FEDERAL RESERVE BANK of ST. LOUIS

The Local-Spillover Decomposition of an Aggregate Causal Effect

Timothy G. Conley, Bill Dupor, Mahdi Ebsim, Jingchao Li, and Peter B. McCrory

Abstract

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

JEL codes: E1

Federal Reserve Bank of St. Louis, *Review*, Second Quarter 2026, Vol. 108, No. 6, pp. 1-12

https://doi.org/10.20955/r.2026.06

1 INTRODUCTION

"No man is an island..."
—John Donne, 1624

Over the past few decades, there has been a movement in economics toward 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 it 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

Timothy G. Conley is a professor at Western University, Bill Dupor is a senior economic policy advisor at the Federal Reserve Bank of St. Louis, Mahdi Ebsim is an economics PhD student at New York University, Jingchao Li is an associate professor at East China University of Science and Technology, and Peter B. McCrory is head of economics at Anthropic. Conley thanks the Social Science and Humanities Research Council of Canada for support. The authors thank the editor and a referee for helpful comments.

Michael Owyang and Juan Sánchez are editors in chief of the *Review*. They are supported by Research Division economists and research fellows, who provide input and referee reports on the submitted articles.

© 2026, Federal Reserve Bank of St. Louis. The views expressed in this article are those of the author(s) and do not necessarily reflect the views of the Federal Reserve System, the Board of Governors, or the regional Federal Reserve Banks. Articles may be reprinted, reproduced, published, distributed, displayed, and transmitted in their entirety if copyright notice, author name(s), and full citation are included. Abstracts, synopses, and other derivative works may be made only with prior written permission of the Federal Reserve Bank of St. Louis.

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.

Using this panel-based estimate of the aggregate multiplier, we substantially improve precision relative to one based on aggregate data alone, without taking a stand on the form of the spillover. We develop a technique to estimate the local and spillover effects of region-level treatments. Our parameterization allows us to infer the effects of aggregate-level treatments from these region-level estimates. We illustrate this approach by studying how regional defense spending affects regional income, decomposing 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*). Moreover, 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.

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]

For many papers in the fiscal multiplier literature that rely upon panel data—and not just those jointly estimating local and spillover effects—the inclusion of time fixed effects is an essential building block of the analysis, particularly since in many instances both the treatment and outcome variables are correlated over time with other, unobserved aggregate shocks. For example, this be a prominent issue if applying the approach we develop in this paper to government spending as a whole, since both automatic and discretionary changes in aggregate government spending introduce a countercyclical component and thereby bias our estimates.

By contrast, we make progress relative to the literature by employing an identification strategy that does not require the inclusion of time fixed effects. In particular, our approach relies, in part, upon an identifying assumption that aggregate changes in government defense spending are uncorrelated over time with the region-specific business cycle. This identifying assumption is analogous to the assumption that military build up and draw downs are orthogonal to aggregate business cycle fluctuations but extended to a panel setting under the more conservative view that the defense policy in the U.S. is not determined by economic conditions in any particular region in the U.S.[2]

We estimate the local and spillover multipliers of government defense spending over a four-year horizon using efficient generalized method of moments (GMM). This allows us to parse the local and spillover channels of government spending shocks while 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 on 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 is 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 ordinary least squares (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 the initial one-year change in each region's share of defense 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] For example, see Suárez-Serrato and Wingender (2016) (spillovers among U.S. counties), and McCrory (2020) (trade spillovers between U.S. states).

[2] This latter point is similar to what Nakamura and Steinsson (2014) write: "Our identifying assumption is that the United States does not embark on military buildups—such as those associated with the Vietnam War and the Soviet invasion of Afghanistan—because states that receive a disproportionate amount of military spending are doing poorly relative to other states."

2 THE LOCAL-SPILLOVER DECOMPOSITION

We use annual data at the state level from 1964 through 2006.[3] 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 equal-sized divisions. Our basic units of analysis are $N = 9$ divisions.[4] 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 as g_t and q_t, respectively.

