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

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

February 26, 2023

Abstract

This paper presents a method to decompose causal effects of U.S. defense spending on income, using panel data, into: *(i)* a local effect, and *(ii)* a spillover effect. We estimate positive local and spillover multipliers. By construction, the sum of the local and spillover multipliers provides an estimate of the aggregate multiplier, which is slightly greater than one in our benchmark specification. Enlisting disaggregate data improves precision relative to using aggregate data alone. More generally, we provide a template 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.

[§] New York University, msebsim@gmail.com.

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

[‖] LinkedIn, peter.mccrory@gmail.com.

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 effects 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 estimate local and spillover effects of region-level treatments. Our parameterization allows us to infer effects of aggregate-level treatments from these region-level local and spillover estimates. We illustrate this approach by studying the local and spillover effects of regional defense spending on regional income. We decompose the causal effect of government defense spending into: (*i*) a local (or direct) effect, and (*ii*) a spillover (or indirect) effect.

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, most papers using regional data do not estimate either the aggregate or spillover effects of treatment (*ii*). Morever, as we detail in the following section, existing papers that do estimate spillover effects have not linked the spillover and aggregate effects of a regional treatment.

We estimate the local and spillover multipliers of government defense spending over a four year horizon 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 two different 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. Our first specification maintains an exogeneity assumption for four-year cumulative growth in defense spending at the regional level—this specification is analogous to an OLS regression. Our second specification instruments for four-year regional spending growth with a component of it that is plausibly exogenous due to timing. Specifically, we use initial one-year change in regional share spending. This short-horizon change in defense spending in each region is plausibly exogenous to the region-level error term because it is infeasible for the federal government to reallocate contracts, personnel and other spending without some implementation delay.

Our four main results are: (1) The local and spillover multipliers are positive; (2) The sum of the two, i.e. the aggregate multiplier, is slightly greater than one; (3) This panel-based estimate of the aggregate multiplier is 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 2006. Since we have fewer time periods than states, we mitigate estimation problems via dimension reduction. We partially aggregate the data into N sets of states, creating a panel based on geographic proximity and roughly equally-sized divisions. Our basic units of analysis are $N = 9$ divisions.[1] We use $g_{i,t}$ and $q_{i,t}$ to denote real military spending and real income, 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 income. Each region's normalized cumulative change in real income over horizon h is defined as:

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

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

Each division receives both a local and spillover treatment. The local treatment for division i

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

is the cumulative change in its defense spending:

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

We again normalize using lagged national income, q_{t-1}. Scaling by lagged income in both the outcome and treatment ensures that in the regressions of $y_{i,t+h}^{h}$ upon $x_{i,t+h}^{h}$ described below, the estimated coefficient can be interpreted as a multiplier: the dollar change in local income in response to one dollar of local or spillover defense spending.

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 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 divisions $j \neq i$. We use $\tilde{x}_{i,t+h}^{h}$ to denote this leave-out mean, with:

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

Equation (3) defines a division's spillover treatment to be the change in average defense spending outside of that division. This leave off mean spillover definition is motivated by the observation that regions within the U.S. share common monetary and federal fiscal policies, suggesting a very broad form of spillover. For example, suppose defense spending in one region drives up inflation, then the central bank might raise its target interest rate, which could influence economic activity similarly across all other regions. As a second example, an increase in one region's federal defense spending—if tax financed—would affect tax rates and potentially economic activity similarly across the other regions.

In related work, researchers have instead constructed spillover measures impacting a region using weighted averages of other regions' fiscal treatments. For example, Auerbach and Gorodnichenko (2013) study output growth and government spending across countries and use import shares to construct weights. Suárez Serrato and Wingender (2016) study U.S. county-level employment and transfer spending data and use geographic closeness to construct weights. McCrory (2020) studies the cross-state spillovers from government spending associated with trade flows between U.S. states. Dupor and McCrory (2018) study U.S. regional employment changes and stimulus spending and use worker commuting flows to build weights. From a macroeconomist's perspective, our local-spillover decomposition has an advantage relative to these approaches because it facilitates examining the aggregate effects implied by our estimated local and spillover effects. In contrast, the above papers do not jointly estimate the aggregate effects implied by their estimated local effects and spillover effects.

