The *Sine Aggregatio* Approach to Applied Macro[*]

Timothy G. Conley[†], Bill Dupor[‡] and Mahdi Ebsim[§]

November 11, 2022

We develop a method to use disaggregate data to conduct causal inference in macroeconomics. The approach permits one to infer the aggregate effect of a macro treatment using regional outcome data and a valid instrument. We estimate a macro effect without (*sine*) the aggregation (*aggregatio*) of the outcome variable. We exploit cross-equation parameter restrictions to increase precision relative to traditional, aggregate series estimates and provide a method to assess robustness to departures from these restrictions. We illustrate our method via estimating the jobs effect of oil price changes using regional manufacturing employment data and an aggregate oil supply shock.

1 Introduction

This paper presents a new method for using disaggregate data to conduct causal inference about a macroeconomic treatment effect. With a properly constructed statistical model, we can conduct estimation of an aggregate effect without aggregating the outcome variable. By exploiting the information in the disaggregate panel, our

[*]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. Preliminary and not for citation.

[†]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.

method can dramatically improve the precision of estimates relative to using solely aggregate data.

As a starting point, consider the problem of estimating the effect of an aggregate variable, $\mathcal{X}_t$, on an aggregate outcome, $\mathcal{Y}_t$, via an instrumental variable (IV) strategy. The outcome depends on the treatment according to a linear relationship given by

$$\mathcal{Y}_t = B\mathcal{X}_t + \nu_t \tag{1}$$

where t represents time and ν_t is an error term.[1] Identification of B relies on an instrument, $\mathcal{Z}_t$, that a researcher can defensibly assume is orthogonal to ν_t.

Suppose $\mathcal{Y}_t$ is the sum of some regional, sectoral or other disaggregate series.[2] That is, we have series for N mutually exclusive, disaggregate groups $\mathcal{Y}_{i,t}$, and $\mathcal{Y}_t = \sum_{i=1}^{N} \mathcal{Y}_{i,t}$. This allows the possibility that disaggregate data may be informative about the aggregate effect and motivates our method aimed at utilizing this information.

Group-level data are very likely to be helpful in learning about B when the group-level effect of $\mathcal{X}$ is the same across all groups. In this case, a group-level analog of (1) with common parameters is well-specified:

$$\mathcal{Y}_{i,t} = \frac{B}{N}\mathcal{X}_t + u_{i,t}. \tag{2}$$

Summing (2) across i for each t reproduces the aggregate regression (1). There could be large efficiency gains in estimating (2) relative to (1), with the caveat that its common parameter assumption may be too restrictive. There is a fundamental tension between a potential need for group-level treatment heterogeneity and a potentially large efficiency gain from imposing (valid) common parameter restrictions across groups. Our method is motivated by the concern that in many applications a common parameter assumption may be too restrictive, so a method to systematically

[1] Here assume we have concentrated out any conditioning information such as lags of the outcome or other variables.

[2] Appendix A provides 11 examples of publicly available U.S. data sets that satisfy this requirement and have been used (extensively in most cases) in macroeconomic analyses.

examine departures from it is useful.

Consider a group-level analog of (1) with parameter heterogeneity:

$$\mathcal{Y}_{i,t} = \beta_i \mathcal{X}_t + \nu_{i,t}. \tag{3}$$

We consider an aggregate effect of interest that is a weighted sum of the β_i. With a given set of weights, $\{s_i\}$, our aggregate effect is $B = \sum_i s_i \beta_i$. There may be large efficiency gains from exploiting restrictions on sets of common parameters in the β_i.

We estimate B, the macro effect of an aggregate treatment, using the group-level responses to the treatment via GMM with group-level moment conditions and cross-group parameter restrictions. Furthermore, we provide a simple way to examine robustness to the common-parameter restrictions underlying our potential efficiency gains.

Our approach complements a very large body of existing, related work with disaggregate data which is of two types. First, much of this literature estimates relative or local effects, rather than macro effects, by regressing regional outcomes on regional treatments (e.g., Chodorow-Reich et al. (2012), Clemens and Miran (2012) and Mian, et.al. (2013)). In addition, a similarly large body of work estimates regional impacts of aggregate treatments but neither imposes group-level moment conditions nor relevant parameter restrictions (e.g Carlino and DeFina (1998), Mumtaz, et.al. (2018) and Owyang and Zubairy (2013)).

