A Full-Information Approach to Granular Instrumental Variables

Christiane Baumeister, University of Notre Dame
James D. Hamilton, UCSD

Differences between local and aggregate outcomes can be an important source of identification.

Examples:

- Bartik instruments
- Granular instrumental variables (Gabaix and Koijen, 2023)

- Our paper shows how to exploit the power of this idea using full-information maximum likelihood estimation.
- We illustrate with an analysis of the world oil market.

A model of the world oil market

Data from 1973:M1 to 2023:M2 (drop COVID)

q_{it} = growth rate of country i oil production

s_{qi} = share of country i in world total

$\sum_{i=1}^{n} s_{qi} q_{it}$ = approximate growth in global oil production

Our empirical analysis will use the three biggest producers (U.S., Saudi Arabia, Russia) plus the rest of the world $(n = 4)$

c_{jt} = growth rate of country j oil consumption

s_{cj} = share of country j in world total

$\sum_{j=1}^{m} s_{cj} c_{jt}$ = approximate growth in global oil consumption

Our empirical analysis will use the three biggest consumers (U.S., Japan, Europe) plus the rest of the world $(m = 4)$

Supply curve of country i

$$q_{it} = \phi_{qi} p_t + \mathbf{b}'_{qi} \mathbf{x}_{t-1} + u_{qit} + u_{\chi it}$$

ϕ_{qi} = country i short-run supply elasticity

$\mathbf{x}_{t-1}$ contains 12 lags production and consumption of every country in world plus 12 lags of world price

u_{qit} = supply shock for country i

$u_{\chi it}$ = error in measuring country i production

Demand curve of country j

$$c_{jt} = \phi_{cj} p_t + \mathbf{b}'_{cj} \mathbf{x}_{t-1} + u_{cjt} + u_{\psi jt}$$

ϕ_{cj} = country j short-run demand elasticity

u_{cjt} = demand shock for country j

$u_{\psi jt}$ = error in measuring country j consumption

Inventory demand

$$v_t = \phi_v p_t + \mathbf{b}_v' \mathbf{x}_{t-1} + u_{vt}$$

This equals difference between correctly measured production and consumption

$$v_t = \sum_{i=1}^{n} s_{qi}(q_{it} - u_{\chi it}) - \sum_{j=1}^{m} s_{cj}(c_{jt} - u_{\psi jt})$$

Structural model:

$$q_{it} = \phi_{qi} p_t + \mathbf{b}'_{qi} \mathbf{x}_{t-1} + u_{qit} + u_{\chi it} \qquad i = 1, \ldots, n$$

or $\underset{(n\times 1)}{\mathbf{q}_t} = \underset{(n\times 1)}{\boldsymbol{\phi}_q} \; p_t + \underset{(n\times k)}{\mathbf{B}_q} \; \mathbf{x}_{t-1} + \underset{(n\times 1)}{\mathbf{u}_{qt}} + \underset{(n\times 1)}{\mathbf{u}_{\chi t}}$

$$c_{jt} = \phi_{cj} p_t + \mathbf{b}'_{cj} \mathbf{x}_{t-1} + u_{cjt} + u_{\psi jt} \qquad j = 1, \ldots, m$$

or $\underset{(m\times 1)}{\mathbf{c}_t} = \underset{(m\times 1)}{\boldsymbol{\phi}_c} \; p_t + \underset{(m\times k)}{\mathbf{B}_c} \; \mathbf{x}_{t-1} + \underset{(m\times 1)}{\mathbf{u}_{ct}} + \underset{(m\times 1)}{\mathbf{u}_{\psi t}}$

$$(\mathbf{s}'_q \boldsymbol{\phi}_q - \mathbf{s}'_c \boldsymbol{\phi}_c - \phi_v) p_t =$$

$$(\mathbf{s}'_c \mathbf{B}_c - \mathbf{s}'_q \mathbf{B}_q + \mathbf{b}'_v) \mathbf{x}_{t-1} + \mathbf{s}'_c \mathbf{u}_{ct} - \mathbf{s}'_q \mathbf{u}_{qt} + u_{vt}$$

These are $n + m + 1$ equations to determine the $n + m + 1$ variables $\mathbf{q}_t, \mathbf{c}_t, p_t$ in terms of the structural shocks $(\mathbf{u}_{qt}, \mathbf{u}_{\chi t}, \mathbf{u}_{ct}, \mathbf{u}_{\psi t}, u_{vt})$

