ECONOMIC RESEARCH
FEDERAL RESERVE BANK OF ST. LOUIS

WORKING PAPER SERIES

# The Local-Spillover Decomposition of an Aggregate Causal Effect

| | |
|---|---|
| **Authors** | Timothy Conley, Bill Dupor, Mahdi Ebsim, Jingchao Li, and Peter B. McCrory |
| **Working Paper Number** | 2021-006A |
| **Creation Date** | May 2021 |
| **Citable Link** | https://doi.org/10.20955/wp.2021.006 |
| **Suggested Citation** | Conley, T., Dupor, B., Ebsim, M., Li, J., McCrory, P.B., 2021; The Local-Spillover Decomposition of an Aggregate Causal Effect, Federal Reserve Bank of St. Louis Working Paper 2021-006. URL https://doi.org/10.20955/wp.2021.006 |

Federal Reserve Bank of St. Louis, Research Division, P.O. Box 442, St. Louis, MO 63166

# The Local-Spillover Decomposition of an Aggregate Causal Effect[*]

Timothy G. Conley[†], Bill Dupor[‡], Mahdi Ebsim[§], Jingchao Li[¶], and Peter B. McCrory[||]

May 12, 2021

**Abstract**

This paper presents a method to decompose the causal effect of government spending into: (*i*) a local (or direct) effect, and (*ii*) a spillover (or indirect) effect. Each effect is measured as a multiplier: the unit change in output of a one unit change in government spending. We apply this method to study the effect of U.S. defense spending on output using regional panel data. We estimate a positive local multiplier and a negative spillover multiplier. By construction, the sum of the local and spillover multipliers provides an estimate of the aggregate multiplier. The aggregate multiplier is close to zero and precisely estimated. We show that enlisting disaggregate data improves the precision of aggregate effect estimates, relative to using aggregate time series alone. Our paper provides a template for researchers to conduct inference about local, spillover and aggregate causal effects in a unified framework.

Keywords: local and spillover effects, aggregate fiscal multiplier

[*]The analysis set forth does not reflect the views of the Federal Reserve Bank of St. Louis or the Federal Reserve System. Conley thanks the Social Science and Humanities Research Council of Canada for support.

[†]Western University, tconley3@uwo.ca.

[‡] Federal Reserve Bank of St. Louis, william.d.dupor@stls.frb.org, billdupor@gmail.com.

[§] Federal Reserve Bank of St. Louis, mahdi.ebsim@stls.frb.org.

[¶] School of Business, East China University of Science and Technology, jingchaoli@ecust.edu.cn.

[||] University of California–Berkeley, pbmccrory@berkeley.edu.

*No man is an island, [...]*
John Donne, 1692

Over the past few decades, there has been a movement in economics towards applying disaggregate—particularly regional—data to answer macroeconomic questions. For example, Chodorow-Reich (2020) cites 50 papers published between 2012 and 2018 in top economics journals that combine cross-regional variation in exogenous shocks and regional outcomes in an attempt to infer the macroeconomic effect of such shocks. For these papers, the statistical unit of observation is a region.

This contrasts with the standard approach to causal inference in empirical macroeconomics, embodied in early work by Sims (1972), in which the unit of observation is an entire economy, which is sampled repeatedly over time. Variation in the treatment, which for Sims is an exogenous change in monetary policy, occurs along the time dimension.

When the treatment of interest is defined at the regional level, spillovers across regions must be considered—regions are not "islands." If region $A$ receives a treatment, region $B$ may be affected by that treatment even if $B$ receives no treatment on its own. A cross-region "spillover" might arise, for example, from regional trade in goods or movements in factors of production, and constitutes a classic violation of the Stable Unit Treatment Value Assumption (SUTVA) which requires that potential outcomes be unaffected by the treatment status of other observational units. The local effect of a treatment need not equal the treatment's aggregate effect in the presence of spillovers. For example, if there are negative spillovers across regions, then a positive local effect will overstate the aggregate effect of the treatment.

This paper develops a technique to decompose an aggregate effect of an exogenous shock into its local and spillover components. We illustrates this approach via an application studying the effects of of local and spillover defense spending on output. We decompose the causal effect of government defense spending into: (*i*) a local (or direct) effect, and (*ii*) a spillover (or indirect) effect. Using regional defense spending data, we estimate a positive local multiplier and a negative spillover multiplier. The sum of the local and spillover effects, i.e., the aggregate government spending multiplier, is precisely estimated (when based on regional data) and close to zero.