For simplicity of exposition, we assume that the same aggregate instrument $\mathcal{Z}_t$ is valid and strong at the group level so that under standard regularity conditions GMM estimation using a stacked set of the group-level moments $E(\mathcal{Z}_t \nu_{i,t}) = 0$ will generate reliable confidence intervals. The approach is straightforward to apply with group-specific instruments, weak instrument robust estimators, nonlinear models or other criterion functions.

Our method lets the researcher decide how much parameter flexibility to allow across groups. A wide variety of parameter restrictions can be examined; the number of alternative restrictions is limited only by computational feasibility. There are many reasons researchers may want to examine the possibility of cross-group heterogeneity in treatment effects. Groups of states for example may have different industrial composition or labor market conditions that underlie treatment effect heterogeneity.

To illustrate our method, we examine restrictions that allow a small number of groups, K, to have differing group-level treatment slopes, while the rest share a common value. That is, we restrict at most K groups to have individual β_i. Thus the most extreme restriction in this illustration, with $K = 0$, is that all groups' β_i are the same. As K increases the sets of restrictions become more flexible. In Appendix B, we estimate the model using a different set of restrictions: common coefficients within clusters of state groups.

We begin with estimating the model under each possible restriction. In our example, each restriction has K groups with individual β_i and the remainder with a common value. For each of these N choose K parameterizations, we obtain a confidence interval (CI) for B. We take the union of these CIs to obtain an interval estimate for B that is conservative. This basic approach is closely related to that followed by Conley, Hansen and Rossi (2012) in the context of potential deviations from a IV exclusion restriction and Hansen, Kozbur and Misra (2021) in a high-dimensional model selection context.

We apply our method to estimate the response of manufacturing employment to an exogenous oil price shock. We study an oil price shock because it is viewed as a highly plausible "supply" factor that varies in part for reasons that are exogenous to the business cycle. Our IV approach utilizes the Känzig (2021) monthly oil supply news series as an instrument for world oil price changes. Känzig (2021) uses short-run changes in oil futures prices around a tight window of OPEC announcements to construct an instrument that is correlated with exogenous changes in oil prices. We study manufacturing employment because U.S. manufacturing is highly oil-dependent and the sector's monthly employment data are available at the regional level. In the remainder of the paper, all references tp employment should be understood to mean manufacturing employment. We divide U.S. states into nine disjoint groups, each with roughly the same group employment totals.[3]

[3]The number of groups is limited by the panel's time dimension and our need for a precise estimate of the moment variance matrix.

Using the standard aggregative approach yields point estimates that suggest a contractionary oil supply shock reduces employment but the effect is imprecisely estimated. We estimate the aggregate effect using our panel-based approach for $K = 0$ through $K = 3$. For each of these values, the midpoint of the CI corresponds to a twenty percent increase in the real price of oil reducing employment by about one percent. For each value of K, the corresponding CI is substantially smaller than that from the aggregative approach. For example, the $K = 1$ CI is $(-2.5, 0.40)$ and aggregate-based analogue is $(-4.8, 0.26)$. This implies a CI shrinkage of 42 percent.

2 Related Research

Our paper complements the existing literature on using disaggregate data to answer macroeconomic questions. A first generation used disaggregate (e.g., regional or sectoral) differences in exposure to a macro shock and the resulting differences in outcomes to identify causal effects.[4]

However, several authors have noted that, with cross-regional spillovers, these local (or relative) effect estimates are potentially biased estimates of the aggregate effects of a policy (e.g., Cochrane (2012), Nakamura and Steinsson (2014) and Ramey (2011)). One advance (which speaks to this issue) has been to relate a particular local effect empirical estimate to an analogous partial equilibrium object from a structural economic model (e.g., Kaplan and Violante (2014) and Berger, et.al. (2017)). Relatedly, several papers have used those local estimates to help calibrate parameters from a fully-specified equilibrium macro models (e.g., Beraja, Hurst and Ospina (2019), Dupor, et.al. (2020) and Nakamura and Steinsson (2014)).[5]

In contrast, we use disaggregate data to answer macro questions without taking a stand on the specific preferences, technology, endowment and market mechanisms that underlie the data generating process used to approximate an actual economy.