It is possible to estimate structural parameters like $\boldsymbol{\phi}_q$ and $\boldsymbol{\phi}_c$ if we make assumptions about the correlations between these structural shocks

Example 1: Granular instrumental variables

Suppose all countries have the same demand elasticity ϕ_c and that the demand shock for country j has an idiosyncratic component and a common global factor

$$u_{cjt} = f_{ct} + \eta_{cjt}$$

$$c_{jt} = \phi_c p_t + \mathbf{b}'_{cj}\mathbf{x}_{t-1} + f_{ct} + \eta_{cjt} + u_{\psi jt}$$

Take arithmetic average over $j = 1, \ldots, m$

$$\bar{c}_t = \phi_c p_t + \bar{\mathbf{b}}'_c\mathbf{x}_{t-1} + f_{ct} + \bar{\eta}_{ct} + \bar{u}_{\psi t}$$

$$c_{jt} = \phi_c p_t + \mathbf{b}'_{cj}\mathbf{x}_{t-1} + f_{ct} + \eta_{cjt} + u_{\psi jt}$$

$$\bar{c}_t = \phi_c p_t + \bar{\mathbf{b}}'_c\mathbf{x}_{t-1} + f_{ct} + \bar{\eta}_{ct} + \bar{u}_{\psi t}$$

Difference between country j and average,

$$c_{jt} - \bar{c}_t = (\mathbf{b}'_{cj} - \bar{\mathbf{b}}_c)'\mathbf{x}_{t-1} + (\eta_{cjt} - \bar{\eta}_{ct}) + (u_{\psi jt} - \bar{u}_{\psi t}),$$

depends only on idiosyncratic shocks

η_{cjt} for $j = 1, \ldots, m$ and measurement errors

If these are uncorrelated with supply shocks, $c_{jt} - \bar{c}_t$ is valid instrument for estimating any supply curve

More powerful instruments by multiplying by s_{cj} and summing over j:

$$c_t - \bar{c}_t = (\mathbf{b}_c' - \bar{\mathbf{b}}_c)'\mathbf{x}_{t-1} + (\eta_{ct} - \bar{\eta}_{ct}) + (u_{\psi t} - \bar{u}_{\psi t})$$

Implication: difference between share-weighted consumption $c_t = \sum_{j=1}^{m} s_{cj} c_{jt}$ and unweighted average $\bar{c}_t = m^{-1} \sum_{j=1}^{m} c_{jt}$ is valid instrument for estimating supply elasticities.

Example 2: Uncorrelated supply and demand shocks

If $\mathbf{u}_{qt} + \mathbf{u}_{\chi t}$ is uncorrelated with $\mathbf{u}_{ct} + \mathbf{u}_{\psi t}$ then

$$E(\mathbf{q}_t - \boldsymbol{\phi}_q p_t - \mathbf{B}_q \mathbf{x}_{t-1})(\mathbf{c}_t - \boldsymbol{\phi}_c p_t - \mathbf{B}_c \mathbf{x}_{t-1})' = \mathbf{0}_{nm}$$

$\hat{\varepsilon}_{qit}$ = residual from OLS regression of q_{it} on $\mathbf{x}_{t-1}$

$$T^{-1} \sum_{t=1}^{T} (\hat{\varepsilon}_{qit} - \phi_{qi} \hat{\varepsilon}_{pt})(\hat{\varepsilon}_{cjt} - \phi_{cj} \hat{\varepsilon}_{pt}) = 0$$

$$i = 1, \ldots, n; \; j = 1, \ldots, m$$

This gives (mn) equations to estimate

$n + m$ values of $\boldsymbol{\phi}_q$ and $\boldsymbol{\phi}_c$.

However, these overidentifying restrictions are rejected in our dataset.

