Fiscal Stimulus: Evidence

Emi Nakamura and Jón Steinsson

UC Berkeley

April 2019
Neoclassical models:

- Positive but small multipliers (less than one)
  - Due to negative wealth effect
  - Hours go up but consumption goes down

- Exception:
  - Persistent spending with “flexible capital”
  - Investment rises a lot but consumption still falls
New Keynesian models:

- Multiplier highly dependent on monetary policy response
  - Constant real rate: multiplier 1
  - Lean against the wind: multiplier less than 1
  - Zero lower bound: multiplier larger than 1
- Multipliers larger with credit constrained agents
  (old Keynesian multiplier logic)
Other important issues:
- Multiplier dependent of tax response (varies across episodes)
- Multiplier dependent on type of spending
- Etc., Etc.

No single multiplier!
- Challenging to use aggregate multiplier estimates to distinguish between models
Government purchases multiplier:

When government purchases of goods and services go up by $1, how many dollars does output go up by?

\[(Y_t - Y_{t-1}) = \alpha + \beta(G_t - G_{t-1}) + \epsilon_t\]
Government purchases multiplier:

- When government purchases of goods and services go up by $1, how many dollars does output go up by?

\[
(Y_t - Y_{t-1}) = \alpha + \beta (G_t - G_{t-1}) + \epsilon_t
\]

- Usually divide through by \( Y_{t-1} \):

\[
\frac{Y_t - Y_{t-1}}{Y_{t-1}} = \alpha + \beta \frac{G_t - G_{t-1}}{Y_{t-1}} + \epsilon_t
\]

- Why?
Government purchases multiplier:

When government purchases of goods and services go up by $1, how many dollars does output go up by?

\[ (Y_t - Y_{t-1}) = \alpha + \beta (G_t - G_{t-1}) + \epsilon_t \]

Usually divide through by \( Y_{t-1} \):

\[ \frac{Y_t - Y_{t-1}}{Y_{t-1}} = \alpha + \beta \frac{G_t - G_{t-1}}{Y_{t-1}} + \epsilon_t \]

Why?

- Reduces heteroskedasticity.
Multiplier regression:

\[
\frac{Y_t - Y_{t-1}}{Y_{t-1}} = \alpha + \beta \frac{G_t - G_{t-1}}{Y_{t-1}} + \epsilon_t
\]

Different from

\[
\frac{Y_t - Y_{t-1}}{Y_{t-1}} = \alpha + \beta \frac{G_t - G_{t-1}}{G_{t-1}} + \epsilon_t
\]
Multiplier regression:

\[
\frac{Y_t - Y_{t-1}}{Y_{t-1}} = \alpha + \beta \frac{G_t - G_{t-1}}{Y_{t-1}} + \epsilon_t
\]

Different from

\[
\frac{Y_t - Y_{t-1}}{Y_{t-1}} = \alpha + \beta \frac{G_t - G_{t-1}}{Y_{t-1}} + \epsilon_t
\]

Second specification estimates an elasticity as opposed to a multiplier.

Some papers estimate elasticity and then convert to multiplier by multiplying by average value of \( Y/G \)

(Ramey-Zubairy 18 argue this is not a good practice)
Suppose we seek to estimate:

\[
\frac{Y_t - Y_{t-1}}{Y_{t-1}} = \alpha + \beta \frac{G_t - G_{t-1}}{Y_{t-1}} + \epsilon_t
\]

An important empirical problem is endogeneity of \( G_t \)

What is the likely nature of the endogeneity?
Countercyclical spending:

- Governments might systematically spend more when output is low due to other shocks in an effort to counteract these other shocks and stabilize economy
- In this case, OLS would be downward biased
Countercyclical spending:
- Governments might systematically spend more when output is low due to other shocks in an effort to counteract these other shocks and stabilize economy
- In this case, OLS would be downward biased

Procyclical spending:
- Balanced budget rules or credit constraints may lead government to spend more when things are good for other reasons
- In this case, OLS would be upward biased
Methods for identification:

- Wars (Barro-Redlick 11; Hall 09; Ramey 11)
- VARs (Blanchard-Perotti 02; Gali, et al. 07, Perotti 07)
- Regional shocks (Chodorow-Reich et al. 12, Shoag 13, Nakamura-Steinsson 14, Acconcia-Corsetti-Simonelli 14)
Pros and Cons of Looking at Wars

War yield large changes in spending. Military spending easy to model. Infrequent. Is military spending associated with wars exogenous? Often easy to rule out reverse causality (war not due to state of US business cycle). But does "exclusion restriction" hold? (i.e., do wars only affect output through spending?) Important confounding factors: Patriotism, Rationing, price controls, Mismeasurement of prices of tanks and wages of soldiers. Barro-Redlick think war-time multiplier overestimate true multipliers. Hall 09 thinks they are underestimates. Nakamura-Steinsson (UC Berkeley)
Pros and Cons of Looking at Wars

- War yield large changes in spending
- Military spending easy to model
- Infrequent
Pros and Cons of Looking at Wars

- War yield large changes in spending
- Military spending easy to model
- Infrequent
- Is military spending associated with wars exogenous?

Barro-Redlick 11 think war-time multiplier overestimate true multipliers
Hall 09 thinks they are underestimates
Nakamura-Steinsson (UC Berkeley)

Fiscal Stimulus: Evidence
April 2019 10 / 69
Pros and Cons of Looking at Wars

- War yield large changes in spending
- Military spending easy to model
- Infrequent
- Is military spending associated with wars exogenous?
  - Often easy to rule out reverse causality
    (war not due to state of US business cycle)
  - But does “exclusion restriction” hold?
    (i.e., do wars only affect output through spending?)
PROS AND CONS OF LOOKING AT WARS

- War yield large changes in spending
- Military spending easy to model
- Infrequent

Is military spending associated with wars exogenous?
- Often easy to rule out reverse causality
  (war not due to state of US business cycle)
- But does “exclusion restriction” hold?
  (i.e., do wars only affect output through spending?)

Important confounding factors:
- Patriotism
- Rationing, price controls
- Mismeasurement of prices of tanks and wages of soldiers
Pros and Cons of Looking at Wars

- War yield large changes in spending
- Military spending easy to model
- Infrequent
- Is military spending associated with wars exogenous?
  - Often easy to rule out reverse causality
    (war not due to state of US business cycle)
  - But does “exclusion restriction” hold?
    (i.e., do wars only affect output through spending?)
- Important confounding factors:
  - Patriotism
  - Rationing, price controls
  - Mismeasurement of prices of tanks and wages of soldiers
- Barro-Redlick 11 think war-time multiplier overestimate true multipliers
- Hall 09 thinks they are underestimates
Look at variation in military spending from 1917-2006
Dominated by WWI, WWII, Korean War

The figure shows the change in per capita real government purchases (nominal purchases divided by the GDP deflator), expressed as a ratio to the prior year’s per capita real GDP. The black graph is for defense purchases, and the gray graph is for nondefense purchases by all levels of government. The data on government purchases since 1929 are from Bureau of Economic Analysis and, before that, from Kendrick (1961). The GDP data are described in the online appendix.

In comparison to these three large wars, the post-1954 period features much more modest variations in defense spending. The changes in defense outlays were again mainly exogenous with respect to GDP.

Source: Barro and Redlick (2011)
<table>
<thead>
<tr>
<th>Year</th>
<th>DG (% GDP)</th>
<th>Year</th>
<th>DG (% GDP)</th>
</tr>
</thead>
<tbody>
<tr>
<td>WWI</td>
<td></td>
<td>Korea</td>
<td></td>
</tr>
<tr>
<td>1917</td>
<td>3.5</td>
<td>1951</td>
<td>5.6</td>
</tr>
<tr>
<td>1918</td>
<td>14.9</td>
<td>1952</td>
<td>3.3</td>
</tr>
<tr>
<td>1919</td>
<td>-7.9</td>
<td>1953</td>
<td>0.5</td>
</tr>
<tr>
<td>1920</td>
<td>-8.2</td>
<td>1952</td>
<td>-2.1</td>
</tr>
<tr>
<td>WWII</td>
<td></td>
<td>Vietnam</td>
<td></td>
</tr>
<tr>
<td>1941</td>
<td>10.6</td>
<td>1966</td>
<td>1.2</td>
</tr>
<tr>
<td>1942</td>
<td>25.8</td>
<td>1967</td>
<td>1.1</td>
</tr>
<tr>
<td>1943</td>
<td>17.2</td>
<td></td>
<td></td>
</tr>
<tr>
<td>1944</td>
<td>3.6</td>
<td>Reagan</td>
<td></td>
</tr>
<tr>
<td>1945</td>
<td>-7.1</td>
<td>1982-1985</td>
<td>0.4-0.5</td>
</tr>
<tr>
<td>1946</td>
<td>-25.8</td>
<td>Bush II</td>
<td></td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>2002-2004</td>
</tr>
</tbody>
</table>

Source: Barro and Redlick (2011)
Empirical specification:

\[
\frac{Y_t - Y_{t-1}}{Y_{t-1}} = \alpha + \beta_1 \frac{G_t - G_{t-1}}{Y_{t-1}} + \beta_2 \frac{G_{t-1} - G_{t-2}}{Y_{t-2}} + \beta_3 \frac{G^*_t - G^*_{t-2}}{Y_{t-2}} + \text{controls} + \epsilon_t
\]

- \(G^*_t\) captures news at time \(t\) about future spending from Ramey 11
  - Gathered from Business Week estimates of changes in spending over next 3 to 5 years
equaled 0). Then we assumed that the timing of the news corresponded to the one found by Ramey (2011, Table II) for World War II: run-up period for 1914–16 corresponding to 1939–40, war buildup of 1917–18 corresponding to 1941–43, and wind-down for 1919–20 corresponding to 1944–46. The resulting measure of defense news for World War I is a rough approximation, and it would be valuable to extend Ramey's analysis formally to this period.