[4]For example, Chodorow-Reich (2020) cites 50 papers published between 2012 and 2018 in top economics journals that attempt to infer causal macro impact using cross-regional variation in exogenous shocks or policy changes and regional outcomes. For these papers, the statistical unit of observation is a region.

[5]See also Guren, et.al. (2019) and Wolf (2019) for related approaches.

Even though we use disaggregate outcome data, our basic identification is based on time series variation.

3 The Econometric Model

We present our method in a case matching our application where the researcher has balanced panel data on a group-level outcome $Y_{i,t}$, with T periods and N groups.[6] In addition, the researcher is interested a treatment effect upon the long difference of $Y_{i,t}$ denoted as: $y^{\delta}_{i,t+\delta} = Y_{i,t+\delta} - Y_{i,t-1}$. The analogous aggregate treatment is denoted $x^{\delta}_{t+\delta}$ and z_t denotes the aggregate instrument. We consider a (slightly) generalized version of an aggregate long difference by taking a weighted sum, using the notation $y^{\delta}_{t+\delta} \equiv \sum_i s_i y^{\delta}_{i,t+\delta}$ where the s_i are weights.

We work with a local projection estimation equation:

$$y^{\delta}_{i,t+\delta} = \gamma_i + \beta_i x^{\delta}_{t+\delta} + \epsilon^{\delta}_{i,t+\delta} \tag{4}$$

where γ_i is a group fixed effect. The parameter of interest is the effect of $x^{\delta}_{t+\delta}$ on the aggregate outcome $y^{\delta}_{t+\delta}$, denoted by B, with $B \equiv \sum_j s_j \beta_j$.

The paramter B could be estimated via an aggregate regression. Taking an s_i weighted sum of (4) in the cross section yields an aggregate regression equation:

$$y^{\delta}_{t+\delta} = \gamma + B x^{\delta}_{t+\delta} + \epsilon^{\delta}_{t+\delta} \tag{5}$$

The typical approach is to estimate equation (5) via instrumental variables. We aim to improve efficiency by using our panel data to estimate (4).

Our IV moment conditions are

$$E\left(z_t \epsilon^{\delta}_{i,t+\delta}\right) = 0 \tag{6}$$

for $i = 1, \ldots, N$. Our identification assumption is that, at the group-level, the error term of the group-level equation is uncorrelated with the aggregate instrument z_t.

[6] The method is straightfoward to apply to unbalanced panels.

We first estimate a set of models, with each imposing a restriction upon the β_i and estimating parameters via GMM using moment condition (6). We consider R total restrictions and index each restriction using r. We also refer to the N by one parameter vector collecting all β_i, subject to r as β_r with its $\{s_i\}$ weighted sum denoted B_r. For each r we obtain a CI for B_r, CI(B_r), implied by the usual, strong-instrument large-sample distribution approximation for the GMM estimator of B_r. We want our method to be flexible enough to handle the possibility that a restriction r may be at odds with the data and result in nonsense estimates and CI. To immunize our procedure from such implausible restrictions, we construct a 'mixture' CI that uses a model specification test for each r. We use a standard GMM over-identification test statistic under the restriction r, which we denote J_r. If a model is rejected by the over-identification test, we replace its standard CI with an empty set. Our mixture CI will be denoted $\widetilde{CI}(B_r)$. For example with a 1% critical value for the over-identification test, to get a 90% $\widetilde{CI}(B_r)$ we use:

$$90\% \ \widetilde{CI}(B_r) = \begin{cases} 91\% \text{ CI}(\text{B}_r) \text{ if } \text{J}_r < 1\% \text{ critical value} \\ \emptyset \text{ else} \end{cases} \tag{7}$$

Thus, we replace the estimated standard 91% CI(B_r) with an empty set if the model is rejected by the 1% level specification test. Under the null of a correctly specified model, our specification test rejects 1% of the time and our resulting 90% $\widetilde{CI}(B_r)$ will of course not cover the true B because it is empty. The worst case scenario from a coverage point of view is that every time the specification test rejects, the 91% CI(B_r) would have covered B. Therefore, our mixture 90% $\widetilde{CI}(B_r)$ will still have at least 90% coverage.