A less restrictive model

We allow for $\mathbf{u}_{qt}$ and $\mathbf{u}_{ct}$ to be correlated through common dependence on a single global factor f_t

We also allow $\mathbf{u}_{ct}$ to depend on idiosyncratic factors and a global demand shock f_{ct} with different loadings for each country

$$\mathbf{u}_{ct} = \mathbf{h}_c f_t + \boldsymbol{\gamma}_c f_{ct} + \boldsymbol{\eta}_{ct}$$

Reduced form is a VAR(12)

$$\underset{(9\times 1)}{\mathbf{y}_t} = (\mathbf{q}_t', \mathbf{c}_t', p_t)'$$

$$\mathbf{y}_t = \Pi \mathbf{x}_{t-1} + \boldsymbol{\epsilon}_t$$

$$\mathbf{x}_{t-1} = (1, \mathbf{y}_{t-1}', \ldots, \mathbf{y}_{t-12}')'$$

$$
\begin{bmatrix} \mathbf{I}_n & \mathbf{0}_{nm} & -\boldsymbol{\phi}_q \\ \mathbf{0}_{mn} & \mathbf{I}_m & -\boldsymbol{\phi}_c \\ \mathbf{0}_{1n} & \mathbf{0}_{1m} & 1 \end{bmatrix}
\begin{bmatrix} \boldsymbol{\epsilon}_{qt} \\ \boldsymbol{\epsilon}_{ct} \\ \varepsilon_{pt} \end{bmatrix} =
$$

$$
\begin{bmatrix} \mathbf{h}_q f_t + \boldsymbol{\gamma}_q f_{qt} + \boldsymbol{\eta}_{qt} + \mathbf{u}_{\chi t} \\ \mathbf{h}_c f_t + \boldsymbol{\gamma}_c f_{ct} + \boldsymbol{\eta}_{ct} + \mathbf{u}_{\psi t} \\ \alpha \mathbf{s}'_c(\mathbf{h}_c f_t + \boldsymbol{\gamma}_c f_{ct} + \boldsymbol{\eta}_{ct}) - \alpha \mathbf{s}'_q(\mathbf{h}_q f_t + \boldsymbol{\gamma}_q f_{qt} + \boldsymbol{\eta}_{qt}) + \alpha u_{vt} \end{bmatrix}
$$

$$
\mathbf{A}\boldsymbol{\epsilon}_t = \mathbf{u}_t
$$

$$
E(\mathbf{u}_t \mathbf{u}'_t) = \mathbf{D}
$$

Assumptions

Idiosyncratic shocks uncorrelated

$$E\begin{bmatrix} \boldsymbol{\eta}_{qt} \\ \boldsymbol{\eta}_{ct} \end{bmatrix}\begin{bmatrix} \boldsymbol{\eta}'_{qt} & \boldsymbol{\eta}'_{ct} \end{bmatrix}$$

$$= \begin{bmatrix} \boldsymbol{\Sigma}_q & \mathbf{0}_{nm} \\ \mathbf{0}_{mn} & \boldsymbol{\Sigma}_c \end{bmatrix}$$ (diagonal)

Factor normalization

$$E\begin{bmatrix} f_t \\ f_{qt} \\ f_{ct} \end{bmatrix}\begin{bmatrix} f_t & f_{qt} & f_{ct} \end{bmatrix} = \mathbf{I}_3$$

Measurement errors have common variance

$$E\begin{bmatrix} \mathbf{u}_{\chi t} \\ \mathbf{u}_{\psi t} \end{bmatrix}\begin{bmatrix} \mathbf{u}'_{\chi t} & \mathbf{u}'_{\psi t} \end{bmatrix} = \begin{bmatrix} \sigma_\chi^2 \mathbf{I}_n & \mathbf{0}_{nm} \\ \mathbf{0}_{mn} & \sigma_\psi^2 \mathbf{I}_m \end{bmatrix}$$

