrier, Gust)

Outline

- Explain SVAR procedure
- Show SVAR procedure leads to large errors
- Derive errors analytically and discuss special cases
- Show adding more variables does not help
- Show basic insights hold up in small sample
- Put small sample results in context of literature

SVAR Procedure

What You Get from SVAR Procedure

- Structural MA for ϵ=[‘technology shock’, ‘demand shock’]$'$

$$X_t = A_0\epsilon_t + A_1\epsilon_{t-1} + A_2\epsilon_{t-2} + \ldots,\ E\epsilon_t\epsilon_t' = \Sigma$$

What You Get from SVAR Procedure

- Structural MA for ϵ=[‘technology shock’, ‘demand shock’]$'$

$$X_t = A_0\epsilon_t + A_1\epsilon_{t-1} + A_2\epsilon_{t-2} + \ldots,\ E\epsilon_t\epsilon_t' = \Sigma$$

where X_t = [Δ Log labor productivity, $(1-\alpha L)$ Log hours]$'$

- Specifications with different α's:
    - DSVAR: $\alpha = 1$
    - LSVAR: $\alpha = 0$
    - QDSVAR: $\alpha \in (0,1)$

What You Get from SVAR Procedure

- Structural MA for ϵ=[‘technology shock’, ‘demand shock’]$'$

$$X_t = A_0\epsilon_t + A_1\epsilon_{t-1} + A_2\epsilon_{t-2} + \ldots,\ E\epsilon_t\epsilon_t' = \Sigma$$

where $X_t = [\Delta$ Log labor productivity, $(1-\alpha L)$ Log hours$]'$

- Identifying assumptions:
  - technology and demand shocks uncorrelated ($\Sigma = I$)
  - demand shock has no long-run effect on productivity

Impulse Responses and Long-Run Restriction

- Impulse response from structural MA:

  Blip ϵ_1^d for response of productivity to demand

$$\log(y_1/l_1) - \log(y_0/l_0) = A_0(1,2)$$

$$\log(y_2/l_2) - \log(y_0/l_0) = A_0(1,2) + A_1(1,2)$$

$$\vdots$$

$$\log(y_t/l_t) - \log(y_0/l_0) = A_0(1,2) + A_1(1,2) + \ldots + A_t(1,2)$$

Impulse Responses and Long-Run Restriction

- Impulse response from structural MA:

  Blip ϵ_1^d for response of productivity to demand

$$\log(y_1/l_1) - \log(y_0/l_0) = A_0(1,2)$$

$$\log(y_2/l_2) - \log(y_0/l_0) = A_0(1,2) + A_1(1,2)$$

$$\vdots$$

$$\log(y_t/l_t) - \log(y_0/l_0) = A_0(1,2) + A_1(1,2) + \ldots + A_t(1,2)$$

- Long-run restriction:

  Demand shock has no long run effect on level of productivity

$$\sum_{j=0}^{\infty} A_j(1,2) = 0$$

Deriving Structural MA from VAR

- OLS regressions on bivariate VAR: $B(L)X_t = v_t$

$$X_t = B_1 X_{t-1} + B_2 X_{t-2} + B_3 X_{t-3} + B_4 X_{t-4} + v_t, \quad E v_t v_t' = \Omega$$

Deriving Structural MA from VAR

- OLS regressions on bivariate VAR: $B(L)X_t = v_t$

$$X_t = B_1X_{t-1} + B_2X_{t-2} + B_3X_{t-3} + B_4X_{t-4} + v_t, \quad Ev_tv_t' = \Omega$$

- Invert to get MA: $X_t = B(L)^{-1}v_t = C(L)v_t$

$$X_t = v_t + C_1v_{t-1} + C_2v_{t-2} + \dots$$

Identifying Assumptions

- Work from MA: $X_t = v_t + C_1 v_{t-1} + C_2 v_{t-2} + \ldots,\ E v_t v_t' = \Omega$

- Structural MA for ϵ=[‘technology shock’, ‘demand shock’]$'$

$$X_t = A_0 \epsilon_t + A_1 \epsilon_{t-1} + A_2 \epsilon_{t-2} + \ldots,\ E \epsilon_t \epsilon_t' = \Sigma$$

with $A_0 \epsilon_t = v_t$, $A_j = C_j A_0$, $A_0 \Sigma A_0' = \Omega$

Identifying Assumptions

- Work from MA: $X_t = v_t + C_1 v_{t-1} + C_2 v_{t-2} + \ldots,\ E v_t v_t' = \Omega$
- Structural MA for ϵ=[‘technology shock’, ‘demand shock’]$'$

$$X_t = A_0 \epsilon_t + A_1 \epsilon_{t-1} + A_2 \epsilon_{t-2} + \ldots,\ \ E \epsilon_t \epsilon_t' = \Sigma$$

  with $A_0 \epsilon_t = v_t$, $A_j = C_j A_0$, $A_0 \Sigma A_0' = \Omega$

- Identifying assumptions determine parameters in A_0, Σ
  - Structural shocks ϵ are orthogonal, $\Sigma = I$
  - Demand shocks have no long-run effect on labor productivity

Identifying Assumptions

- Work from MA: $X_t = v_t + C_1 v_{t-1} + C_2 v_{t-2} + \ldots,\ E v_t v_t' = \Omega$

- Structural MA for ϵ=[‘technology shock’, ‘demand shock’]$'$

$$X_t = A_0 \epsilon_t + A_1 \epsilon_{t-1} + A_2 \epsilon_{t-2} + \ldots,\ E\epsilon_t \epsilon_t' = \Sigma$$

with $A_0 \epsilon_t = v_t$, $A_j = C_j A_0$, $A_0 \Sigma A_0' = \Omega$

- Identifying assumptions determine parameters in A_0, Σ
  - Structural shocks ϵ are orthogonal, $\Sigma = I$
  - Demand shocks have no long-run effect on labor productivity

$\Rightarrow$ 4 equations, 4 parameters in A_0 ($A_0 A_0' = \Omega$, $\sum_j A_j(1,2) = 0$)

The Theoretical Laboratory

Use Model Satisfying Key SVAR Assumptions

- Shocks are orthogonal

- Technology shock has long-run effect on y/l, demand shock does not

Use Model Satisfying Key SVAR Assumptions

- Shocks are orthogonal

- Technology shock has long-run effect on y/l, demand shock does not

This is a best case scenario for SVARs

RBC Model

- Households choose c, x, l to solve:

$$\max\ E_0 \sum_{t=0}^{\infty} \beta^t U(c_t, l_t)$$

$$\text{s.t.} \qquad c_t + x_t = (1-\tau_{lt})w_t l_t + r_t k_t + T_t$$

$$k_{t+1} = (1-\delta)k_t + x_t$$

- Technology: $y_t = F(k_t, z_t l_t)$

- Resource constraint: $c_t + x_t = y_t$

Shocks in RBC Model

- Technology shocks

$$\log z_t = \mu_z + \log z_{t-1} + \eta_{zt}$$

- "Demand" shocks

$$\tau_{lt} = (1 - \rho_l)\bar{\tau}_l + \rho_l \tau_{lt-1} + \eta_{lt}$$

Shocks in RBC Model

- Technology shocks

$$\log z_t = \mu_z + \log z_{t-1} + \eta_{zt}$$

- "Demand" shocks

$$\tau_{lt} = (1 - \rho_l)\bar{\tau}_l + \rho_l \tau_{lt-1} + \eta_{lt}$$

- Model satisfies the key SVAR assumptions
  - Shocks orthogonal ($\eta_z \perp \eta_l$)
  - Technology shock has long-run effect on y/l, demand shock doesn't

Shocks in RBC Model

- Technology shocks

$$\log z_t = \mu_z + \log z_{t-1} + \eta_{zt}$$

- "Demand" shocks

$$\tau_{lt} = (1 - \rho_l)\bar{\tau}_l + \rho_l \tau_{lt-1} + \eta_{lt}$$

- Model has theoretical impulse response

$$X_t = D(L)\eta_t$$

where $X_t = [\Delta \log y_t/l_t, \Delta \log l_t]'$ or $X_t = [\Delta \log y_t/l_t, \log l_t]'$

Model's Functional Forms and Parameters

- Assume
  - $U(c, l) = \log c + \psi \log(1 - l)$
  - $F(k, zl) = k^{\theta}(zl)^{1-\theta}$
- Derive general analytical results
- Since interested in robustness, when displaying quantitative results,
  - start with MLE parameter estimates for US as baseline
  - also show for **wide regions** of parameter space

Our Evaluation of SVAR Procedure

- Use RBC model satisfying SVAR's key assumptions
- Model has **theoretical** impulse response

$$X_t = D(L)\eta_t$$

Our Evaluation of SVAR Procedure

- Use RBC model satisfying SVAR's key assumptions

- Model has **theoretical** impulse response

$$X_t = D(L)\eta_t$$

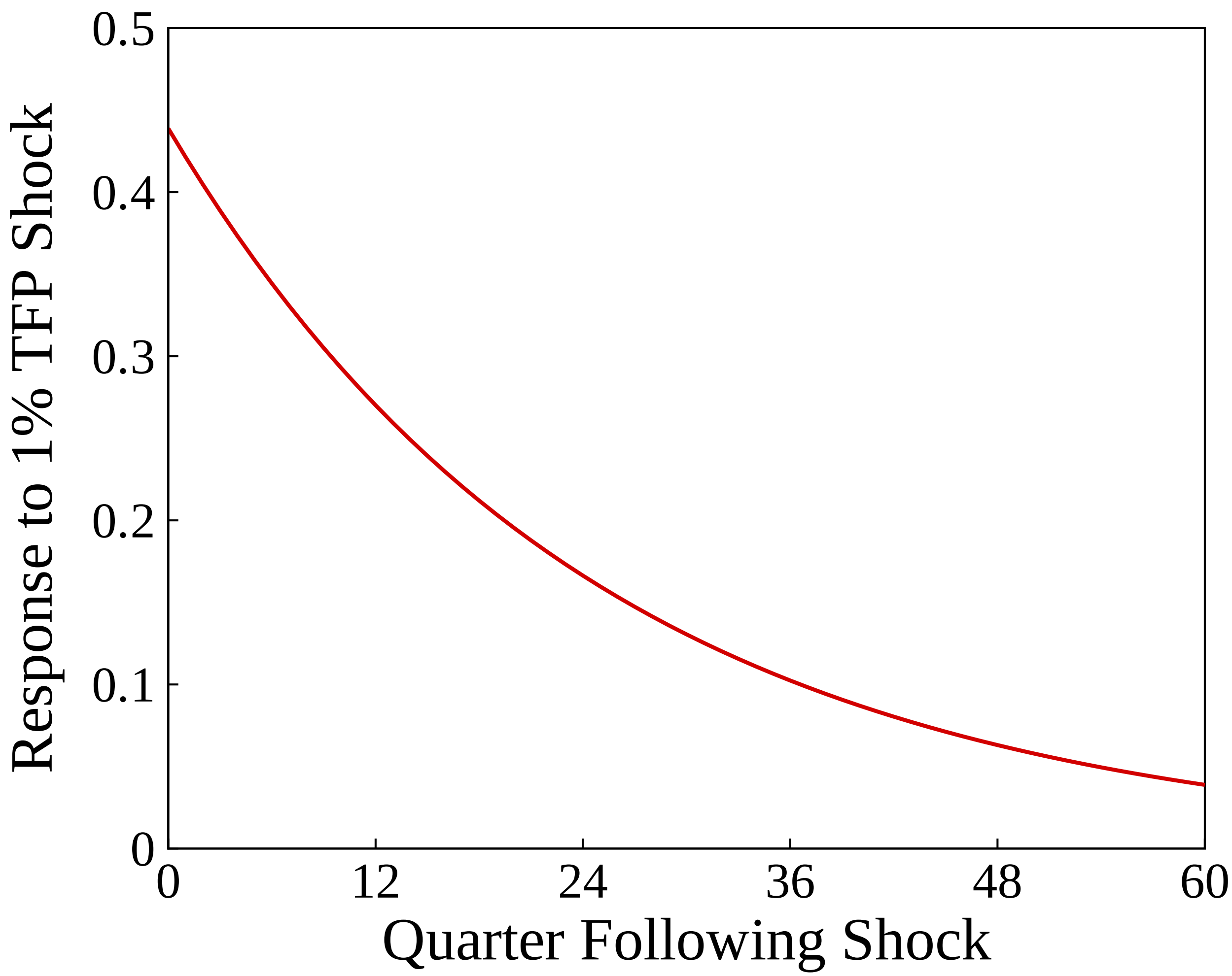
- Derive SVAR's **empirical** impulse response

$$X_t = A(L)\epsilon_t$$

- Compare model impulse responses with SVAR responses

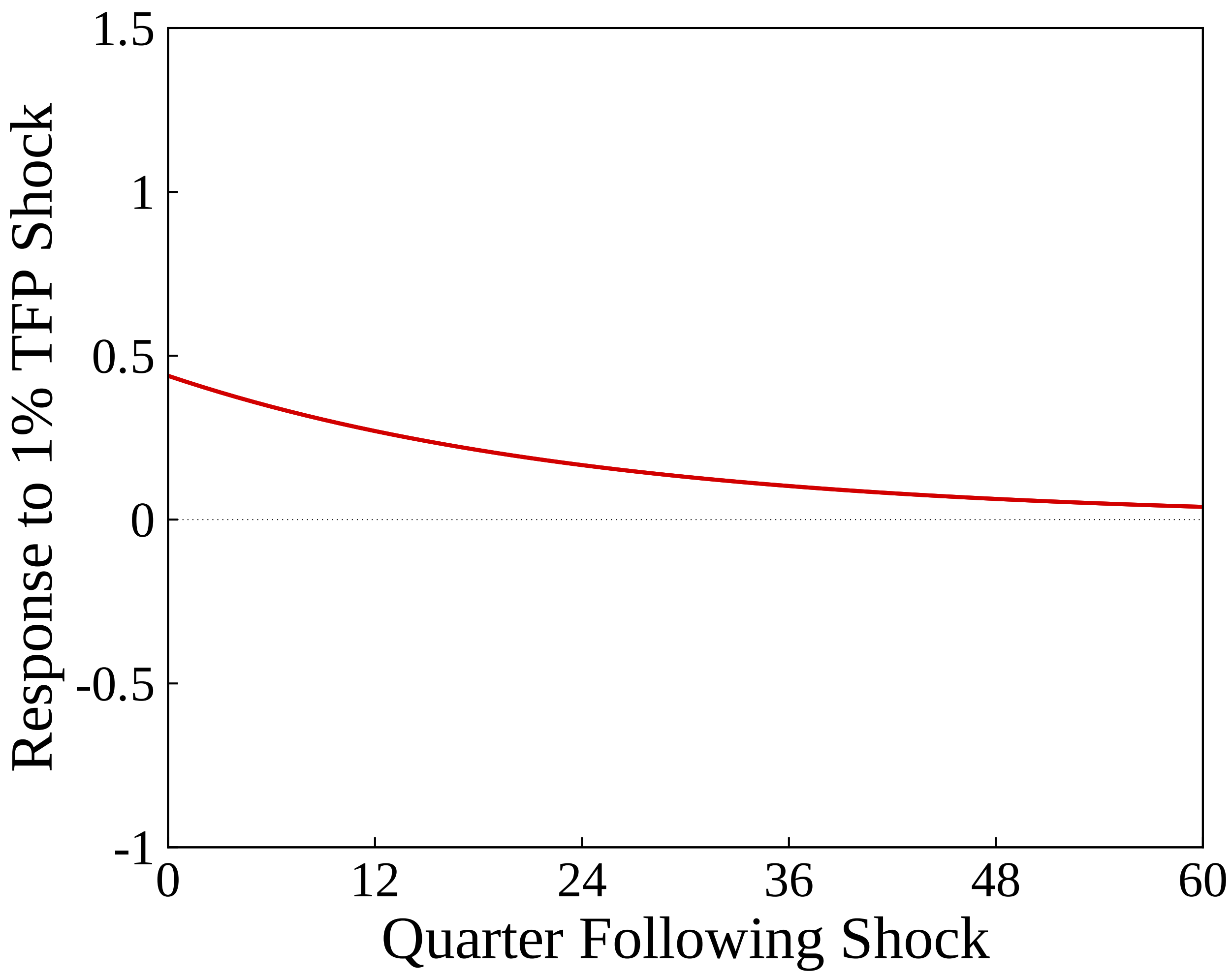
Specification Error – Population Moments

Model Impulse Response of Hours



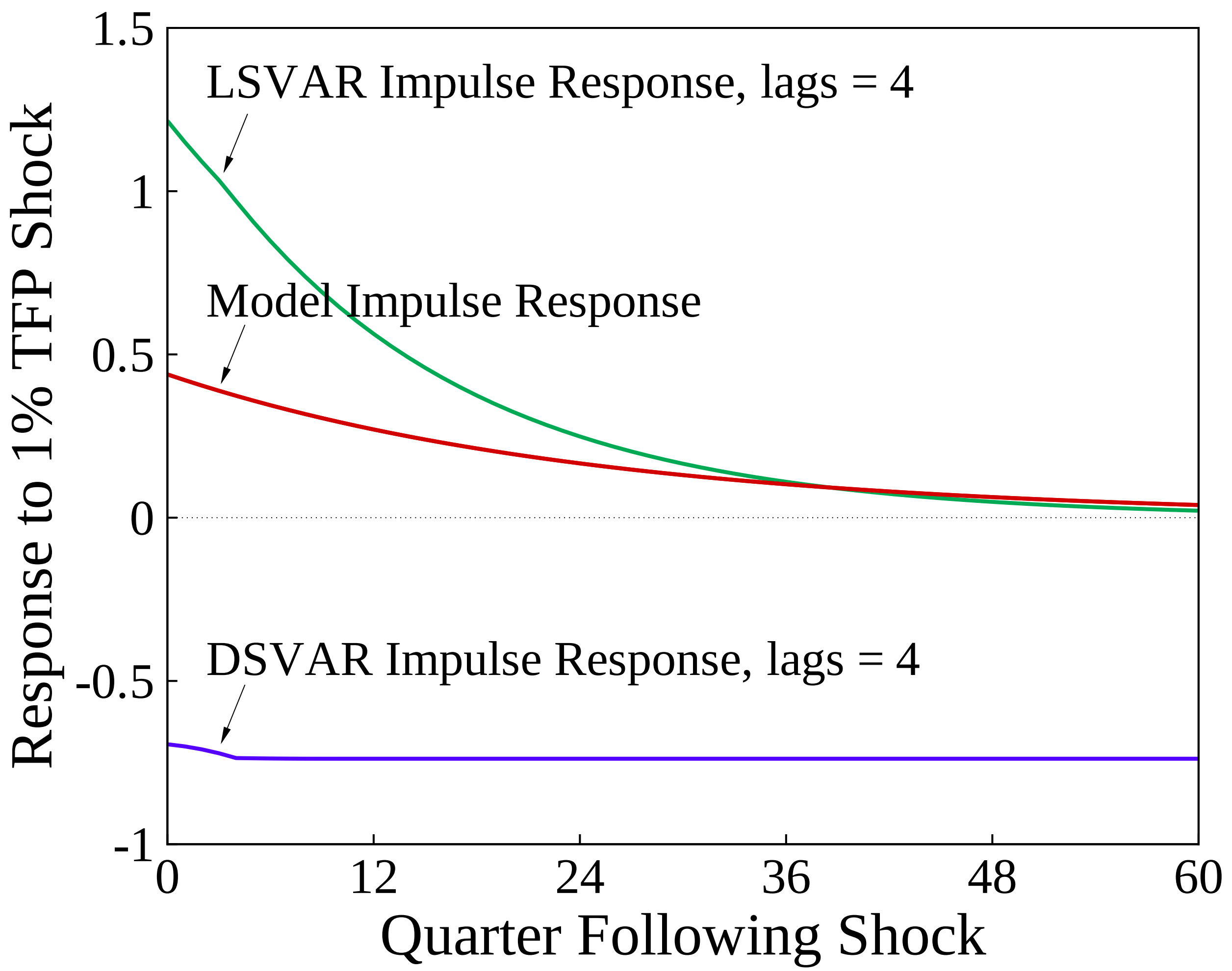
After a 1% TFP shock, hours rise .44%. The half-life of the response is 18 quarters.