Finally, to obtain a conservative CI for B we take the union of the $\widetilde{CI}(B_r)$ over all $r \in R$. Our method can be implemented with any set of restrictions R that is small enough to be computationally feasible. We anticipate researchers will want to examine results for more than one set of restrictions, e.g. for a set of values of K.

Appendix C restates the complete method in step-by-step "cookbook" form. In the next section, we execute the procedure using data on oil supply shocks and employment.

4 Application: Oil Price News Shocks

We utilize monthly US state-level data from January 1991 to January 2017.[7] For dimension reduction, we define groups to be collections of states, with each state belonging to one group. We aggregate employment within group and $Y_{i,t}$ is taken to be log employment in group-month (i, t), so $y^{\delta}_{i,t+\delta}$ is the long difference in the log of group-level employment. Our treatment variable $x^{\delta}_{t+\delta}$ is the analogous long difference in the log real price of oil.[8] We scale $x^{\delta}_{t+\delta}$ such that a unit treatment corresponds to a 20 percent increase in the real price of oil between $t-1$ and $t+\delta$. Because of the delayed effect of oil shocks on economic activity found in existing research, we investigate 18 and 24 month horizons for our local projection estimation (i.e., $\delta = 17, 23$).

We use nine groups, chosen so that the groups have similar shares of employment, given in Table 1. In all but one case, the groups are geographically contiguous.[9] We collapse the state-level data into groups because we want to ensure we have reliable covariance matrix estimates, which will be used to construct the efficient GMM weighting matrix as well to construct confidence intervals.

We use the Känzig (2021) oil shock series as our instrument. Känzig (2021) uses variation in oil futures prices around a tight window of OPEC announcements to construct an instrument. Although oil prices and oil price futures are in general endogenous to the world macroeconomy, world economy factors should already be priced into oil futures and are plausibly unchanged within the announcement window. Känzig (2021) estimates that a contractionary shock immediately increases the oil spot price, reduces oil production gradually and has a delayed, negative effect on industrial production.

Our instrument series, z_t, is constructed as the set of positive (i.e., contractionary) values of the Känzig (2021) oil supply news shock. All non-negative values are set

[7] Our state-level employment data are from the Bureau of Labor Statistics.

[8] Oil price data are from the U.S. Energy Information Administration. The oil price is transformed from a nominal to real variable using the CPI.

[9] New York is a single group and splits a set of northeastern states that together make up another group, which include Vermont and Pennsylvania.

Table 1: Classification of states into nine groups

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 manufacturing employment and also maintain geographic proximity within groups.

equal to zero in our application.[10]

We obtain efficient estimates via iterated GMM assuming second-moment independence.[11] We estimate the long-run covariance matrix of our moment conditions via a Bartlett covariance matrix estimator that places non-zero weight on the sample autocovariances at up to 20 month (for $\delta = 17$) and 26 month (for $\delta = 23$) leads and lags.[12]

The parameter of interest, B, is the percentage change in national employment in response to a twenty percent oil price increase. We estimate this parameter over both the 18 month and 24 month horizon.

Our identification restriction—that the group-level error term is uncorrelated with the oil supply shock—is somewhat stronger than the typical restriction that the

[10]We limit attention to contractionary supply shocks. Whereas contractionary supply shocks have strong predictive power for oil price changes (at the horizons we consider and over our sample), expansionary ones do not.

[11]We iterate on the GMM procedure until the absolute value of the change in J statistic across one iteration is less than 10^{-3}. Convergence is typically achieved after three iterations. Relative to two-step GMM, using iterated GMM lessens the arbitrary dependence on the initial weighting matrix.

[12]Our results are robust to using alternate Bartlett weights from 21 to 27 months (for $\delta = 17$) and 27 to 33 months (for $\delta = 23$) leads and lags.