Figure II shows the estimates for the present value of the expected addition to nominal defense spending when expressed as a ratio to the prior year's nominal GDP. World War II stands out, including the run-up values of 0.40 in 1940, 1.46 in 1941, and 0.75 in 1942, and the wind-down values of –0.07 in 1944 and –0.19 in 1945. The peak at the start of the Korean War (1.16 in 1950) is impressive, signaling that people were concerned about the potential...
TABLE II
Equations for GDP Growth, Various Samples

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starting date</td>
<td>1950</td>
<td>1939</td>
<td>1930</td>
<td>1930</td>
<td>1917</td>
<td>1954</td>
</tr>
<tr>
<td>Δg: defense</td>
<td>0.68*</td>
<td>0.44**</td>
<td>0.46**</td>
<td>0.48**</td>
<td>0.47**</td>
<td>0.98</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td>(0.06)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.65)</td>
</tr>
<tr>
<td>Δg: defense</td>
<td>0.01</td>
<td>0.20**</td>
<td>0.21*</td>
<td>0.25**</td>
<td>0.16</td>
<td>-0.54</td>
</tr>
<tr>
<td>(−1)</td>
<td>(0.28)</td>
<td>(0.06)</td>
<td>(0.09)</td>
<td>(0.08)</td>
<td>(0.08)</td>
<td>(0.56)</td>
</tr>
<tr>
<td>Δg*: defense news</td>
<td>0.026</td>
<td>0.039**</td>
<td>0.034*</td>
<td>0.034*</td>
<td>0.034*</td>
<td>-0.120</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.011)</td>
<td>(0.015)</td>
<td>(0.014)</td>
<td>(0.017)</td>
<td>(0.112)</td>
</tr>
<tr>
<td>U(−1)</td>
<td>0.50**</td>
<td>0.58**</td>
<td>0.61**</td>
<td>0.58**</td>
<td>0.47**</td>
<td>0.51**</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.14)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.10)</td>
<td>(0.18)</td>
</tr>
<tr>
<td>Δτ(−1)</td>
<td>-0.54**</td>
<td>-0.16</td>
<td>-0.26</td>
<td>-0.52*</td>
<td>-0.19</td>
<td>-0.48*</td>
</tr>
<tr>
<td></td>
<td>(0.21)</td>
<td>(0.16)</td>
<td>(0.22)</td>
<td>(0.23)</td>
<td>(0.25)</td>
<td>(0.22)</td>
</tr>
<tr>
<td>Yield spread</td>
<td>-43.9*</td>
<td>-37.8</td>
<td>-101.5**</td>
<td>-103.4**</td>
<td>-73.6**</td>
<td>-43.1*</td>
</tr>
<tr>
<td>squared</td>
<td>(20.7)</td>
<td>(22.0)</td>
<td>(12.8)</td>
<td>(12.4)</td>
<td>(12.2)</td>
<td>(21.8)</td>
</tr>
<tr>
<td>p-value, defense variables</td>
<td>0.030</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.000</td>
<td>0.47</td>
</tr>
<tr>
<td>R²</td>
<td>0.48</td>
<td>0.82</td>
<td>0.75</td>
<td>0.77</td>
<td>0.66</td>
<td>0.45</td>
</tr>
<tr>
<td>σ</td>
<td>0.017</td>
<td>0.019</td>
<td>0.027</td>
<td>0.026</td>
<td>0.030</td>
<td>0.018</td>
</tr>
</tbody>
</table>

Source: Barro and Redlick (2011)
If variation in government spending is truly random, what is the role of the controls in the regressions?
If variation in government spending is truly random, what is the role of the controls in the regressions?

- They soak up noise in the regression and make the estimates more precise
- But if they end up affecting the point estimates substantially, this suggests that spending may not be truly random
TABLE III
Nondefense Government Purchases and Transfers

<table>
<thead>
<tr>
<th></th>
<th>(1)</th>
<th>(2)</th>
<th>(3)</th>
<th>(4)</th>
<th>(5)</th>
<th>(6)</th>
</tr>
</thead>
<tbody>
<tr>
<td>Starting date</td>
<td>1950</td>
<td>1930</td>
<td>1950</td>
<td>1930</td>
<td>1950</td>
<td>1950</td>
</tr>
<tr>
<td>Δg: defense</td>
<td>0.89**</td>
<td>0.46**</td>
<td>0.34</td>
<td>0.51**</td>
<td>0.84**</td>
<td>0.46</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td>(0.08)</td>
<td>(0.32)</td>
<td>(0.10)</td>
<td>(0.24)</td>
<td>(0.26)</td>
</tr>
<tr>
<td>Δg: defense (−1)</td>
<td>−0.13</td>
<td>0.21*</td>
<td>0.08</td>
<td>0.18*</td>
<td>−0.36</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(0.27)</td>
<td>(0.09)</td>
<td>(0.28)</td>
<td>(0.09)</td>
<td>(0.25)</td>
<td>(0.26)</td>
</tr>
<tr>
<td>Δg*: defense news</td>
<td>0.040**</td>
<td>0.036*</td>
<td>0.028</td>
<td>0.033*</td>
<td>0.014</td>
<td>0.016</td>
</tr>
<tr>
<td></td>
<td>(0.016)</td>
<td>(0.016)</td>
<td>(0.016)</td>
<td>(0.015)</td>
<td>(0.013)</td>
<td>(0.014)</td>
</tr>
<tr>
<td>U(−1)</td>
<td>0.64**</td>
<td>0.60**</td>
<td>0.43*</td>
<td>0.62**</td>
<td>0.26*</td>
<td>0.55**</td>
</tr>
<tr>
<td></td>
<td>(0.17)</td>
<td>(0.11)</td>
<td>(0.18)</td>
<td>(0.10)</td>
<td>(0.16)</td>
<td>(0.16)</td>
</tr>
<tr>
<td>Δτ(−1)</td>
<td>−0.45*</td>
<td>−0.25</td>
<td>−0.56**</td>
<td>−0.25</td>
<td>−0.26</td>
<td>−0.38</td>
</tr>
<tr>
<td></td>
<td>(0.20)</td>
<td>(0.23)</td>
<td>(0.21)</td>
<td>(0.22)</td>
<td>(0.19)</td>
<td>(0.20)</td>
</tr>
<tr>
<td>Yield spread squared</td>
<td>−31.2</td>
<td>−100.9**</td>
<td>−28.4</td>
<td>−102.3**</td>
<td>−38.9*</td>
<td>−21.6</td>
</tr>
<tr>
<td></td>
<td>(20.0)</td>
<td>(13.3)</td>
<td>(25.4)</td>
<td>(13.0)</td>
<td>(18.1)</td>
<td>(20.5)</td>
</tr>
<tr>
<td>Δg: nondefense</td>
<td>2.65**</td>
<td>0.12</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td>(0.93)</td>
<td>(0.63)</td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δ(transfers)</td>
<td>—</td>
<td>—</td>
<td>−1.53</td>
<td>0.64</td>
<td>—</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td>(0.92)</td>
<td>(0.68)</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Δ(GM sales)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>3.66**</td>
<td>—</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(0.86)</td>
<td></td>
</tr>
<tr>
<td>Δ(GE sales)</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>—</td>
<td>17.6**</td>
</tr>
<tr>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>(4.7)</td>
<td></td>
</tr>
<tr>
<td>R²</td>
<td>0.54</td>
<td>0.75</td>
<td>0.51</td>
<td>0.75</td>
<td>0.63</td>
<td>0.57</td>
</tr>
<tr>
<td>σ</td>
<td>0.017</td>
<td>0.027</td>
<td>0.017</td>
<td>0.027</td>
<td>0.015</td>
<td>0.016</td>
</tr>
</tbody>
</table>