Model Impulse Response of Hours



After a 1% TFP shock, hours rise .44%. The half-life of the response is 18 quarters.

Model and SVAR Impulse Responses of Hours



SVARs miss significantly on impact and and half-life; outside range of current models.

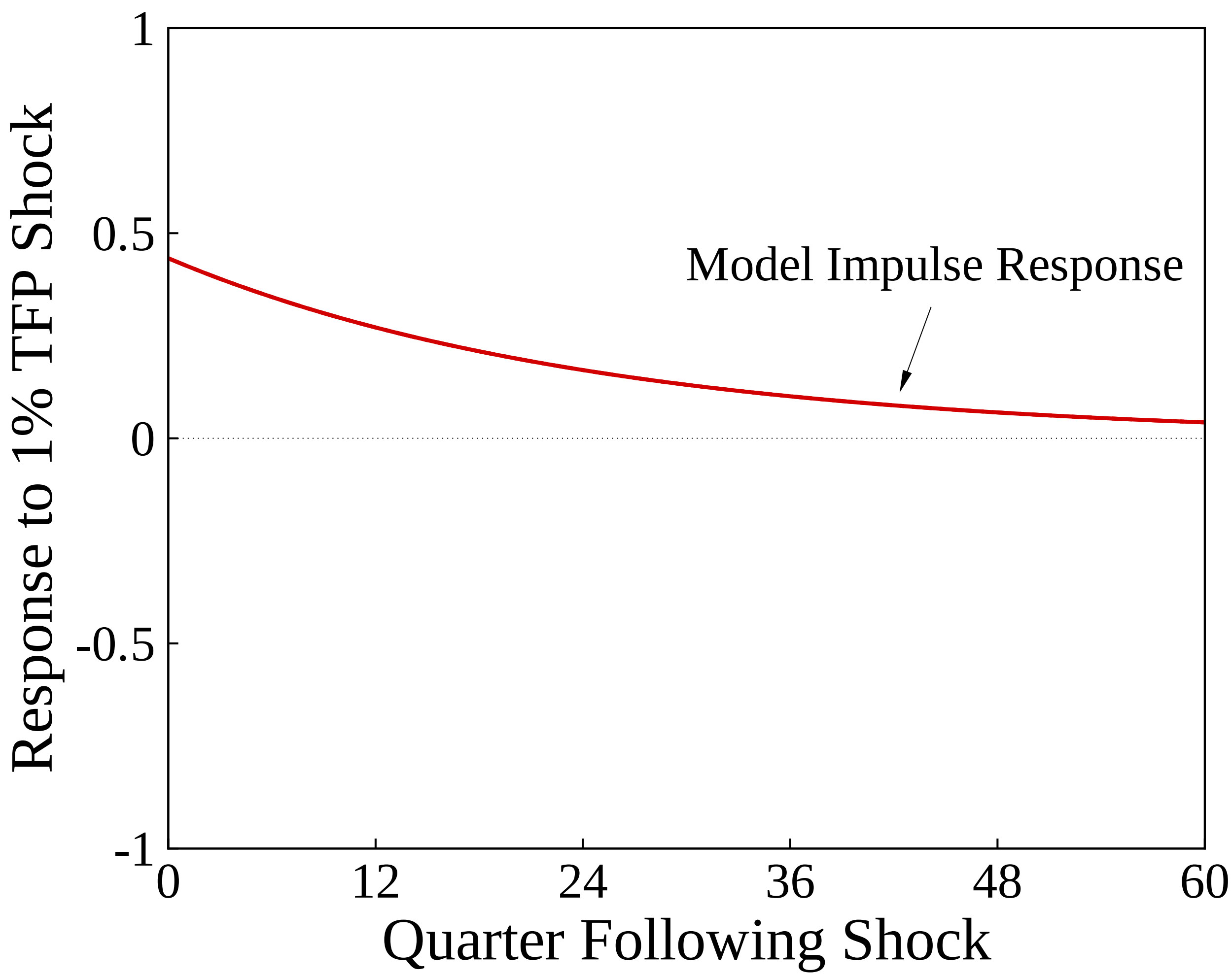
Recap

- Used model satisfying SVAR key assumptions
- Compared SVAR responses and model responses
- SVAR procedures lead to large errors
- Now explain why

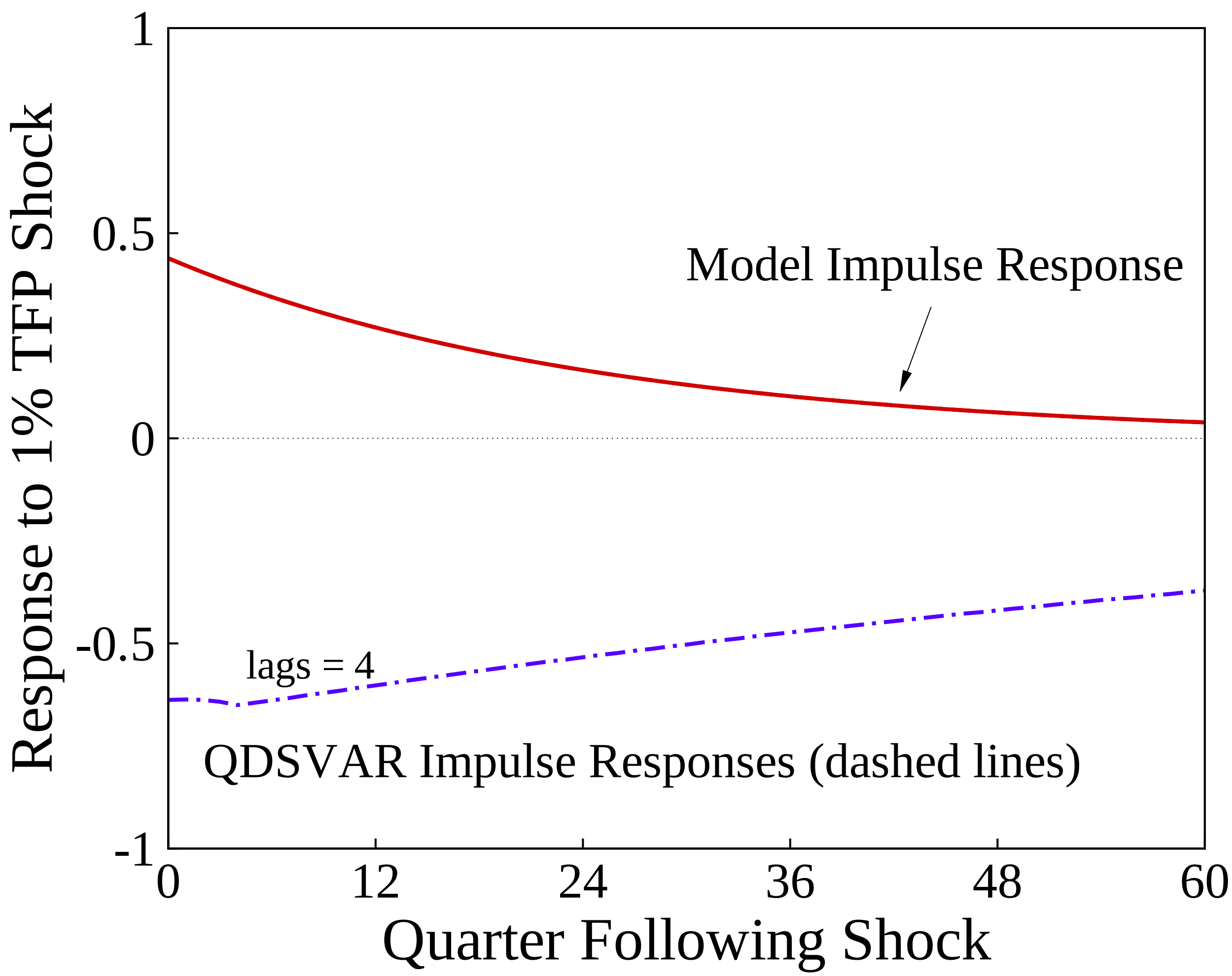
Analyze Specification Error

- Implicit assumptions of SVAR
  - MA representation is invertible
  - Few AR lags enough (e.g., 4)
- Violation of either leads to error
  - Easy to get around invertibility by quasi-differencing $(1-\alpha L)\log l_t$
  - Hard to get around short lag length problem

The Short Lag Length Problem (QDSVAR)

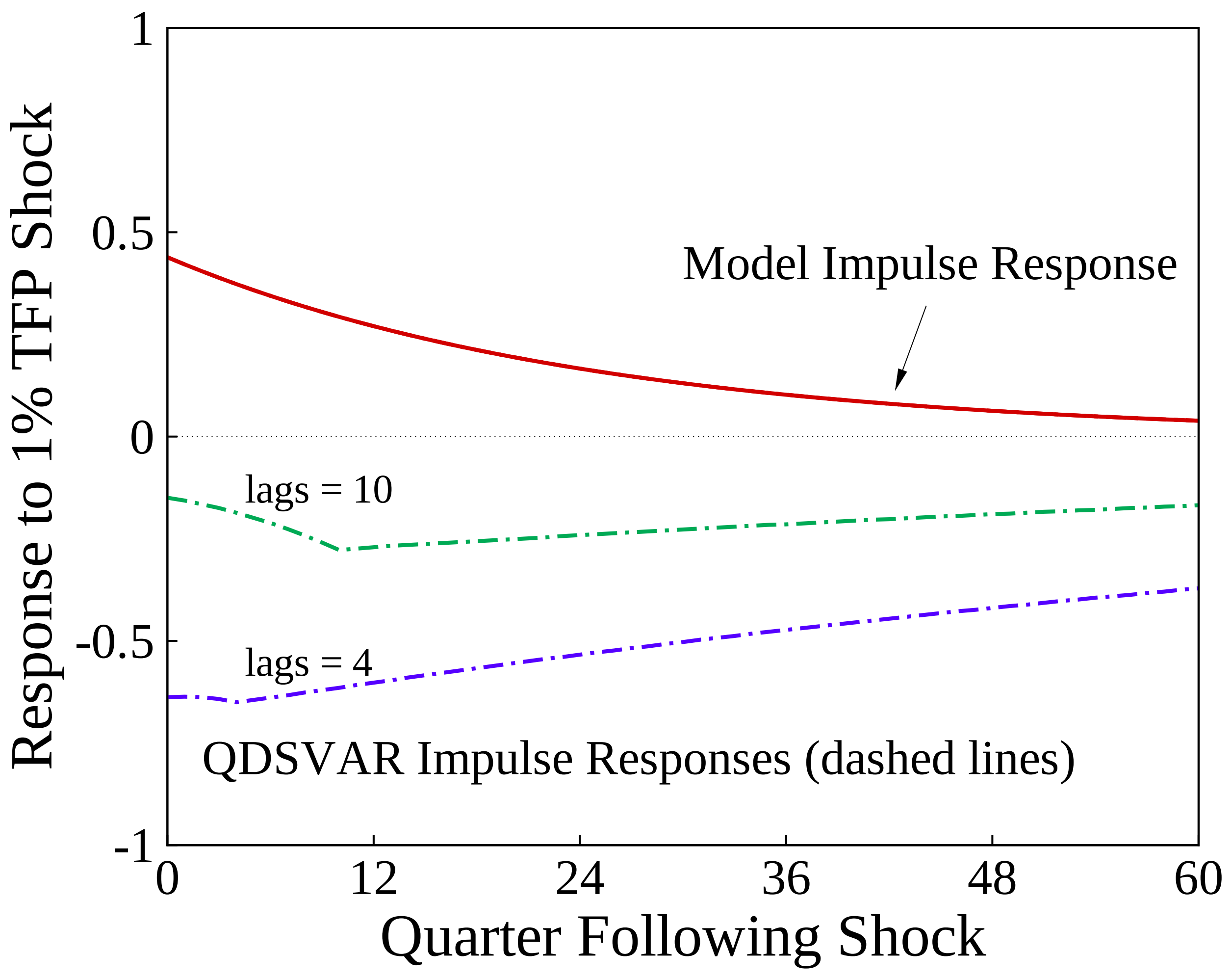


Short Lag Length Leads to Large Errors

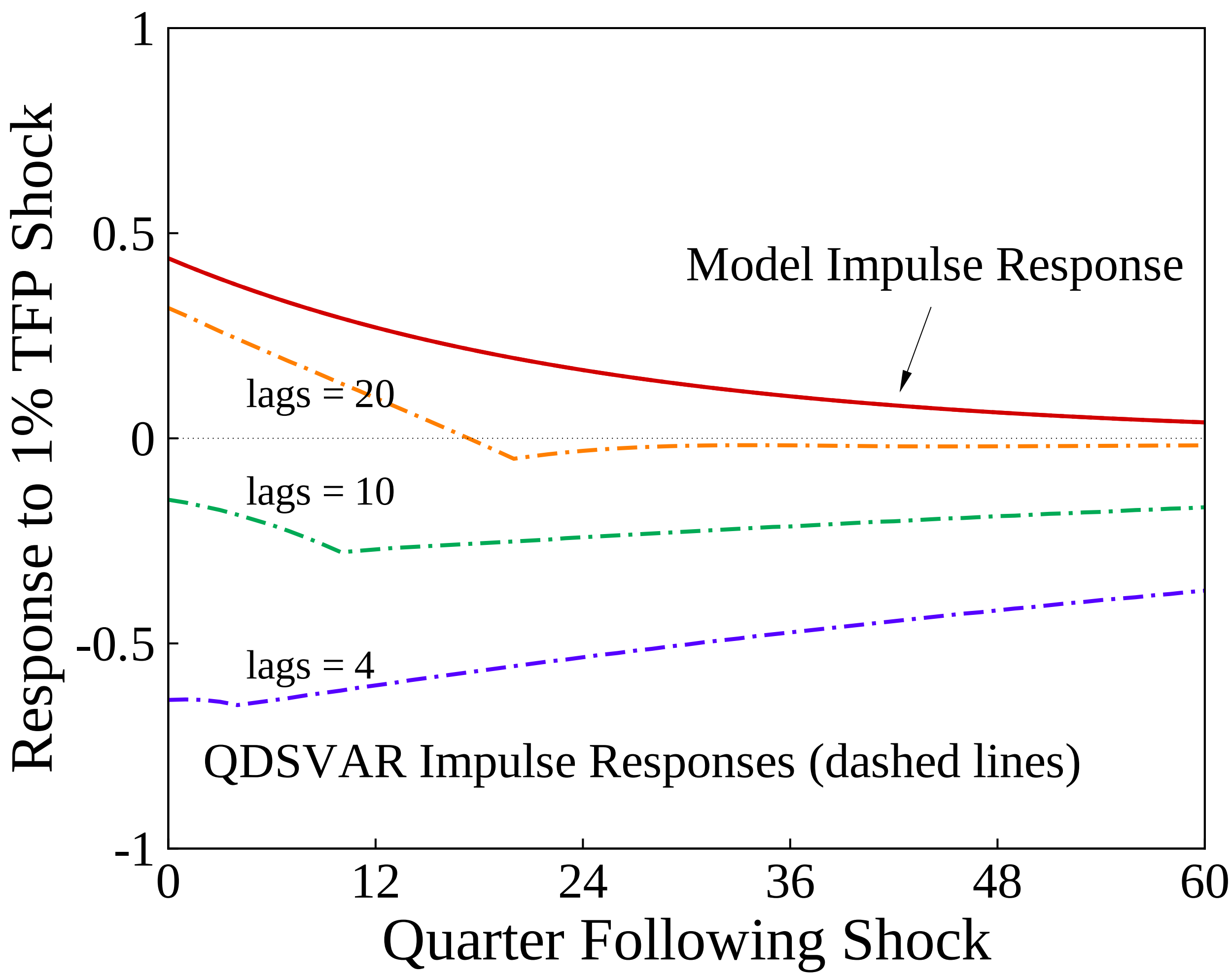


QDSVAR with $\alpha = .99$ doesn't uncover model's impulse response.