We focus on the effect of cumulative changes in local and spillover military spending on cumulative changes in real income. Each region's normalized cumulative change in real income 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 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 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 income, q_{t-1}. Scaling by lagged income in both the outcome and treatment ensures that in the regressions of $y^h_{i,t+h}$ on $x^h_{i,t+h}$ described below, the estimated coefficient can be interpreted as a multiplier—the dollar change in local income in response to \$1 of local or spillover defense spending.

We focus on a four-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; while at longer horizons, we are limited by the time length of our sample. Additionally, 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 $\tilde{x}^h_{i,t+h}$ to denote this leave-out mean, with

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

Equation (3) defines a division's spillover treatment as the change in average defense spending outside of that division. This leave-out 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, if defense spending in one region drives up inflation, the central bank might raise its target interest rate, which could influence economic activity similarly across all other regions. As another example, an increase in one region's federal defense spending—if tax-financed—would affect tax rates and potentially economic activity in similar ways 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) examine U.S. county-level employment and transfer spending data, constructing weights based on geographic closeness. McCrory (2020) analyzes cross-state spillovers from government spending associated with trade flows between U.S. states, while Dupor and McCrory (2018) study U.S. regional employment changes and stimulus spending, constructing weights from worker commuting flows. 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 article, 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.

[3] Our start year is based on the availability of state-level output, and the final year is chosen for comparability with existing studies.

[4] The precise division-state assignments are presented in Table Appendix B.1.

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:

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

where $W_{i,t-1}$ is a vector of conditioning information containing $y^0_{i,t-1}$, $x^0_{i,t-1}$, and x^0_{t-1}.[5] We define ψ_h as the local multiplier at horizon h for a region. The local multiplier gives the response of own-division income to a one-unit increase in own-division defense spending, 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 $W_{i,t-1}$).

We construct an aggregate econometric equation by summing the division-level equation across divisions. Taking sums of both sides over i using equation (4), we obtain

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

We define $\psi_h + \omega_h$ as the aggregate multiplier because it gives the response of aggregate income to a one-unit increase in aggregate defense spending. On the left-hand side is the aggregate outcome, and on the right-hand side is the aggregate treatment variable multiplied by a coefficient equal to the sum of the local and spillover multipliers. Because the treatment variable on the right-hand side is expressed as a ratio, aggregation by summation is valid 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.

In addition to providing a parsimonious form of spillover, the leave-out mean definition offers two additional advantages. First, to construct the local-spillover decomposition, we require the division-level estimation equation to aggregate to the national level, and the leave-out mean provides a simple form that ensures this aggregation. Second, because we have many divisions, the leave-out mean is highly correlated across them. As a result, identification of the spillover effect in our analysis comes almost entirely from time-series variation. This stands in stark contrast to studies 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), which rely on cross-sectional variation in the spillover treatment.

As implied by equation (3), we define the spillover as the change in income in a division due to total defense spending outside of the division, holding own-division defense spending fixed. In contrast, existing related papers define the spillover more narrowly: as a weighted sum of the treatment (defense spending in our case) in other regions. 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. Dupor and McCrory (2018) study U.S. regional employment changes and stimulus spending and use worker commuting flows to build weights.

For each of the above cited papers, if the specified weights accurately define the spillover, the model is correctly identified and the issues we discuss in this paper are less relevant. If, however, the specified-weights are incorrect or there are one or more additional spillover channels, then a bias is likely to arise. One particular interesting case is a "general spillover," related more closely to common monetary and fiscal policy, rather than trade linkages or geographic closeness. Our method, which relies on time series variation for identification, allows for general spillovers, in contrast to the above papers, where identification works through cross-sectional variation.[6]

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 $\psi_h + \omega_h$. Using this panel-based estimate of the aggregate multiplier, we substantially improve precision relative to using one based on aggregate data alone, without taking a stand on the form of the spillover.