Source: Barro and Redlick (2011)
**Blanchard and Perotti (2002)**

- Structural VAR based evidence for fiscal stimulus:
  \[ X_t = A(L)X_{t-1} + U_t \]

- \( X_t = [T_t, G_t, Y_t] \)
- Four lags (and quarter dependence of coefficients)
- Various different detrending methods plus some dummy variables
They argue:

- VAR methods better suited for study of fiscal policy than monetary policy
- Variation in government spending occurs for many reasons other than output stabilization
- Implementation lags implies no response of spending to output within, say, a quarter
Consider the relationship between $\Delta G$ and $\Delta Y$

Blanchard and Perotti’s “identifying assumption” for $\Delta G$:
- Output does not affect government spending contemporaneously
Consider the relationship between $\Delta G$ and $\Delta Y$

Blanchard and Perotti’s “identifying assumption” for $\Delta G$:

- Output does not affect government spending contemporaneously

Given this identifying assumption, would it work to simply estimate:

$$\Delta Y_t = \alpha + \beta \Delta G_t + \epsilon_t$$
Two things can go wrong in causal inference:

1. Reverse causality: Causality can go “opposite” way (simultaneity bias)
2. Omitted variable bias: A third factor can cause movements in both variables
Two things can go wrong in causal inference:

1. Reverse causality: Causality can go “opposite” way (simultaneity bias)
2. Omitted variable bias: A third factor can cause movements in both variables

Blanchard and Perotti’s “identifying assumption” deals with reverse causality, but not the omitted variables bias
What is Blanchard and Perotti’s strategy for dealing with omitted variables bias?
What is Blanchard and Perotti’s strategy for dealing with omitted variables bias?

- By controlling for four lags of $\Delta Y_t$, $\Delta G_t$, and $\Delta T_t$
What is Blanchard and Perotti’s strategy for dealing with omitted variables bias?

- By controlling for four lags of $\Delta Y_t$, $\Delta G_t$, and $\Delta T_t$

General feature of “structural” VARs: identification by controlling for lags
Examples of Omitted Variable

News shocks about future output that are not captured by lags:

- Terrorist attacks
- Wars
- Financial crises
- Oil price shocks
- Regime shifts in monetary policy

Each one may only matter for a few data points. But they can add up.
tic process as the post-1960 period. Thus, our strategy is to proceed in two steps. For most of the paper we run regressions starting from 1960:1. In Section VIII we extend the sample back to include the 1950s and look at what we can learn from this longer sample.

Note that our benchmark sample still includes one large net tax cut episode, the 1975:2 tax cut. This episode is a well-identified, isolated, temporary, tax cut. Hence, it can be easily and

Source: Blanchard and Perotti (2002)
They exclude 1950’s because “difficult to think of the early 1950’s as being generated by the same stochastic process as the post-1960 period.”

But is this an important disadvantage?
They exclude 1950’s because “difficult to think of the early 1950’s as being generated by the same stochastic process as the post-1960 period.”

But is this an important disadvantage?

- Not obvious
  - Different perspective: Large variation very valuable for identification

Also dummy out tax cut in 1975:II
### TABLE IV
**Responses to Spending Shocks**

<table>
<thead>
<tr>
<th></th>
<th>1 qrt</th>
<th>4 qrts</th>
<th>8 qrts</th>
<th>12 qrts</th>
<th>20 qtrs</th>
<th>peak</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>DT</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>0.84*</td>
<td>0.45</td>
<td>0.54</td>
<td>1.13*</td>
<td>0.97*</td>
<td>1.29*</td>
</tr>
<tr>
<td>GCN</td>
<td>1.00*</td>
<td>1.14*</td>
<td>0.95*</td>
<td>0.70*</td>
<td>0.42*</td>
<td></td>
</tr>
<tr>
<td>TAX</td>
<td>0.13</td>
<td>0.14</td>
<td>0.17</td>
<td>0.43*</td>
<td>0.52*</td>
<td></td>
</tr>
<tr>
<td><strong>ST</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>GDP</td>
<td>0.90*</td>
<td>0.55</td>
<td>0.65</td>
<td>0.66</td>
<td>0.66</td>
<td>0.90*</td>
</tr>
<tr>
<td>GCN</td>
<td>1.00*</td>
<td>1.30*</td>
<td>1.56*</td>
<td>1.61*</td>
<td>1.62*</td>
<td></td>
</tr>
<tr>
<td>TAX</td>
<td>0.10</td>
<td>0.18</td>
<td>0.33</td>
<td>0.36</td>
<td>0.37</td>
<td></td>
</tr>
</tbody>
</table>

Source: Blanchard and Perotti (2002)
FIGURE V
Response to a Spending Shock

Notice the use of one standard deviation bands

Evidently not much information in 1960-1997 sample
Notice the use of one standard deviation bands
Evidently not much information in 1960-1997 sample
Also, max response of output is after 15 quarters, and after a “wavy response”
Estimates further out rely heavily on iteration of VAR system
Notice the use of one standard deviation bands
Evidently not much information in 1960-1997 sample
Also, max response of output is after 15 quarters, and after a "wavy response"
Estimates further out rely heavily on iteration of VAR system
Alternative way to report results:
  - Cumulative response of output divided by cumulative response of spending (over some horizon)
  - Ramey-Zubairy 18 report such results using Blanchard-Perotti 02 identification for sample period including Korean war and find estimate below one
VAR with government spending “ordered first”
(i.e., is not contemporaneously affected by other variables in VAR)
- Large: Government spending, GDP, hours, consumption of non-durables and services, private nonresidential investment, real wage, budget deficit, personal disposable income.
- Small: Government spending, GDP, consumption, deficit

Quarterly data, four lags

Table 1. Estimated effects of government spending shocks.

<table>
<thead>
<tr>
<th>Estimated Fiscal Multipliers</th>
<th>1stQ</th>
<th>4thQ</th>
<th>8thQ</th>
<th>1stQ</th>
<th>4thQ</th>
<th>8thQ</th>
<th>Implied Fiscal Parameters</th>
</tr>
</thead>
<tbody>
<tr>
<td>1948:I–2003:IV</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>ρₜ₁</td>
</tr>
<tr>
<td>Baseline spending</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td>0.85</td>
</tr>
<tr>
<td>Small VAR</td>
<td>0.51</td>
<td>0.31</td>
<td>0.28</td>
<td>0.04</td>
<td>0.09</td>
<td>0.19</td>
<td></td>
</tr>
<tr>
<td>Larger VAR</td>
<td>0.41</td>
<td>0.31</td>
<td>0.68</td>
<td>0.07</td>
<td>0.11</td>
<td>0.49</td>
<td>0.80</td>
</tr>
<tr>
<td>Excluding military</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>Small VAR</td>
<td>0.15</td>
<td>-0.12</td>
<td>0.34</td>
<td>-0.11</td>
<td>0.24</td>
<td>0.32</td>
<td>0.95</td>
</tr>
<tr>
<td>Larger VAR</td>
<td>0.36</td>
<td>0.62</td>
<td>1.53</td>
<td>0.03</td>
<td>0.51</td>
<td>0.68</td>
<td>0.94</td>
</tr>
</tbody>
</table>

| 1954:I–2003:IV              |      |      |      |      |      |      | ρₜ₁ | φₜ₁ | φₕ₁ |
| Baseline spending           |      |      |      |      |      |      | 0.95 | 0.13 | 0.20 |
| Small VAR                   | 0.74 | 0.75 | 1.22 | 0.14 | 0.46 | 0.73 |      |      |      |
| Larger VAR                  | 0.68 | 0.70 | 1.74 | 0.17 | 0.29 | 0.95 | 0.95 | 0.10 | 0.30 |
| Excluding military          |      |      |      |      |      |      |      |      |      |
| Small VAR                   | 0.63 | 1.95 | 2.60 | 0.25 | 1.41 | 1.12 | 0.95 | 0.05 | 0.50 |
| Larger VAR                  | 0.74 | 2.37 | 3.50 | 0.37 | 1.39 | 1.76 | 0.95 | 0.01 | 0.50 |

| 1960:I–2003:IV              |      |      |      |      |      |      | ρₜ₁ | φₜ₁ | φₕ₁ |
| Baseline spending           |      |      |      |      |      |      | 0.95 | 0.13 | 0.20 |
| Small VAR                   | 0.91 | 1.05 | 1.32 | 0.19 | 0.59 | 0.84 |      |      |      |
| Larger VAR                  | 0.81 | 0.44 | 0.76 | 0.20 | 0.25 | 0.45 | 0.95 | 0.08 | 0.20 |
| Excluding military          |      |      |      |      |      |      |      |      |      |
| Small VAR                   | 0.72 | 1.14 | 1.19 | 0.17 | 0.78 | 0.68 | 0.94 | 0.03 | 0.50 |
| Larger VAR                  | 1.13 | 1.89 | 2.08 | 0.40 | 1.14 | 1.07 | 0.98 | 0.01 | 0.55 |