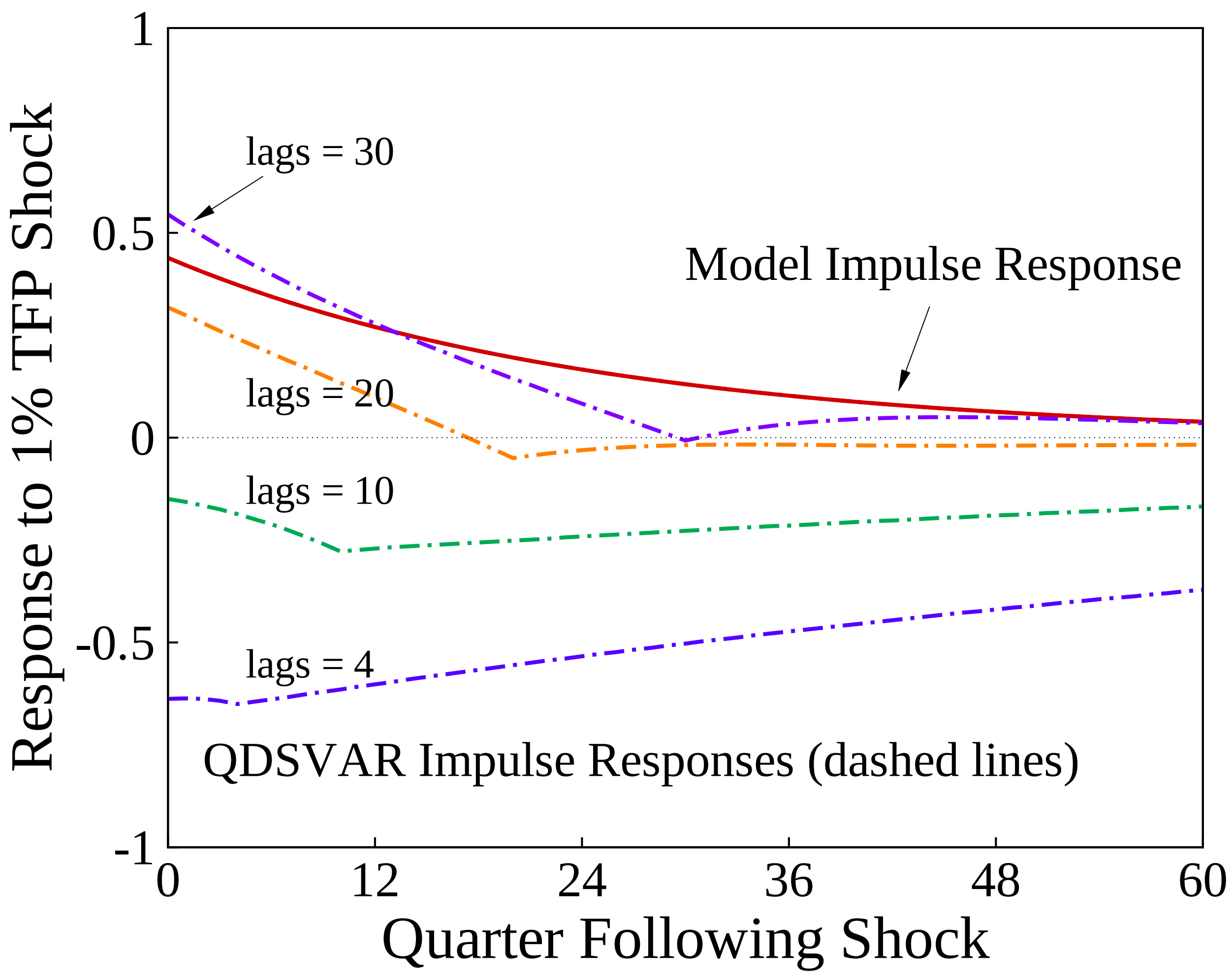
Short Lag Length Leads to Large Errors



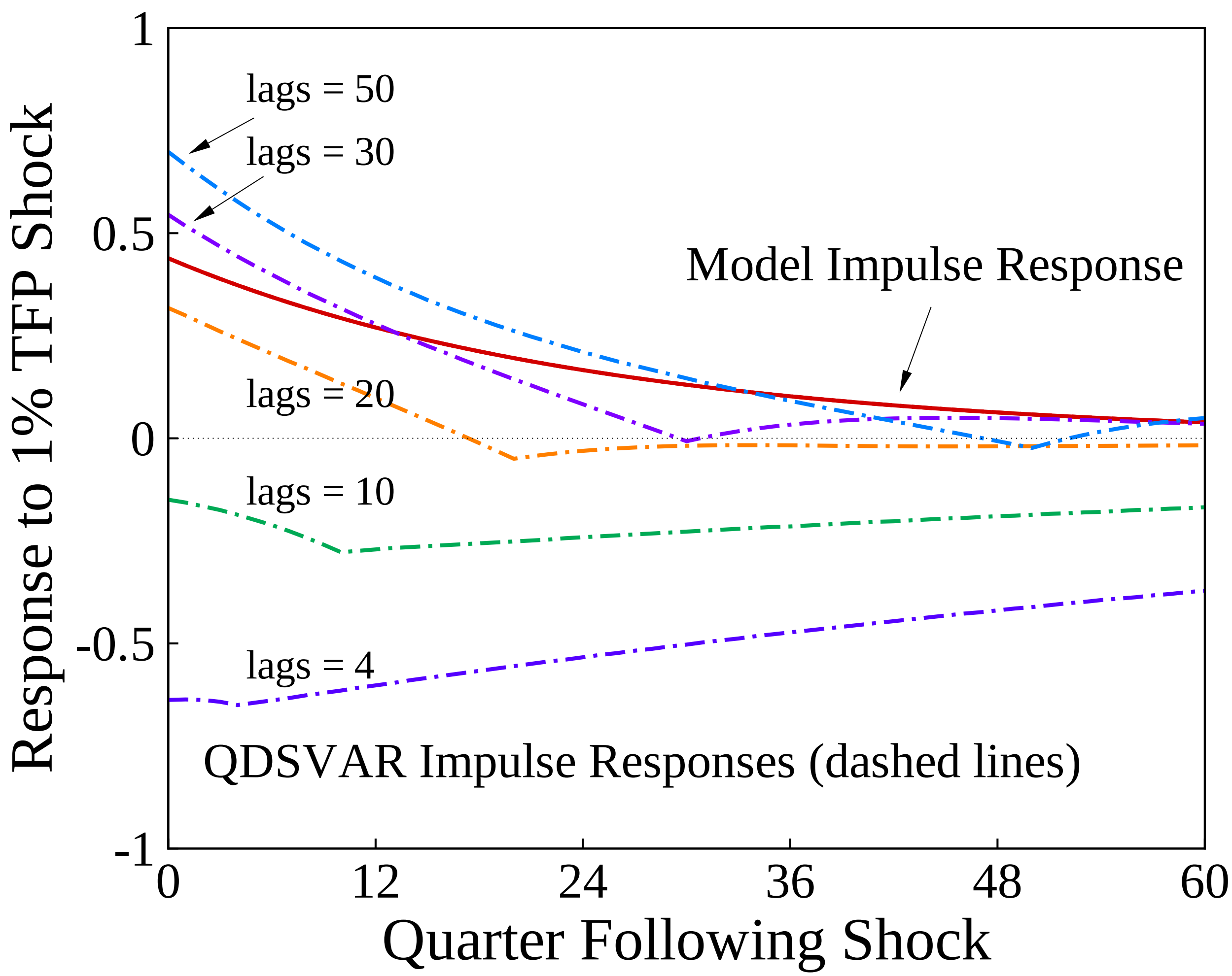
Short Lag Length Leads to Large Errors



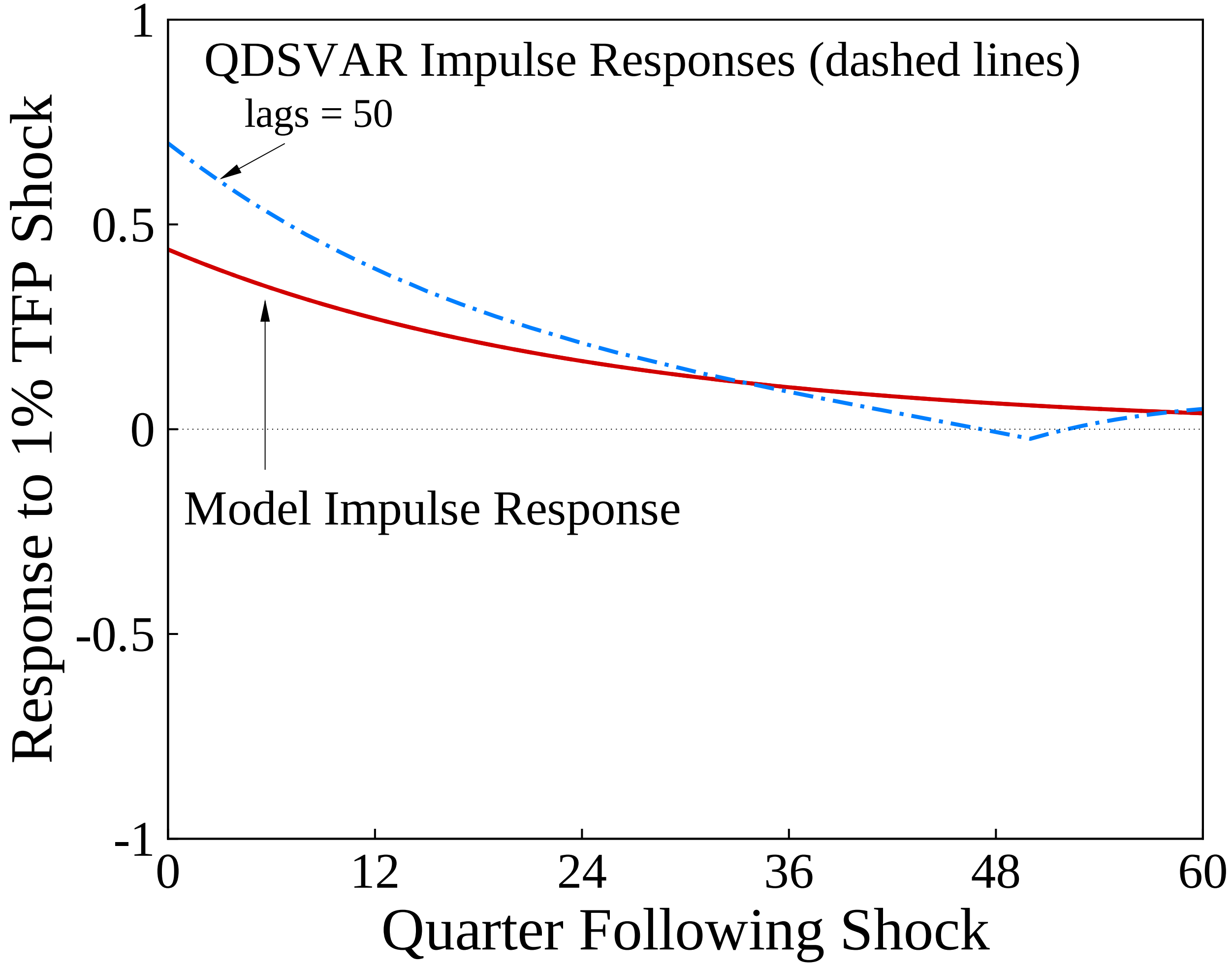
Short Lag Length Leads to Large Errors



Short Lag Length Leads to Large Errors

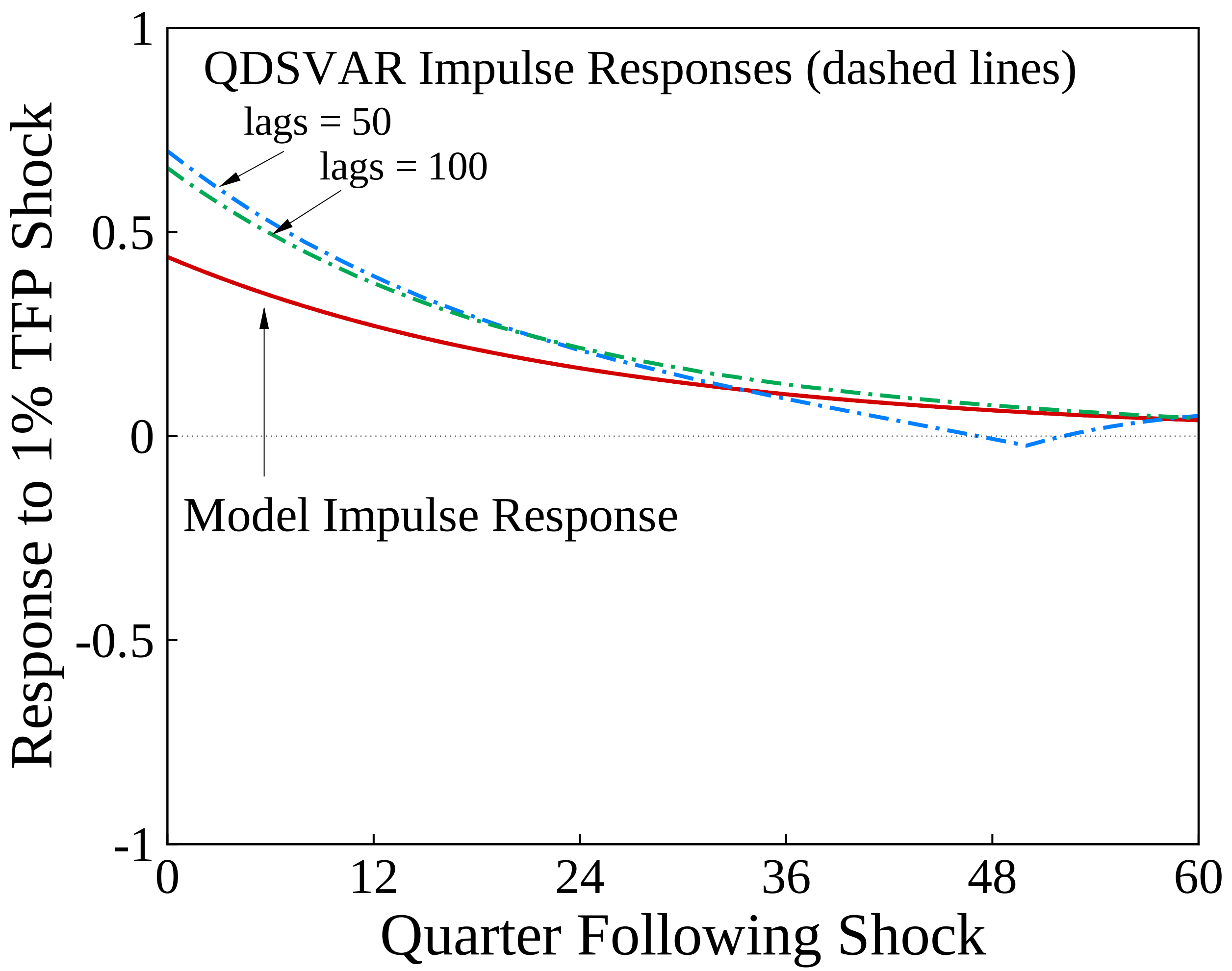


Short Lag Length Leads to Large Errors

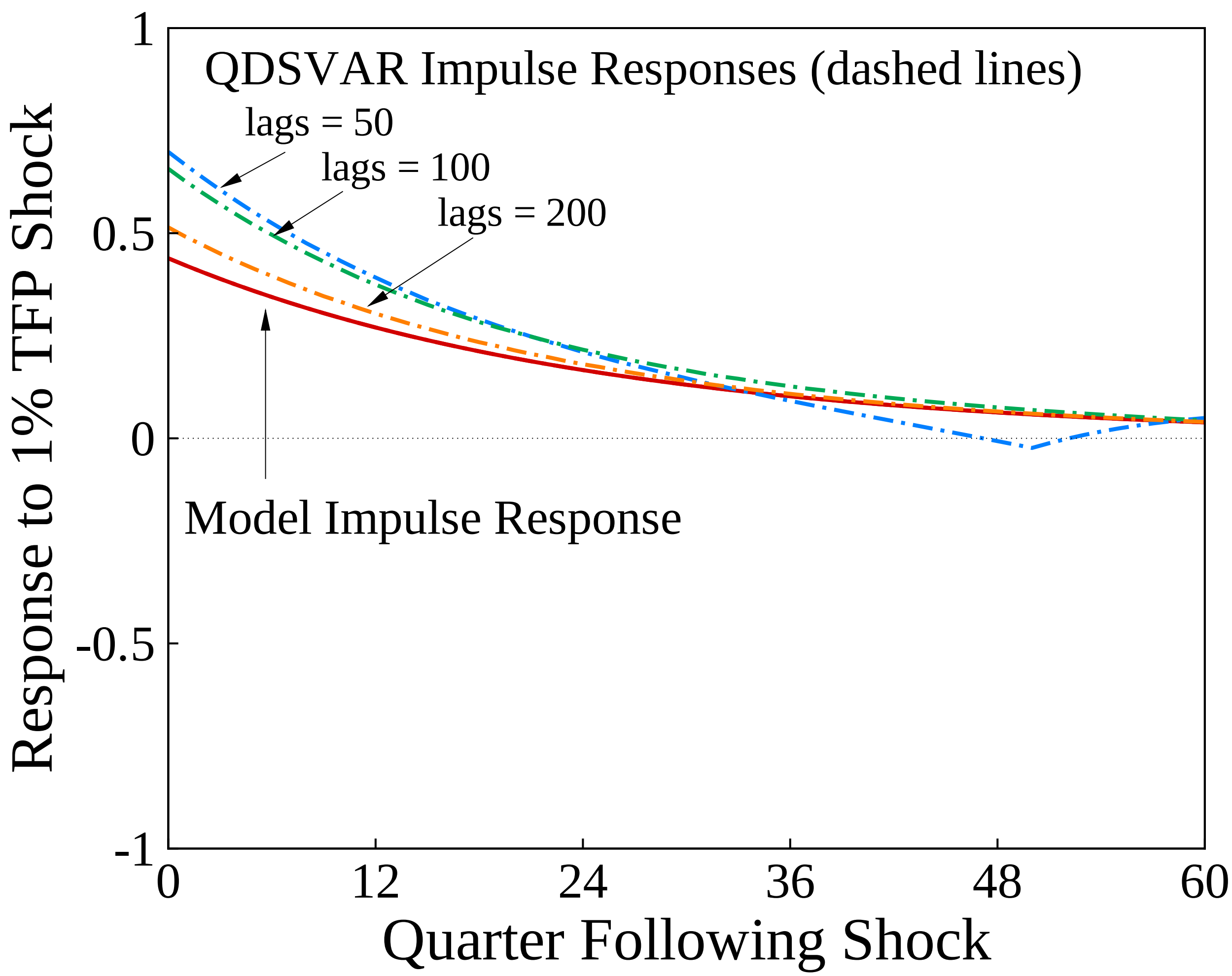


Overshoots and then comes back...