Typically, macroeconomists employ aggregate, time-series variation to estimate the combined effects of (*i*) and (*ii*). With exogenous variation in the aggregate treatment, this approach can estimate the treatment's aggregate effect but cannot distinguish between its local and spillover effects. With panel or cross-sectional data, researchers often only estimate the local effect of treatment (*i*). While this object may be of interest on its own, these papers do not typically estimate either the aggregate or spillover effects of treatment (*ii*).

Our local-spillover decomposition respects the relationship that the sum of the local and spillover effects equals the aggregate effect. In contrast, papers from the fiscal multiplier literature that do estimate both local and spillover effects using cross-sectional data (e.g., geographic regions) fail to impose this adding-up relationship—oftentimes as a direct consequence of the inclusion of time fixed effects.[1]

We estimate the local and spillover multipliers of government defense spending using efficient GMM. This allows us to parse the local and spillover channels of government spending shocks while at the same time testing parameter restrictions across regions. We find support for a parameter restriction that yields sharper inference on the aggregate effects of defense spending than when relying solely upon aggregate, time-series variation.

The GMM moment conditions we use exploit three identification strategies. First, throughout, we rely on the common argument that national defense spending is exogenous because it is determined by international geopolitical factors and national security concerns and thus orthogonal to local economic conditions in any region, over time. We present one specification that maintains an exogeneity assumption for defense spending at the regional level—this specification is analogous to a naive OLS regression. As regional spending is more prone (than aggregate spending) to endogeneity concerns, we investigate two alternative specifications that combine different regional instruments with aggregate spending. One relies upon a timing assumption by using lagged regional spending as an instrument and the other uses a subset of regional defense expenditures that are largely pre-determined as an instrument. Our quantitative results agree across all three specifications.

Our four main results are: (1) The local multiplier is positive and the spillover multiplier is negative; (2) The sum of the two, i.e. the aggregate multiplier, is close to zero; (3) This panel-based estimate of the aggregate multiplier is substantially more precise than one based on aggregate data alone; and (4) Our exercise provides a template for researchers to conduct inference about local, spillover and aggregate causal effects in a unified framework.

# 1 The Local-Spillover Decomposition

We use annual data at the state level from 1964 through 2011. Since we have fewer time periods than states, we adopt two strategies to mitigate estimation problems via dimension reduction. First, we partially aggregate into $N$ sets of states, creating a panel based on census divisions. Our basic units of analysis are $N = 9$ census divisions.[2] Our second strategy for dimension reduction is to explore further aggregation of moments across sets of census divisions in some specifications. We use $g_{i,t}$ and $q_{i,t}$ to denote real military spending and real output, respectively, for division $i$ in year $t$. National aggregates of $g_{i,t}$ and $q_{i,t}$ are denoted $g_t$ and $q_t$, respectively.

We focus on the effect of cumulative changes in local and spillover military spending upon cumulative changes in real output. Each region's normalized cumulative change in real output

[1]For example, see Suárez Serrato and Wingender (2016) (spillovers among U.S. counties), and McCrory (2020) (trade spillovers between U.S. states).

[2]The precise division-state assignments are in Table 3.

over horizon $h$ is defined as:

$$y^h_{i,t+h} = \sum_{j=0}^{h} \left(q_{i,t+j} - q_{i,t-1}\right) / q_{t-1} \tag{1}$$

We use initial period national output, $q_{t-1}$, to normalize changes to facilitate aggregation.

Each census division receives both a local and spillover treatment. The local treatment for census division $i$ is the cumulative change in its defense spending:

$$x^h_{i,t+h} = \sum_{j=0}^{h} \left(g_{i,t+j} - g_{i,t-1}\right) / q_{t-1} \tag{2}$$

We again normalize using lagged national output, $q_{t-1}$. Scaling by lagged output in both the outcome and treatment ensures that in the regressions of $x^h_{i,t+h}$ upon $x^h_{i,t+h}$ described below, the estimated coefficient can be interpreted as a multiplier: the percentage change in output in response to a government spending increase equal to one percent of output.

We focus on a 4-year horizon, i.e., the on-impact effect of the spending plus that over the three following years, by setting $h = 3$. At shorter horizons, spillovers arising from reallocation may be less apparent. At longer horizons, we are limited by the time length of our sample. Also, a longer horizon moves us away from our motivating interest in the business cycle frequency stabilization of fiscal policy.