Note: Large VAR corresponds to the 8-variable VAR described in the text; Small VAR estimates are based on a 4-variable VAR including government spending, output, consumption, and the deficit. Government spending excluding military was obtained as GFNEH + GSEH + GFNIH + GSIH. For each specification ρₜ₁ is the AR(1) coefficient that matches the half-life of the estimated government spending response. Parameter φₜ₁ is obtained as the difference of the VAR-estimated impact effects of government spending and deficit, respectively. Finally, given ρₜ₁ and φₜ₁, we calibrate the parameter φₕ₁ such that the dynamics of government spending (21) and debt (37) are consistent with the horizon at which the deficit is back to steady state, matching our empirical VAR responses of the fiscal deficit.

Source: Gali, Lopez-Salido, Valles (2007)
• Multipliers much smaller if Korean war included  
  (perhaps due to large tax increases)
• Multiplier bigger for non-defense spending
  Barro-Redlick argue this is endogenous
Multipliers much smaller if Korean war included (perhaps due to large tax increases)
Multiplier bigger for non-defense spending
Barro-Redlick argue this is endogenous

Measure of multipliers: \( \frac{dY_{t+k}}{dG_t} \)
- Tricky to interpret
- Not only \( dG_t \) that is affecting \( dY_{t+k} \), also \( dG_{t+1} \ldots dG_{t+k} \)
- Alternative: Ratio of cumulative impulse responses
- Fahri and Werning 16 has a nice discussion of this
- Use narrative approach to identify shocks to government spending
- Dates when Business Week suddenly began forecasting large increases in defense spending
- Ramey 11 adds: 2001:III (9/11)
Identifying Government Spending Shocks

Figure I

Real Government Spending Per Capita (in thousands of chained dollars, 2005)

Source: Ramey (2011)
War dates “Granger cause” VAR shocks

- Government spending “shocks” are anticipated!
- Doesn’t mean they are necessarily endogenous
- Invalidates VAR method for constructing impulse response
  - Some of the effects occur when news arrives
  - Some of the effects occur when spending occurs
  - VAR misses effects that occur prior to spending
  - VAR misspecified, impulse response potentially way off

(My discussion here is somewhat different than Ramey’s)
FIGURE V
Comparison of VAR Defense Shocks to Forecasts: Korea and Vietnam

Notes. The top and middle panels are based on log per capita real defense spending on a quarterly calendar year basis. The bottom panels are nominal, annual data on a fiscal year basis.

three quarters was anticipated as of August and September of 1950. The bottom graph shows Business Week’s forecasts of defense spending. The June 1950 forecast, made before the Korean War started, predicted that defense spending would remain at about $15 billion per year. Two months later in August 1950, Business Week correctly predicted the rise in defense spending through fiscal year 1952. By September 1950, it had correctly predicted the rise through fiscal year 1954. Thus, it is clear that the positive VAR shocks are several quarters too late. It is also interesting to note that while Business Week was predicting a future decline in defense spending as early as April 1953 when a truce seemed imminent, the VAR records a negative defense spending shock in the first quarter of 1954. Thus, the VAR shocks are not accurately reflecting news about defense spending.

Source: Ramey (2011)
Comparison of VAR Defense Shocks to Forecasts: Carter–Reagan and 9/11

Notes. The top and middle panels are based on log per capita real defense spending on a quarterly calendar-year basis. The bottom panels are nominal, annual data on a fiscal year basis.

Forecasts were not as accurate for Vietnam. As of August 1965, several noted senators were forecasting much higher expenditures than the Johnson Administration was quoting. The forecasts kept rising steadily for some time. Thus, while it is true that there were a number of positive spending shocks in the first years of the Vietnam War, it is not clear that the VAR gets the timing right.

In Figure VI, the VARs show many positive shocks during the Carter–Reagan build-up through 1985. The bottom panel shows, however, that as of January 1981, the OMB was very accurately predicting spending in fiscal years 1981–1984. On the other hand, the October 1981 forecast over-predicted defense spending in fiscal years 1985 and 1986. However, all of the forecast error for 1985 and 1986 can be attributed to the fact that inflation fell much...

Source: Ramey (2011)
How should anticipation of spending affect results?

- Suppose $G \uparrow$ is announced one period in advance
- What happens upon announcement?
How should anticipation of spending affect results?

- Suppose $G \uparrow$ is announced one period in advance
- What happens upon announcement?
  - Negative wealth effect: $C \downarrow, H \uparrow$
  - Anticipatory investment: $I \uparrow$
- If you measure shock as occurring when spending occurs, you will miss these effects
War dates variable embedded in a VAR — ordered first
Quadratic trend, four lags
Sample period: 1947-2008

War Dates essentially an instrument for spending
Comparison of Identification Methods: Response to a Government Spending Shock (Standard error bands are 68% confidence intervals)

Source: Ramey (2011)
Source: Ramey (2011)
Ramey argues that difference between results based on structural VAR identification and War Dates identification has to do with timing.

War dates recognize that news about spending occurs before spending occurs.
Ramey argues that difference between results based on structural VAR identification and War Dates identification has to do with timing.

War dates recognize that news about spending occurs before spending occurs.

VARs miss initial drop in consumption.

Delaying War dates yields VAR type results.

FIGURE VIII
The Effect of Mistiming the Ramey–Shapiro Dates (Standard error bands are 68% confidence intervals)

Source: Ramey (2011)
Structural VAR studies: $G \uparrow \Rightarrow C \uparrow, W/P \uparrow$

Ramey-Shapiro "war dates": $G \uparrow \Rightarrow C \downarrow, W/P \downarrow$
VARs versus War Dates

- Structural VAR studies: $G \uparrow \Rightarrow C \uparrow, W/P \uparrow$
- Ramey-Shapiro "war dates": $G \uparrow \Rightarrow C \downarrow, W/P \downarrow$
- Massive focus on whether $C \uparrow$ or $C \downarrow$ in literature
VARs versus War Dates

- Structural VAR studies: $G \uparrow \Rightarrow C \uparrow, W/P \uparrow$
- Ramey-Shapiro "war dates": $G \uparrow \Rightarrow C \downarrow, W/P \downarrow$

- Massive focus on whether $C \uparrow$ or $C \downarrow$ in literature
- Suggestion that this distinguishes between Neoclassical models and New Keynesian models
- In fact Keynesian models can generate both depending on monetary policy
Regional Multipliers
Since the Great Recession:

- Explosion of empirical work estimating regional multipliers
- Wide array of identification strategies:
  - Windfall returns on state pension plans (Shoag 15)
  - Military buildups (Nakamura-Steinsson 14)
  - Crackdown on Mafia infiltrated municipalities in Italy (Acconcia et al. 14)
  - Spending discontinuities at decadal census population revisions (Suarez Serrato-Wingender 16)
  - Evidence from ARRA (Chodorow-Reich et al. 12, Wilson 12, Dupor-Mehkari 16)
- Survey: Chodorow-Reich 17
Regional Multipliers: Pros and Cons

- A lot more data, a lot more variation
- Allows for difference-in-difference identification
- Allows for powerful class of instruments:
  - Differential regional exposure to aggregate shocks
- Regional multiplier not the same as aggregate multipliers
  - Not answering the “right” question?
  - What do we learn?
\[
\left( \frac{Y_{it} - Y_{it-2}}{Y_{it-2}} \right) = \alpha_i + \gamma_t + \beta \left( \frac{G_{it} - G_{it-2}}{Y_{it-2}} \right) + \epsilon_{it}
\]

- \(G_{it}\) is prime military contract spending
- State fixed effects (state specific trends)
- Year fixed effects (controls for aggregate shocks)
- Variables measured per capita
- Biannual regressions (in lieu of dynamics)
- Government spending potential endogenous and measured with error
spending relative to total output for the United States as a whole. First, notice that most of the variation in national military spending is driven by geopolitical events—such as the Vietnam War, Soviet invasion of Afghanistan, and 9/11. Second, it is clear from the figure that military spending in California is systematically more sensitive to movements in national military spending than military spending in Illinois. The 1966–1971 Vietnam War drawdown illustrates this. Over this period, military procurement in California fell by 2.5 percentage points (almost twice the national average), while military procurement in Illinois fell by only about 1 percentage point (about 2/3 the national average). We use this variation in the sensitivity of military spending across regions to national military buildups and drawdowns to identify the effects of government spending shocks. Our identifying assumption is that the United States does not embark on a military buildup because states that receive a disproportionate amount of military spending are doing poorly relative to other states. This assumption is similar—but weaker than—the common identifying assumption in the empirical literature on the effects of national military spending, that variation in national military spending is exogenous to the US business cycle.