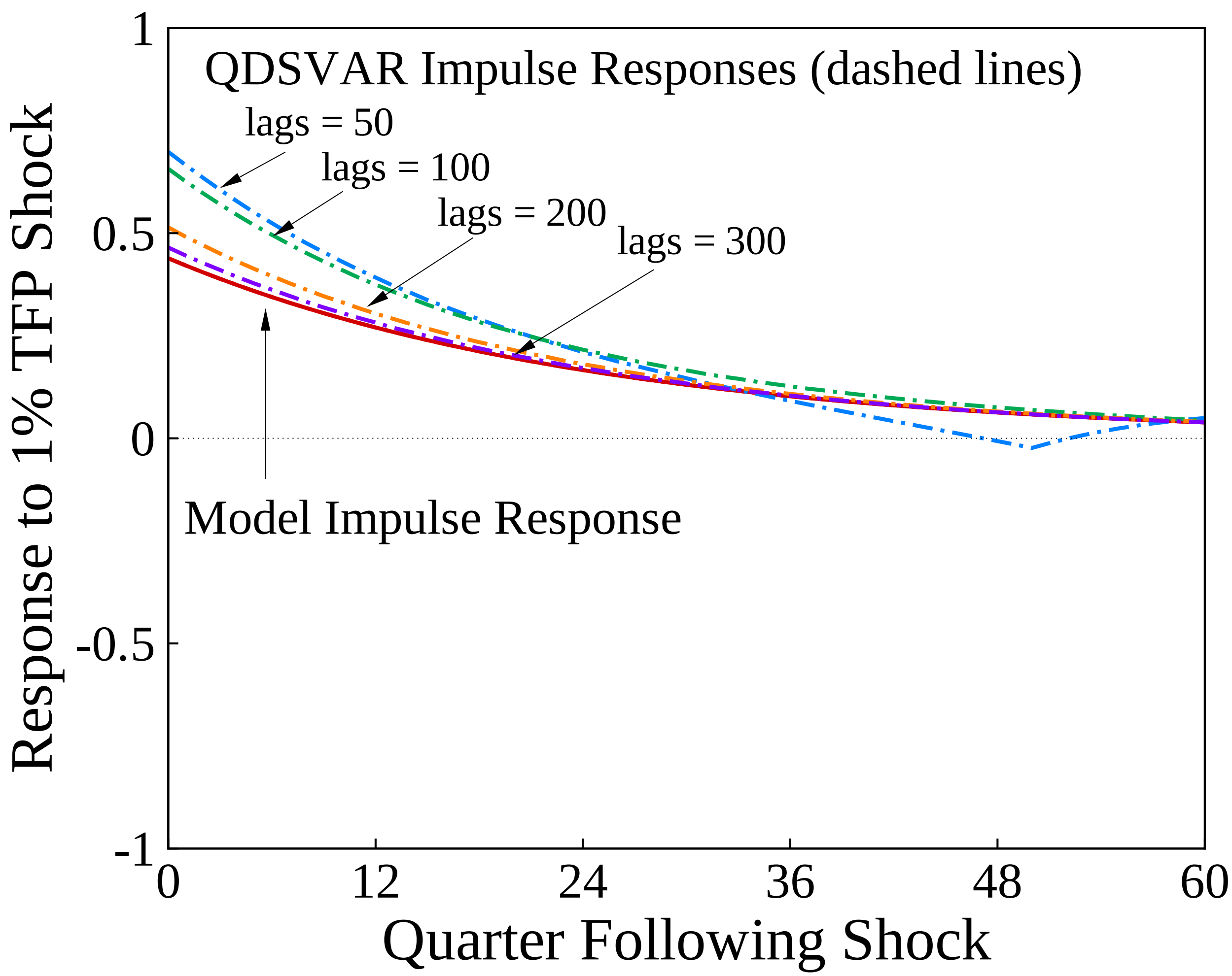
Short Lag Length Leads to Large Errors



Short Lag Length Leads to Large Errors



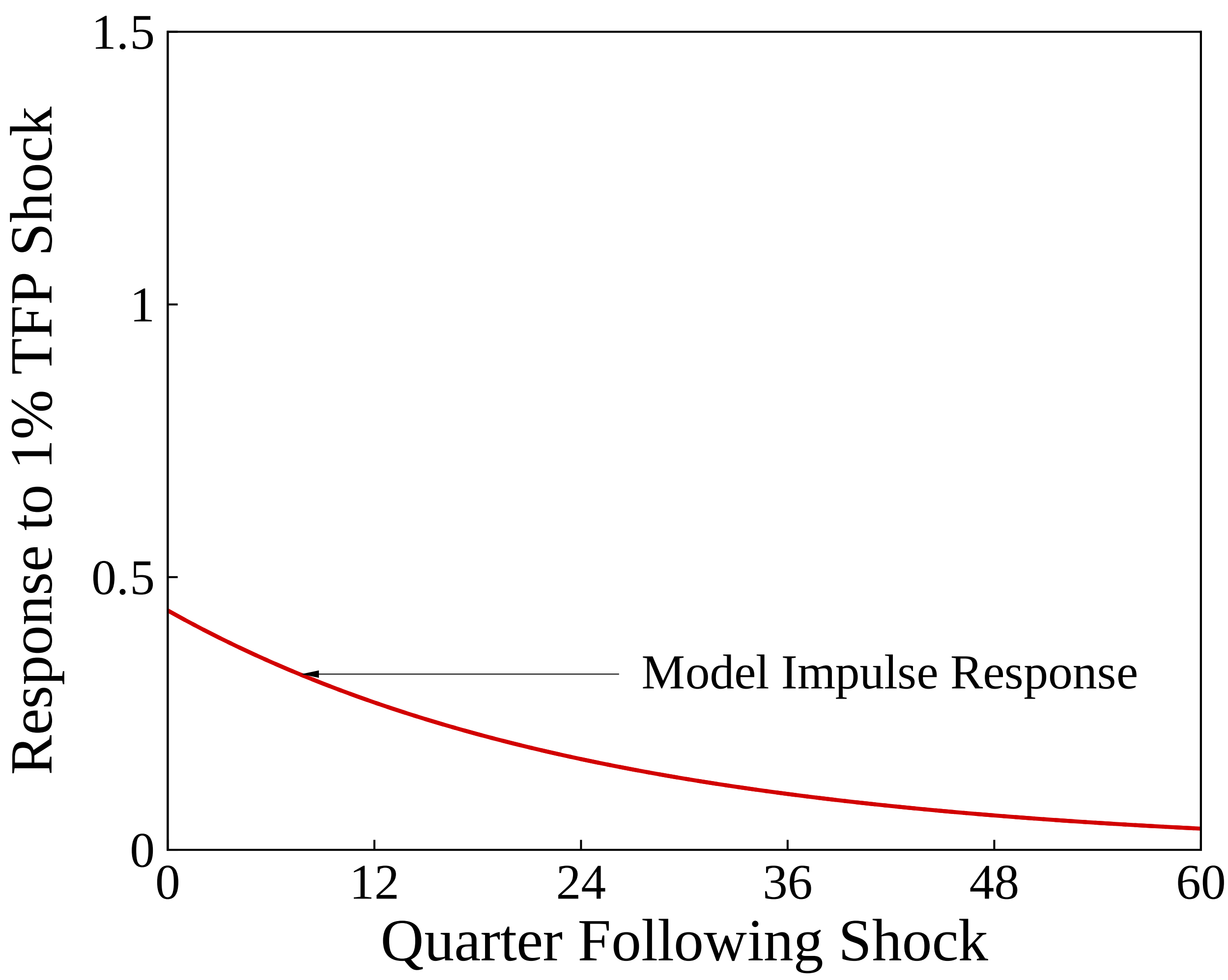
Short Lag Length Leads to Large Errors



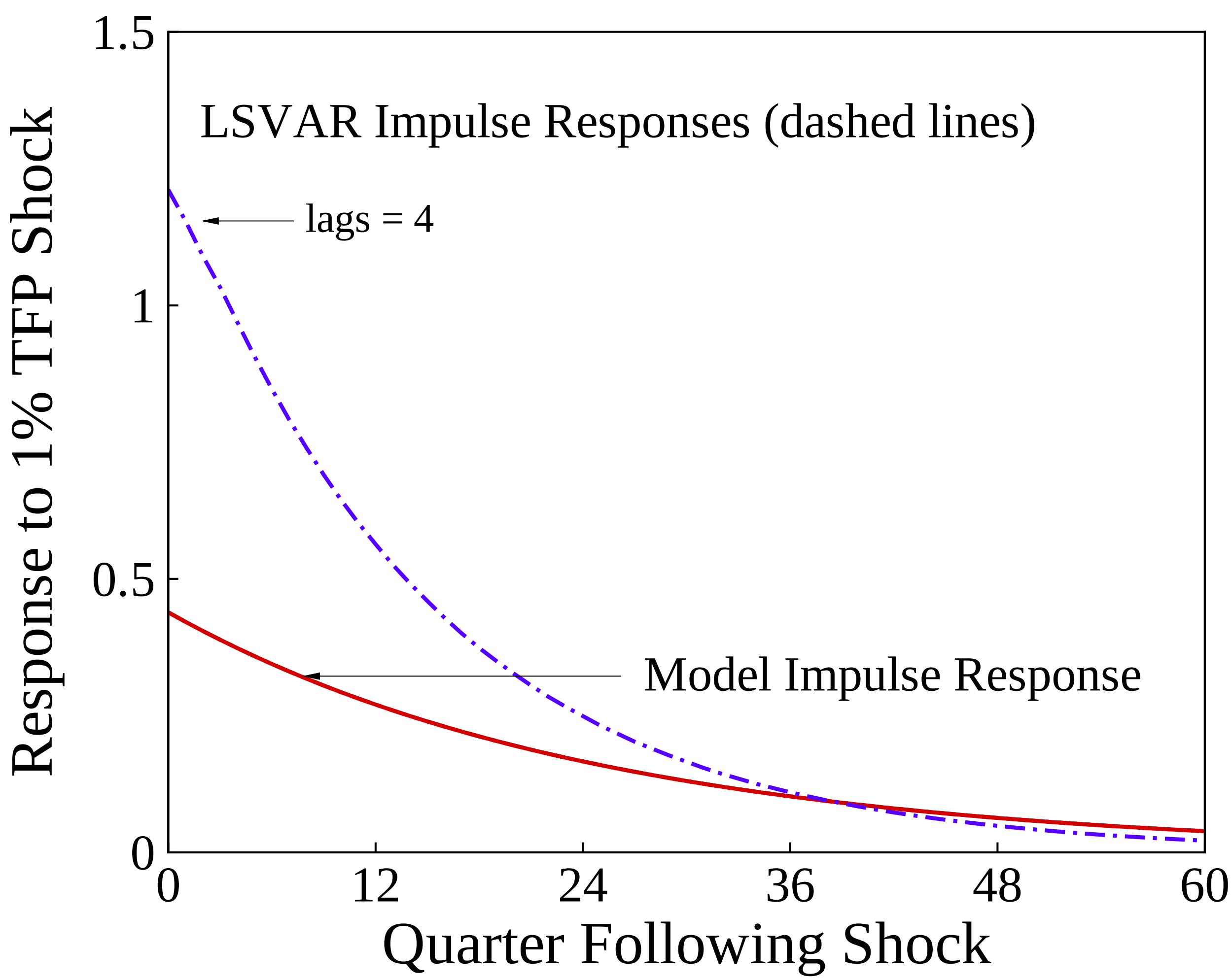
Finally get close with 300 lags, which is more than the number of periods in our datasets.

What if we had used hours in levels?

Short Lag Length Problem (LSVAR)

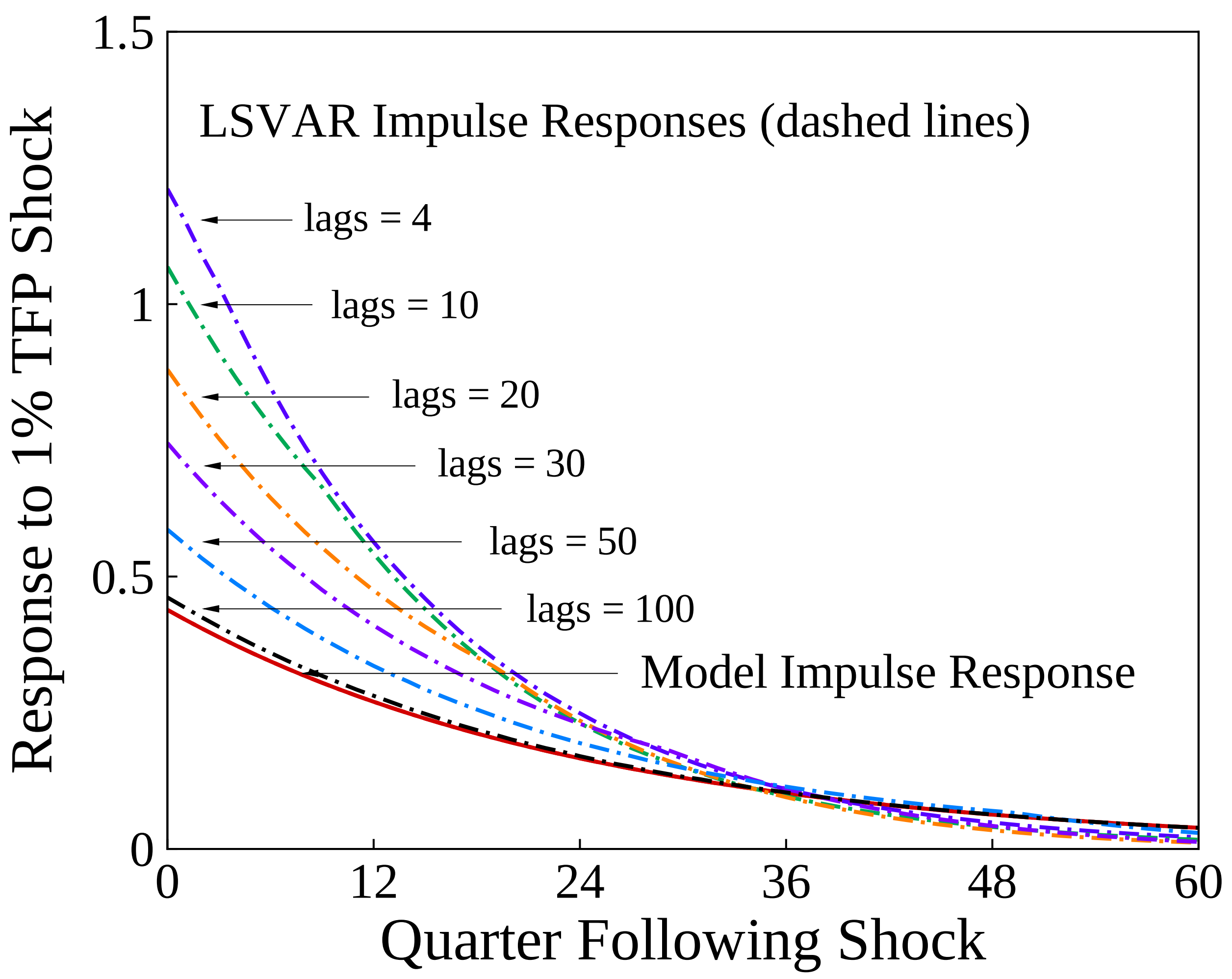


Short Lag Length Problem (LSVAR)



LSVAR doesn't uncover model's impulse response

Short Lag Length Problem (LSVAR)



Finally get close with 100 lags, which is still too many given the number of periods in our datasets.

Long AR Needed Because of Capital

- Capital decision rule, with $\hat{k}_t = k_t/z_{t-1}$:

$$\log \hat{k}_{t+1} = \gamma_k \log \hat{k}_t + \gamma_z \eta_{zt} + \gamma_\tau \tau_{lt}$$

- So others, like l_t, have ARMA representation

$$\log l_t = \gamma_k \log l_{t-1} + \phi_z (1 - \kappa_z L) \eta_{zt} + \phi_\tau (1 - \kappa_\tau L) \tau_{lt}$$

- What does the AR representation, $B(L) X_t = v_t$, look like?