[5] We omit an intercept in equation (4) because all variables are de-meaned.

[6] Each spillover paper in the entire existing literature, at least of which we are aware, uses only cross-sectional variation because it uses a cross-section or else a panel with time fixed effects.

We focus on cumulative multipliers that 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 the spending costs over a given horizon. Ramey and Zubairy (2018) compellingly explain that cumulative multipliers are more valuable from a policy perspective than other (sometimes reported) statistics, such as impact multipliers and peak multipliers.

We estimate these equations via 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 stack these moment conditions for the nine divisions and estimate the model via GMM.

Our benchmark interval estimates use conventional, strong-instrument approximations; that is, 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 on $W_{i,t-1}$—are orthogonal to past and future fundamental shocks to the economy, which Stock and Watson (2018) call lead-lag exogeneity. To assess whether conditioning on $W_{i,t-1}$ is adequate for this assumption 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 article, we find this to be the case for all instruments defined below—evidence consistent with their being lead-lag exogenous, given $W_{i,t-1}$.

We consider two alternative instrument sets $z^h_{i,t+h}$ to estimate equation (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. The exogeneity of x^h_{t+h} (national military spending) relies on the common argument that aggregate defense spending is determined by geopolitical factors and hence orthogonal to regional business cycle shocks. We anticipate that it will be useful in identifying our spillover coefficient, as it should be strongly correlated with $\tilde{x}^h_{i,t+h}$ because of the substantial overlap in their components. Next, $x^h_{i,t+h}$ serves as our division-level instrument. Thus, the first set of instruments relies on exogeneity in defense spending at both the division and aggregate levels. 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. This instrumental variable (IV) specification requires 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 differences in 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, this specification replaces $x^h_{i,t+h}$ with $x^0_{i,t}$, the one-year-scaled change in defense spending in division i in the impact year t. The validity of $x^0_{i,t}$ as an instrument reflects a realistic timing restriction.[7] While it might be possible for the federal government to reallocate resources across geographic areas over several years, it is implausible that at a one-year horizon it could undertake any significant repositioning of contracts, troops, or newly purchased equipment and structures in response to political motivations working through local business cycles.

We consider two alternative instrument sets $z^h_{i,t+h}$. The first is analogous to OLS and the second employs external instruments. In both cases, the first elements are our conditioning information $W_{i,t-1}$, lagged, scaled one-year changes in division-level income and defense spending as well as the one-year change in national defense spending. In addition, both sets contain the scaled cumulative change in national military spending, x^h_{t+h}. As discussed above, we rely on the common argument that aggregate defense spending is exogenous 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.

Our alternatives $z^h_{i,t+h}$ assumption differ only in their last elements: an instrument with division-level local variation, which we anticipate will be useful for identifying our local coefficient. In the first case, we use $x^h_{i,t+h}$ as 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, 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. We refer to this, and other cases below, where there are no excluded instruments using the abbreviation NE.

[7] For example, Blanchard and Perotti (2002) identify general government spending using a timing restriction.

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. Our aggregation to the division level partly mitigates this concern, but we are still motivated to utilize an alternative instrument strategy.

Our second instrument set (IV-S) replaces $x^h_{i,t+h}$ with $x^0_{i,t}$, which is simply the impact (i.e., year t) one-year scaled change in defense spending in division i. This amounts to a timing restriction. While it might be possible for the federal government to reallocate resources across geographic areas over the period of a few years, at a one-year horizon is implausible that the federal government could undertake any significant repositioning of resources. The remaining instruments are unchanged.

We obtain efficient GMM estimates via iterated GMM.[8] We estimate the long-run variance-covariance matrix of our moment conditions via a Bartlett/Newey-West covariance matrix estimator that places nonzero weight on the sample autocovariances up to four lags.[9] To reduce dimension, we also impose an assumption of second-moment independence between z and u. Sample moments of z are multiplied by sample moments of the estimated u and then combined via Bartlett weights to estimate the long-run variance-covariance matrix of the moment conditions.