Our operational definition of a spillover variable is a 'leave-out' mean of $x_{j,t+h}$ across all census divisions $j \neq i$. We use $\tilde{x}^h_{i,t+h}$ to denote this leave-out mean, with:

$$\tilde{x}^h_{i,t+h} \equiv \frac{1}{N-1} \sum_{j \neq i} x^h_{j,t+h}. \tag{3}$$

Next, $y^h_{i,t+h}$, $x^h_{i,t+h}$ and $\tilde{x}^h_{i,t+h}$ are each de-meaned by the census division averages of the corresponding variables. For the remainder of the paper, we work with the de-meaned variables, reusing the notation $y^h_{i,t+h}$, $x^h_{i,t+h}$ and $\tilde{x}^h_{i,t+h}$. National aggregates of $x^h_{i,t}$ and $y^h_{i,t}$ are referred to as $x^h_t$ and $y^h_t$, respectively.

We estimate a local-spillover decomposition via a regression of (de-meaned) normalized cumulative output changes on (de-meaned) local and spillover normalized defense spending changes:

$$y^h_{i,t+h} = \psi_h x^h_{i,t+h} + \omega_h \tilde{x}^h_{i,t+h} + u^h_{i,t+h} \tag{4}$$

We call $\psi_h$ the local multiplier at horizon $h$ for a region. The local multiplier gives the response of own-division output to a one unit own-division defense spending increase holding fixed defense spending in other divisions. The spillover multiplier, $\omega_h$, gives the response of own-division output to other areas' defense spending holding fixed own spending.

Taking sums of both sides over $i$ using (4), we have

$$y_{t+h}^{h} = (\psi_h + \omega_h)\, x_{t+h}^{h} + u_{t+h}^{h} \tag{5}$$

We call $\psi_h + \omega_h$ the aggregate multiplier because it gives the response of aggregate output to a one unit increase in the aggregate defense spending.

We also examine an alternative estimator of the aggregate multiplier via a set of regressions with division-level outcomes $y_{i,t+h}^{h}$ and an aggregate level regressor:

$$y_{i,t+h}^{h} = \gamma_h \frac{x_{t+h}^{h}}{N} + \epsilon_{i,t+h}^{h} \tag{6}$$

By construction, the value of $\gamma_h$ is equal to the aggregate multiplier $\psi_h + \omega_h$.

We focus on cumulative multipliers which give the accumulated change in output over a specific horizon with respect to the accumulated change in military spending over the same horizon. As such, it reflects both the output benefits and spending costs added up over a given horizon. Ramey and Zubairy (2018) explain compellingly that cumulative multipliers are more valuable from a policy perspective than other (sometimes reported) statistics, such as impact multipliers and peak multipliers.

We estimate equation (4) via the Generalized Method of Moments (GMM) using a 2 by 1 instrument $z_{i,t+h}^{h}$. For each census division $i$, we have the vector moment condition:

$$E\left(z_{i,t+h}^{h} u_{i,t+h}^{h}\right) = 0 \tag{7}$$

In our benchmark specifications we simply stack these moment conditions for the nine census divisions. In alternative specifications, we reduce dimensions by aggregating these moment conditions within a set of non-overlapping groups indexed with $r = 1, \ldots, R$. Using $S_r$ to refer to the set of census divisions contained in group $r$, we have:

$$E\left(\sum_{i \in S_r} z_{i,t+h}^{h} u_{i,t+h}^{h}\right) = 0 \tag{8}$$

When $R = 4$, we use groups that correspond to the four census regions.[3] For estimation, we simply stack these moment conditions for all four census regions and estimate via GMM. We note that when using group moment conditions, we maintain the same definition of spillovers and parameter interpretation as those at the division level. Our four census region moments merely provide a reduction in dimension with eight versus 18 moments in the (nine) census division moments case. We expect this to improve the quality of large-sample approximations, albeit at a

[3] See Appendix Table 4.

potential loss of information.

We consider three alternative instrument sets $z^h_{i,t+h}$, each of which is a two by one vector. In all three cases, the first element is the scaled cumulative change in national military spending, $x^h_{t+h}$. National defense spending can be treated as exogenous by following the geo-political factors argument described above. The three alternative instrument sets differ from each other only in their second element: the instrument for local spending.

In the first case, we use $x^h_{i,t+h}$ as our local instrument. Thus, the first set of instruments relies on exogeneity in defense spending at both the division and aggregate level. This case is analogous to naively estimating the model by OLS, since these two instruments span the same space as the local and spillover variables in equation (4). Nevertheless, we include this specification as as useful comparison.

A salient endogeneity concern is that, even if national defense spending is exogenous over time, the federal government may reallocate spending across states in response to current or anticipated economic conditions. For example, congressional representatives for a state experiencing bad shocks may be able to boost spending impacting their constituents. Our aggregation to the division level partly mitigates this concern, but we are still motivated to utilize two alternative instrument strategies. Our second instrument set pairs $x^1_{i,t+1}$ (rather than $x^h_{i,t+h}$) with $x^h_{t+h}$. We call $x^1_{i,t+1}$ the cumulative change in short-horizon own spending. Short-horizon own spending is plausibly exogenous to the error term because redeploying resources across areas takes significant time to plan and then to implement. On the other hand, regional business cycles, driven by the error term, happen at a relatively high frequency and are largely unanticipated.