Model Has Infinite-order AR

- Proposition 1: Model has VAR coefficients B_j such that

$$B_j = MB_{j-1}, \quad j \geq 2,$$

where M has eigenvalues equal to α (the differencing parameter) and

$$\left(\frac{\gamma_k - \gamma_l \phi_k / \phi_l - \theta}{1 - \theta}\right)$$

γ_k, γ_l are coefficients in the capital decision rule

ϕ_k, ϕ_l are coefficients in the labor decision rule

- Eigenvalues of M are α and .96 for the baseline parameters

What Happens with Too Few Lags

- Recall
  - VAR: $B(L)X_t = v_t, Ev_tv_t' = \Omega \quad \Rightarrow X_t = C(L)v_t$
  - Empirical impulse response: $X_t = A(L)\epsilon_t$

- Response of hours to technology on impact:

  $A_0(2,1)$ is a function of Ω and $\bar{C} = I + C_1 + C_2 + \ldots$

What Happens with Too Few Lags

- Theoretical and empirical impulse responses different if
  - $\Omega_m \neq \Omega$
  - $\bar{C}_m \neq \bar{C}$

- Proposition 2: If VAR has 1 lag, SVAR recovers

$$\Omega = \Omega_m + M\left(\Omega_m - \Omega_m V(X)^{-1}\Omega_m\right)M'$$

$$\bar{C}^{-1} = \bar{C}_m^{-1} + M(I-M)^{-1}C_{m,1} + M\left(\Omega_m - V(X)\right)V(X)^{-1}$$

Notice that M is important factor in garbled terms!

Two Special Cases

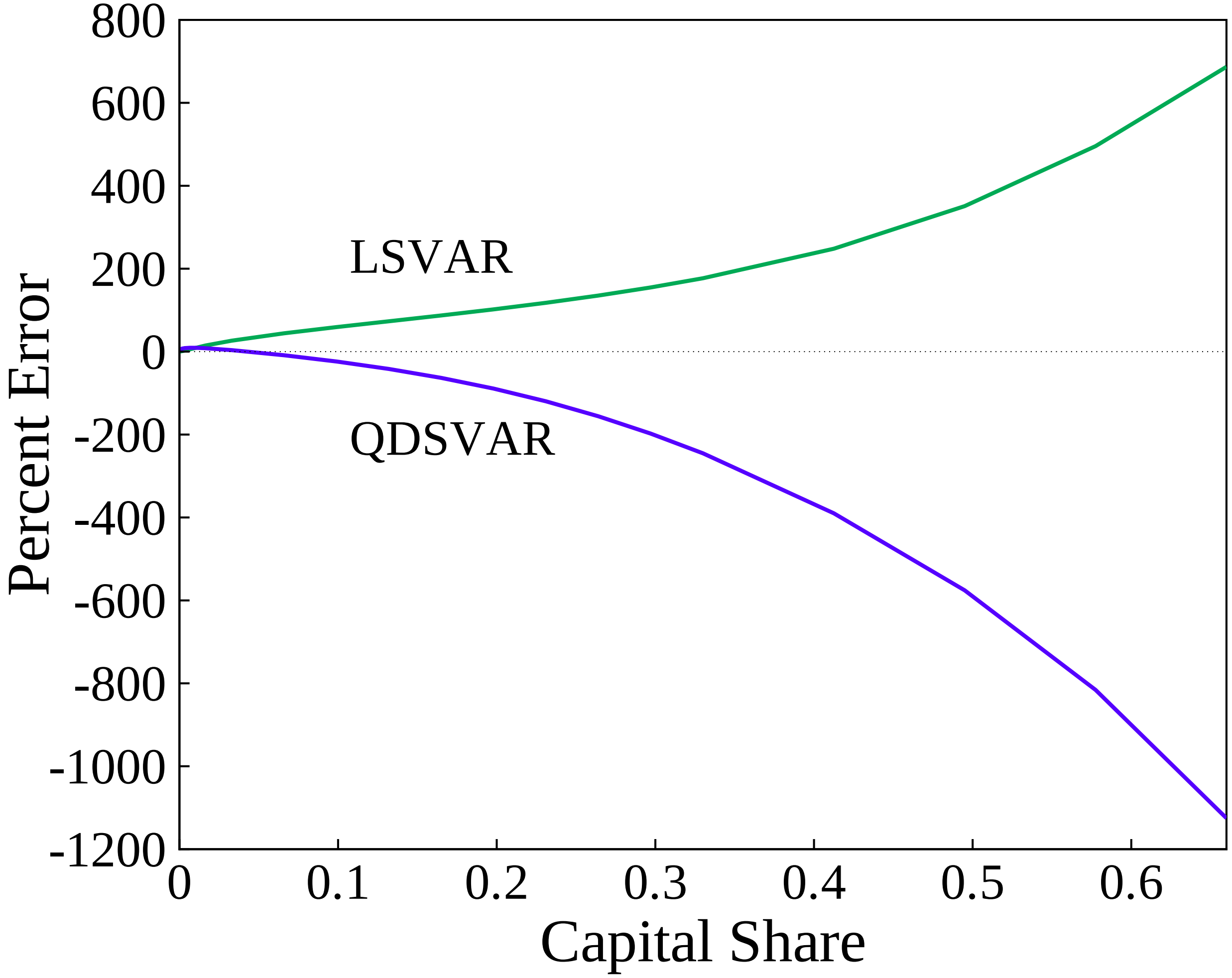
SVAR Procedure in Two Special Cases

- Proposition 3: If model has
  - no capital ($\theta = 0$) or
  - only one shock ($\sigma_l = 0$),

  then SVAR uncovers model's impulse response.

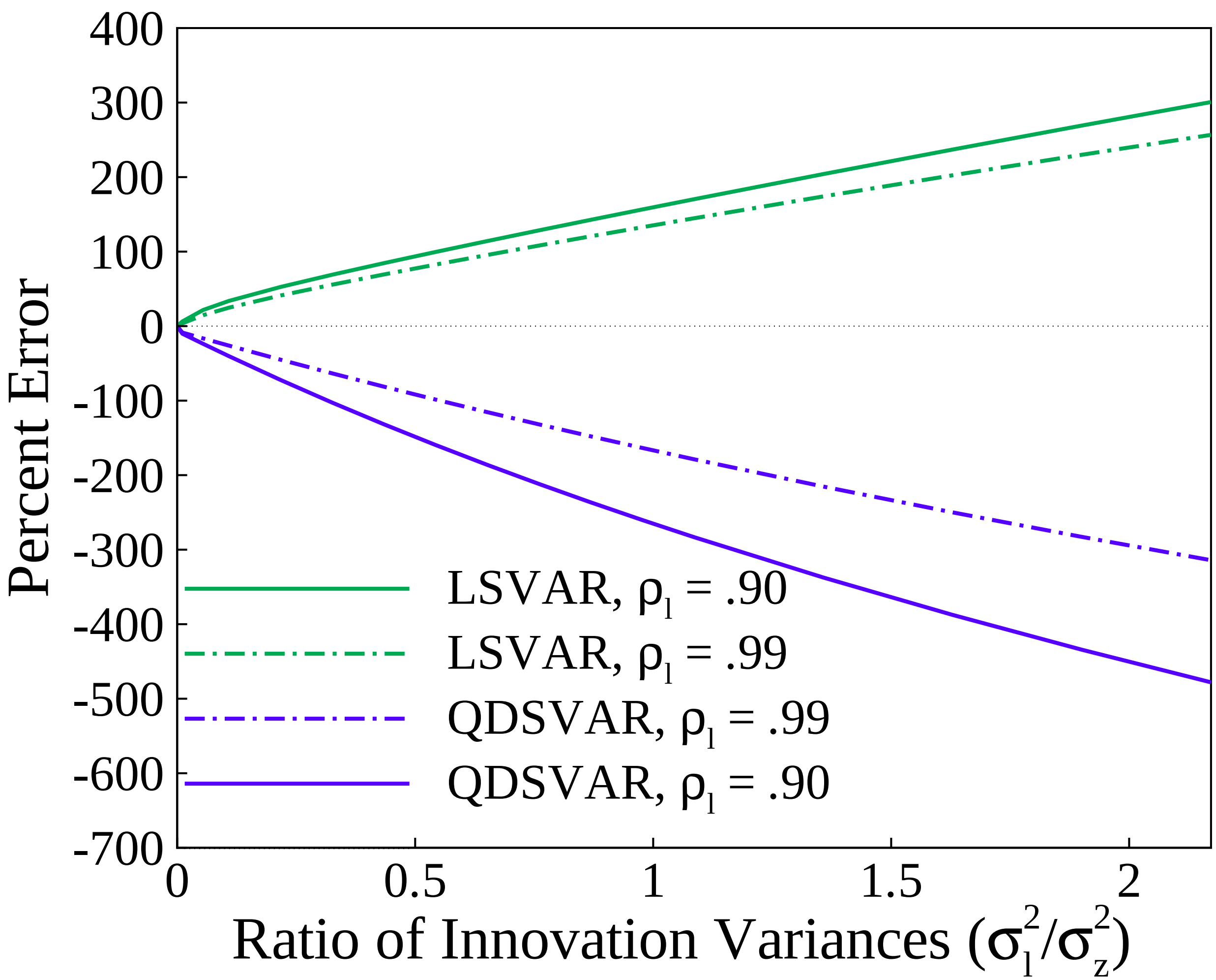
- Next, explore size of errors as we vary these parameters

Vary the Capital Share



% error in 4-lag SVAR response on impact relative to the model's response

Vary the Stochastic Processes



% error in 4-lag SVAR response on impact relative to the model's response is LARGE for a wide range of processes.

Less Consensus on Shock Processes

- Wide range of estimates for fraction of output variance due to technology
  - McGrattan (1994): total output variance due to technology

    41% with standard error 46%

  - Eichenbaum (1991): model HP variance/data HP variance

    5% to 200% for model with technology shocks only

  - Gali-Rabanal (2004): business cycle component due to technology

    LSVAR: 3% to 37%

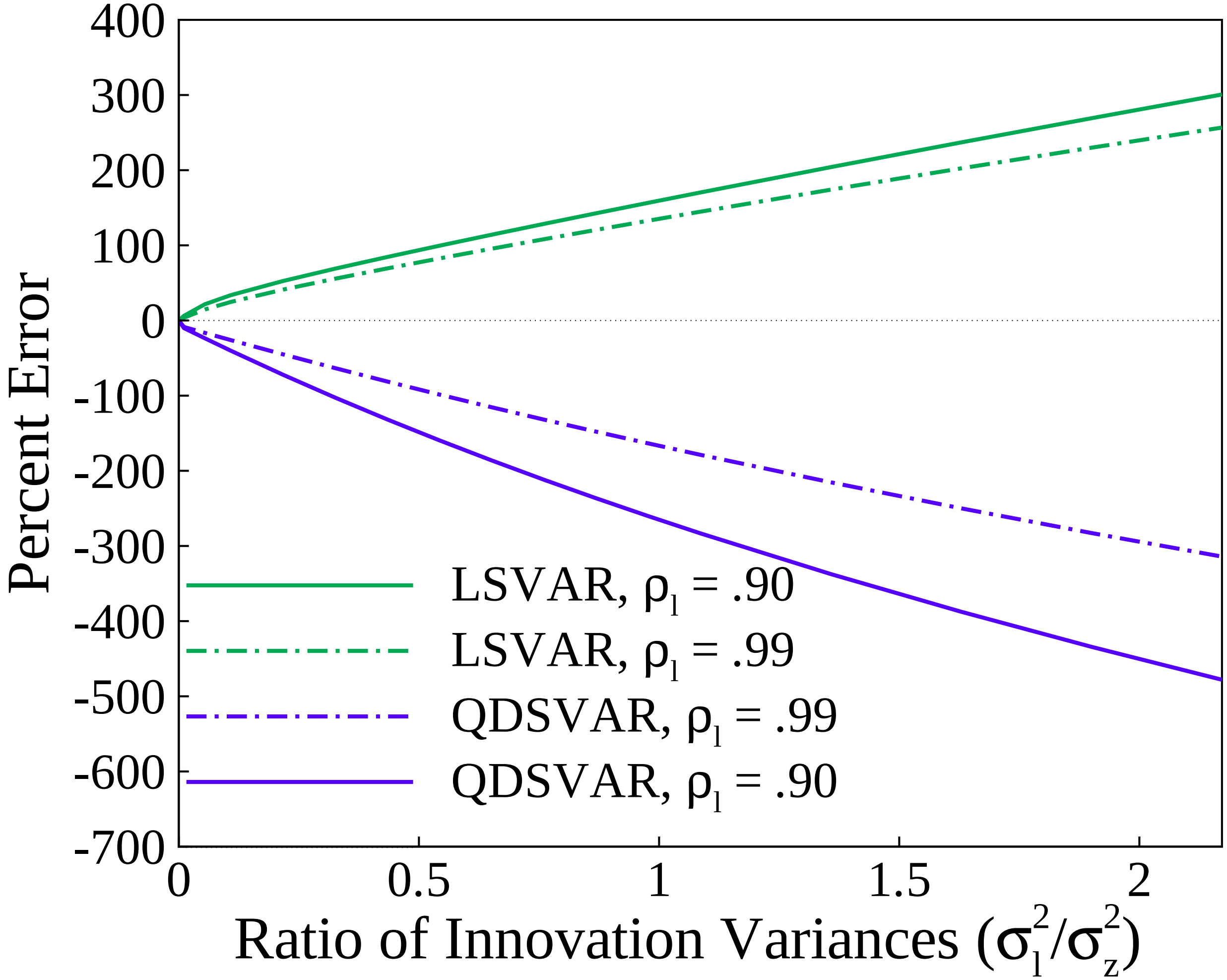
    DSVAR: 6% to 31%

Summary of the Evidence

What the data are actually telling us is that, while technology shocks almost certainly play some role in generating the business cycle, there is simply an enormous amount of uncertainty about just what percent of aggregate fluctuations they actually do account for. The answer could be 70% as Kydland and Prescott (1989) claim, but the data contain almost no evidence against either the view that the answer is really 5% or that the answer is 200%.

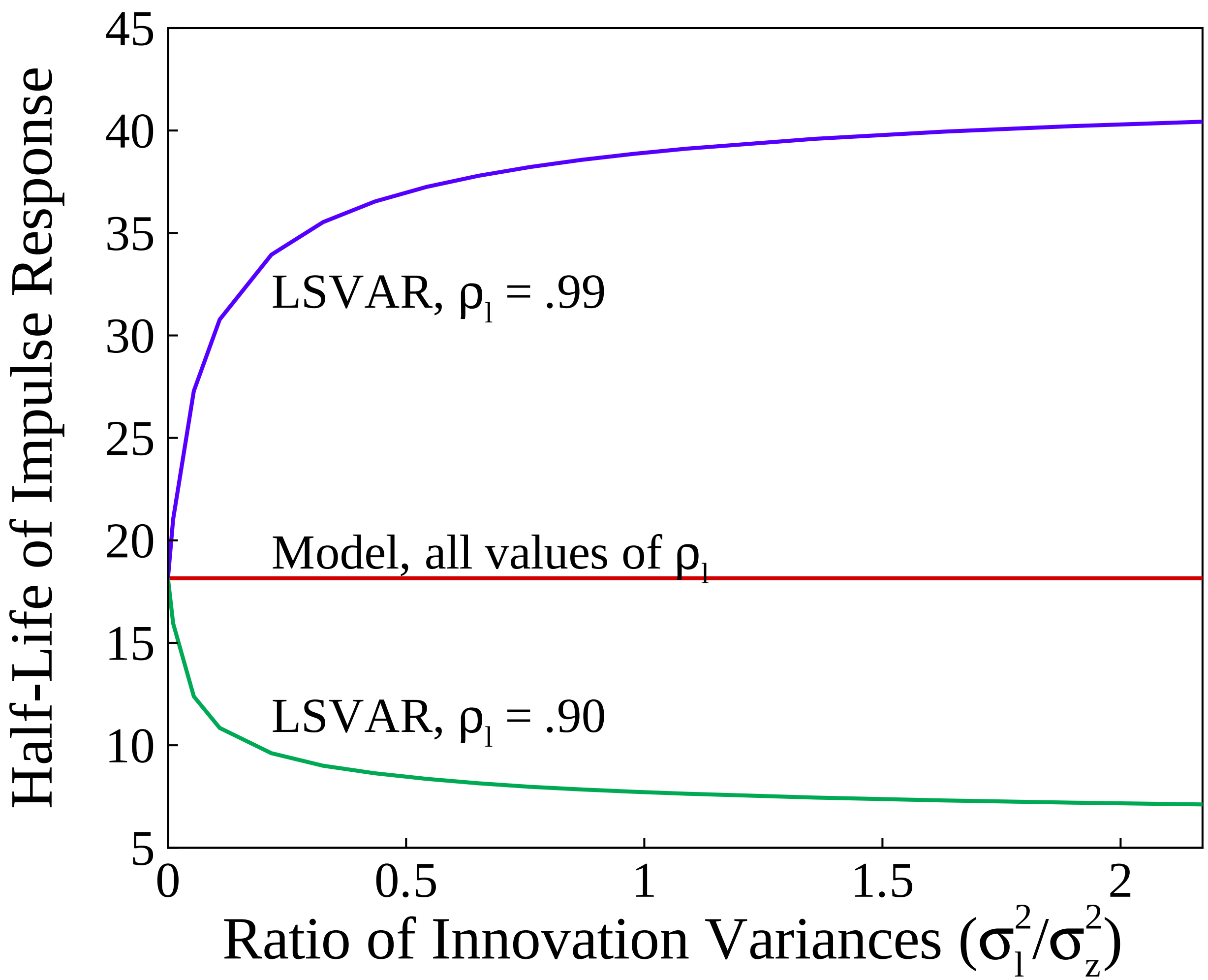
— Martin Eichenbaum, 1991, J. of Economic Dynamics and Control

Vary the Stochastic Processes



% error in 4-lag SVAR response on impact relative to the model's response is LARGE for a wide range of processes.

Half-Life Error Also Large for Wide Range of Processes



Wide range for 4-lag LSVAR half-life estimates which should not vary in theory (some not shown).