Table 2: Aggregate and *sine aggregatio* confidence intervals for the employment effect of a 20 percent increase in the real price of oil, 24 month horizon

Estimation Method	Low CI	High CI	Midpt.	CI rel length (%)	Num Models
Aggregated data	-3.85	0.32	-1.77		
Sine aggregatio					
K = 0	-1.71	0.13	-0.79	44	1
K = 1	-2.13	0.50	-0.81	63	9
K = 2	-2.68	0.96	-0.86	87	36
K = 3	-2.74	1.16	-0.79	93	84

Note: Dependent variable=24 month change in log employment. 90 percent CIs reported. K=number of response coefficients allowed to vary across regions. For the $K = 0$ case, the J-statistic and associated p-value for the over-identifying restriction are 7.5 and 0.50.

Table 3: Aggregate and *sine aggregatio* confidence intervals for the employment effect of a 20 percent increase in the real price of oil, 18 month horizon

Estimation Method	Low CI	High CI	Midpt.	CI rel length (%)	Num Models
Aggregated data	-4.79	0.26	-2.26		
Sine aggregatio					
K = 0	-2.02	0.01	-1.01	40	1
K = 1	-2.53	0.40	-1.07	58	9
K = 2	-2.70	0.76	-0.97	69	36
K = 3	-3.10	0.96	-1.07	81	84

Note: Dependent variable=18 month change in log employment. 90 percent CIs reported. K=number of response coefficients allowed to vary across regions. For the $K = 0$ case, the J-statistic and associated p-value for the over-identifying restriction are 8.0 and 0.43.

aggregate error term is orthogonal to the supply shock. Simply put, our identification assumption requires that the Känzig national-level identification carries over to the group level.

Table 2 reports the effect on employment of a 20 percent increase in the real price of oil at a 24 month horizon. It contains six columns: the estimation method (and value of K), the two CI bounds, the midpoint of those bounds, the relative length of CI and the number of models estimated.

The first row of estimates corresponds to the standard macro approach: a univariate estimate using the fully aggregated data (i.e., equation (5)). The 90 percent

CI equals $(-3.9, 0.32)$ with a midpoint equal to -1.8. It indicates that the oil price increase–driven by a contractionary news shock–has a lower bound of decreasing employment by 3.9 percent to an upper bound estimate of increasing employment by 0.3 percent. Notably, the effect is imprecisely estimated.[13]

The remaining rows contain the estimates using the panel outcome variable and the *sine aggregatio* approach. Each row corresponds to a case of allowing K slopes to differ, with the remaining slopes restricted to be identical. Note that the confidence intervals often shrink dramatically relative to the estimate based on the standard approach using aggregate data. Comparing results across values of K allows the researcher to assess robustness to the associated parameter restrictions.

The midpoint of the CI (which in this case is also the point estimate) when $K = 0$ is -0.79, which is of the same sign but somewhat closer to zero than the aggregated-data based estimate of -1.8. The entry for the "CI rel length (%)" column indicates that the $K = 0$ *sine aggregatio* CI length is 44 percent of the aggregate estimate CI length.[14]

The next row reports the $K = 1$ estimate. Here the reported CI is union of nine models;[15] each model corresponds to one of the nine groups being allowed to have a distinct slope. The midpoint of this estimate equals -0.81, which is slightly lower than the $K = 0$ case. As one might anticipate, the $K = 1$ CI expands slightly relative to the common coefficients case and equals $(-2.1, 0.50)$.

The final two rows report estimates for K equal 2 and 3 cases. For $K = 2$, we

[13]This estimate is nearly identical to that if we instead using a long difference in log aggregate employment, denoted $Y^A_{t+\delta} - Y^A_{t-1}$, as our outcome and estimated:

$$Y^A_{t+\delta} - Y^A_{t-1} = \gamma^A + B^A x^\delta_{t+\delta} + \eta^\delta_{t+\delta} \tag{8}$$

This follows from the high quality of the first-order approximation for national employment growth rate based on group-level employment growth rates, the correlation between $y^\delta_{t+\delta}$ and $(Y^A_{t+\delta} - Y^A_{t-1})$ us greater than .999.

[14]We also report the J test of overidentification restrictions, which embeds the coefficient restrictions on β_i. The J-statistic equals 7.5 and has eight degrees of freedom. The associated p-value equals 0.50. Thus, the data do not reject the model at conventional significant levels.