We employ two separate approaches to constructing instruments that capture the differential sensitivity of military spending across regions to national military buildups and drawdowns. Our baseline approach is to instrument for state or region military procurement using total national procurement interacted with a state or region dummy. The “first stage” in the two-stage least squares interpretation of this procedure is to regress changes in state spending on changes in aggregate spending and fixed effects allowing for different sensitivities across different states. This

---

22 Below, we will sometimes refer to spending relative to GDP simply as spending and the change in spending divided by GDP simply as the change in spending, for simplicity.

23 Murtazashvili and Wooldridge (2008) derive conditions for consistency of the fixed effects instrumental variables estimator we employ for a setting in which the multiplier varies across states.

---

**Figure 1. Prime Military Contract Spending as a Fraction of State GDP**

Source: Nakamura and Steinsson (2014)
National military buildups exogenous to relative conditions in states receiving disproportionate procurement spending

Use differential sensitivity to national shocks across states to identify effects on state output

Intuition:

- When $\Delta G_{US} > 0$, $\Delta G_{CA} > \Delta G_{IL}$
- What is effect on $\Delta Y_{CA}$ vs $\Delta Y_{IL}$?

Identifying assumption:

- No other shock $\alpha_i \mathcal{E}_t$ correlated with $\Delta G_{US}$ in the time series and differentially affects same set of states as our instrument (i.e., $\alpha_i$ correlated with differential cross-sectional sensitivity of our instrument)
Baseline instrument:

- National spending interacted with state dummy
- In effect, we estimate sensitivity of state spending to national spending in “first stage”
Instruments

Baseline instrument:
- National spending interacted with state dummy
- In effect, we estimate sensitivity of state spending to national spending in “first stage”

Bartik (1991) type instrument:
- National spending scaled by each state’s average spending in the first five years of sample
- Idea: Spending varies more in states with a lot of spending
state or region. We present results for output both deflated by national CPI and our measure of state CPI. The point estimates of $\beta$ for the output regression range from 1.4 to 1.9, while the point estimates of $\beta$ for the employment regression range from 1.3 to 1.8. The estimates using regional data are, in general, slightly larger than those based on state data, though the differences are small and statistically insignificant. The standard errors for the state regressions range from 0.3–0.4, while those for the region regressions range from 0.6–0.9. As is clear from Figure 1, the variation we use to estimate the multiplier is dominated by a few military buildups and drawdowns. These results control for short-term movements in population associated with government spending by running the regressions on per capita variables. The last column of Table 2 looks directly at population movements by estimating an analogous specification to equation (1) where the left-hand-side variable is $(\text{Pop}_t - \text{Pop}_{t-2})/\text{Pop}_{t-2}$ and the right-hand-side government spending variable is constructed from the level of government spending and output rather than per capita government spending and output. We find that the population responses to government spending shocks are small and cannot be distinguished from zero for the two year time horizon we consider. We also present estimates of the effects of military spending on consumer prices. These are statistically insignificantly different from zero, ranging from close to zero to a small positive number.

Figure 3 gives a visual representation of our main specification for output. The figure plots averages of changes in output against predicted military spending (based on our first-stage regression), grouped by 30 quantiles of the predicted military spending variable. Both variables are demeaned by year and state fixed effects. The

### Table 2—The Effects of Military Spending

<table>
<thead>
<tr>
<th>Prime military contracts</th>
<th>1.43</th>
<th>1.85</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(0.36)</td>
<td>(0.58)</td>
</tr>
<tr>
<td>Prime contracts plus</td>
<td>1.62</td>
<td>1.62</td>
</tr>
<tr>
<td>military compensation</td>
<td>(0.40)</td>
<td>(0.84)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,989</td>
<td>390</td>
</tr>
</tbody>
</table>

**Notes:** Each cell in the table reports results for a different regression with a shorthand for the main regressor of interest listed in the far left column. A shorthand for the dependent variable is stated at the top of each column. The dependent variable is a two-year change divided by the initial value in each case. Output and employment are per capita. The regressor is the two-year change divided by output. Military spending variables are per capita except in Population regression. Standard errors are in parentheses. All regressions include region and time fixed effects, and are estimated by two-stage least squares. The sample period is 1966–2006 for output, employment, and population, and 1969–2006 for the CPI. Output is state GDP, first deflated by the national CPI and then by our state CPI measures. Employment is from the BLS payroll survey. The CPI measure is described in the text. Standard errors are clustered by state or region.

Source: Nakamura and Steinsson (2014)
vast majority of points in the figure are located in the northeast and southwest quadrants, leading to a positive coefficient in our IV regression. To assess the robustness of our results to outliers, we have experimented with dropping states and regions with especially large or small estimated sensitivity of spending to national spending and this slightly raises the estimated open economy relative multiplier.

In Table 3, we report results for the simpler “Bartik” approach to constructing instruments. For output, this approach yields an open economy relative multiplier of roughly 2.5 for the states and 2.8 for the regions. For employment, this approach also yields larger open economy relative multipliers than our baseline specification—1.8 for states and 2.5 for regions. Our estimates using the Bartik-type instruments are somewhat less precise than those using our baseline instruments. This arises because, in constructing this instrument, we use the level of spending in each state as a proxy for the sensitivity of state spending to national spending—but it is an imperfect proxy.

Table 3 also reports a number of alternative specifications for the effects of military procurement on output and employment designed to evaluate the robustness of our results. We report the output multiplier when per capita output is constructed using a measure of the working age population as opposed to the total population. We add the price of oil interacted with state dummies as controls to our baseline regression.

28 Missouri and Connecticut have substantially higher estimated sensitivity of spending to national spending than other states and North Dakota has a substantially negative estimated sensitivity (alone among the states). Dropping any combination of these states from our baseline regression slightly raises our multiplier estimate. Dropping all three yields 1.88 (0.57).

29 State-level measures of population by age group are available from the Census Bureau starting in 1970. We define the working age population as the population between the ages of 19 and 64.

Figure 3. Quantiles of Change in Output versus Predicted Change in Military Spending

Notes: The figure shows averages of changes in output and predicted military spending (based on our first-stage regression), grouped by 30 quantiles of the predicted military spending variable. Both variables are demeaned by year and state fixed effects.

We add the real interest rate interacted with state dummies as controls to our baseline regression. We estimate the employment regression using the BEA’s employment series (available from 1969) instead of BLS payroll employment. Table 3 shows that these specifications all yield similar results to our baseline estimates.

We have extensively investigated the small-sample properties of our estimation approach using Monte Carlo simulations. These simulations indicate that neither the regional regressions nor the regressions using the Bartik-type instruments suffer from bias associated with weak or many instruments. However, our estimates of the state regressions using our baseline instruments are likely to be conservative in the sense of underestimating the open economy relative multiplier for states by roughly 10 percent (implying that the true state-level open economy relative multiplier is 1.65 rather than 1.43). Intuitively, this downward bias arises because instrumental variables does not fully correct for endogeneity in small samples when instruments are weak or when many instruments are used—i.e., IV is biased in the direction of OLS.30 Table 3 also reports results using the LIML estimator, which is

30 See Stock, Wright, and Yogo (2002) for an overview of this issue. The concern is that the first stage of the IV procedure may pick up some of the endogenous variation in the explanatory variable in the presence of a large num-