Our third instrument set exploits the fact that one component of defense spending, fulfillment of previously arranged contracts, is predetermined. The inherent "stickiness" in contracts makes reallocating contract-based spending across regions infeasible. This instrument set pairs $x^h_{t+h}$ with the contracts component of division-level spending, also at a short horizon, which we call $x^{1,Contracts}_{i,t+1}$.

We also estimate equation (6) via the Generalized Method of Moments (GMM). For each census division $i$, we use the scalar moment condition:

$$E\left(z^h_{t+h}\epsilon^h_{i,t+h}\right) = 0 \tag{9}$$

where the instrument $z^h_{t+h}$ is equal to $x^h_{t+h}$, exploiting the assumed exogeneity of aggregate defense spending over time for each division. These moment conditions are simply stacked for all nine divisions.

We obtain efficient GMM estimates via Iterated GMM.[4] We estimate the long-run variance-

[4]For all the estimates presented, the Iterated GMM procedure begins with a Two-step GMM procedure starting from an identity weighting matrix. The efficient two-step GMM point estimate is then used to re-estimate the weighting matrix. This new weighting matrix is used to obtain third-step GMM point estimate used to re-estimate the weighting

covariance matrix of our moment conditions via a Bartlett/Newey-West covariance matrix estimator that places non-zero weight on the sample autocovariances up to nine lags. To reduce dimension, we also impose an assumption of second-moment independence. Sample moments of $z$ are multiplied by sample moments of estimated $u$ and then combined via Bartlett weights to estimate the long-run variance-covariance matrix of our moment conditions.[5]

Our data sources are as follows. State-level output data are from the Bureau of Economic Analysis. Defense wages and contracts apportioned to state geographies are from Dupor and Guerrero (2017). This variable, summed across regions, is less than NIPA-measured defense spending in each year of the sample because the federal government does not provide geographic identifiers for every dollar of military spending. To construct a regional defense spending variable that aggregates to a series consistent with the NIPA data, we calculate $g_{i,t}$ as the year $t$ NIPA-measured aggregate defense spending multiplied by the year $t$, division $i$ apportionment of the defense wage and contract amounts.

## 2 Results

Table 1 first reports the estimate of $\psi_h + \omega_h$ from (5) with aggregate data via OLS (with Bartlett/Newey-West standard errors). Recall this specification estimates the aggregate multiplier using the aggregate time series and does not decompose the multiplier into its local and spillover components. The aggregate multiplier estimate is 0.98 (SE=1.07). The point estimate implies that a dollar increase in national defense spending causes national output to increase by 98 cents. Note that the estimate is extremely imprecise. One cannot reject a multiplier that is zero or large in absolute value and is either negative or positive.

The next row contain the estimate of $\gamma_h$ using (6). This specification has slope parameter equal to that for the first row ($\psi_h + \omega_h$) but exploits our 9 regional-level time series. The model is over-identified with 8 ($= 9 - 1$) degrees of freedom. The $p$-value associated with this specification's $J$-statistic equals 0.36. Thus, the data do not reject the common slope restriction across divisions. The aggregate multiplier estimate equals -0.22 (SE=0.12). Note that the point estimate from the panel-based estimate lies well within any conventionally-sized confidence interval of the aggregate-data based one. Importantly, we see a dramatic increase in the precision from using the regional moment conditions relative to using a single national moment condition. The standard error falls from 1.07 to 0.12. The large increase in precision is due to our use of efficient GMM weighting, both its inverse-variance weighting aspect and the exploitation of covariances across divisions appear to play a role.

Next, we present the paper's key deliverable: a decomposition of the aggregate effect of de-

matrix, again used to get another point estimate, etc. This iteration continues until the distance between GMM objective iterates is less than $10^{-3}$. All estimates converge in less than 20 iterations.

[5]Our results are robust to using alternate Bartlett weights including 2 to 13 lags.

