The *Sine Aggregatio* Approach to Applied Macro*

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

July 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-series parameter restrictions to increase precision relative to traditional, aggregate series estimates and provide a method to assess robustness to modest 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 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

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

*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.

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.

In this paper, we focus on a common scenario where the available time series are of modest length, resulting in noisy estimates based solely on aggregate data. Furthermore, in our scenario, $\mathcal{Y}_t$ is by construction the sum of some regional, sectoral or other disaggregate series.[2] That is, we have series for N mutually exclusive, disaggegate 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. Thus, 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 examine departures from it is useful. Examining relaxations of a common parameter assumption is also well-motivated even in situations where a test of a restriction like that in equation (2) is not rejected, given the potential for tests involving many parameters to have low power against economically relevant alternatives where a few parameters differ.

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

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

The aggregate effect of interest, B, is the sum across groups of the β_i.[3]

We present an empirical method that exploits efficiency gains from restrictions on sets of common parameters in the β_i and allows researchers to tailor these restrictions to match their view of the needed flexibility. For simplicity of exposition, we assume that the same ag-

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

[2] Examples include regional employment and sectoral output.

[3] In Section 3, we generalize the definition of B to allow for weighted sums.

gregate 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. Our method is straightforward to apply with group-specific instruments, weak instrument robust estimators, or other criterion functions.

Our method lets the researcher decide how much parameter flexibility to allow across groups. We consider restrictions that allow a small number of groups, K, to have differing group-level treatment slopes, i.e. β_i; 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 framework, with $K = 0$, is that all groups' β_i are the same. As K increases the sets of restrictions become more flexible.[4]

The method begins with estimating the model under each possible restriction that has K groups with individual β_i and the rest with a common value. For each of these N choose K parameterizations, we obtain a confidence interval (CI) for β. We take the union of these CIs to obtain an interval estimate for β 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 illustrate our method by estimating 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.[5] We divide U.S. states into nine disjoint groups, each with roughly the same group employment totals.[6]

In our benchmark specification, using the standard aggregative approach yields point estimates that imply a contractionary oil supply shock reduces employment but the effect is imprecisely estimated. We next estimate the aggregate effect using our panel-based approach

[4]Our method can be easily adapted to other sets of candidate restrictions as long as they are few enough in number to be computationally feasible.

[5]In the remainder of the paper, all references to employment should be understood to mean manufacturing employment.

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

for $K = 0$ through $K = 3$. For each of these values, the midpoint of the CI implies a twenty percent increase in the real price of oil driven by the shock drives down 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 $(-4.8, 0.38)$. This implies a CI shrinkage of 43 percent.

2 Related Research

Our paper offers a new path beyond 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.[7] 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., Nakamura and Steinsson (2014)). One advance 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 dynamic equilibrium macro models (e.g., Beraja, Hurst and Ospina (2019), Dupor, et.al. (2020) and Nakamura and Steinsson (2014)).[8]

In contrast to this path, 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. Even though we use disaggregate data, our basic identification is based on time series variation in our instrument(s). In that aspect, our method follows the Sims (1972) structural vector autoregression approach, along with the later developments of narrative and external instruments.

[7] 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.

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

3 The Econometric Model

There are T periods and N groups. Let $Y_{i,t}$ denote a group-level observable outcome.[9] Next, let $x^{\delta}_{t+\delta}$ be the aggregate treatment and z_t be the aggregate instrument. We define the the long difference of $Y_{i,t}$ as: $y^{\delta}_{i,t+\delta} = Y_{i,t+\delta} - Y_{i,t-1}$. Also, we define $y^{\delta}_{t+\delta} \equiv \sum_i s_i y^{\delta}_{i,t+\delta}$ where the s_i are share weights.[10] In our application, $Y_{i,t}$ equals log employment in group-month (i,t).

Consider the following panel, 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. Let this aggregate response equal B, with $B \equiv \sum_j s_j \beta_j$.[11] To see this, for each t multiply equation (4) by s_i for each i and then sum over i. This delivers:

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

Note that to a first-order approximation, $y^{\delta}_{t+\delta} \approx Y^{A}_{t+\delta} - Y^{A}_{t-1}$, where Y^{A}_{t} is the log of national employment.[12]

Our moment conditions are

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

for $i = 1, \ldots, N$. In our specific application, z_t is the oil price news shock described in the introduction and $x^{\delta}_{t+\delta}$ is the change in the log real price of oil between $t-1$ and $t+\delta$. Our identification assumption is that, at the group-level, the error term of the group-level equation is uncorrelated with the oil shock instrument.

We are concerned with applications where N is large enough that the researcher is motivated to find efficiency gains from imposing common parameter restrictions on subsets of the β_i. We also want a method that can be applied with several candidate restrictions, so that we can investigate robustness to variation in restrictions. We focus on restrictions that are relaxations of a common parameter assumption. Our restrictions take the form: All but

[9] In our application, each group corresponds to a collection of U.S. states, in which each U.S. state belongs to one group, and a period equals one month.

