NBER WORKING PAPER SERIES

QUASI-EXPERIMENTAL SHIFT-SHARE RESEARCH DESIGNS

Kirill Borusyak
Peter Hull
Xavier Jaravel

Working Paper 24997
http://www.nber.org/papers/w24997

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
September 2018, Revised December 2019

We are grateful to Rodrigo Adão, Joshua Angrist, David Autor, Moya Chin, Andy Garin, Ed Glaeser, Paul Goldsmith-Pinkham, Larry Katz, Michal Kolesár, Gabriel Kreindler, Jack Liebersohn, Eduardo Morales, Jack Mountjoy, Jörn-Steffen Pischke, Brendan Price, Isaac Sorkin, Jann Spiess, Itzchak Tzachi Raz, various seminar participants, and four anonymous referees for helpful comments. We thank David Autor, David Dorn, and Gordon Hanson, as well as Paul Goldsmith-Pinkham, Isaac Sorkin, and Henry Swift, for providing replication code and data. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

Quasi-Experimental Shift-Share Research Designs
Kirill Borusyak, Peter Hull, and Xavier Jaravel
NBER Working Paper No. 24997
September 2018, Revised December 2019
JEL No. C18,C21,C26,F16,J21

ABSTRACT

Many studies use shift-share (or “Bartik”) instruments, which average a set of shocks with exposure share weights. We provide a new econometric framework for such designs in which identification follows from the quasi-random assignment of shocks, allowing exposure shares to be endogenous. This framework is centered around a numerical equivalence: conventional shift-share instrumental variable (SSIV) regression coefficients are equivalently obtained from a transformed regression where the shocks are used directly as an instrument. This equivalence implies a shock-level translation of the SSIV exclusion restriction, which holds when shocks are as-good-as-randomly assigned and large in number, with sufficient dispersion in their average exposure. We discuss and illustrate several practical insights delivered by this framework.

Kirill Borusyak
Drayton House
30 Gordon Street
London
WC1H 0AX
United Kingdom
k.borusyak@ucl.ac.uk

Peter Hull
Becker Friedman Institute
University of Chicago
Saieh Hall 215
5757 South University Avenue
Chicago, IL 60637
and NBER
hull@uchicago.edu

Xavier Jaravel
London School of Economics
Houghton Street
London
WC2A 2AE
UK
x.jaravel@lse.ac.uk

A Replication programs and data is available at https://github.com/borusyak/shift-share

1 Introduction

A large and growing number of empirical studies use shift-share instruments: weighted averages of a common set of shocks, with weights reflecting heterogeneous shock exposure. In many settings, such as those of Bartik (1991), Blanchard et al. (1992) and Autor et al. (2013), a regional instrument is constructed from shocks to industries with local industry employment shares measuring the shock exposure. In other settings, researchers may combine shocks across countries, income groups, or foreign markets to instrument for treatments at the regional, individual, or firm level.[1]

The claim for instrument validity in shift-share instrumental variable (SSIV) regressions must rely on some assumptions about the shocks, exposure shares, or both. This paper develops a novel framework for understanding SSIV regressions as leveraging exogenous variation in shocks, even when variation in exposure shares is endogenous. Our approach is centered around a simple numerical equivalence: we show that SSIV regression coefficients are identically obtained from a transformed IV regression, estimated at the level of shocks. In this equivalent regression the outcome and treatment variables are first averaged, using exposure shares as weights, to obtain shock-level aggregates. The shocks then directly instrument for the aggregated treatment. Importantly, this equivalence only relies on the structure of the shift-share instrument and thus applies to outcomes and treatments that are not typically computed at the level of shocks. It follows that the SSIV exclusion restriction holds if and only if shocks are uncorrelated, in large samples, with a particular residual: the average unobserved determinants of the original outcome among observations most exposed to a given shock.