Table 3—Alternative Specifications for Effects of Military Spending

<table>
<thead>
<tr>
<th></th>
<th>1. Output level instr.</th>
<th>2. Employment level instr.</th>
<th>3. Output per working age</th>
<th>4. Output OLS</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>States</td>
<td>Regions</td>
<td>States</td>
<td>Regions</td>
</tr>
<tr>
<td>Prime military contracts</td>
<td>2.48</td>
<td>2.75</td>
<td>1.81</td>
<td>2.51</td>
</tr>
<tr>
<td></td>
<td>(0.94)</td>
<td>(0.69)</td>
<td>(0.41)</td>
<td>(0.31)</td>
</tr>
<tr>
<td>Prime contracts plus</td>
<td>4.79</td>
<td>2.60</td>
<td>2.07</td>
<td>1.97</td>
</tr>
<tr>
<td>military compensation</td>
<td>(2.65)</td>
<td>(1.18)</td>
<td>(0.67)</td>
<td>(0.98)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,989</td>
<td>390</td>
<td>1,989</td>
<td>390</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>5. Output with oil controls</th>
<th>6. Output with real int. controls</th>
<th>7. Output LIML</th>
<th>8. BEA employment</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>States</td>
<td>Regions</td>
<td>States</td>
<td>Regions</td>
</tr>
<tr>
<td>Prime military contracts</td>
<td>1.32</td>
<td>1.89</td>
<td>1.40</td>
<td>1.80</td>
</tr>
<tr>
<td></td>
<td>(0.36)</td>
<td>(0.54)</td>
<td>(0.35)</td>
<td>(0.59)</td>
</tr>
<tr>
<td>Prime contracts plus</td>
<td>1.43</td>
<td>1.72</td>
<td>1.61</td>
<td>1.59</td>
</tr>
<tr>
<td>military compensation</td>
<td>(0.39)</td>
<td>(0.66)</td>
<td>(0.40)</td>
<td>(0.84)</td>
</tr>
<tr>
<td>Observations</td>
<td>1,989</td>
<td>390</td>
<td>1,989</td>
<td>390</td>
</tr>
</tbody>
</table>

Source: Nakamura and Steinsson (2014)
Cyclical Sensitivity

- Potential threat to identification:
  - Sensitivity of military spending correlated with overall cyclical sensitivity
- In fact cyclical sensitivity uncorrelated with military sensitivity
- Consider:
  \[ \Delta Y_{it} = \alpha_i + \gamma_t + \beta s_i \Delta Y_t + \epsilon_{it} \]
  where
  - \( s_i \) is average level of military spending in state \( i \)
- If states with high \( s_i \) are more cyclically sensitive, \( \beta > 0 \)
- In fact \( \beta < 0 \)
Baseline results have 50 instruments
Potential weak / many instrument problem
When instruments are weak / many, IV is biased towards OLS
  Intuition: Overfitting – i.e., fitting endogenous noise in 1st stage
  Good read: Stock-Wright-Yogo 02
Rule of thumb: First stage F-stat of excluded instruments > 10
  In our case, state reg with baseline instruments: First state F-stat = 5
  Multiplier biased by about 10% towards OLS
  (we ran extensive Monte Carlo simulations)
IV versus OLS

- Large difference between IV (1.4-2.8) and OLS (0.1-0.6)
- Why?
IV versus OLS

- Large difference between IV (1.4-2.8) and OLS (0.1-0.6)
- Why?
  - Endogeneity: States doing badly get more spending
  - Measurement error in spending variable

Sample period 1966-1982

Results: 1.3 (0.5), versus OLS of 0.2 (0.2) and Bartik of 2.0 (0.4)
**IV versus OLS**

- Large difference between IV (1.4-2.8) and OLS (0.1-0.6)
- Why?
  - Endogeneity: States doing badly get more spending
  - Measurement error in spending variable
- Eliminate only measurement error by instrumenting for prime contract spending with shipments data
  - Sample period 1966-1982
  - Results: 1.3 (0.5), versus OLS of 0.2 (0.2) and Bartik of 2.0 (0.4)
MULTIPLIER IN TIMES OF SLACK

Is the multiplier larger in times of slack?

\[
\frac{Y_{it} - Y_{it-2}}{Y_{it-2}} = \alpha_i + \gamma_t + \beta_h \frac{G_{it} - G_{it-2}}{Y_{it-2}} + (\beta_l - \beta_h) I_{it} \frac{G_{it} - G_{it-2}}{Y_{it-2}} + \epsilon_{it}
\]

- \( I_{it} \) is an indicator for periods of low slack
  - Based on unemployment at the start of interval
  - National slack: National unemployment rate is below its median
  - State slack: State unemployment rate is below its median

- \( \beta_h \): Multiplier in high slack periods
- \( \beta_l - \beta_h \): Difference in multiplier between low and high slack periods
our estimates for the time period as a whole. Given the limited number of business cycles in our sample, we are not, however, able to estimate these effects with much statistical precision. The difference in the multiplier in the high and low spending periods is only moderately statistically significant (with \( p \)-values of 0.06 and 0.07).

For employment, the multiplier estimates for the high slack periods are close to those for the period as a whole and the difference in the multiplier between the high and low spending periods is relatively small and statistically insignificant.

### Table 5—Effects of Military Spending in High versus Low Unemployment Periods

<table>
<thead>
<tr>
<th></th>
<th>Output</th>
<th></th>
<th>Employment</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>National slack</td>
<td>State slack</td>
<td>National slack</td>
<td>State slack</td>
</tr>
<tr>
<td>( \beta_h )</td>
<td>3.54</td>
<td>4.31</td>
<td>1.85</td>
<td>1.32</td>
</tr>
<tr>
<td></td>
<td>(1.55)</td>
<td>(1.80)</td>
<td>(0.87)</td>
<td>(0.81)</td>
</tr>
<tr>
<td>( \beta_l - \beta_h )</td>
<td>-2.80</td>
<td>-3.37</td>
<td>-0.75</td>
<td>0.03</td>
</tr>
<tr>
<td></td>
<td>(1.49)</td>
<td>(1.84)</td>
<td>(0.89)</td>
<td>(0.84)</td>
</tr>
</tbody>
</table>

Notes: A shorthand for the dependent variable is stated at the top of each column. The dependent variable is a two-year change divided by the initial value in each case. All variables are per capita. Standard errors are in parentheses. The unit of observation is US states for all regressions in the table. The two regressors are (i) the two-year change in military spending and (ii) the two-year change in military spending interacted with a dummy indicating low slackness. We employ two different measures of slackness: “National slack” refers to whether the national unemployment rate is below its median value over the sample period; “State slack” refers to whether the state unemployment rate is below its median value over the sample period. This yields the effect of spending during high unemployment periods (\( \beta_h \)) and the difference between the effect of spending during low and high unemployment periods (\( \beta_l - \beta_h \)). The national slack regressions include state and time fixed effects. The state slack regressions include state and time fixed effects interacted with the low slackness dummy. The regression are estimated by two-stage least squares. The sample period is 1966–2006. Output is state GDP. Employment is from the BLS payroll survey.

Source: Nakamura and Steinsson (2014)
Relative multiplier we estimate, not the same as aggregate multiplier

- States don’t have to pay for spending (financed federally)
- Monetary policy can’t react in cross-section
- Spillovers to other states
What Do We Learn?

- Relative multiplier we estimate, not the same as aggregate multiplier
  - States don’t have to pay for spending (financed federally)
  - Monetary policy can’t react in cross-section
  - Spillovers to other states
- One reaction:
  - Not so useful since this it not what we are really interested in
    (which is aggregate multiplier)
What Do We Learn?

- Relative multiplier we estimate, not the same as aggregate multiplier
  - States don’t have to pay for spending (financed federally)
  - Monetary policy can’t react in cross-section
  - Spillovers to other states

One reaction:
- Not so useful since this it not what we are really interested in
  (which is aggregate multiplier)

Different reaction:
- Perhaps relative multiplier is a powerful statistic in distinguishing between different models (e.g., RBC vs. New Keynesian)
- Aggregate multiplier is actually not very strong on that front
We can use relative multiplier estimate as a moment to distinguish between competing structural models.

To this end, we write down a two-region macro model that nests competing models (RBC and New Keynesian).

Calculate relative multiplier in different versions of the model.
The Model

- Two regions
  - Home and foreign goods imperfect substitutes
    - Use commodity flow data to estimate “openness” (US regions ≈ Spain)
  - Labor immobile (regressions in per capita terms)
  - Common monetary policy
  - Common tax policy

- Households consume and supply labor

- Firms hire labor and set prices
  - Neoclassical model: Prices adjust frictionlessly, economy responds efficiently to shocks
  - New Keynesian model: Sluggish price response, output may be inefficiently low
Consider several different cases:

- **Aggregate Monetary Policy:**

  \[ \hat{i}_t = \rho \hat{i}_{t-1} + (1 - \rho) (\phi_{\pi} \hat{\pi}_t^{ag} + \phi_{\gamma} \hat{y}_t^{ag} + \phi_{g} \hat{g}_t^{ag}) \]

  - Volcker-Greenspan: \( \rho = 0.8, \phi_{\pi} = 1.5, \phi_{\gamma} = 0.5, \phi_{g} = 0 \)
  - Constant real interest rate (\( r \) unresp. to \( G \))
  - Constant nominal interest rate (\( i \) unresp. to \( G \))

- **Aggregate Tax Policy:**

  - Constant labor income tax (lump-sum taxes vary)
  - Labor income tax balances budget
data from the model described in Section III, time-aggregating it up to an annual frequency, and running the regression (26) on this data.