[10] In our application, we use the initial period $Y_{i,t}$ to construct these weights.

[11] Here, we abuse notation slightly by changing the definition of B from the raw sum of β_i, defined in the introduction, to the corresponding share-weighted sum.

[12] In our data set, the correlation between $y^{\delta}_{t+\delta}$ and $Y^{A}_{t+\delta} - Y^{A}_{t-1}$ is greater than 0.999, indicating a very close approximation.

a small number K of the β_i are the same.[13]

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 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 test of GMM over-identification restrictions under the restriction r, which we denote J_r. If a model is rejected, 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 J_r test, to get a 90% $\widetilde{CI}(B_r)$ we use:

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

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. Rather than pick a specific value for K, we anticipate researchers will want to examine results for a set of values of K from zero to some small number.

4 Application: Oil Price News Shocks

First, let G_i be the set of states that belong to group i. Let Emp_t^j denote the employment level in state j in month t. Let $Y_{i,t} = \log\left(\sum_{j \in G_i} Emp_t^j\right)$. Then our outcome variable is the horizon δ long difference of the group-level log employment.

[13]However, our method can readily be applied with any other type of restrictions, e.g. clusters within which groups have common parameters, as long as the total number of options is not too large.

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.

We use nine groups, chosen so that the groups have similar shares of employment. In all but one case, the groups are geographically contiguous.[14]

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.

Because of the delayed effect of oil shocks on economic activity found in existing research, we choose a 18 month horizon for our local projection estimation (i.e., $\delta = 17$). Our state-level employment data are from the Bureau of Labor Statistics. 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.

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 equal to zero in our application.[15] Finally, we scale $x^{\delta}_{t+\delta}$ such that a unit shock is associated with a 20

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

[15] 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.

percent increase in the real price of oil between $t-1$ and $t+\delta$. Our sample covers January 1991 to January 2017.

Next, we obtain efficient estimates via iterated GMM assuming second-moment independence.[16] 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 leads and lags.[17]

Table 2: 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	-4.81	0.38	-2.21		
Sine aggregatio					
K = 0	-2.02	0.01	-1.01	39	1
K = 1	-2.53	0.40	-1.07	57	9
K = 2	-2.70	0.76	-0.97	67	36
K = 3	-3.10	0.96	-1.07	78	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-stat and associated p-value for the over-identifying restriction are 8.0 and 0.43.

Table 3: 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	-3.92	0.45	-1.73		
Sine aggregatio					
K = 0	-1.73	0.11	-0.81	42	1
K = 1	-2.11	0.48	-0.81	59	9
K = 2	-2.69	0.94	-0.87	83	36
K = 3	-2.74	1.16	-0.79	89	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-stat and associated p-value for the over-identifying restriction are 7.5 and 0.50.

[16] 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. In contrast to two-step GMM, using iterated GMM lessens the arbitrary dependence on the initial weighting matrix.

[17] Our results are robust to using alternate Bartlett weights from 10 to 30 month lags.

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.[18]

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 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.

We compare our approach to estimation using just the aggregate time series. That is, we estimate

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

In this case, the model is exactly identified.

Table 2 reports the effect on employment of a 20 percent increase in the real price of oil, both at an 18 month horizon and 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 (8)). The 90 percent CI equals $(-4.8, 0.38)$ with a midpoint equal to -2.2. It indicates that the oil price increase–driven by a contractionary news shock–has a lower bound of decreasing employment by 4.8 percent to an upper bound estimate of increasing employment by 0.38 percent. Notably, the effect is imprecisely estimated.

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 -1.0, which is of the same sign but somewhat closer to zero than the aggregated-data based estimate of -2.2. The entry for the "CI rel length (%)" column indicates that the $K = 0$ *sine aggregatio* CI length is 39 percent of the aggregate estimate CI length.[19]

[18] Here we use the approximation $Y^A_{t+\delta} - Y^A_{t-1} \approx \sum_i s_i y^\delta_{i,t+\delta}$.

[19] We also report the J test of overidentification restrictions, which embeds the coefficient restrictions on β_i. The J-statistic equals 8.0 and has eight degrees of freedom. The associated p-value equals 0.43. Thus, the data do not reject the model at conventional significant levels.

The next row reports the $K = 1$ estimate. Here the reported CI is union of nine models;[20] each model corresponds to one of the nine groups being allowed to have a distinct slope. The midpoint of this estimate equals -1.04, 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 $(-2.5, 0.40)$.

The final two rows report estimates for K equal 2 and 3 cases. For $K = 2$, we 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 substantial improvements relative to the aggregate series CI coming in at 67% and 78% 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 24 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 42 percent of the analogue aggregate-based CI length. The corresponding value for $K = 1$ equals 59 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 an macroeconomy responds to macro treatments.

Given the novelty of our approach, it is worthwhile to restate some of the requirements

[20] 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.

necessary to usefully apply the method. 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.

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.

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

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.

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.

Plagborg-Møller, M. and C. Wolf (2021), "Local Projections and VARs Estimate the Same Impulse Responses." *Econometrica*, 89: 955-980.

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

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