Table 1: Cumulative output multiplier of defense spending, baseline results

| | (1) Local | (2) Spillover | (3) Aggregate | (4) Overid. Test |
|---|---|---|---|---|
| Aggregate, 1 Nationwide Region | | | 0.98 | |
| | | | (1.07) | |
| Aggregate, 9 Divisions | | | -0.22* | 0.36 |
| | | | (0.12) | |
| Decomposition (IV #1) | 0.52*** | -0.36*** | 0.16 | 0.80 |
| | (0.13) | (0.11) | (0.17) | |
| Decomposition (IV #2) | 0.57*** | -0.32*** | 0.26 | 0.78 |
| | (0.16) | (0.11) | (0.20) | |
| Decomposition (IV #3) | 0.53*** | -0.34*** | 0.19 | 0.59 |
| | (0.17) | (0.11) | (0.21) | |

Notes: $T = 43$, $h = 3$. For the aggregated-data based estimate, dependent variable=four year cumulative growth in national output. For all other rows, moment conditions are evaluated at the nine census divisions, estimation is done via Iterative GMM, $N = 9$, and the dependent variable=four year cumulative change in division output scaled by lagged national output. Column (4) contains the over-identification test $p$-values. * $p < .1$, ** $p < .05$, *** $p < .01$

fense spending into a local and spillover component. We begin with IV#1 results which, as explained above, rely on exogeneity of both aggregate and division-level defense spending. These provide a useful comparison with the other two IV specifications, which utilize short-horizon and contracts-only spending to correct for potential endogeneity.

The row labelled "Decomposition (IV #1)" gives estimates of (4), the local-spillover decomposition. First, the $p$-value for the over-identification test equals 0.80, indicating that the data once again do not reject the model. The local multiplier equals 0.52 (SE=0.13), which implies that a one-dollar increase in a region's defense spending (holding fixed defense spending in other regions) increases that region's output by 52 cents. The spillover multiplier equals -0.36 (SE=0.11), indicating that a one-dollar increase in other regions' spending (holding own-region defense spending fixed) reduces that region's output by 36 cents. Both coefficients are statistically different from zero.

Next, the row labelled "Decomposition (IV #2)" estimates (4) but with our second instrument set. Recall this instrument vector contains the scaled cumulative change in aggregate defense spending and short-horizon (one-year) local spending. Compared to the first set, point estimates are similar and their standard errors increase slightly. The local multiplier equals 0.57 (SE=0.16) and the spillover multiplier equals -0.32 (SE=0.11).

The row labelled "Decomposition (IV #3)" presents estimates of (4) using our third instrument set. This set pairs aggregate cumulative defense spending with short-horizon local spending resulting only from defense procurement. Procurement is spending related to shipments fulfilling

Table 2: Cumulative output multiplier of defense spending, additional specifications

| | (1) Local | (2) Spillover | (3) Aggregate | (4) Overid. Test |
|---|---|---|---|---|
| Aggregate, 1 Nationwide Region | | | 0.98 | |
| | | | (1.08) | |
| 4 Regions | | | 0.11 | 0.23 |
| | | | (0.67) | |
| 9 Divisions | | | -0.22$^{*}$ | 0.36 |
| | | | (0.12) | |
| Decomposition (IV # 1), 4 Regions | 0.86$^{***}$ | -0.67 | 0.19 | 0.28 |
| | (0.16) | (0.63) | (0.69) | |
| 9 Divisions | 0.52$^{***}$ | -0.36$^{***}$ | 0.16 | 0.80 |
| | (0.13) | (0.11) | (0.17) | |
| Decomposition (IV # 2), 4 Regions | 0.95$^{***}$ | -0.49 | 0.46 | 0.35 |
| | (0.36) | (0.76) | (0.70) | |
| 9 Divisions | 0.57$^{***}$ | -0.32$^{***}$ | 0.26 | 0.78 |
| | (0.16) | (0.11) | (0.20) | |
| Decomposition (IV # 3), 4 Regions | 0.72$^{*}$ | -0.60 | 0.12 | 0.37 |
| | (0.43) | (0.80) | (0.69) | |
| 9 Divisions | 0.53$^{***}$ | -0.34$^{***}$ | 0.19 | 0.59 |
| | (0.17) | (0.11) | (0.21) | |

Notes: Dependent variable=four year cumulative change in census division output scaled by lagged national output. $N = 9$, $T = 43$, $h = 3$ and $R = 1, 4, 9$ across specifcations. The aggregate multiplier for "Decomposition" specifications equals the sum of the local and spillover coefficients. Column (4) contains the over-identification test $p$-values. * $p < .1$, ** $p < .05$, *** $p < .01$

contracts made by the Department of Defense. We see qualitatively the same story as with Decomposition IV #2, similar point estimates and slightly larger standard errors relative to Decomposiiton IV-1. The local multiplier is 0.53 (SE=0.17) and the spillover multiplier is -0.34 (SE=0.11).

Qualitatively, our decomposition estimates of local and spillover parameters are the same across our three IV specifications. This is reflected in the similarities in the aggregate multiplier in column (3) across each of our last three rows of estimates. The point estimates lie within a tight band of 0.16 and 0.26, with standard errors that are substantially smaller than one would find using aggregate data alone.