Sensitivity of SVAR Error to Estimation Procedure

- Estimation Procedures

  1. Baseline: data on (y, l, x, g) in 2-shock model

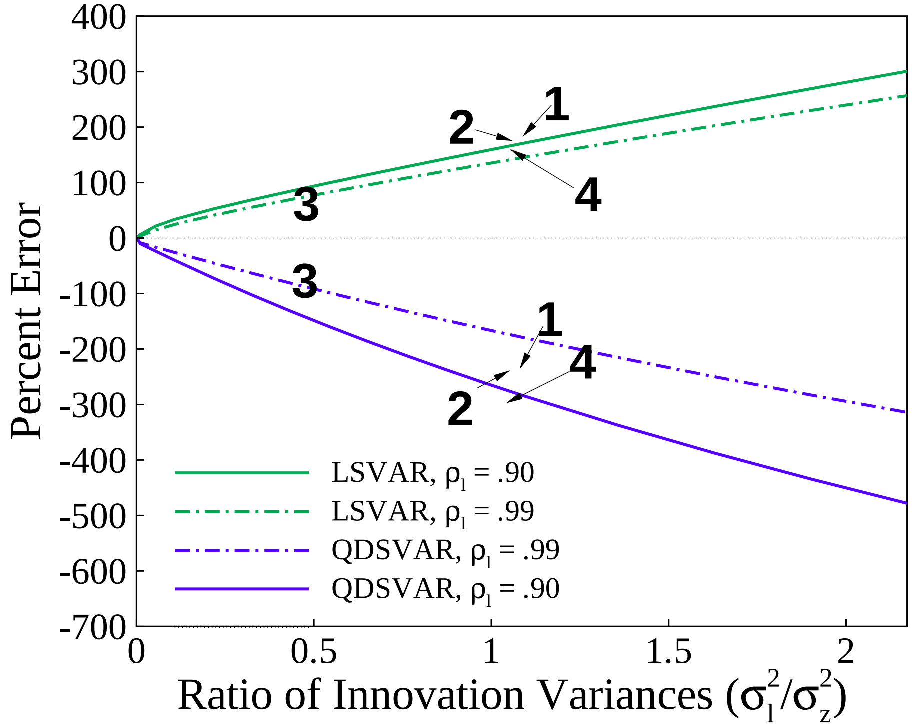
  2. Tiny measurement error: $\frac{1}{1000} \times$ observed variance rather than $\frac{1}{100}$

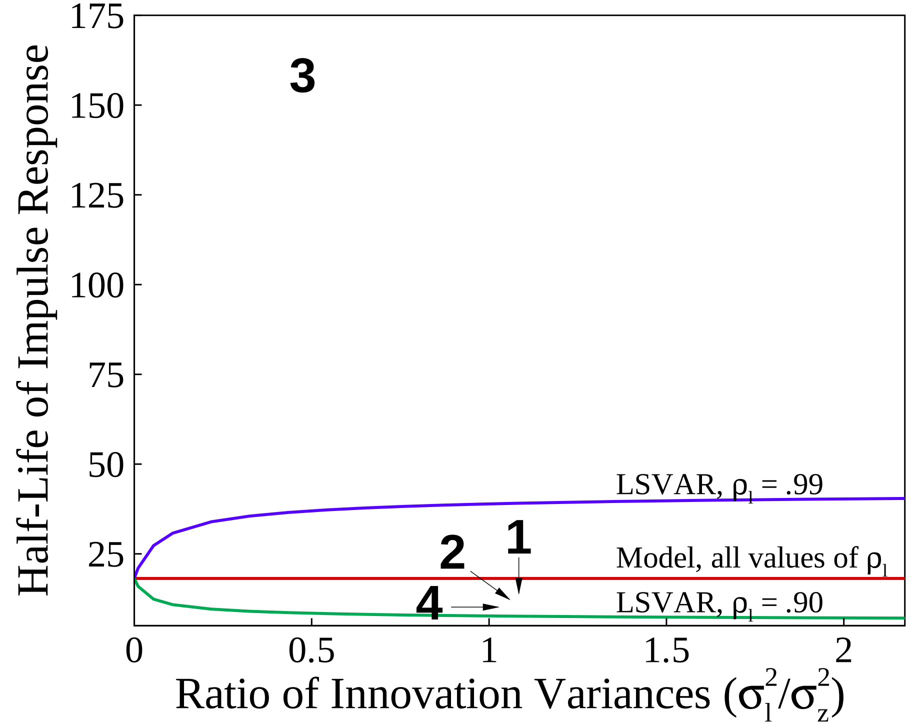
  3. Restricted observer: data on (y, l) in 2-shock model

  4. Govt. consumption: data on (y, l, x, g) in 3-shock model

- Are the errors large in all cases?

Yes, the Errors Are Large for All MLE Estimates





Recap

- Used model satisfying SVAR key assumptions
- SVAR procedures doesn't robustly uncover model's impulse responses
- Short lag length is the problem

Conjectured Solutions to Short Lag Length Problem

Use More Theory

- Business cycle models have state space representations:

$$\log \hat{k}_{t+1} = \gamma_k \log \hat{k}_t + \gamma_s' s_t$$
$$s_{t+1} = P s_t + Q \eta_{t+1}$$

  with shocks s_t (e.g., $\log z_t$, τ_{lt}). Write as VAR:

$$S_{t+1} = F S_t + G \eta_{t+1}$$

- with the measurement equation

$$X_t = H S_t + \omega_t$$

  X_t are observations, S_t is the state, and ω_t is measurement error

- Why not estimate this system as opposed to various VARs?

Throw in More Variables

- Conjecture: adding investment-output ratio fixes short lag problem

- We find: It does not

- Demonstrate this with 3-variable system:

  - Add investment-output to SVAR variables

  - Add orthogonal AR(1) for government spending or investment tax

Long AR Needed Since Model Has Infinite-order AR

- Proposition 4: Model has VAR coefficients B_j such that

$$B_j = MB_{j-1}, \quad j \geq 2,$$

  M has eigenvalues equal to 0, α (the differencing parameter), and

$$\lambda = \frac{1-\delta}{1+g_y}$$

  where g_y is growth rate of output.

- Eigenvalue $\lambda = .98$ for our parameters

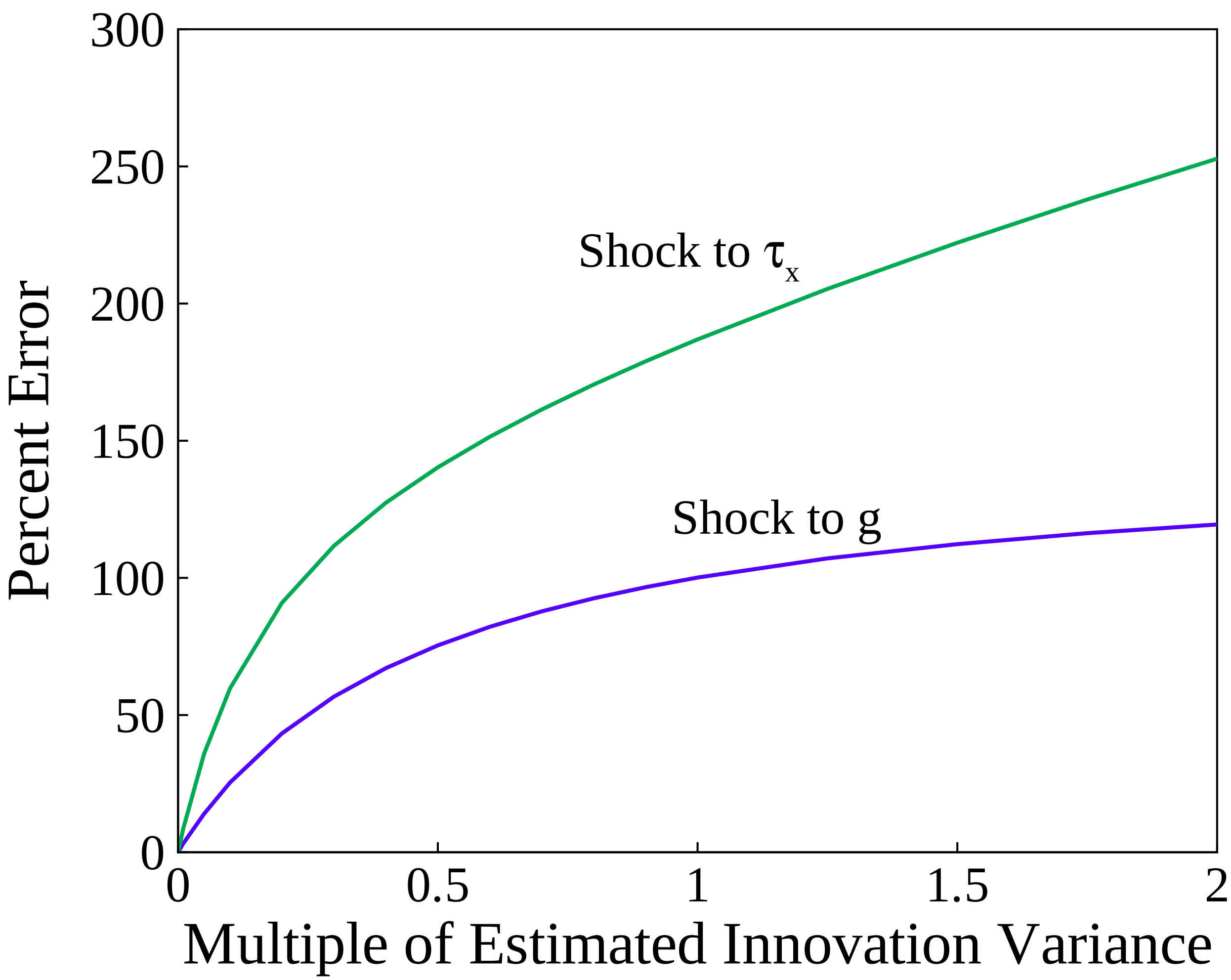
Generalization of Special Cases When Works

- Recall special cases for bivariate SVAR: no capital or one shock

- Singularity rule of thumb in general:
  - SVAR with few lags uncovers truth if

$$
\begin{aligned}
&\#\text{ of singularities in decay matrix } M \\
&\qquad + \#\text{ of singularities in shock covariance } \Omega \\
&\qquad\qquad \geq \#\text{ of variables in VAR}
\end{aligned}
$$

Need Close to Singular Variance-Covariance



Error large even if third shock contributes little to business cycles.

Recap

- Short lag length leads to large errors in SVARs

- Problem not fixed easily given available data

- Adding investment-output ratio does not fix the problem

- More variables than shocks only works if shock **variances** negligible.

Small Sample Properties

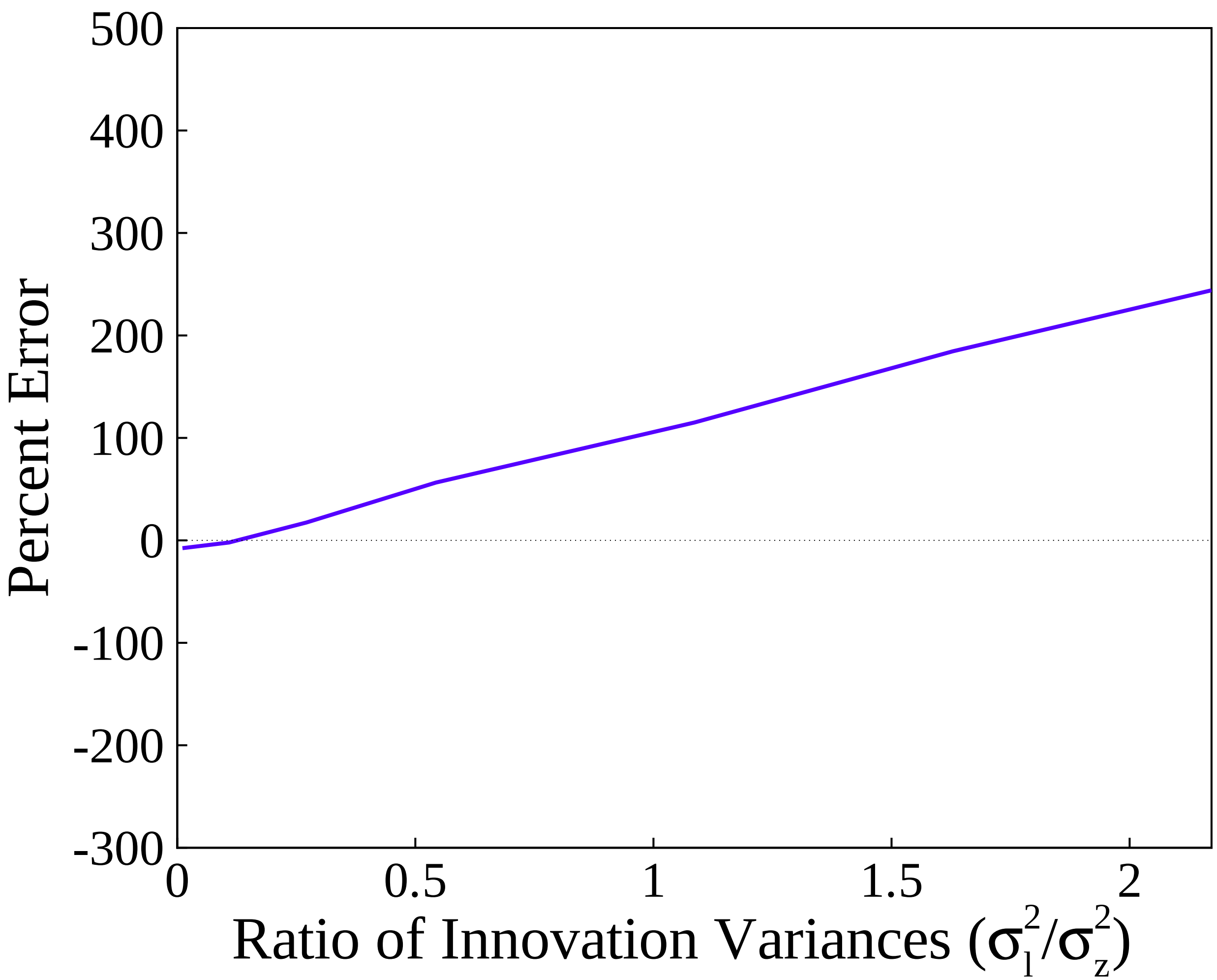
SVAR Procedure in Small Sample

- Found large errors in population for wide region of parameter space

- What about in small sample?

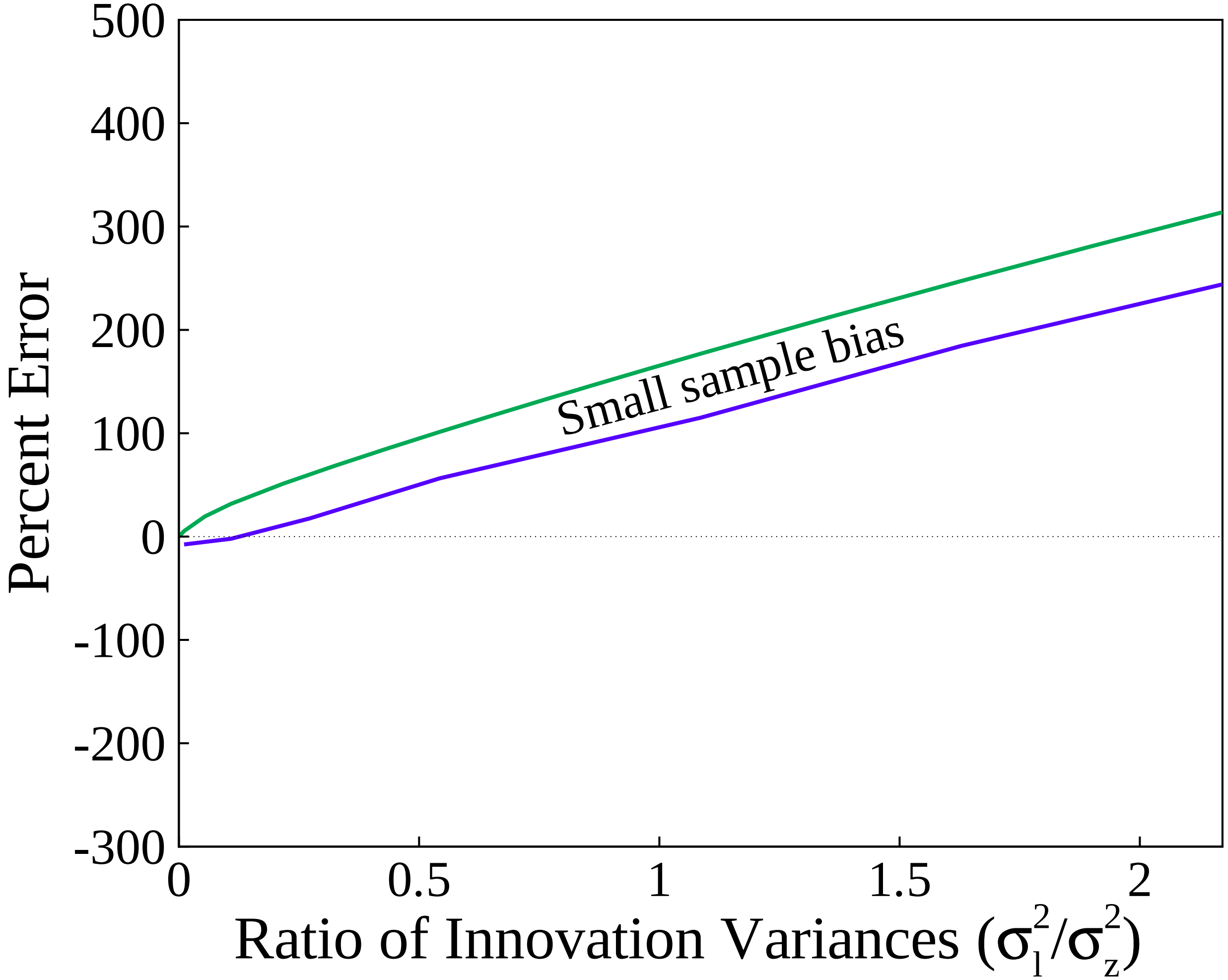
  - for RBC model, draw 1000 sequences of η of length 180
  - use model to derive 1000 sequences for productivity and hours
  - apply SVAR procedure to each dataset
  - compute bootstrapped confidence bands
  - repeat for **wide region** of parameter space