The first column of Table 6 reports results on the closed economy aggregate multiplier. These results clearly indicate that the closed economy aggregate multiplier is highly sensitive to aggregate monetary and tax policy—a point emphasized by Woodford (2011); Eggertsson (2010); Christiano, Eichenbaum, and Rebelo (2011); and Baxter and King (1993). In the New Keynesian model with a Volcker-Greenspan monetary policy, it is quite low—only 0.20. The low multiplier arises because the monetary authority reacts to the inflationary effects of the increase in government spending by raising real interest rates. This counteracts the expansionary effects of the spending shock. For monetary policies that respond less aggressively to inflationary shocks, the closed economy multiplier can be substantially larger. For the constant real-rate policy, the multiplier is one (Woodford 2011). Intuitively, since the real interest rate remains constant rather than rising when spending increases there is no “crowding out” of consumption, implying that output rises one-for-one with government spending. For the constant nominal-rate policy, the multiplier is larger than one and can become very large depending on parameters. It is 1.70 if the government spending shock is relatively transient (half-life of one year, \( \rho_g = 0.85 \)). With more persistent government spending shocks (\( \rho_g = 0.933 \)) it becomes infinite. However, it should be kept in mind that the case we are considering is effectively assuming that the economy stays at the zero lower bound indefinitely. If the economy is expected to revert to, e.g., a Volcker-Greenspan monetary policy before some fixed future point the multiplier is finite.44 The intuition for the large multipliers with a constant nominal-rate policy is that the government spending shock raises inflationary expectations, which lowers the real interest rate and thereby “crowds in” private demand.

### Table 6—Government Spending Multiplier in Separable Preferences Model

<table>
<thead>
<tr>
<th>Panel</th>
<th>Closed economy aggregate multiplier</th>
<th>Open economy relative multiplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Panel A. Sticky prices</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Volcker-Greenspan monetary policy</td>
<td>0.20</td>
<td>0.83</td>
</tr>
<tr>
<td>Constant real rate</td>
<td>1.00</td>
<td>0.83</td>
</tr>
<tr>
<td>Constant nominal rate</td>
<td>( \infty )</td>
<td>0.83</td>
</tr>
<tr>
<td>Constant nominal rate (( \rho_g = 0.85 ))</td>
<td>1.70</td>
<td>0.90</td>
</tr>
<tr>
<td>Panel B. Flexible prices</td>
<td></td>
<td></td>
</tr>
<tr>
<td>Constant income tax rates</td>
<td>0.39</td>
<td>0.43</td>
</tr>
<tr>
<td>Balanced budget</td>
<td>0.32</td>
<td>0.43</td>
</tr>
</tbody>
</table>

Notes: The table reports the government spending multiplier for output deflated by the regional CPI for the model presented in the text with the separable preferences specification. Panel A presents results for the model with sticky prices, while panel B presents results for the model with flexible prices. The first three rows differ only in the monetary policy being assumed. The fourth row varies the persistence of the government spending shock relative to the baseline parameter values. The fifth and sixth rows differ only in the tax policy being assumed.

Source: Nakamura and Steinsson (2014)
**Unresponsive to Policy**

- Key advantage of relative multiplier: Not sensitive to changes in monetary and tax policy
- Intuition: Aggregate policy is “differenced out”
- Yields multiplier for relatively “unresponsive” monetary/tax policy
- Same as multiplier for small open economy with fixed exchange rate
Relative nominal interest rate fixed

May seem analogous to zero lower bound situation

- Stimulus lowers short-term real interest rate
Relative nominal interest rate fixed

May seem analogous to zero lower bound situation
  - Stimulus lowers short-term real interest rate

Crucial difference:
  - Long-term real interest rate doesn’t fall
  - Purchasing power parity must hold
  - Any rise in relative price level will be reversed
  - Demand determined by long-term real rate
undone by a fall in relative prices in that region later on. In fact, after their initial jump, relative prices are anticipated to fall more in the long run than they are anticipated to rise further in the short run. This implies that the relative long-term real interest rate actually rises slightly in the home region in response to an increase in government spending.

To more clearly see the intuition for this result, Figure 4 presents the impulse response of the price level and the real interest rate in the home region relative to the foreign region after a government spending shock in our model. The home price level rises for several periods, but then falls back to its original level. This movement in prices implies that the real interest rate in the home region initially falls, but then rises above its steady state level for a prolonged period. Figure 5 shows what happens to consumption in the home region relative to the foreign region after a government spending shock. Despite the short-run fall in the real interest rate, consumption falls. This is because households anticipate high real rates in the future—equivalently, they face a high current long-term real interest rate—and therefore cut their consumption.

Since the relevant interest rate for consumption decisions—the long-term real interest rate—actually rises slightly in response to an increase in government spending irrespective of the persistence of the shock and other parameters, the fixed relative nominal interest rate policy in a monetary union is fundamentally different from

---

46 Parsley and Wei (1996) present evidence for rapid convergence of relative prices following regional shocks using data for US regions.

47 Corsetti, Kuester, and Muller (2011) show that the same logic holds for the case of a small open economy with a fixed exchange rate.

---

**Figure 4. Prices and Real Interest Rates after a Government Spending Shock**

*Note:* The figure plots the relative price level and the relative real interest rate in the two regions for the model with separable preferences after a positive government spending shock to the home region.

**Source:** Nakamura and Steinsson (2014)
Introduce “ultra-Keynesian” features

- Consumption and work are complements
  (Aguiar-Hurst, 2005; Schmitt-Grohe-Uribe, 2010)
- Such complementarities can raise fiscal multiplier
  (Monacelli-Perotti, 2008; Bilbii, 2009; Hall, 2009)
- Intuition:
  - Higher output raises marginal utility of consumption
  - This leads to even higher output
Previous work by Monacelli and Perotti (2008), Bilbiie (2011), and Hall (2009) has shown that allowing for complementarities between consumption and labor can have powerful implications for the government spending multiplier. The basic intuition is that, in response to a government spending shock, households must work more to produce the additional output. This raises consumption demand since consumption is complementary to labor. But to be able to consume more, still more production must take place, further raising the effects on output.

The second column of Table 7 presents estimates of the open economy relative multiplier for the model with GHH preferences. The New Keynesian model with GHH preferences can match our empirical findings in Section II of an open economy multiplier of roughly 1.5 (assuming a quarterly persistence of $\rho_g = 0.933$ as in the military spending data). As in the model with separable preferences, this statistic is entirely insensitive to the specification of aggregate policies. For the case of more transitory government spending shocks ($\rho_g = 0.50$), the open economy relative multiplier rises to 2.0. The Neoclassical model, however, continues to generate a low multiplier ($0.3$) in this model.

Figure 6 plots relative output and consumption in the New Keynesian model with GHH preferences after a positive shock to home government spending. Both output and consumption rise on impact by a little more than twice the amount of the shock. They then both fall more rapidly than the shock. The fact that the initial rise in consumption is as large as the rise in output—which is partly fulfilling increased orders from the government—implies that the home region responds to the shock by running a trade deficit in the short run. Consumption eventually falls below its steady state level for a period of time. During this time, the home region is running a trade surplus. Intuitively, the complementarity between consumption and labor implies that home households want to shift their consumption toward periods of high work effort associated with positive government spending shocks.

Table 7—Government Spending Multiplier in GHH Model

<table>
<thead>
<tr>
<th>Panel A. Sticky prices</th>
<th>Closed economy aggregate multiplier</th>
<th>Open economy relative multiplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Volcker-Greenspan monetary policy</td>
<td>0.12</td>
<td>1.42</td>
</tr>
<tr>
<td>Constant real rate</td>
<td>7.00</td>
<td>1.42</td>
</tr>
<tr>
<td>Constant nominal rate</td>
<td>$\infty$</td>
<td>1.42</td>
</tr>
<tr>
<td>Constant nominal rate ($\rho_g = 0.50$)</td>
<td>8.73</td>
<td>2.04</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th>Panel B. Flexible prices</th>
<th>Closed economy aggregate multiplier</th>
<th>Open economy relative multiplier</th>
</tr>
</thead>
<tbody>
<tr>
<td>Constant income tax rates</td>
<td>0.00</td>
<td>0.30</td>
</tr>
<tr>
<td>Balanced budget</td>
<td>$-0.18$</td>
<td>0.30</td>
</tr>
</tbody>
</table>

Notes: The table reports the government spending multiplier for output deflated by the regional CPI for the model presented in the text with the GHH preferences specification. Panel A presents results for the model with sticky prices, while panel B presents results for the model with flexible prices. The first three rows differ only in the monetary policy being assumed. The fourth row varies the persistence of the government spending shock relative to the baseline parameter values. The fifth and sixth rows differ only in the tax policy being assumed.