Next, $y^h_{i,t+h}$, $x^h_{i,t+h}$ and $\tilde{x}^h_{i,t+h}$ are each de-meaned by the 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 income changes on (de-meaned) local and spillover normalized defense spending changes:

$$y^h_{i,t+h} = \psi_h x^h_{i,t+h} + \omega_h \tilde{x}^h_{i,t+h} + \phi'_h W_{i,t-1} + u^h_{i,t+h} \tag{4}$$

$W_{i,t-1}$ is a vector of conditioning information containing $y^0_{i,t-1}$, $x^0_{i,t-1}$ and x^0_{t-1}.

We construct an aggregate econometric equation by summing this division-level equation across divisions. On the left-hand side, we have the aggregate outcome. On the right side, we have the aggregate treatment variable multiplied by a coefficient, equal to the sum of the local and spillover. Because the right-hand side treatment variable is a ratio, aggregation by summing works if the division-level treatments share a common denominator. Using an aggregate, rather than a division-level variable, is advantageous for this purpose, but nonstandard in existing research.

We call ψ_h the local multiplier at horizon h for a region. The local multiplier gives the response of own-division income to a one unit own-division defense spending increase holding fixed average defense spending in other divisions. The spillover multiplier, ω_h, gives the response of own-division income to other areas' average defense spending holding fixed own spending (and $W_{i,t-1}$).

Besides providing a parsimonious form of spillover, the leave-off mean definition of the spillover has two additional advantages. First, to construct the local-spillover decomposition, we require the division-level estimation equation to aggregate to the national level. The leave-off mean is a simple form that ensures this aggregation. Second, since we have many divisions, the leave-off mean is highly correlated across divisions. As such, identification of the spillover effect in our paper will come almost entirely from the time series variation in the spillover. This provides a stark contrast with papers that estimate spillovers using a cross-section (e.g., Dupor and McCrory (2018)) or a panel with time fixed effects (e.g., Auerbach and Gorodnichenko (2013)). These papers rely on cross-sectional variation in the spillover treatment. Combining our and the existing approaches merits future research.

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

$$y^h_{t+h} = (\psi_h + \omega_h)\, x^h_{t+h} + \sum_{j=1}^{N} \phi'_h W_{t-1} + u^h_{t+h}. \tag{5}$$

We call $\psi_h + \omega_h$ the aggregate multiplier because it gives the response of aggregate income 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 $y^h_{i,t+h}$ and an aggregate level regressor:

$$y^h_{i,t+h} = \gamma_h \frac{x^h_{t+h}}{N} + \xi'_h W_{i,t-1} + \epsilon^h_{i,t+h} \tag{6}$$

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

We focus on cumulative multipliers which give the accumulated change in income over a specific horizon with respect to the accumulated change in military spending over the same horizon. As such, it reflects both the income 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 5 by 1 instrument $z^h_{i,t+h}$. For each division i, we have the vector moment condition:

$$E\left(z^h_{i,t+h} u^h_{i,t+h}\right) = 0 \tag{7}$$

We then simply stack these moment conditions for the nine divisions and estimate via GMM.

Our benchmark interval estimates use conventional, strong instrument approximations, i.e. that instruments satisfy the usual relevance requirement. We investigate whether weak instruments are an issue by also examining results using weak instrument robust methods (See Appendix A). Furthermore, because we use a local projection method to identify the causal effect of interest, we need instruments that—conditional upon $W_{i,t-1}$—are orthogonal to past and future fundamental shocks to the economy, what Stock and Watson (2018) call lead-lag exogeneity. To investigate whether conditioning upon $W_{i,t-1}$ is adequate for lead-lag exogeneity to be plausible, we follow Stock and Watson's suggested diagnostic of checking whether, conditional on $W_{i,t-1}$, current period instruments are unforecastable by lagged outcome variables. In results not reported in the paper, we find this to be the case for all instruments defined below, evidence consistent with their being lead-lag exogeneous, given $W_{i,t-1}$.

We consider two alternative instrument sets $z^h_{i,t+h}$ to estimate (7). The first set, denoted (NE), contains $W_{i,t}$, x^h_{t+h} and $x^h_{i,t+h}$. That is, there are no excluded instruments. Exogeneity of x^h_{t+h}, national military spending relies on the common argument that aggregate defense spending because it is determined by geo-political factors and hence orthogonal to regional business cycle shocks. We anticipate that it will be useful in identifying our spillover coefficient as it should have a strong correlation with $\tilde{x}^h_{i,t+h}$ due to the substantial overlap in their components. Next, $x^h_{i,t+h}$ is our division-level 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 estimating the model by OLS (with cross-division parameter restrictions), since the last two instruments span the same space as the local and spillover variables in equation (4).