LSVAR Small Sample Specification Error



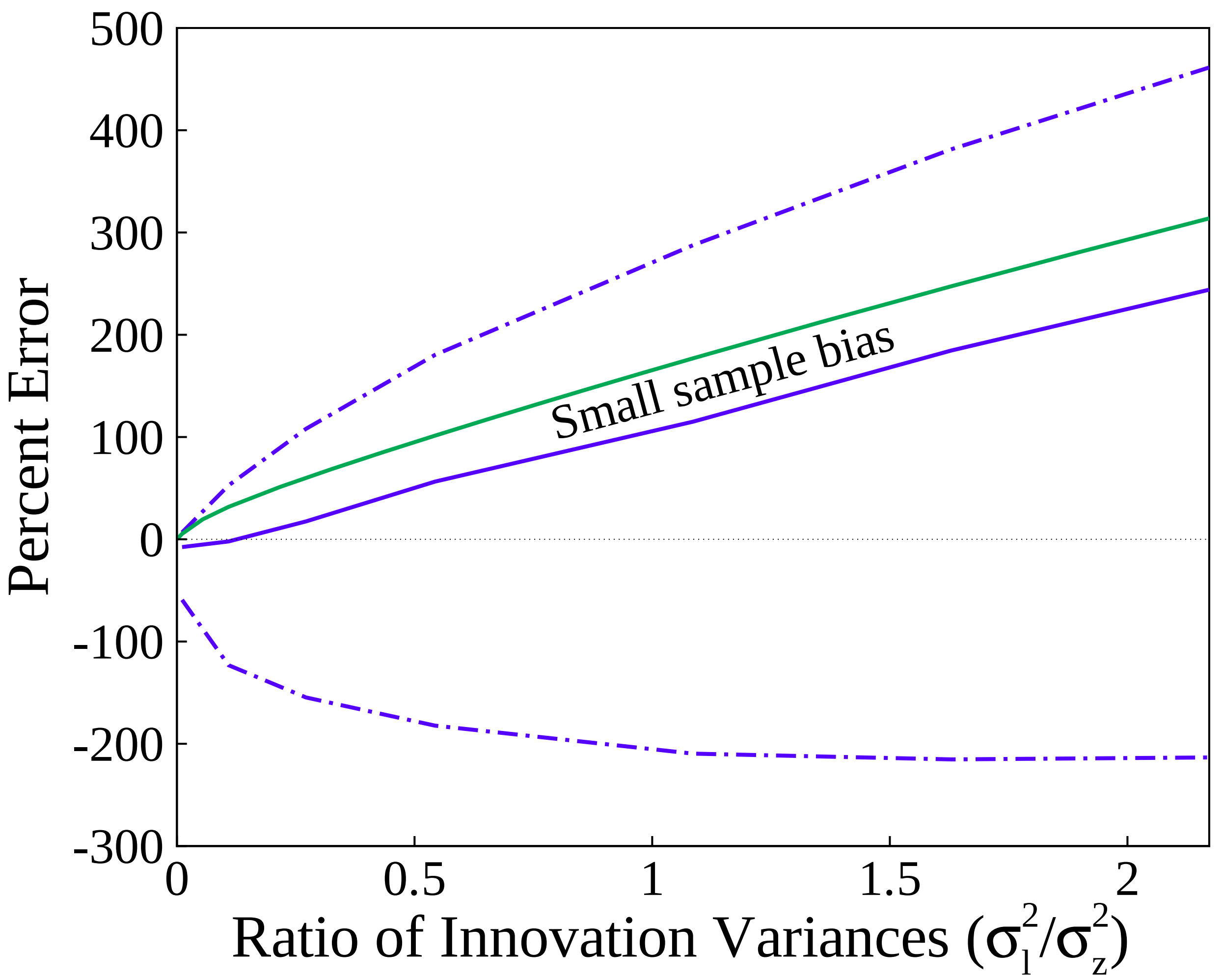
% error in 4-lag LSVAR response on impact relative to model's response, averaged over 1000 impulse responses.

LSVAR Small Sample Specification Error



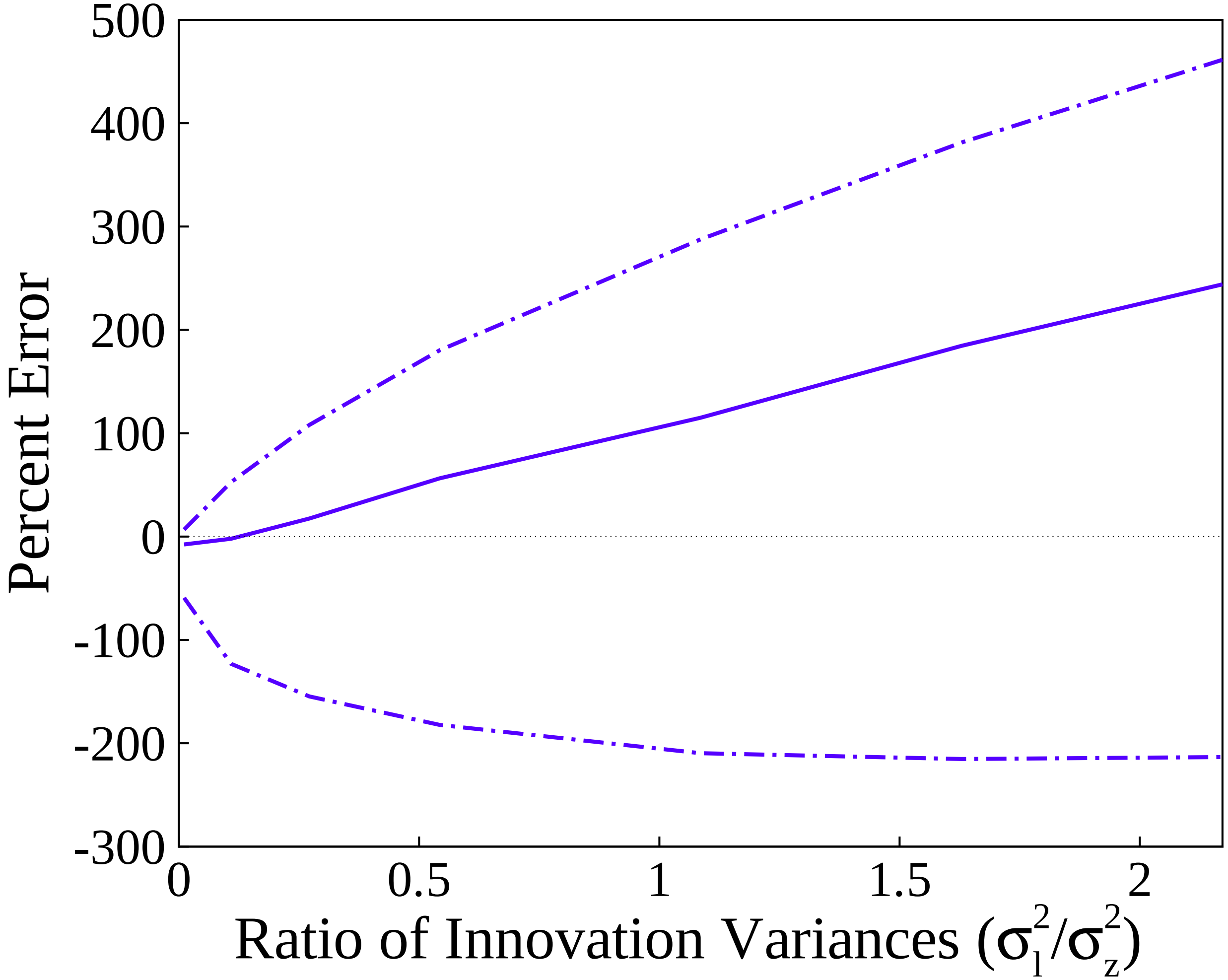
Compare to the population specification error — both large.

LSVAR Small Sample Specification Error



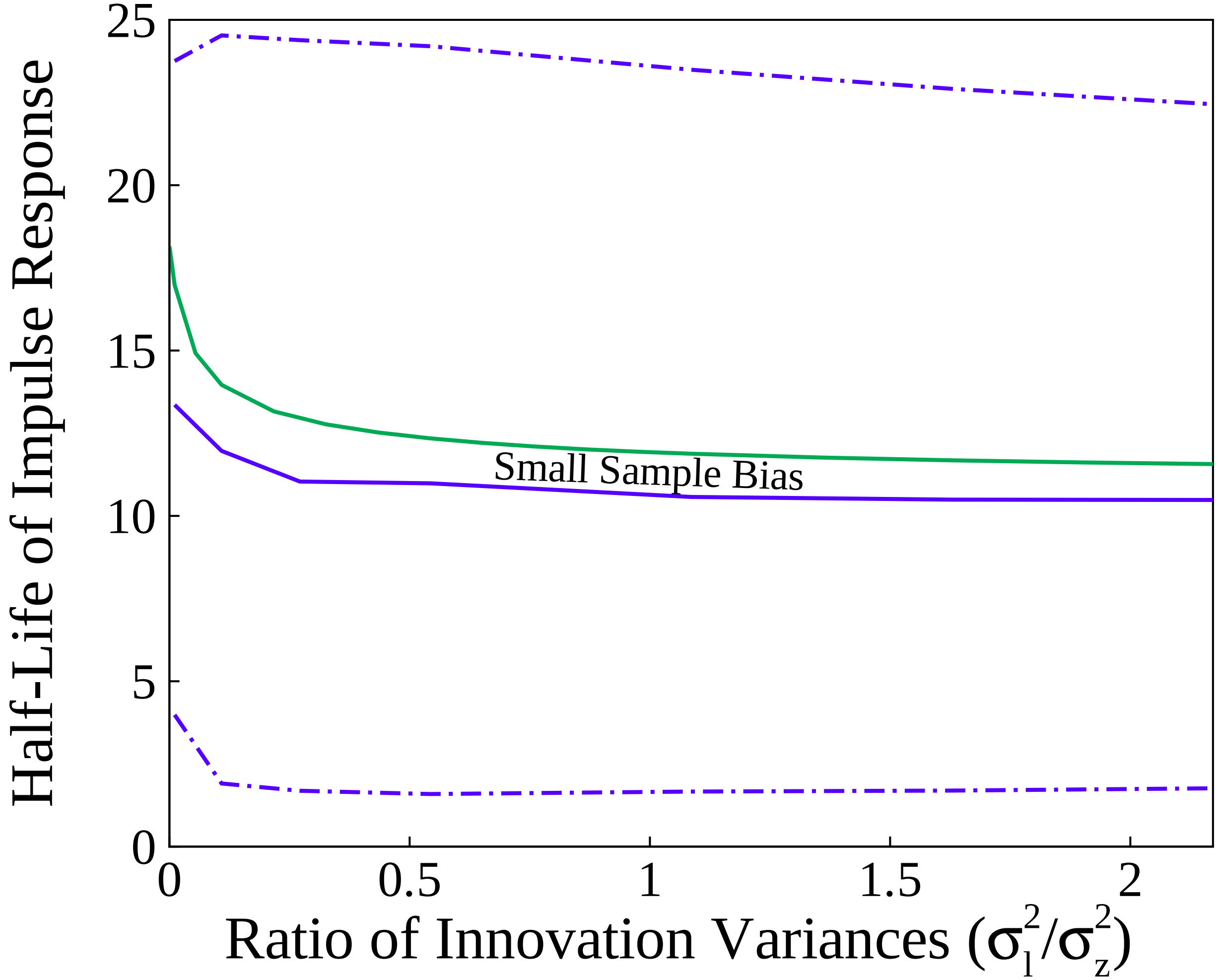
Now include confidence bands.

LSVAR Small Sample Specification Error



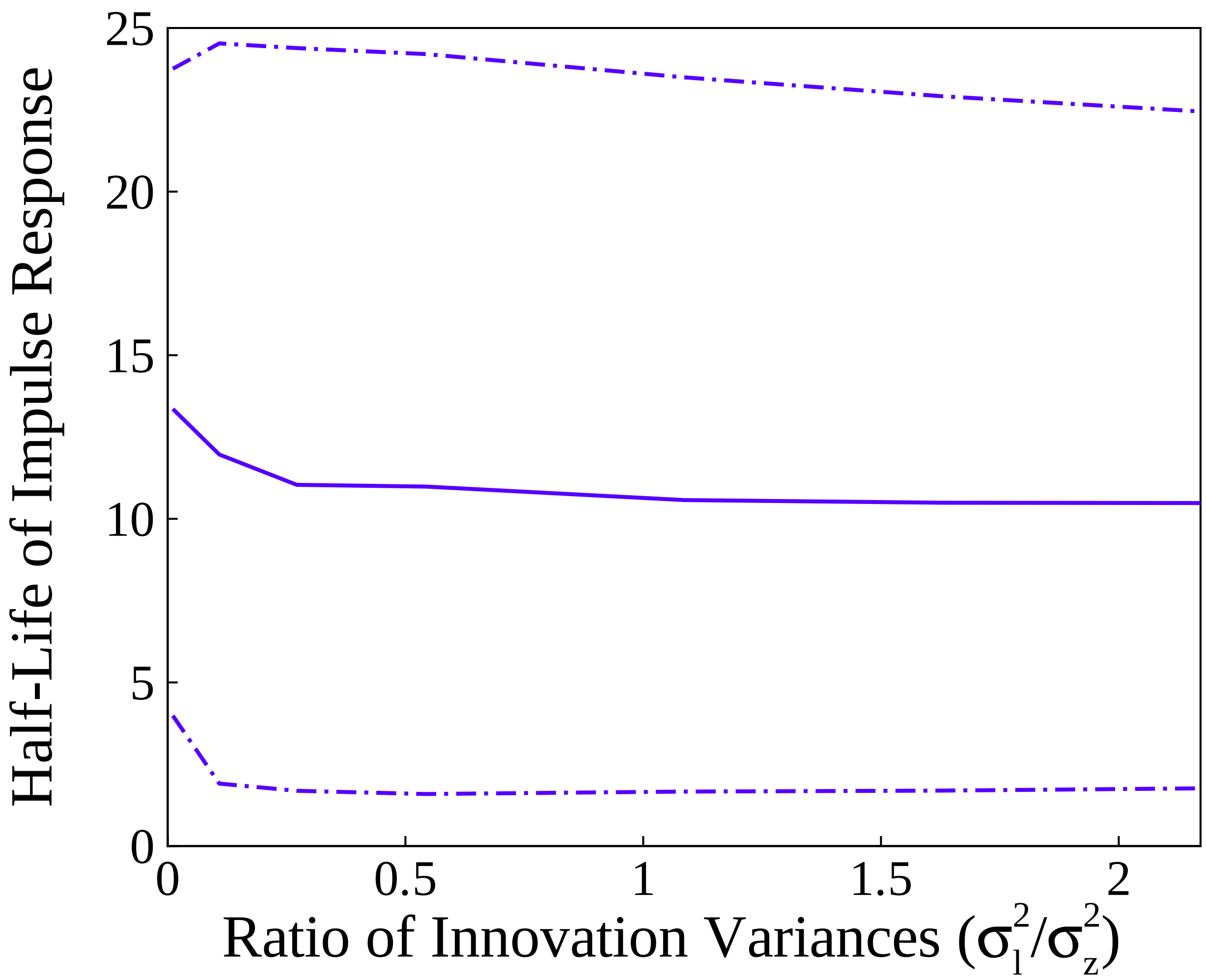
% error in 4-lag LSVAR response on impact relative to model's response is LARGE for wide range of variances.

Half-Life Error Also Large for Wide Range



Do the same thing for LSVAR half-life.

Half-Life Error Also Large for Wide Range



LARGE range (1/2 to 6 years) for LSVAR half-life estimates which should not vary in theory.