Our finding of an aggregate defense spending multiplier that is less than one is in line with existing research based solely on macro data, given our level of precision. For example, Hall (2009) finds a defense spending multiplier equal 0.47 using aggregate, annual data between 1948 and 2008. Applied to a more comparable period (1960 – 2008) to our sample, he finds a multiplier 0.13. Barro (1981) estimates defense spending multipliers of about 0.6 for increases in spending associated with WWI, WWII and the Korean War.

Our findings of positive local and negative spillover multipliers are consistent with a factor reallocation explanation in line with evidence presented in Blanchard and Katz (1992).[6] We do note that alternative explanations for the negative spillover multiplier may be equally if nor more important.

We conclude with an illustration of the benefits of using dis-aggregated data at the census division level. Table 2 presents results with different levels of aggregation, values for R. For each of the specifications (i.e., aggregate, decomposition (IV #1), decomposition (IV #2), decomposition (IV #3)), we see that increasing the number of groups tends to substantially reduce the corresponding estimates' standard errors. Consider, for example, how the spillover coefficient for the "Decomposition (IV #1)" case changes as we increase the number of groups from the census region level (four) to the census division level (nine). The point estimate increases from -0.67 to -0.36. The standard error falls from 0.63 to 0.11. The increase in precision going from four regions to nine divisions results from our ability with divisions to exploit relatively low variance divisions via efficient GMM.

Finally, a potential issue with our approach, as with any IV method, is instrument strength for IV#2 and IV#3 results.[7] The standard errors reported above are based on strong-instrument asymptotics. In the Appendix, we assess the potential for weak instrument concerns by directly comparing our results to weak-instrument robust confidence sets constructed following Chernozhukov and Hansen (2008) (CH). Confidence regions based on strong instrument asymptotics line up closely with analogous CH weak instrument robust regions. [8] Our CH confidence regions qualitatively agree with those derived from our strong instrument asymptotics, supporting the use of strong instrument approximations in our results tables.

## 3 Conclusion

When researchers employ cross-sectional or panel data to estimate the causal effect of some treatment, they typically only identify a local effect. In the presence of spillovers between observational units, this local effect is generally different from an aggregate (or average) effect.

However, in many contexts, particularly in macroeconomics, policymakers are interested primarily in an aggregate effect. In this paper, we show how to augment a standard local effect regression to account for potential spillovers between observational units. By exploiting the panel structure of the data, we show how to jointly identify the local and spillover effects of the treatment in a way that allows for easy conversion to treatment's aggregate effect.

To apply our methodology in a straightforward manner, a suitable application should have

[6]Intuitively, an area that receives defense dollars sees an increase in economic activity locally; however, as capital and labor reallocate towards that area, other areas see declining economic activity.

[7]Instruments IV#1 is equivalent with using regressors as instruments, so weak instrument issues do not arise.

[8]The CH procedure is described and results are presented in the Appendix.

two crucial ingredients, beyond simply an orthogonality condition. First, one requires sufficiently long disaggregate (individual/group) time series data on both the outcome and treatment variables of interest. This is because our methodology relies upon moment conditions that are satisfied along the time dimension. Second, one needs some level of treatment effect similarity across individuals/groups. This is important to facilitate aggregate interpretations of our panel-based estimates. A fruitful extension of our methodology would be to allow for some treatment effect heterogeneity in terms of both local and spillover effects

## References


Barro, R. (1981), "Output Effects of Government Purchases," *Journal of Political Economy*, 89, 6, 1086–1121.

Blanchard, O. and L. Katz L (1992), "Regional Evolutions," *Brookings Papers on Economic Activity*. 1992(1), 1-75.

Chernozhukov, V. and C. Hansen (2008), "The Reduced Form: A Simple Approach to Inference with Weak Instruments," *Economic Letters*, 100, 1, 68-71.

Chodorow-Reich, G. (2020), "Regional Data in Macreconomics: Some Advice for Practioners," *Journal of Economic Dynamics and Control*, 115.

Dupor, B. and R. Guerrero (2017), "Local and Aggregate Fiscal Policy Multipliers," *Journal of Monetary Economics*, 12, 16-30.

Hall, R. (2009), "By How Much Does GDP Rise if the Government Buys More Output?" *Brookings Papers on Economic Activity*, 40, 183-249.

McCrory, P. (2020), "Tradable Spillovers of Fiscal Policy: Evidence from the 2009 Recovery Act," working paper.

Ramey, V. and S. Zubairy (2018), "Government Spending Multipliers in Good Times and in Bad: Evidence from U.S. Historical Data," *Journal of Political Economy*.