$$\mathbf{D} = \begin{bmatrix} \mathbf{h}_q\mathbf{h}_q' + \boldsymbol{\gamma}_q\boldsymbol{\gamma}_q' + \boldsymbol{\Sigma}_q + \sigma_\chi^2\mathbf{I}_n & \mathbf{h}_q\mathbf{h}_c' & -\alpha\left(\mathbf{h}_q\mathbf{h}_q' + \boldsymbol{\gamma}_q\boldsymbol{\gamma}_q' + \boldsymbol{\Sigma}_q\right)\mathbf{s}_q + \alpha\mathbf{h}_q\mathbf{h}_c'\mathbf{s}_c \\ \mathbf{h}_c\mathbf{h}_q' & \mathbf{h}_c\mathbf{h}_c' + \boldsymbol{\gamma}_c\boldsymbol{\gamma}_c' + \boldsymbol{\Sigma}_c + \sigma_\psi^2\mathbf{I}_m & -\alpha\mathbf{h}_c\mathbf{h}_q'\mathbf{s}_q + \alpha(\mathbf{h}_c\mathbf{h}_c' + \boldsymbol{\gamma}_c\boldsymbol{\gamma}_c' + \boldsymbol{\Sigma}_c)\mathbf{s}_c \\ -\alpha\mathbf{s}_q'\left(\mathbf{h}_q\mathbf{h}_q' + \boldsymbol{\gamma}_q\boldsymbol{\gamma}_q' + \boldsymbol{\Sigma}_q\right) + \alpha\mathbf{s}_c'\mathbf{h}_c\mathbf{h}_q' & -\alpha\mathbf{s}_q'\mathbf{h}_q\mathbf{h}_c' + \alpha\mathbf{s}_c'(\mathbf{h}_c\mathbf{h}_c' + \boldsymbol{\gamma}_c\boldsymbol{\gamma}_c' + \boldsymbol{\Sigma}_c) & \alpha^2\left[\mathbf{s}_q'\left(\mathbf{h}_q\mathbf{h}_q' + \boldsymbol{\gamma}_q\boldsymbol{\gamma}_q' + \boldsymbol{\Sigma}_q\right)\mathbf{s}_q - 2\mathbf{s}_c'\mathbf{h}_c\mathbf{h}_q'\mathbf{s}_q + \mathbf{s}_c'(\mathbf{h}_c\mathbf{h}_c' + \boldsymbol{\gamma}_c\boldsymbol{\gamma}_c' + \boldsymbol{\Sigma}_c)\mathbf{s}_c + \sigma_v^2\right] \end{bmatrix}$$

Model has 15 testable overidentifying assumptions.

These are not rejected in our dataset.

Maximum likelihood estimates of elasticities and their standard errors

U.S. supply	0.019	(0.017)
Saudi supply	0.259	(0.056)
Russia supply	0.029	(0.011)
ROW supply	0.043	(0.029)
U.S. demand	-0.094	(0.031)
Japan demand	-0.018	(0.037)
Europe demand	-0.225	(0.045)
ROW demand	-0.161	(0.045)
Inventory demand	-0.314	(0.060)

Global supply elasticity: 0.064 (0.021)

Global demand elasticity: -0.139 (0.037)

Loadings on global demand factor

U.S.	1.415	(0.444)
Japan	1.548	(0.525)
Europe	2.044	(0.564)
Rest of world	0.967	(0.364)