[15]As explained above, the method also requires that we "throw out" any model estimate and associated CI with a sufficiently high J-statistics; however, none of the models are rejected by the overidentifying restrictions in any specification in Table 2 and 3.

estimate the model for all sets of restrictions where two groups have group-specific slopes and the remaining seven have identical slopes. Likewise for $K = 3$, all sets of three groups are allowed group-specific slopes. The CIs are of course wider than in the $K = 1$ case and widen as K grows. However, both $K = 2$ and $K = 3$ CIs offer less improvement relative to the aggregate series CI, equaling 87% and 93% of its length, respectively. Comparing results across values of K allows the researcher to assess robustness to the associated parameter restrictions.

Table 3 estimates the analog of Table 2 except we use the 18 month horizon. The pattern for the CIs is very similar. The employment effect using the standard aggregate approach is estimated imprecisely. Using the *sine aggregatio* approach dramatically improves the precision of the employment effect estimate. The $K = 0$ CI length is 40 percent of the analogue aggregate-based CI length. The corresponding value for $K = 1$ equals 58 percent.

Tables 2 and 3 illustrate both the potential precision gains from using disaggregate with parameter restrictions and a way to examine robustness of results to modest relaxations of such restrictions. Relative to aggregate series estimates, CI length is drastically lower with a $K = 0$ fully common parameter restriction. Importantly, these tables also illustrate that much of these gains in precision can still occur for at least some levels of relaxation in the restrictions, for $K = 1$ or more. We anticipate that comparing results for a variety of K values will allow researchers to easily assess the robustness of their results to parameter restrictions.

5 Conclusion

We develop a new method to use disaggregate data to answer questions about how a macroeconomy responds to macro treatments.

In this section, we review some of the requirements necessary to usefully apply the method and describe when the method is likely to achieve significant efficiency gains.

First, a researcher needs time varying instrument(s). It also requires adequately long disaggregate (e.g., individual, region, sector) time series for a sufficiently large number of disaggregate groups that together are representative of all or nearly all of an entire macroeconomy. Our approach relies on moment conditions that are satisfied along the time dimension. Finally, to achieve efficiency gains, some commonality of the parametric response across groups is needed, although our approach allows for some departure from identical regional responses.

The potential efficiency gains from using our method dis-aggregated data relative to aggregates will depend in large part on two things. First, the extent of common parameters across groups, the more commonality the greater the scope for efficiency improvement as more groups' series will be informative about common parameters. Second, the potential for efficiency gains will depend on the variance-covariance structure across the elements of the vector in our moment condition (6), the collection of $(z_t\epsilon^{\delta}_{i,t+\delta})$ across groups $i = 1, \ldots, N$. We use efficient GMM weighting, with an estimate of the inverse of the long-run variance-covariance matrix of our sample moment conditions. Thus, e.g., if there are groups with low variance moments, they will be weighted heavily in estimation with potentially big efficiency gains. The correlations across sample moments of course also matter with gains from multiple moments perhaps being largely undone in cases where moments are highly correleated across groups.

References

Beraga, M., E. Hurst and J. Ospina (2019), "The Aggregate Implications of Regional Business Cycles," *Econometrica*, 87(6), 1789-1833.

Berger, D., V. Guerrieri, G. Lorenzoni and J. Varva (2017), "House Prices and Consumption Spending," *Review of Economic Studies*, 85(3),1502-1542.

Carlino, G. and R. DeFina (1998), "The Differential Regional Effects of Monetary Policy," *Review of Economics and Statistics*, 80(4), 572-587.

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

Chodorow-Reich, G., L. Feiveson, Z. Liscow and W. Woolston (2012), "Does State Fiscal Relief During Recessions Increase Employment? Evidence from the American Recovery and Reinvestment Act," *American Economic Journal: Economic Policy*, 4, 118-145.

Clemens, J. and S. Miran (2012), "Fiscal Policy Multipliers on Subnational Government Spending," *American Economic Journal: Economic Policy,* 4, 46-68.

Cochrane, J. (2012), "Manna From Heaven: The Harvard Stimulus Debate," Accessed August 14, 2015, http://johnhcochrane.blogspot.com/2012/03/manna-from-heaven-harvardstimulus.html

Conley, T., C. Hansen and P. Rossi (2012), "Plausibly Exogenous," *Review of Economics and Statistics*, 94(1): 260-272.