Sims, C. (1972), "Money, Income and Causality," *American Economic Review*, 62 (4), 540-552.

Suárez Serrato, J. C. and P. Wingender (2016), "Estimating Local Fiscal Multipliers," NBER Working Paper No. 22425.

# A Appendix: Weak-instrument Robust Inference

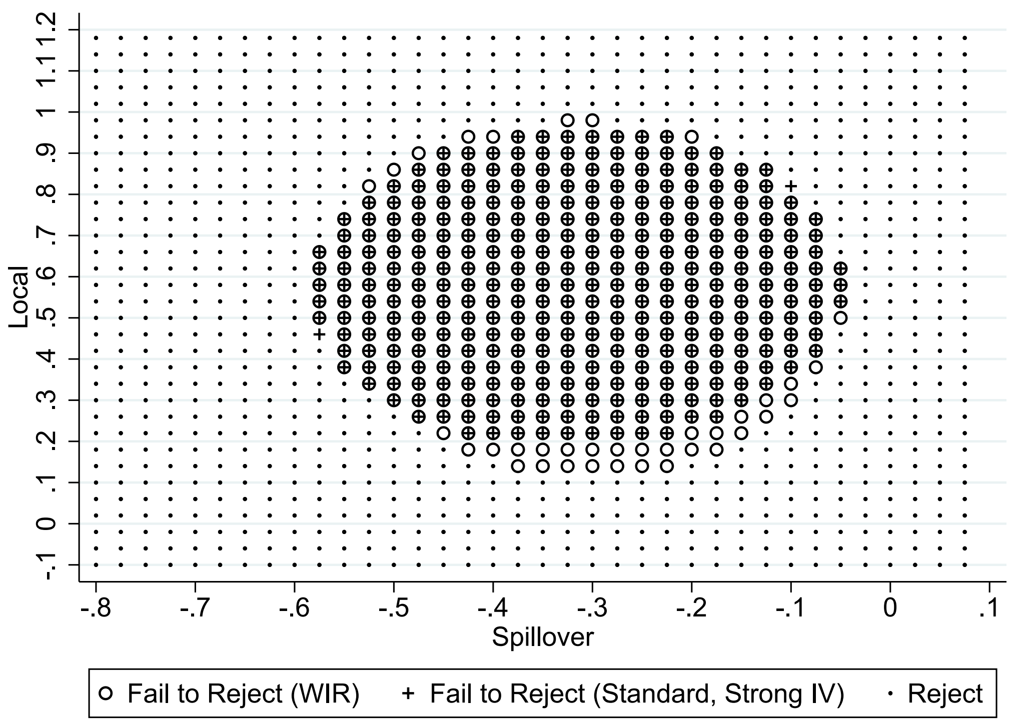
In this Appendix, we present Weak Instrument Robust (WIR) confidence sets following Chernuzukov and Hansen (2008) (CH). The CH method exploits that fact that under the null hypothesis that the true parameter values are $(\psi_0, \omega_0)$ one can construct the error terms in equation (4) via:

$$u^h_{i,t+h} = y^h_{i,t+h} - \psi_0 x^h_{i,t+h} - \omega_0 \tilde{x}^h_{i,t+h}. \tag{10}$$

Then this null hypothesis can be tested via a regression of $u^h_{i,t+h}$ upon the instrument vector $z^h_{i,t+h}$ with the exclusion restriction implying they should be orthogonal. The coefficients in a regression of $u^h_{i,t+h}$ upon $z^h_{i,t+h}$ should jointly be zero. To implement this across our census divisions we jointly estimate a stacked set of division-specific moment conditions and allow for dependence just as in our benchmark specification with R=9, and use a Wald test for the $z^h_{i,t+h}$ coefficients being zero. If this Wald test fails to reject at the 5% level, then the point $(\psi_0, \omega_0)$ is in our 95% WIR confidence set. We then repeat for a grid of values for $(\psi_0, \omega_0)$.