Example 1: Effects of a one-standard-deviation increase in global demand

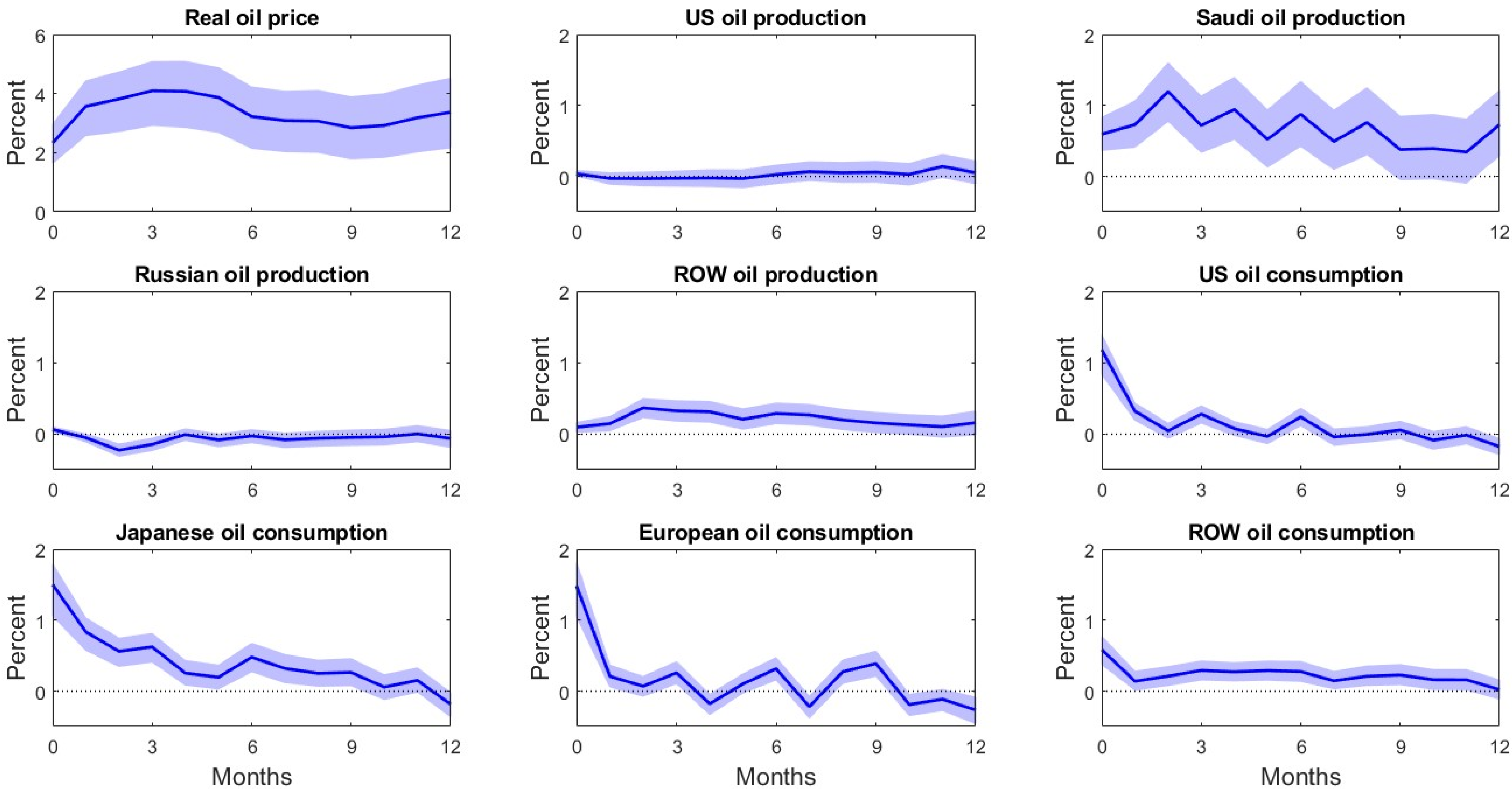
- On impact, this shifts demand curve for each country by magnitudes in previous slide (e.g., U.S. consumption demand increases by 1.4%)
- In equilibrium this causes oil price to go up 2.3%
- Supplies and net demand adjust in response

Impact effects of one-standard-deviation increase in world demand

Variable	as % of country			% of world
	direct effect (1)	response to price (2)	net effect (3)	net effect (4)
p	2.330			
q_{US}	0	0.043	0.043	0.005
q_{Saudi}	0	0.604	0.604	0.072
q_{Russia}	0	0.068	0.068	0.010
q_{ROW}	0	0.101	0.101	0.061
q				0.149
c_{US}	1.415	-0.218	1.197	0.299
c_{Japan}	1.548	-0.043	1.505	0.105
c_{Europe}	2.044	-0.524	1.520	0.122
c_{ROW}	0.967	-0.375	0.592	0.355
c				0.882
v				0.733

- Inventory adjustment plays a big role in mitigating the effect of the demand shock
- Temporary increase in demand met by sales out of inventories
- This response keeps the overall price increase modest
- Increased Saudi Arabian production is another important stabilizing factor