Our data sources are as follows: State-level income data are from the Bureau of Economic Analysis. Data on defense wages and contracts apportioned to state geographies come from Dupor and Guerrero (2017). This variable, summed across regions, is less than defense spending measured in the National Income and Product Accounts (NIPA) 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.

3 RESULTS

Table 1 reports our benchmark results. Column (1) reports estimates of equation (4), where we use own-division spending to instrument for itself (i.e., no excluded instruments) and restrict the coefficient on other-division spending (i.e., the spillover multiplier) to equal zero. Both the treatment (defense spending) and the outcome (income) are measured over a cumulative four-year horizon, as explained previously. The coefficient equals 0.23 (SE = 0.13), indicating that a \$1 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-division 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 the 1 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 article's key deliverable: a decomposition of the aggregate effect of defense spending into a local and spillover coefficient. First, column (3) uses own-division spending and average other-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), and the spillover multiplier equals 0.78 (SE = 0.44). According to the point estimate, a \$1 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

[8] 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 the third-step GMM point estimate, which is then used to re-estimate the weighting matrix to obtain another point estimate, etc. This iteration continues until the Euclidean distance between the current and previous value of the estimated parameter vector is less than 10^{-3}. All estimates converge in less than 30 iterations.

[9] Our results are robust to using alternate Bartlett weights, including up to ten 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 spending	0.23* (0.13)	0.26*** (0.08)	0.30*** (0.07)	0.31*** (0.07)	-	-
Other-division spending	-	-	0.78* (0.44)	0.78* (0.46)	-	-
Aggregate multiplier	-	-	1.08** (0.47)	1.09** (0.48)	0.71 (0.67)	1.39*** (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

NOTE: For columns (1) through (4) and (6), moment conditions are evaluated at the division level, estimation is done via iterated GMM, and the dependent variable is the four-year cumulative change in division income, scaled by lagged national income. For column (5), the dependent variable is the four-year cumulative growth in national income. P-value is from the overidentification test. NE denotes no excluded instruments. See the text for the description of IV-Agg and IV-S. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

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 spending	0.44 (0.30)	0.33 (0.23)	0.23*** (0.09)	0.10 (0.40)	-
Other-division spending	-	-	0.54 (0.63)	0.61 (0.91)	-
Aggregate multiplier	-	-	0.77 (0.68)	0.71 (0.61)	0.71 (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

NOTE: For columns (1) through (4), the dependent variable is the four-year cumulative change in division income, scaled by lagged national income. Estimation is done via pooled least squares or pooled IV. Driscoll-Kraay standard errors are reported. For column (5), the dependent variable is the four-year cumulative growth in national income. NE denotes no excluded instruments. See the text for the description of IV-Agg and IV-S. * $p < 0.1$, ** $p < 0.05$, *** $p < 0.01$

spending increases. It is statistically different from zero at the 5 percent level; however, the associated confidence interval includes values greater than and less than one.

Column (4) modifies column (3)'s estimate 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 divisions based on differences in the divisional business cycles. Both the local and spillover multipliers increase somewhat, and 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; that is, we estimate the aggregate multiplier with the traditional macro approach. Thus, we directly estimate equation (5) using aggregate data, and this estimate is presented in column (5). The first two rows of this column are empty because the local and spillover multipliers are not separately identified, as can be seen by inspecting equation (5). 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), which 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), indicating 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 five of the six specifications above, except that we use standard econometric routines that rely on exactly identified, rather than overidentified, models. Each column of the table is comparable to the correspondingly numbered column of Table 1.[10] 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 Table 1. The local and spillover multipliers are positive and less than one, and the aggregate multiplier is positive and close to one but somewhat smaller than the corresponding values 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 a positive but less than one-for-one response of division income to own-division defense spending. This estimate is somewhat larger than the analogous one 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 (referred to as IV-Agg in the table). The multiplier is again less than one; however, it is very imprecisely estimated and not statistically different from zero. Importantly, the estimate of the multiplier is substantially larger than the corresponding one from Table 1.