Dupor, B., M. Karabarbounis, M. Kudlyak, and M.S. Mehkari (2020), "Regional Consumption Responses and the Aggregate Fiscal Multiplier," working paper.

Guren, A., A. McKay, E. Nakamura and J. Steinsson (2019), "What Do We Learn from Cross-Sectional Empirical Estimates in Macroeconomics?," UC-Berkeley, working paper.

Hansen, C., D. Kozbur and S. Misra (2021), "Targeted Undersmoothing: Sensitivity Analysis for Sparse Estimators," *Review of Economics and Statisitcs*.

Känzig, D. (2021), "The Macroeconomic Effects of Oil Supply News: Evidence from OPEC Announcments," *American Economic Review*, 111(4), 1092-1125.

Kaplan, G. and G. Violante (2014), "A Model of the Consumption Response to Fiscal Stimulus Payments," *Econometrica*, 82(4), 1199-1239.

Mian, A., K. Rao and A. Sufi (2013), "Household Balance Sheets, Consumption, and the Economic Slump," *Quarterly Journal of Economics* 128, 1687-1726.

Mumtaz, H., L. Sunder-Plassmann, A. Theophilopoulou (2018), "The State-Level Impact of Uncertainty Shocks," *Journal of Money, Credit and Banking* 50, 1879-1899.

Nakamura, E. and J. Steinsson (2014), "Fiscal Stimulus in a Monetary Union: Evidence from US Regions," *American Economic Review*, Vol. 104, No. 3, pp. 753-92.

Owyang, M. T., and Zubairy, S. (2013). "Who Benefits from Increased Government Spending? A State-level Analysis," *Regional Science and Urban Economics*, 43(3), 445-464.

Ramey, V. (2011), "Can Government Purchases Stimulate the Economy?" *Journal of Economic Literature*, 49, 673-85.

Wolf, C. (2019), "The Missing Intercept: A Demand Equivalence Approach," Princeton University, working paper.

Appendix for "The *Sine Aggregatio* Approach to Applied Macro"

A Sample of Applicable Data Sets

Table 4 provides a list of data sets in which each is disaggregated along at least one dimension (e.g, industry, geography, type of good, function) from which an aggregate variable of interest to macroeconomists is constructed. Each data set consists of a panel. This is a necessary condition for our method to be applied since we rely on time series averages for estimation and inference.

Table 4: Examples of publicly available, U.S. panel data sets suitable for the *sine aggregatio* approach

Data set	**Variable(s)**	**Disaggregation level/type**
Quarterly Census of Employment and Wages	Employment, wages, number of establishments	State, county, MSA, industry
Local Area Unemployment Statistics	Labor force participation, unemployment	State, metropolitan area
BEA Personal Income	Earnings, personal income, transfers	State
Consumer Price Index	Consumer prices	Region, type of good
Personal Income and Outlays	Personal consumption expenditure	Type of good
PCE Price Index	Consumer prices	Type of good
National Income and Product Accounts	Income, compensation, employment, production	Industry
National Income and Product Accounts	Government spending	Function
Industrial Production and Capacity Utilization	Production index, capacity utilization index	Market group, major industry group
IRS Corporate Tax Data	Taxes, credits, payments, net income, income subject to tax	Industry, size of business receipts
IRS Individual Income Tax	Adjusted gross income, exemptions, deductions, tax items	Income percentile, state, county, zip code
BEA Industry Accounts	Value of material inputs by type	Industry

Notes: Each data set contains one or more variables presented at a disaggregate level for which an aggregate variable of interest is constructed and used in macroeconomic research. BEA=Bureau of Economic Analysis, IRS=Internal Revenue Service, PCE=Personal Consumption Expenditure.

B Alternative Departures from a Common Effect

In this section, we consider alternative departures from the common coefficient (i.e., $K = 0$) approach. We suppose that among the nine groups, there are three clusters and each cluster consists of three groups. Within a cluster, the coefficient-of-interest is identical. We impose no cross-cluster restrictions on the coefficient.