Our second instrument set, denoted (IV-S) differs only in its last element: We require an instrument with division-level local variation, which we anticipate will be useful for identifying our local coefficient. A salient endogeneity concern is that, even with exogenous national defense spending over time, the federal government might reallocate spending across states in response to state-level differential economic conditions. For example, congressional representatives for a state experiencing bad shocks may be able to boost spending impacting their constituents. Relative to our first instrument set, it replaces $x_{i,t+h}^{h}$ with $x_{i,t}^{0}$. This is simply the impact (i.e., year t) one-year scaled change in defense spending in division i. The validity of $x_{i,t}^{0}$ as an instrument reflects a realistic timing restriction.[2] While it might be possible for the federal government to reallocate resources across geographic areas over the period of several years, at a one-year horizon is implausible that the federal government could undertake any significant repositioning of contracts, troops, as well as newly purchased equipment and structures in response to political motivations working through local business cycles.

We obtain efficient GMM estimates via Iterated GMM.[3] We estimate the long-run variance-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 four lags.[4] To reduce dimension, we also impose an assumption of second-moment independence of z and u. 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.

Our data sources are as follows. State-level income 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]Blanchard and Perotti (2002), for example, achieve identification of general government spending using a timing restriction.

[3]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 matrix, again used to get another point estimate, etc. This iteration continues until the Euclidean distance between the current and previous value of estimated parameter vector is less than 10^{-3}. All estimates converge in less than 30 iterations.

[4]Our results are robust to using alternate Bartlett weights including up to 10 lags.

Table 1: Cumulative income multiplier of defense spending, benchmark results

	(1) Coef./SE	(2) Coef./SE	(3) Coef./SE	(4) Coef./SE	(5) Coef./SE	(6) Coef./SE
Own division	0.23*	0.26***	0.30***	0.31***	-	-
spending	(0.13)	(0.08)	(0.07)	(0.07)		
Other division	-	-	0.78*	0.78*	-	-
spending			(0.44)	(0.46)		
Aggregate multiplier	-	-	1.08**	1.09**	0.71	1.39***
			(0.47)	(0.48)	(0.67)	(0.41)
Instruments?	NE	IV-Agg	NE	IV-S	NE	NE
P-value	0.44	0.38	0.52	0.67		0.38
Number of division	9	9	9	9	1	9
Number of years	46	46	46	46	46	46

Notes: For columns (1) though (4) and (6), moment conditions are evaluated at the division-level, estimation is done via Iterative GMM, and the dependent variable is the four year cumulative change in division income scaled by lagged national income. For the column (5), dependent variable=four year cumulative growth in national income. P-value is from the test of overidentification. NE denotes no excluded instruments. See the text for the description of IV-Agg and IV-S. * $p < .1$, ** $p < .05$, *** $p < .01$

2 Results

Table 1 provides our benchmark results. Column (1) reports estimates of equation (4) where use own-spending to instrument for itself (i.e., no excluded instruments) and additionally restricts the coefficient on other-division spending (i.e., the spillover multiplier) to equal zero. Both the treatment (defense spending) and outcome (income) are measured over a cumulative four-year horizon as explained previously. The coefficient equals 0.23 (SE=0.13), indicating that a one dollar increase in a division's defense spending raises own division income by 23 cents. Thus, the local multiplier is significantly less than one but statistically different from zero at the 10 percent level. Because we are estimating a system with more moment restrictions than estimation parameters, we report the p-value associated with the overidentification test. For this specification, the statistic equals 0.44, indicating that the model is not rejected at any conventional confidence level.

Column (2) maintains the zero-spillover restriction but uses aggregate spending instead of own-spending as the instrument, denoted IV-Agg in the table. The coefficient equals 0.26 (SE=0.08), which is somewhat greater than the value in column (1). Moreover, the local multiplier is now statistically different from zero at a one percent level. As in column (1), the p-value associated with the overidentifying restriction indicates that the model is not rejected at any conventional confidence level.