LSVAR Extremely Sensitive to Sample Path

- Drives large confidence bands
- Leads to wildly different conclusions
- Reminiscent of the findings in the literature

LSVAR Sensitivity to Small Variations in US Measures

- Three different researchers running an LSVAR with US data:

  - Francis and Ramey using

    Business productivity and demographically adjusted hours

  - Christiano, Eichenbaum, and Vigfusson using

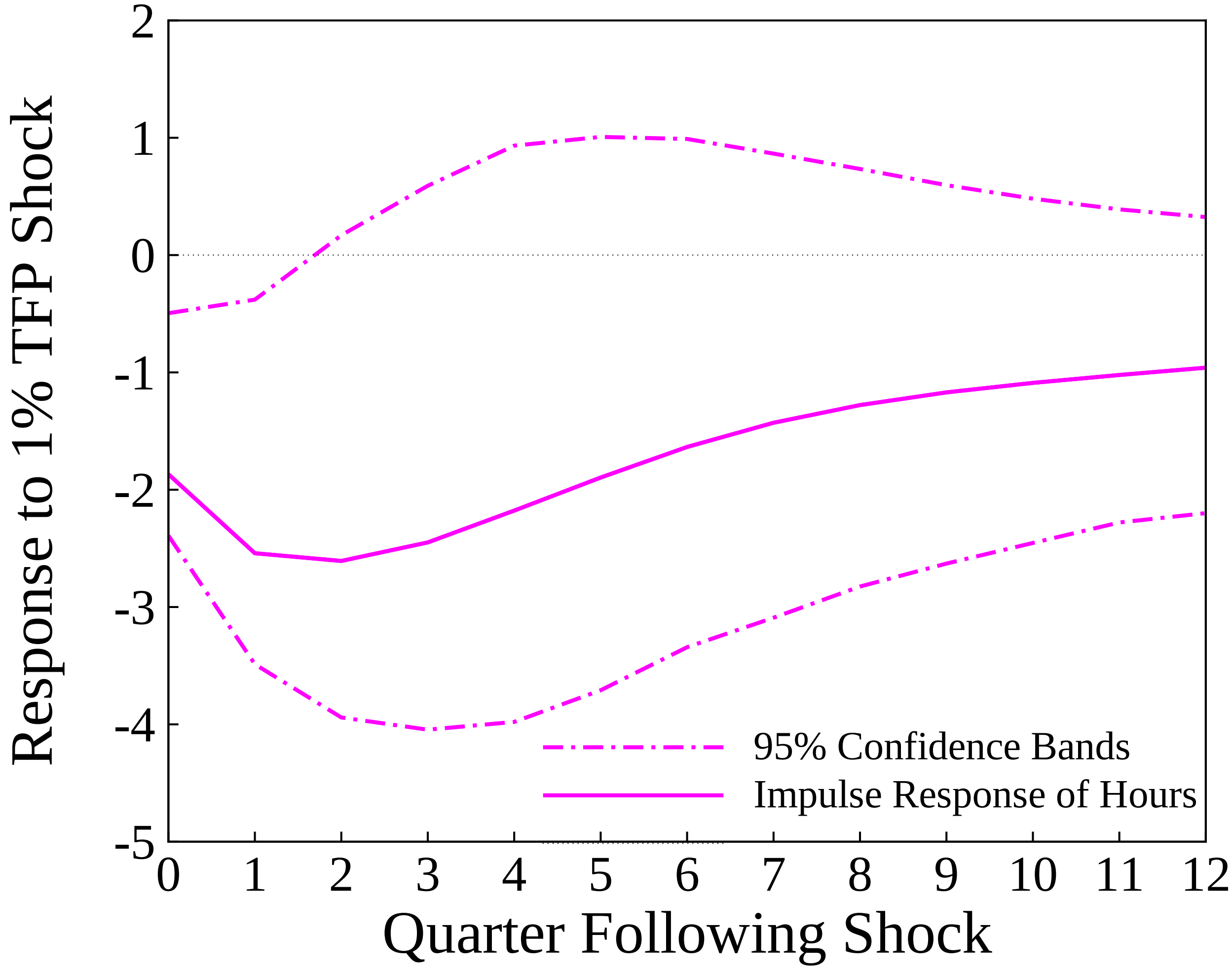
    Business productivity and hours

  - Gali and Rabanal using

    Nonfarm business productivity and hours

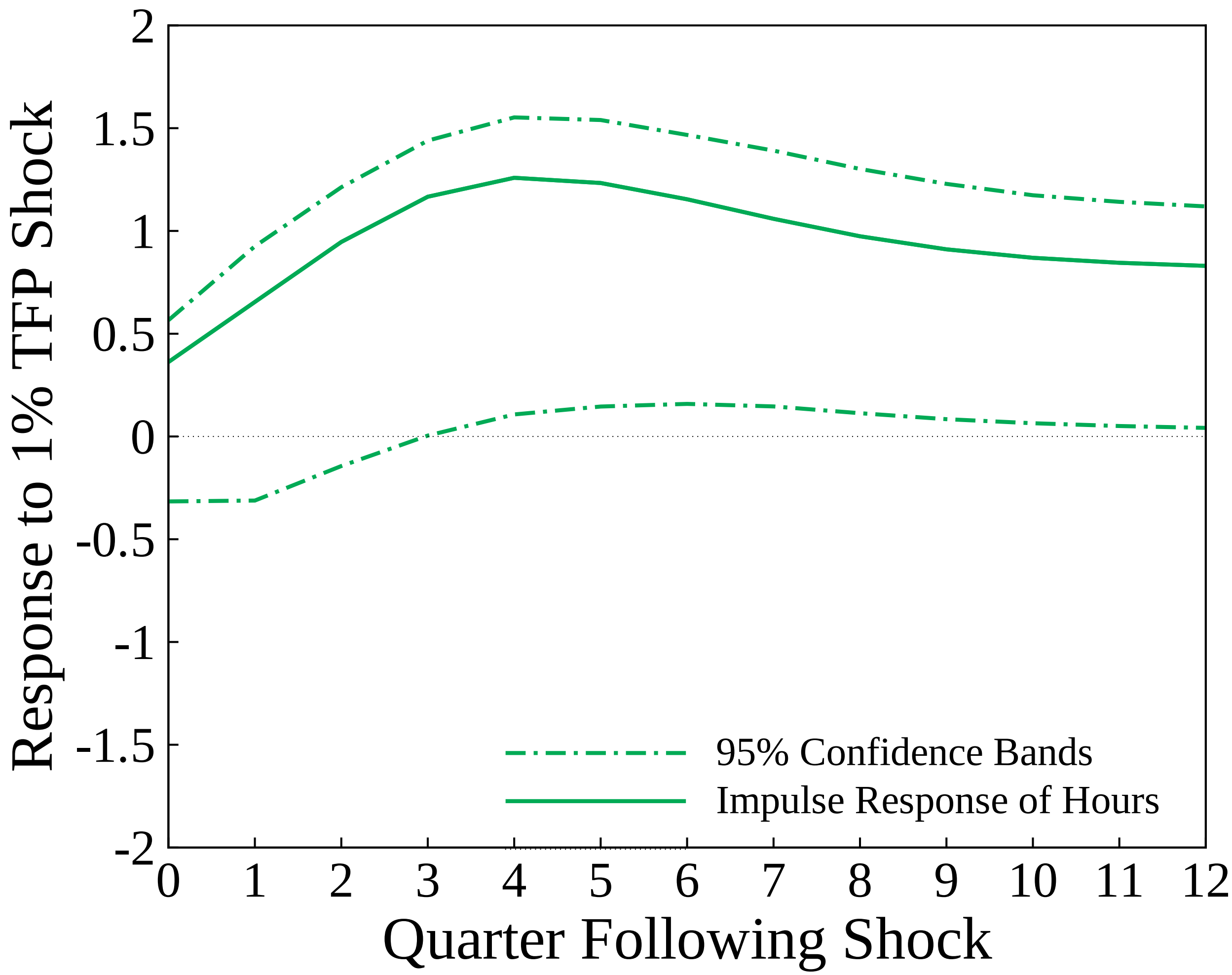
  come to wildly different conclusions

LSVAR Results Not Robust



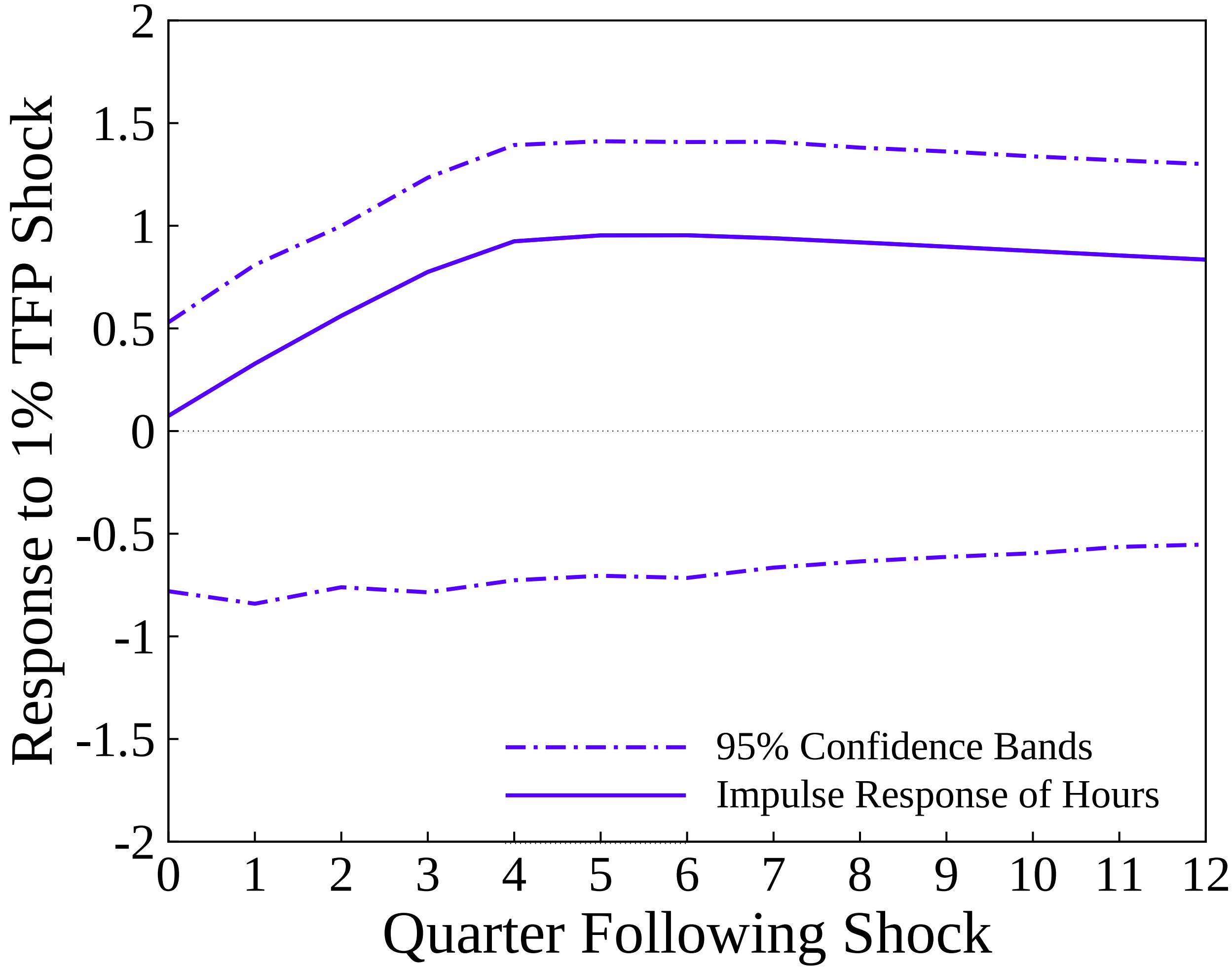
Francis-Ramey infer that the data did not come from an RBC model.

LSVAR Results Not Robust



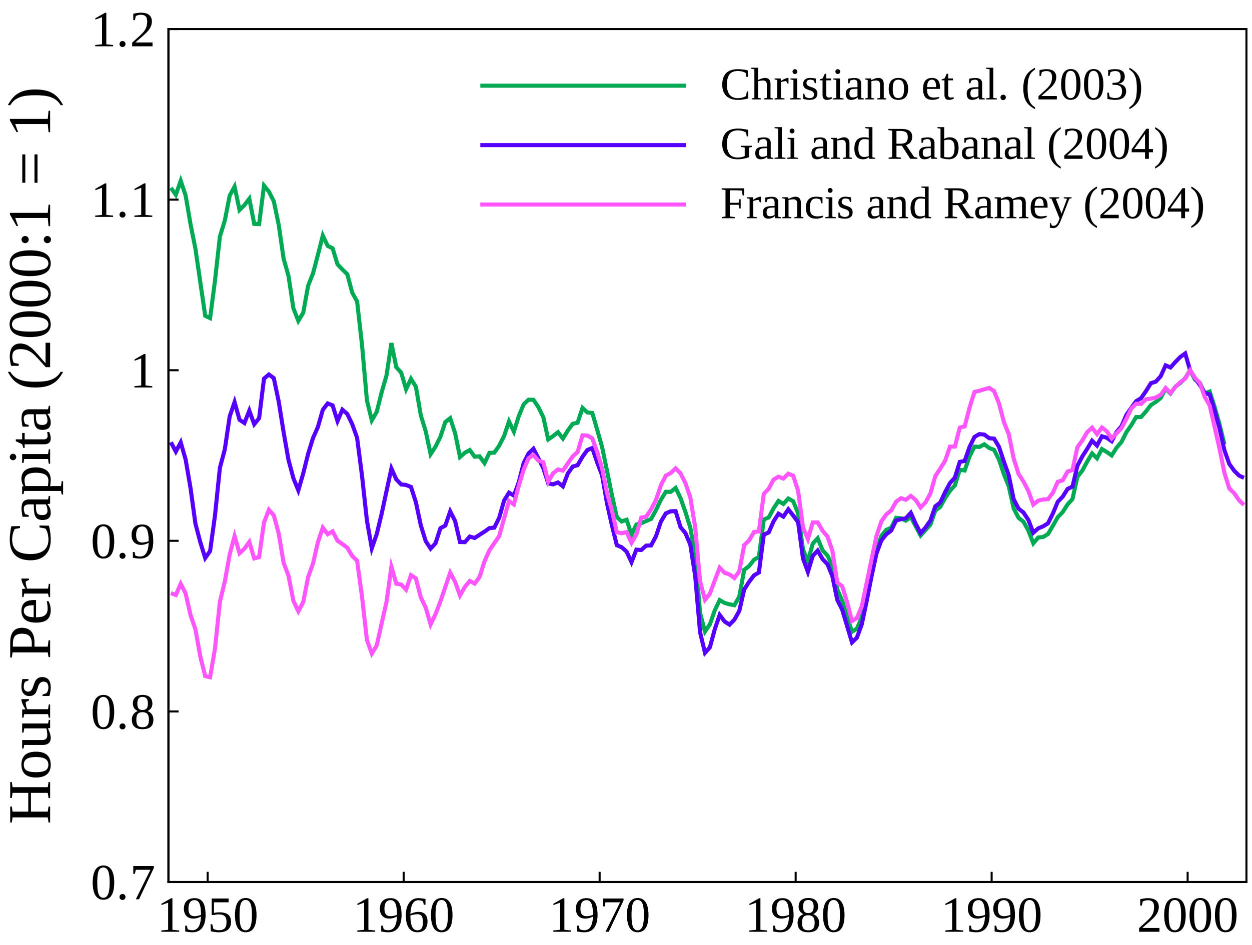
Christiano et al. conclude that RBC theory is alive and well.

LSVAR Results Not Robust



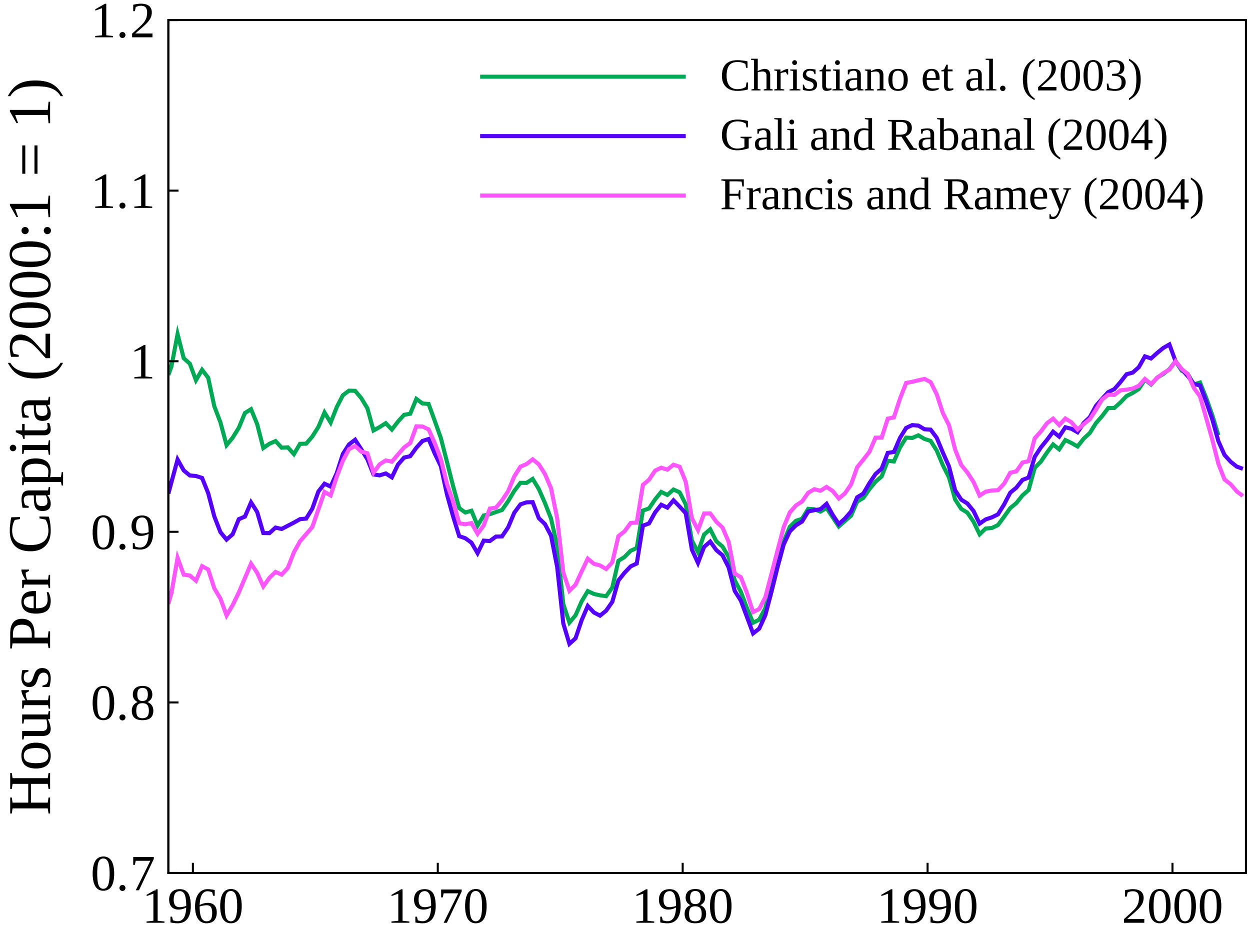
Gali-Rabanal conclude that the results are inconclusive.

LSVAR Sensitive to Data Inputs



Only difference is data inputs, not method.

LSVAR Sensitive Even in Shorter Sample



But even with a shorter sample, reach very different conclusions.

Punchline

- Class of counterexamples to main claim of SVAR literature
  - SVARs fail even weak test
  - Propositions provide conditions when SVAR works

    In models without capital

    In models with singular variance-covariance matrix
  - Quantitatively, errors grow with importance of nontechnology shocks

My New View

- "SVAR fact" is bad language
- SVARs are **not robust** and therefore are not useful guides for theory
- To be useful, must pass strong test

Absent a greater willingness to engage in empirical fragility analysis, structural empirical work will simply cease to be relevant. We may continue to publish, but our influence will surely perish.

Absent a greater willingness to engage in empirical fragility analysis, structural empirical work will simply cease to be relevant. We may continue to publish, but our influence will surely perish.

— Martin Eichenbaum, 1991, J. of Economic Dynamics and Control