Next, we construct clusters according to one of two algorithms. The first builds all possible combinations of groupings into clusters (which we refer to as "all combinations"). This generates a total of 280 possible restrictions.

Table 5: Employment effect of a 20 percent increase in the real price of oil, 24 month horizon: *sine aggregatio* method using clustered departure from the common effect assumption

Estimation Method	Low CI	High CI	Midpt.	CI rel length (%)	Num models
Aggregated data	-3.85	0.32	-1.77		
Sine aggregatio (K=0)	-1.71	0.13	-0.79	44.19	1
3 clusters (all combinations)	-3.27	0.97	-1.15	84.04	280
3 clusters (restricted)	-2.56	0.82	-0.87	66.99	6

Notes: Dependent variable=24 month change in log employment. 90 percent CIs reported. "(all combinations)" row reports our union CI taking the union over mixture CIs from all possible combinations of 3 member-3 cluster partitions of the 9 groups. "(restricted)" row restricts attention to clusters constructed by grouping according to economic variables.

The second algorithm constructs clusters using some measure of economic closeness that potentially relate to differential treatment effects of oil prices on group employment. We use six different state-level variables: the employment share of manufacturing, the output share from the oil and gas extraction industry, the output share of the motor vehicle industry, population, income per capita, and oil usage.[16] For each of the six, we construct clusters using the variable's terciles.

Each of the six variables potentially motivates a common parameter restriction. For example, consider oil usage. It is plausible that oil-intensive clusters experience

[16] Oil usage is measured as British Thermal Units of Petroleum per capita in 1995.

a greater effect on employment of an oil supply disturbance than relatively less oil-intensive clusters. Table 5 presents the results using the clustering assumptions. The first two rows restate the 24-month responses using the aggregated-data and the common coefficient *sine aggregatio* approaches, respectively. The row labelled "3 clusters (all combinations)" reports the union of the confidence intervals for the 280 models case, i.e., all possible three-member clusterings of the nine groups. The midpoint of union is -1.15, which is very close to the common coefficient value. Note that the confidence interval length is slightly smaller than the aggregated data case.. The confidence interval of the "all combinations" case is 84 percent of the confidence interval from the standard, aggregated data method.

The final row of Table 5, labelled "3 clusters (restricted)", uses 6 different clusterings, in which each is determined by the group-level terciles of one of the variables listed above. In this case, there is a substantial improvement of the precision relative to the aggregated-data method. The confidence interval shrinks by about one-third. This final row demonstrates how employing *ex ante* information to restrict the set of potential departures from the common coefficient assumption increases the improvement achieved by the *sine aggregatio* method.

C The Method Expressed Step-by-Step

The following presents a cookbook recipie for the method described in this paper.

1. Select a time series for the aggregate treatment of interest of length T and, if necessary, a valid instrument to ensure that an orthogonality condition will be satisfied.

2. Select an aggregate outcome of interest among those for which the disaggregate panel is available. If the cross-sectional dimension is too large relative to T, some partial cross-sectional aggregation (i.e., collapsing to the group level) may be required. This is because the method relies on a reliable estimate of the second-moment matrix, which will be used to construct the efficient GMM weighting matrix. Let the number of resulting groups equal N.

3. Construct an estimation equation with the group-level analog of the aggregate outcome as the dependent variable. The group-level outcome should be expressed as a function of the aggregate treatment. The equation must "aggregate" such that, if one sums the equation along the cross-sectional dimension, the resulting coefficient on the treatment can be interpreted as an aggregate causal effect.

4. As part of Step 3, specify an equation that parameterizes some departure from a common coefficient assumption across the N groups. For example, allow $N - K$ groups to have common slopes and the remaining K to have different slopes. For another example, allow there to be alternative partitions of the N groups into M clusters. Let Q equal the number of alternative specifications implied by the particular departure from common coefficients.

5. Estimate the equation from Step 3 for the Q alternative models and record the associated confidence interval around the aggregate treatment effect implied by the corresponding group-level parameter estimates. Remove any confidence interval associated that a rejection of that model's overidentification restriction.

6. The union of confidence intervals from the remaining set of models provides a confidence interval for the effect of the aggregate treatment on the aggregate outcome.