Next, we present the paper's key deliverable: a decomposition of the aggregate effect of defense spending into a local and spillover coefficient. First, column (3) uses own and average other

Table 2: Cumulative income multiplier of defense spending, standard estimators

	(1) Coef./SE	(2) Coef./SE	(3) Coef./SE	(4) Coef./SE	(5) Coef./SE
Own-division	0.44	0.33	0.23***	0.10	-
spending	(0.30)	(0.23)	(0.09)	(0.40)	
Other-divisions	-	-	0.54	0.61	-
spending			(0.63)	(0.91)	
Agg multiplier	-	-	0.77	0.71	0.71
			(0.68)	(0.61)	(0.63)
Instruments?	NE	IV-Agg	NE	IV-S	NE
Partial F		164.91		129.36	
Number of division	9	9	9	9	1
Number of years	46	46	46	46	46

Notes: For columns (1) though (4), dependent variable=four year cumulative change in division income scaled by lagged national income. Estimation is done via pooled least-squares or pooled IV. Driscol-Kraay standard errors are reported. For the column (5), dependent variable=four year cumulative growth in national income. NE denotes no excluded instruments. See the text for the description of IV-Agg and IV-S. * $p < .1$, ** $p < .05$, *** $p < .01$

division spending as instruments (i.e., no excluded instruments). Thus, this specification is similar to that of column (1) except we impose no restriction on the other-division spending coefficient. The local multiplier equals 0.30 (SE=0.07), which is similar to the analogous values in columns (1) and (2). The spillover multiplier equals 0.78 (SE=0.44). According to the point estimate, a one-dollar increase in other divisions' spending (holding own-division spending fixed) increases own-division income by 0.78 cents. This indicates substantial cross-division spillovers of defense spending.

By construction, the sum of the local and spillover multipliers is the aggregate multiplier, which equals 1.08 (SE=0.47) in column (3). This implies that the aggregate multiplier is statistically different from zero and the point estimate indicates a slightly greater than one-for-one response of national income to exogenous national spending increases. It is statistically different from zero at a 5 percent level; however, the associated confidence interval includes values greater than one.

Column (4) modifies the estimate from column (3) by changing the instruments. As described previously, the instrument set in this case consists of the cumulative changes in aggregate defense spending and the lagged, share-weighted cumulative changes in aggregate defense spending. The second instrument is constructed to account for the possibility that the federal government reallocates defense spending across division based on differences in the divisional business cycles. Both the local and spillover multipliers increase somewhat. The aggregate multiplier equals 1.09 (SE=0.48) for this specification.

Having estimated the aggregate multiplier using the disaggregate spending data and the divisional moment conditions, we compare these estimates to one that uses the corresponding aggregate treatment and outcome variables to estimate the aggregate multiplier using the traditional macro approach. Thus, we estimate equation (5) directly using aggregate data. This estimate is presented in column (5).

Note that the first two rows of this column are empty because the local and spillover multipliers are not separately identified. This can be seen by inspecting equation (5). Next, note that the number of divisions equals one because we have only one "national" division and the *p*-value entry is empty because the specification is exactly identified. The aggregate multiplier equals 0.71 (SE=0.67). The multiplier is positive and somewhat less than one, but very imprecisely estimated.

Finally, column (6) estimates the aggregate multiplier using equation (6). This corresponds to using the divisional moment conditions but including only aggregate spending as a regressor, rather than both own and other-division spending. The aggregate multiplier estimate equals 1.39 (SE=0.41). Note that the point estimate, standard error and p-value for the overidentification test are close to those from column (4). This indicates the lack of precision of using the traditional aggregate approach relative to using division level moment conditions and efficient GMM. This occurs because we achieve efficiency gains by using nine moment restrictions and assuming common response coefficients.

For comparison, Table 2 estimates the five of the above six specification except we use standard econometric routines that rely on exactly identified, rather than overidentified, models. Each column of Table 2 is comparable to the correspondingly numbered column of Table 1.[5] In each case, we report Driscoll-Kraay standard errors with the identical Bartlett parameter as that from Table 1.

Previewing the main message of Table 2, we find point estimates that are qualitatively very similar to those of the Table 1. The local and spillover multipliers are positive and less than one. The aggregate multiplier is positive and close to one, but somewhat smaller than the corresponding values that are estimated using GMM. The main difference is that the standard errors are much larger when we use the standard routines relative to our new approach.