Dynamic effects of one-standard-deviation increase in world demand



Shaded regions are 68% confidence intervals

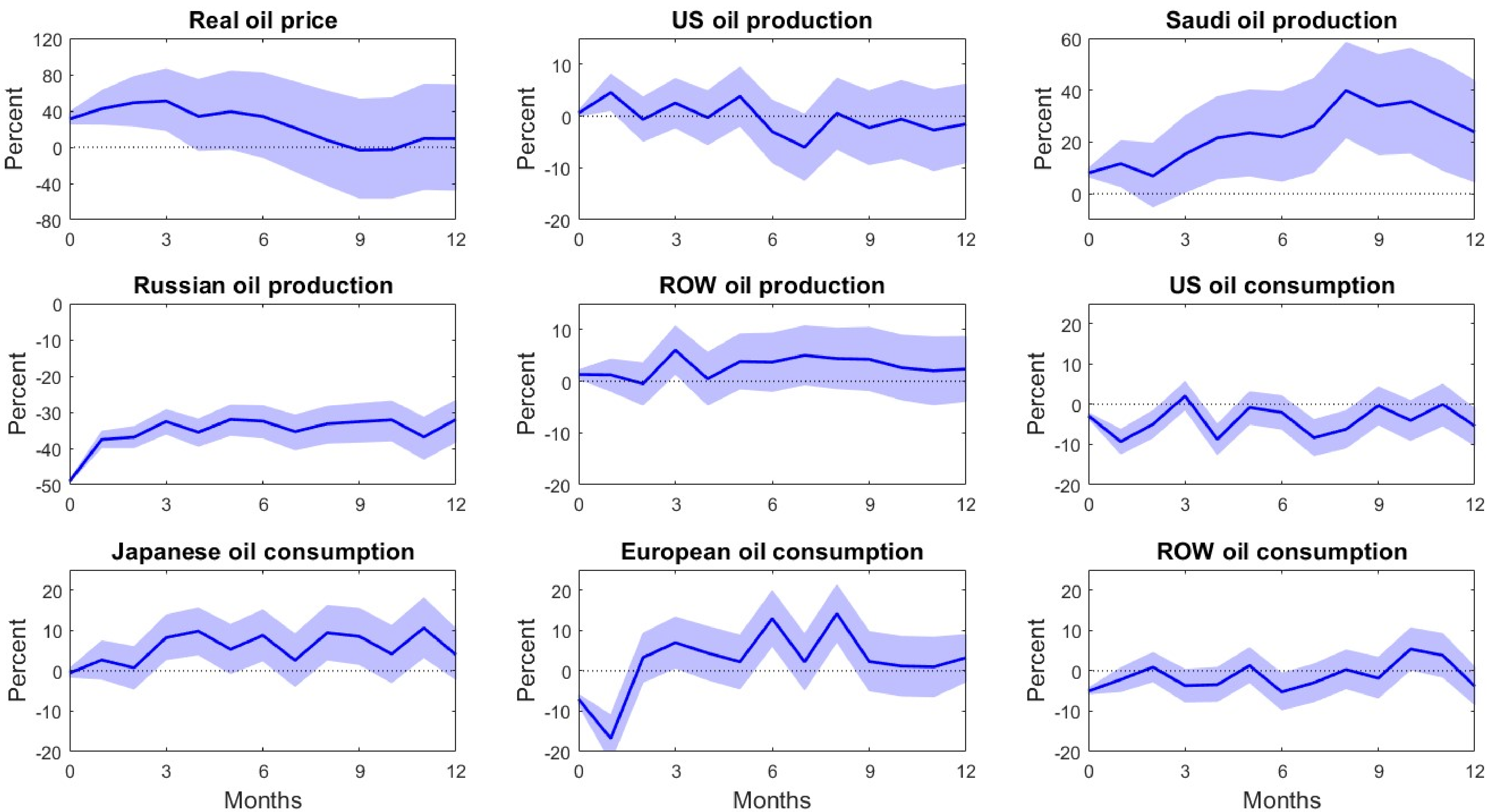
Example 2: Effects of a 50% decrease in Russian oil production

- Suppose geopolitical developments lead to a cut in Russian production of 5.25 mb/d
- For this analysis we impose that inventory sales can not be used to mitigate ($\phi_v = 0$)