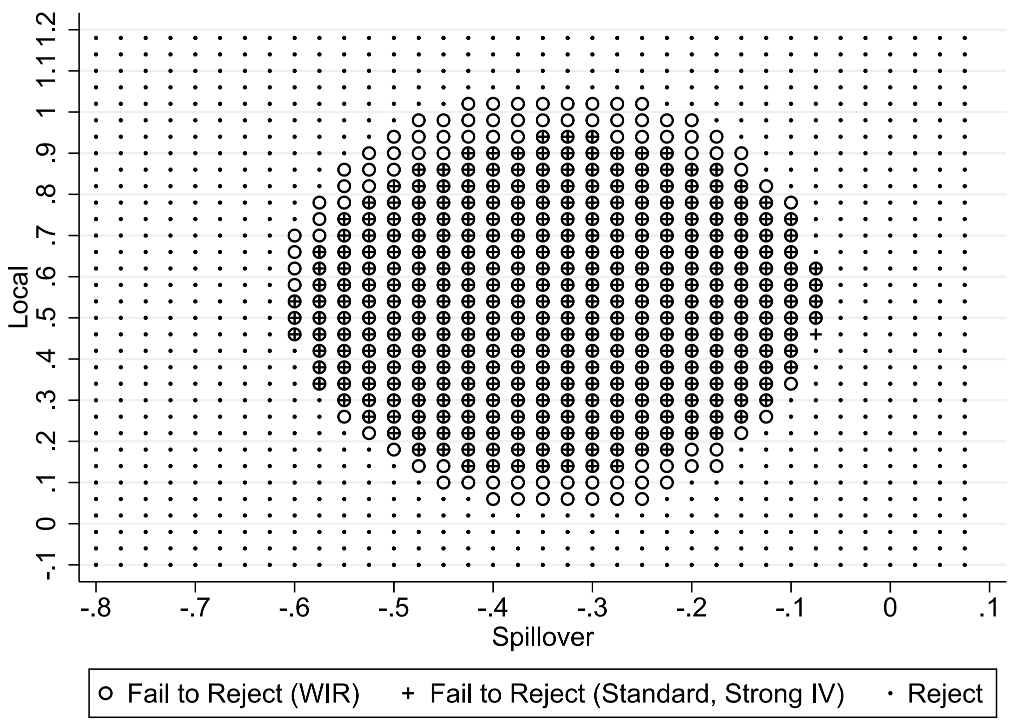
Figure 1 contains a pairwise scatter plot where each location corresponds to one point of a grid of $(\psi_0, \omega_0)$. 95% confidence sets are depicted for both standard, strong-IV asymptotics and WIR confidence sets for IV-based local-spillover decomposition with $R = 9$ (corresponding to the 'Decomposition (IV #2)' row of estimates from Table 1). Solid dots indicate rejection of this parameter value under both standard, strong-IV and WIR approximations. Plusses indicate a confidence set under strong instrument approximations and open circles represent the WIR confidence set. Figure 2 is the analogous scatter plot of the standard and WIR confidence region for our IV#3 specification (corresponding to the 'Decomposition (IV #3)' row of estimates from Table 1). Both Figures indicate a substantial agreement between inference under standard, strong-IV approximations and WIR methods.

Figure 1: Comparison of weak-instrument-robust and standard-asymptotics 95 percent confidence regions, decomposition (IV #2) specification




Note: WIR=weak-instrument robust. $R = 9$, $N = 9$, $T = 43$. The union of the +'s corresponds to the standard, strong-instrument asymptotic 95 percent confidence set. The union of the hollow-circles corresponds to the WIR 95 percent confidence set.

Figure 2: Comparison of weak-instrument-robust and standard-asymptotics 95 percent confidence regions, decomposition (IV #3) specification




Note: WIR=weak-instrument robust. $R = 9$, $N = 9$, $T = 43$. The union of the +'s corresponds to the standard, strong-instrument asymptotic 95 percent confidence set. The union of the hollow-circles corresponds to the WIR 95 percent confidence set.

# B Appendix: Census Divisions Partitions

Table 3: 9 Census Divisions

| | Name | Members |
|---|---|---|
| 1 | East North Central | IL, IN, WI, MI, OH |
| 2 | East South Central | AL, KY, MS, TN |
| 3 | West North Central | IA, KS, MN, MO, ND, NE, SD |
| 4 | West South Central | AR, LA, OK, TX |
| 5 | New England | CT, MA, ME, NH, RI, VT |
| 6 | Middle Atlantic | NJ, PA, NY |
| 7 | South Atlantic | DE, MD, NC, VA, WV, FL, GA, SC |
| 8 | Mountain | AZ, CO, UT, NM, NV, ID, MT, WY, WA, OR |
| 9 | Pacific | CA, OR, WA |

Table 4: 4 Census Regions

| | Name | Members |
|---|---|---|
| 1 | Midwest | IA, KS, MN, MO, ND, NE, SD, IL, IN, WI, MI, OH |
| 2 | South | AL, KY, MS, TN, AR, LA, OK, TX,DE, MD, NC, VA, WV, FL, GA, SC |
| 3 | Northeast | CT, MA, ME, NH, RI, VT, NJ, PA, NY |
| 4 | West | AZ, CO, UT, NM, NV, ID, MT, WY, WA, OR, CA, OR, WA |