Column (1) estimates equation (4) via pooled least squares assuming a zero spillover effect. The coefficient equals 0.44 (SE=0.30), indicating that a positive but less than one-for-one response of division income to own-division defense spending. This is somewhat larger than the analogous estimate from Table 1. Most importantly, the standard error is more than twice as large using the exactly identified model rather than the overidentified model in the previous table.

The column (2) specification instead uses aggregate defense spending rather than own-defense spending as the instrument, which we refer to as IV-Agg in the table. The multiplier is again less than one; however, it is very imprecisely estimated. It is not statistically different from zero.

[5]The only exception is that Table 2 has no column (6), because there is no in that case.

Importantly, the estimate of the multiplier is substantially larger than the corresponding one from Table 1.

Columns (3) and (4) share a similar message. Both imply positive local and spillover multipliers. The implied aggregate multipliers are close to one, 0.77 and 0.71 respectively. These are somewhat smaller than the corresponding values from Table 1, which are 1.08 and 1.09. Each of the multipliers is estimated more precisely using our approach than with aggregates.

Finally, column (5) estimates the aggregate multiplier using aggregated data. The point estimate equals 0.71, which is identical to that from column (5) of Table 1. The standard errors are slightly different across the two (0.67 in Table 1 and 0.63 in Table 2). This occurs because the former estimate imposes second moment independence whereas the latter does not.

Dupor and Guerrero (hereafter, DG) also compare local and aggregate multipliers for defense spending in the U.S. While similar in some respects, the current paper makes and illustrates two methodological advances: a formal decomposition of the aggregate multiplier into local and spillover components; an estimation approach that achieves efficiency gains.

With respect to the aggregate multipliers, DG and the current paper are relatively similar. Both examine 4-year cumulative multipliers using annual data. However, they differ in a few ways. The former uses national income data, a sample that starts in 1951 and a somewhat narrower definition of defense spending, whereas the latter uses GDP, a sample that starts in 1964 and a broader defense spending measure. In the current paper, we are able to use the shorter sample (and thus the broader defense spending definition and broader measure of aggregate activity) because we exploit efficiency gains from using stacked moment conditions and treatment slope coefficient restrictions.

With respect to the state-level based estimates, the results in DG and the current paper are not comparable because of the methodological advance presented in the current paper. Whereas the current paper estimates a state level panel regression that can be summed to a directly comparable aggregate time series regression, the DG state level regression is not aggregable. Finally, DG do attempt to estimate a spillover coefficient which is broadly similar to the goal of this paper; however, rather than using a leave-off sum of national spending to measure the spillover treatment, they use the sum of trade flow weighted defense spending in other states. While trade flows may reflect one potential source of spillovers, it may miss others (e.g., common fiscal and monetary policy and factor market linkages).

3 Conclusion

When researchers employ cross-sectional or panel data to estimate the causal effect of a 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 two crucial ingredients, beyond appropriate instruments. First, one requires sufficiently long disaggregate (individual/group) time series data because we rely on intertemporal moments, allowing for potentially strong cross section dependence. 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

Auerbach, A. and Y. Gorodnichenko (2019), "Output Spillover from Fiscal Policy," *American Economic Review* 103, 141-146.

Blanchard, O. and R. Perotti (2002), "An Empirical Characterization of the Dynamic Effects of Changes in Government Spending and Taxes on Output," *The Quarterly Journal of Economics*, 117(4), 1329-1368.

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 P. McCrory (2018), "A Cup Runneth Over: Fiscal Policy Spillovers from the 2009 Recovery Act," *The Economic Journal*, Vol. 128 (611), 1476-1508

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.

Nakamura, E. and J. Steinsson (2014), "Fiscal Stimulus in a Monetary Union: Evidence from US Regions," *American Economic Review*, 104(3), 753-92.

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.

Stock, J. and M. Watson (2018), "Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments," *Economic Journal*, 128, 917-948.