We use this equivalence result to derive two conditions sufficient for such "shock orthogonality." First, we assume shocks are as-good-as-randomly assigned, as if arising from a natural experiment. This is enough for the SSIV exclusion restriction to hold on average across shock realizations. Second, we assume that a shock-level law of large numbers applies: that the instrument incorporates many sufficiently independent shocks, each with sufficiently small average exposure. Instrument relevance further holds when individual units are mostly exposed to only a small number of shocks, provided those shocks affect treatment. While novel for SSIV, our two quasi-experimental conditions are similar to ones imposed in other settings where the underlying shocks are directly used as instruments, bringing SSIV to familiar econometric territory.[2] We illustrate these ideas in an idealized example, in which a local labor supply elasticity is estimated with a shift-share instrument constructed from quasi-random output subsidy shocks to different industries. We highlight that the SSIV coefficient retains

[1] Observations in shift-share designs may, for example, represent regions impacted by immigration shocks from different countries (Card 2001), firms differentially exposed to foreign market shocks (Hummels et al. 2014), product groups demanded by different types of consumers (Jaravel 2019), or groups of individuals facing different income growth rates (Boustan et al. 2013). Other influential and recent examples of shift-share IVs include Luttmer (2005), Saiz (2010), Kovak (2013), Nakamura and Steinsson (2014), Oberfield and Raval (2014), Greenstone et al. (2014), Diamond (2016), Suárez and Zidar (2016), and Hornbeck and Moretti (2019).

[2] For example, Acemoglu et al. (2016) study the impact of import competition from China on U.S. industry employment using industry (i.e. shock-level) regressions with shocks constructed similarly to those underlying the regional shift-share instrument used in Autor et al. (2013). Our framework shows that both studies can rely on similar econometric assumptions, even though the economic interpretations of the estimates may be different.

its local labor market interpretation, while being identified by exogenous industry-level variation.

This quasi-experimental approach is readily extended to settings where shocks are as-good-as-randomly assigned conditional on shock-level observables and to panel data. For conditional random assignment, we show that quasi-experimental shock variation can be isolated with regression controls that have a shift-share structure; namely, one can control for an exposure-weighted average of the relevant shock-level confounders. In panel data, we show that the SSIV estimator can be consistent both with many shocks per period and with many periods. We also show that unit fixed effects play a special role in SSIV with time-invariant exposure shares, isolating variation in shocks over time. In further extensions we show that it is important to control for the sum of shift-share exposure weights when such sums vary, how SSIV with multiple endogenous variables can also be viewed quasi-experimentally, and how to optimally combine multiple sets of quasi-random shocks with new overidentified shock-level IV procedures.

Our framework also bears practical tools for SSIV inference and testing. Adão et al. (2019) show that conventional standard errors in SSIV regressions are generally invalid in a framework based on ours, with identifying variation in shocks, because observations which similar exposure shares are likely to have correlated residuals. We provide a convenient solution to this problem, deriving assumptions under which conventional standard errors are valid ("exposure-robust") when the SSIV coefficients are estimated at the level of identifying variation, i.e. from equivalent shock-level regressions. Our SSIV inference procedure is asymptotically equivalent to the one proposed by Adão et al. (2019) but has several practical advantages: it can be implemented with standard statistical software, readily extended to various forms of shock dependence (e.g. autocorrelation), and computed in some settings where the estimator of Adão et al. (2019) fails. Appropriate measures of first-stage relevance and valid falsification tests of the quasi-random assignment of shocks can also be obtained in this way.

We illustrate these and other insights from our framework in the setting of Autor et al. (2013) and find evidence which supports the interpretation of their research design as leveraging quasi-random shock variation. This application uses a new Stata package, *ssaggregate*, which we have developed to help practitioners implement the appropriate shock-level analyses.[3]

Our quasi-experimental approach is not the only way to satisfy the SSIV exclusion restriction. In related work, Goldsmith-Pinkham et al. (2019) formalize a different framework based on the exogeneity of the exposure shares, imposing no assumption of shock exogeneity. This approach is motivated by a different numerical equivalence: the SSIV coefficient also coincides with a generalized method of moments estimator, with exposure shares as multiple excluded instruments. Though exposure exogeneity is a sufficient condition for identification (and, as such, implies our shock-level orthogonality condition), we focus on plausible conditions under which it is not necessary.