Source: Nakamura and Steinsson (2014)
Increasingly important in macro:

- Mian-Sufi 14, Nakamura-Steinsson 14, Autor-Dorn-Hansen 13, Baraja-Hurst-Ospina 16, Martin-Phillipon 17, ...

Key challenge:

- How to go from regional responses to aggregate responses
- Cross-sectional responses don’t directly answer key aggregate questions
- GE effects absorbed by time fixed effects
- Common to do “back-of-envelope” calculation
- Typically invalid

Fully specified general equilibrium model needs to translate regional responses to aggregate responses

Regional responses helpful in distinguishing between models
Appendix
Figure 2 illustrates the close relationship between these shipment data and the military procurement data for several states over this period—giving us confidence in the prime military contract data as a measure of the timing and magnitude of regional military production. To summarize this relationship, we estimate the following regression of shipments from a particular state on military procurement,

\[
M_{S_{it}} = \alpha_i + \beta M_{P_{S_{it}}} + \epsilon_{it},
\]

where \(M_{S_{it}}\) is the value of shipments from the Census Bureau data and \(M_{P_{S_{it}}}\) is military procurement spending. This regression yields a point estimate of \(\beta = 0.96\), indicating that military procurement moves on average one-for-one with the value of shipments. The small differences between the two series probably indicate that they both measure regional production with some error. As we discuss below, one advantage of the instrumental variables approach we adopt is that it helps adjust for this type of measurement error.

C. Effects of Government Spending Shocks

The first row of Table 2 reports the open economy relative multiplier \(\beta\) in regression (1) for our baseline instruments. Standard errors are in parentheses and are clustered by states or regions. In the second row of Table 2, we present an analogous set of results using a broader measure of military spending that combines military procurement spending with compensation of military employees for each state.

Figure 2. Prime Military Contracts and Military Shipments

Source: Nakamura and Steinsson (2014). State of prime contractor is where majority of work is done.
Federal vs. Local Financing

- Baseline model has complete markets
  (local vs. federal financing doesn’t matter)

- As robustness, we consider incomplete markets model and compare multipliers with local and federal financing

- Differences are small for our calibration
  (see Fahri-Werning 16 for cases where differences are bigger)

- Multiplier slightly larger with federal financing when prices are sticky
  (demand effect from increased wealth)
preferences raises the open economy relative multiplier when compared to the case of separable preferences is thus that the monetary union implies an accommodative “relative” monetary policy—sufficiently accommodative not to choke off the increase in relative output.

Summing up our results thus far, our estimates of equation (1), based on the military procurement data, yield an open economy relative multiplier of roughly 1.5. This lies far above the open economy relative multipliers for the Neoclassical model—which are below 0.5 for both separable preferences and GHH preferences. Our empirical estimate of 1.5 is also substantially higher than the open economy relative multiplier of 0.83 implied by the New Keynesian model with separable preferences. The New Keynesian model with GHH preferences, however, is able to match the open economy relative multiplier we estimate in the data. Our results are thus consistent with a model in which demand shocks can have large effects on output—if monetary policy is sufficiently accommodative (as it is at the zero lower bound).

D. Model with Incomplete Financial Markets

The model we develop in Section III features complete financial markets across regions of the economy. This implies that all risk associated with differential taxes and labor income across regions—possibly arising from government spending shocks—is perfectly shared. In a recent paper, Farhi and Werning (2012) have shown that in a monetary union with incomplete financial markets across regions, regional government spending multipliers can differ substantially depending on whether the spending is financed by local taxes or federal taxes. Table 8 presents open economy relative multipliers for a version of our model in which the only financial asset that is traded across regions is a noncontingent bond. For this model, we present results for two assumptions about how spending is financed: locally financed spending and federally financed spending.

Table 8 shows that these two versions of the incomplete markets model yield similar results about the open economy relative multiplier to the baseline complete markets model. In the case of federally financed spending, the open economy relative multiplier rises to 0.90 when prices are sticky. The intuition for this is that

<table>
<thead>
<tr>
<th>Table 8—Government Spending Multipliers in Incomplete Markets Model</th>
</tr>
</thead>
<tbody>
<tr>
<td>Closed economy aggregate multiplier</td>
</tr>
<tr>
<td>-------------------------------------</td>
</tr>
<tr>
<td><strong>Panel A. Sticky prices</strong></td>
</tr>
<tr>
<td>Baseline model (complete markets)</td>
</tr>
<tr>
<td>Incomplete markets, locally financed</td>
</tr>
<tr>
<td>Incomplete markets, federally financed</td>
</tr>
<tr>
<td><strong>Panel B. Flexible prices</strong></td>
</tr>
<tr>
<td>Baseline model (complete markets)</td>
</tr>
<tr>
<td>Incomplete markets, locally financed</td>
</tr>
<tr>
<td>Incomplete markets, federally financed</td>
</tr>
</tbody>
</table>

Notes: The table reports the government spending multiplier for output deflated by the regional CPI for a version of the model presented in the text with separable utility in which the only financial asset traded across regions is a noncontingent bond. Panel A presents results for the model with sticky prices, while panel B presents results for the model with flexible prices.

Source: Nakamura and Steinsson (2014)
Baseline model has fixed amount of capital per firm

Does allowing for variable capital change results?

Two versions:

- Firm-specific capital (Woodford, 2003, 2005, Altig et al., 2011)
- Regional capital markets (Christiano et al., 2005)

Firm-specific capital model yields similar results to baseline

Regional capital markets reduce strategic complementarity in price setting (highly unrealistic model)
time and increase investment when the shock occurs. In contrast, the model with regional capital markets yields a smaller multiplier than the baseline model. This occurs despite investment rising as in the firm-specific capital model. The main reason for the fall in the multiplier is that the regional nature of the capital market reduces the degree of strategic complementarity of the price setting decisions across firms relative to the baseline model (since firms that raise their price can costlessly reduce the amount of capital they rent).

Clearly, the assumption that firms rent the capital that they use each period on frictionless regional capital markets is unrealistic. Eichenbaum and Fisher (2007) and Altig et al. (2011) show that adopting the more realistic setting of firm-specific capital helps New Keynesian models with capital match the sluggish response of prices to aggregate disturbances without resorting to unrealistic assumptions about the frequency of price adjustment or the indexing of prices. The final row of Table 9 shows that, with flexible prices, the open economy relative multiplier is close to zero in the firm-specific capital model. Table 9 also presents open economy relative multipliers for CPI inflation. The New Keynesian models generate small increases in relative inflation. This lines up well with our empirical findings on relative inflation. In contrast, the model with flexible prices counterfactually implies a much sharper rise in relative inflation rates.

F. Welfare

The welfare consequences of government spending depend not only on the multiplier, but also on the utility agents derive from the goods and services purchased by the government. Woodford (2011) and Werning (2012) provide an extensive discussion of the welfare consequences of government spending. To illustrate the main forces, suppose household utility can be represented by $U(C_t, L_t, G_t)$ and the production function is $Y_t = f(L_t)^{0.5}$. Household utility may then be written as $U(Y_t - G_t, f^{-1}(Y_t), G_t)$. Following Woodford (2011), we can differentiate this and get

$$dU_\text{d}G = (U_C - U_L f L) dY + (U_G - U_C).$$

51 For simplicity, we abstract from investment, heterogeneous labor markets, and price dispersion due to price rigidity. And we assume that government spending is financed by lump-sum taxes.

Table 9—Open Economy Relative Multiplier in Models with Variable Capital

<table>
<thead>
<tr>
<th>Model</th>
<th>Output</th>
<th>CPI inflation</th>
</tr>
</thead>
<tbody>
<tr>
<td>Baseline model (fixed capital)</td>
<td>1.42</td>
<td>0.17</td>
</tr>
<tr>
<td>Firm-specific capital model</td>
<td>1.47</td>
<td>0.15</td>
</tr>
<tr>
<td>Regional capital market model</td>
<td>0.98</td>
<td>0.09</td>
</tr>
<tr>
<td>Firm-specific capital model, flexible prices</td>
<td>0.25</td>
<td>0.36</td>
</tr>
</tbody>
</table>

*Notes:* The table reports the open economy relative government spending multiplier for output and CPI inflation for our baseline model with GHH preferences and the two models with variable capital, also with GHH preferences. Output is deflated by the regional CPI.

*Source:* Nakamura and Steinsson (2014)