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

A Appendix: Weak-instrument Robust Inference

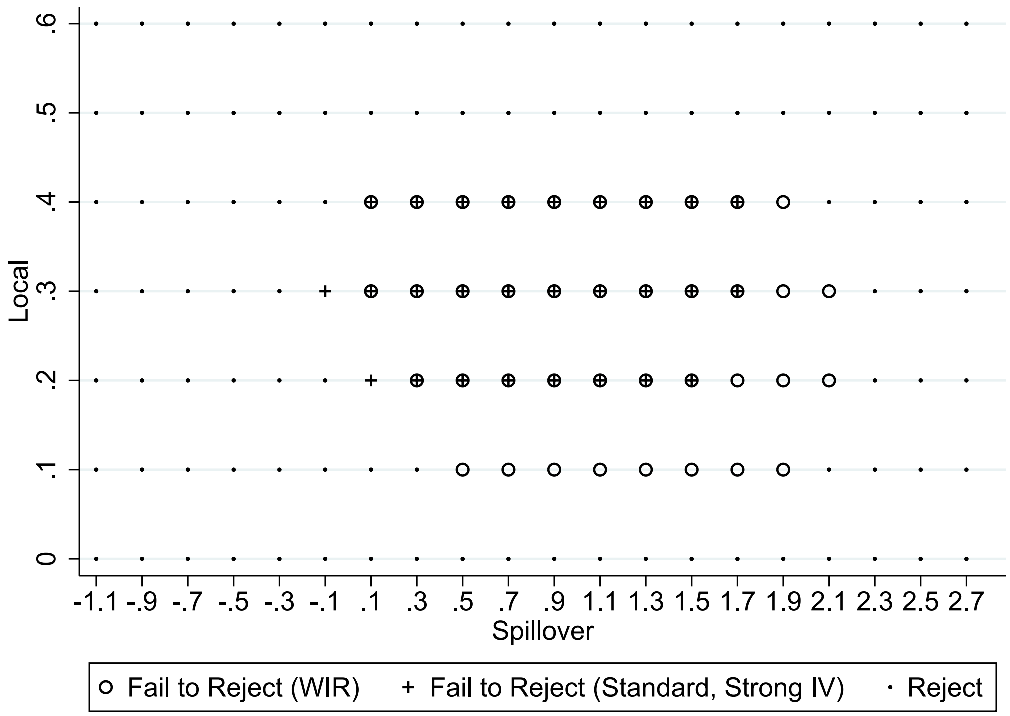
In this Appendix, we present Weak Instrument Robust (WIR) confidence sets for our IV-S specification following Chernuzukov and Hansen (2008) (CH). We first project all variables upon the first three elements of $W_{i,t-1}$ (i.e., the included instruments) and work with residuals. Rather than introduce new notation we recycle notation here and use the same notation as in the main text to denote variables' residuals, e.g. here $x^h_{i,t+h}$ is the residual from a projection of scaled, division i cumulative defense spending upon the included instruments. The CH method exploits the fact that under the null hypothesis that the true parameter values are (ψ_0, ω_0) one can construct the error terms in the (post projection) analog to equation (4) via:

$$u^h_{i,t+h} = y^h_{i,t+h} - \psi_0 x^h_{i,t+h} - \omega_0 \tilde{x}^h_{i,t+h}. \tag{8}$$

Then this null hypothesis can be tested via a regression of $u^h_{i,t+h}$ upon the instrument vector $\left[x^0_{i,t}, x^h_{t+h}\right]$ 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 divisions we jointly estimate a stacked set of division-specific moment conditions and allow for dependence just as in our benchmark specification with $N = 9$, and use a Wald test for the pair of coefficients being zero. If this Wald test fails to reject at the 5% level, then the point (ψ_0, ω_0) is in our 90% WIR confidence set. We then repeat for a grid of values for (ψ_0, ω_0).

Figure 1 contains a pairwise scatter plot where each location corresponds to one point of a grid of (ψ_0, ω_0). 90% confidence sets are depicted for both standard, strong-IV asymptotics and WIR confidence sets for IV-based local-spillover decomposition with $N = 9$, corresponding to column (4) of 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. The figure indicates a substantial overlap between inference under standard, strong-IV approximations and WIR methods. Because of this overlap, in the body of the paper we report standard errors based on strong instrument asymptotics.

Figure 1: Comparison of weak-instrument-robust and standard-asymptotics 90 percent confidence regions, for IV-S specification



Note: WIR=weak-instrument robust. $N = 9$, $T = 43$. The union of the +'s corresponds to the standard, strong-instrument asymptotic 90 percent confidence set. The union of the hollow-circles corresponds to the WIR 90 percent confidence set.

B Appendix: Division Partition

Table 3: Classification of states into nine divisions

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

Notes: States are grouped in order to maintain similar shares of income and also maintain geographic proximity within groups.