Columns (3) and (4) convey 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 the aggregates.

Finally, column (5) estimates the aggregate multiplier using aggregated data. The point estimate equals 0.71, 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), reflecting that 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, this article makes and illustrates two methodological advances: a formal decomposition of the aggregate multiplier into local and spillover components, along with an estimation approach that achieves efficiency gains.

With respect to the aggregate multipliers, DG and this article are relatively similar. Both examine four-year cumulative multipliers using annual data. However, they differ in several ways. The former uses national income data, a sample starting in 1951, and a somewhat narrower definition of defense spending, whereas the latter uses

[10]The only exception is that Table 2 has no column (6), as there is no estimate for that case.

GDP, a sample starting in 1964, and a broader defense spending measure. In this article, 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 state-level estimates, the results in DG and in this article are not comparable because of the methodological advance presented here. Whereas this article 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 attempt to estimate a spillover coefficient that is broadly similar to the goal of this article; however, rather than using a leave-out 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, they may miss others (e.g., common fiscal and monetary policy and factor market linkages).

4 CONCLUSION

When researchers employ cross-sectional or panel data to estimate the causal effect of a treatment, they typically identify only a local effect. In the presence of spillovers between observational units, this local effect generally differs from the aggregate (or average) effect. However, in many contexts—particularly in macroeconomics—policymakers are primarily interested in an aggregate effect. In this article, 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 straightforward conversion to the treatment's aggregate effect.

To apply our methodology in a straightforward manner, a suitable application should have two crucial ingredients, beyond appropriate instruments. First, sufficiently long disaggregate (individual/group) time-series data are required, as our approach relies on intertemporal moments that allow for potentially strong cross-section dependence. Second, there must be some degree of similarity in treatment effects across individuals or groups 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 both local and spillover effects.

REFERENCES

Auerbach, A. and Y. Gorodnichenko, "Output Spillover from Fiscal Policy," *American Economic Review*, 2013, *103*, 141–146.

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

Chodorow-Reich, G., "Regional Data in Macreconomics: Some Advice for Practioners," *Journal of Economic Dynamics and Control*, 2020, *115*, 103875.

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

_ and R. Guerrero, "Local and Aggregate Fiscal Policy Multipliers," *Journal of Monetary Economics*, 2017, *12*, 16–30.

McCrory, P., "Tradable Spillovers of Fiscal Policy: Evidence from the 2009 Recovery Act," *Working Paper*, 2020.

Nakamura, E. and J. Steinsson, "By How Much Does GDP Rise if the Government Buys More Output?," *American Economic Review*, 2014, *104* (3), 753–92.

Ramey, V. and S. Zubairy, "Government Spending Multipliers in Good Times and in Bad: Evidence from U.S. Historical Data," *Journal of Political Economy*, 2018, *126* (2), 850–901.

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

Stock, J. and M. Watson, "Identification and Estimation of Dynamic Causal Effects in Macroeconomics Using External Instruments," *Economic Journal*, 2018, *128* (610), 540–552.

Suárez-Serrato, J. C. and P. Wingender, "Estimating Local Fiscal Multipliers," working paper 22425, National Bureau of Economic Reserarch 2016.