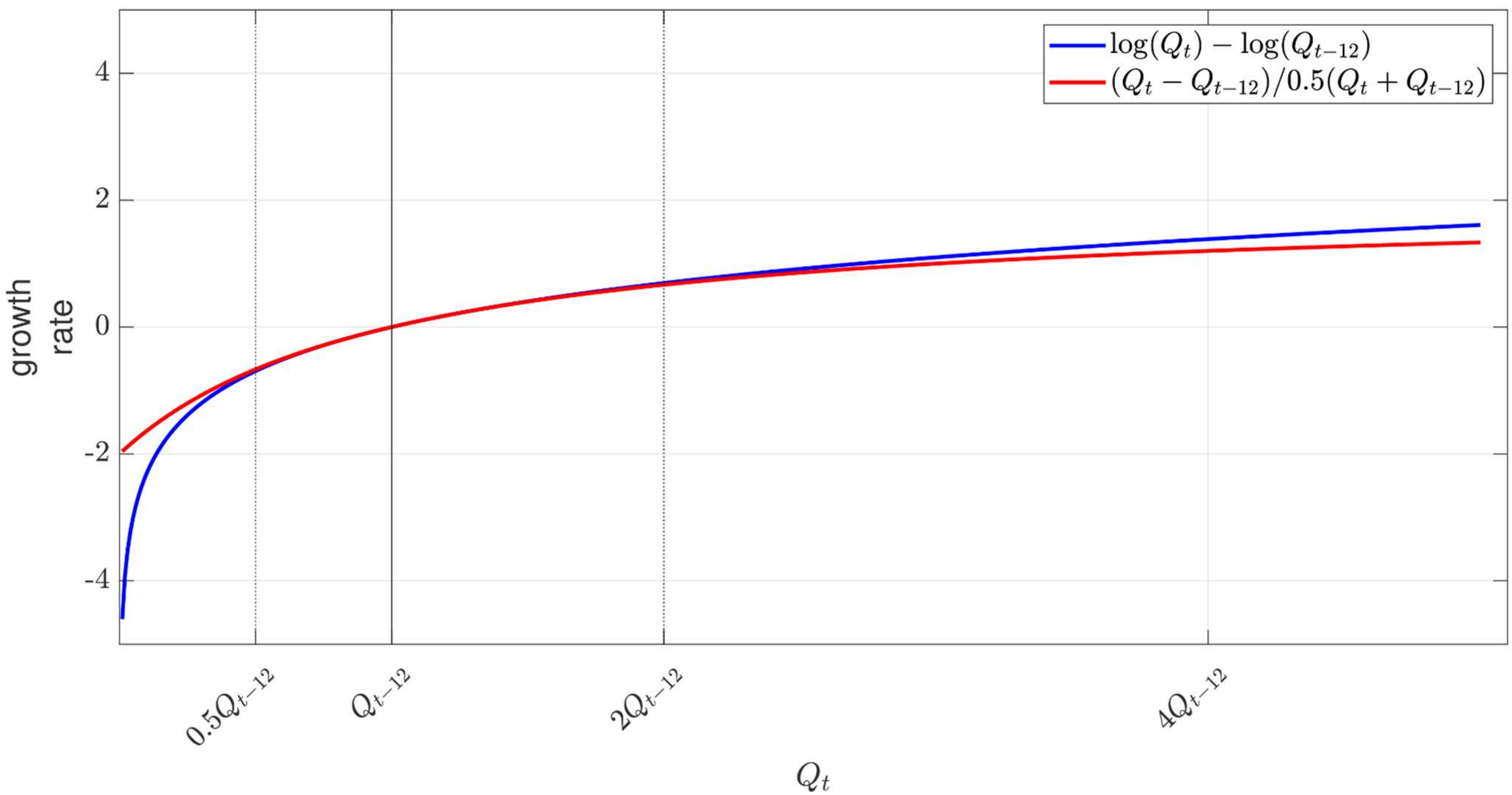
Impact effects of 50% cut in Russian oil production

Variable	as % of country			in mb/d
	direct effect (5)	response to price (6)	net effect (7)	net effect (8)
p	31.185			
q_{US}	0	0.580	0.580	0.072
q_{Saudi}	0	8.084	8.084	0.798
q_{Russia}	-50	0.909	-49.091	-5.252
q_{ROW}	0	1.345	1.345	0.664
q				-3.718
c_{US}	0.000	-2.919	-2.919	-0.480
c_{Japan}	0.000	-0.569	-0.569	-0.019
c_{Europe}	0.000	-7.011	-7.011	-0.289
c_{ROW}	0.000	-5.015	-5.015	-2.930
c				-3.718

Dynamic effects of 50% cut in Russian oil production



Additional slides



$\log(Q_t) - \log(Q_{t-12})$
$(Q_t - Q_{t-12})/0.5(Q_t + Q_{t-12})$
growth rate
4
2
0
-2
-4
$0.5Q_{t-12}$
Q_{t-12}
$2Q_{t-12}$
$4Q_{t-12}$
Q_t