APPENDIX A WEAK-INSTRUMENT-ROBUST INFERENCE

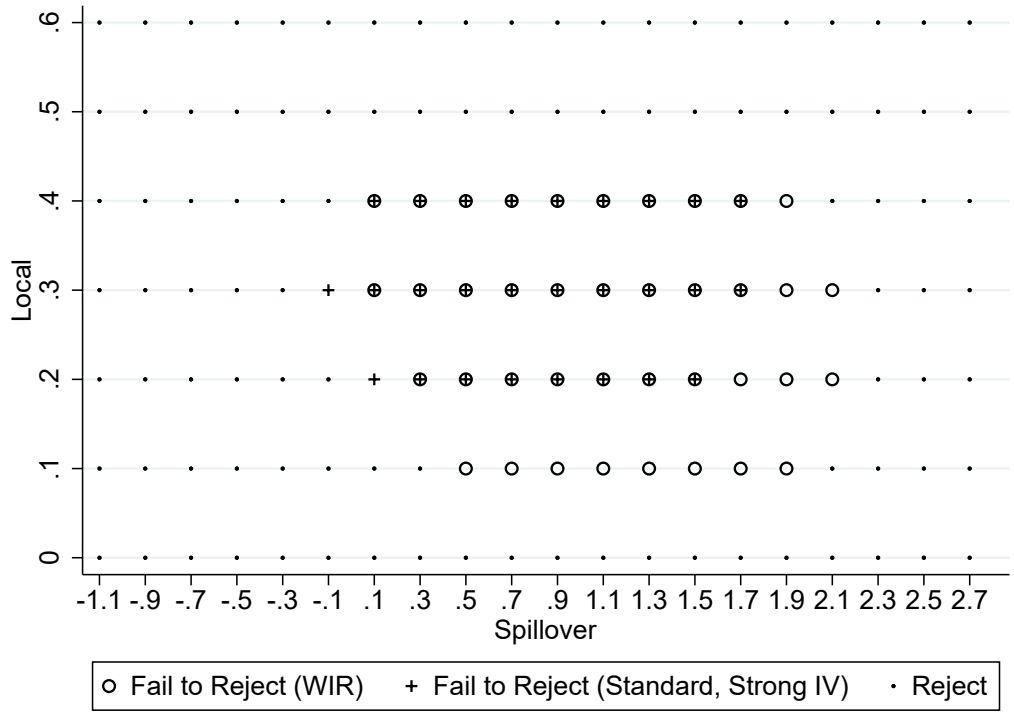
In this appendix, we present weak-instrument-robust (WIR) confidence sets for our IV-S specification following Chernozhukov and Hansen (2008; henceforth CH). We first project all variables on the first three elements of $W_{i,t-1}$ (i.e., the included instruments) and work with residuals. Rather than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

$$u^h_{i,t+h} = y^h_{i,t+h} - \psi_0 x^h_{i,t+h} - \omega_0 \tilde{x}^h_{i,t+h}. \tag{Appendix A.1}$$

Then this null hypothesis can be tested via a regression of $u^h_{i,t+h}$ on the instrument vector $\left[x^0_{i,t}, x^h_{t+h}\right]$ with the exclusion restriction implying that they should be orthogonal. The coefficients in a regression of $u^h_{i,t+h}$ on $z^h_{i,t+h}$ 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 percent level, then the point (ψ_0, ω_0) is in our 90 percent WIR confidence set. We then repeat this for a grid of (ψ_0, ω_0) values.

Figure Appendix A.1 presents a pairwise scatter plot where each location corresponds to one point on a grid of (ψ_0, ω_0). Ninety percent 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. Pluses indicate the confidence set under strong-instrument approximations, and open circles represent the WIR confidence set. The figure indicates substantial overlap between inference under standard strong-IV approximations and WIR methods. Because of this overlap, in the body of the article we report standard errors based on strong-instrument asymptotics.

Figure Appendix A.1
Comparison of Weak-Instrument-Robust and Standard Asymptotic 90 Percent Confidence Regions, for the IV-S Specification



NOTE: WIR; weak-instrument-robust. $N = 9$, and $T = 43$. The union of the plus signs 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.

APPENDIX B DIVISION PARTITION

Table Appendix B.1
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

NOTE: States are grouped to maintain similar shares of income and geographic proximity within groups.