[3]This Stata package creates the shock-level aggregates used in the equivalent regression. Users can install this package with the command *ssc install ssaggregate*. See the associated help file and this paper's replication archive at https://github.com/borusyak/shift-share for more details.

Identification via exogenous shocks seems attractive in many SSIV settings. Consider the Autor et al. (2013; hereafter ADH) shift-share instrument, which combines industry-specific changes in Chinese import competition (the shocks) with local exposure given by the lagged industrial composition of U.S. regions (the exposure shares). In such a setting, exogeneity of industry employment shares may be difficult to justify *a priori*. Indeed, we show that the "shares" view to identification generally fails when there are any unobserved industry shocks that affect regional outcomes through the shares (e.g., unobserved automation trends). Our approach, in contrast, allows researchers to specify a set of shocks that are plausibly uncorrelated with such unobserved factors. Consistent with this general principle, ADH attempt to purge their industry shocks from U.S.-specific confounders by measuring Chinese import growth outside of the United States. Similarly, Hummels et al. (2014) combine country-by-product changes in transportation costs to Denmark (as shocks) with lagged firm-specific composition of intermediate inputs and their sources (as shares). They argue these shocks are "idiosyncratic," which our approach formalizes as "independent from relevant country-by-product unobservables." Other recent examples of where our approach may naturally apply include the exchange rate shocks of Hummels et al. (2011), the education policy shocks of Stuen et al. (2012), the demographic shocks of Jaravel (2019), and the bank health shocks of Xu (2019).

In other shift-share designs, the shocks are equilibrium objects that can be difficult to view as being quasi-experimentally assigned. In the canonical estimation of regional labor supply elasticities by Bartik (1991), for example, the shocks are measured as national industry growth rates. Such growth captures national industry labor demand shocks, which one may be willing to assume are as-good-as-randomly assigned across industries; however, industry growth rates also depend on unobserved regional labor supply shocks. We show that our framework can still apply to such settings by casting the industry employment growth rates as noisy estimates of latent quasi-experimental demand shocks and establishing conditions to ensure the supply-driven estimation error is asymptotically ignorable. These conditions are weaker if the latent shocks are estimated as leave-one-out averages. Although leave-one-out shift-share IV estimates do not have a convenient shock-level representation, we provide evidence that in the Bartik (1991) setting this leave-out adjustment is unimportant.

Formally, our approach to SSIV relates to the analysis of IV estimators with many invalid instruments by Kolesar et al. (2015). Consistency in that setting follows when violations of individual instrument exclusion restrictions are uncorrelated with their first-stage effects. For quasi-experimental SSIV, the exposure shares can be thought of as a set of invalid instruments (per the Goldsmith-Pinkham et al. (2019) interpretation), and our orthogonality condition requires their exclusion restriction violations to be uncorrelated with the shocks. Despite this formal similarity, we argue that shift-share identification is better understood through the quasi-random assignment of a single instrument (shocks), rather than through a large set of invalid instruments (exposure shares) that nevertheless produce a consistent estimate. This view is reinforced by our numerical equivalence, yields a natural shock-level identification condition, and suggests new validations and extensions of SSIV.

Our analysis also relates to other recent methodological studies of shift-share designs, including those of Jaeger et al. (2018) and Broxterman and Larson (2018). The former highlights biases of SSIV due to endogenous local labor market dynamics, and we show how their solution can be implemented in our setting, while the latter studies the empirical performance of different shift-share instrument constructions. As discussed above we also draw on the inferential framework of Adão et al. (2019), who derive valid standard errors in shift-share designs with a large number of idiosyncratic shocks. More broadly, our paper adds to a growing literature studying the causal interpretation of common research designs, including work by Borusyak and Jaravel (2017) and Goodman-Bacon (2018) for event study designs; Hudson et al. (2017) and Chaisemartin and D'Haultfoeuille (2019) for instrumented difference-in-difference designs; and Hull (2018) for mover designs.

The remainder of this paper is organized as follows. Section 2 introduces the environment, derives our numerical equivalence for shift-share IV, and discusses the key shock orthogonality condition. Section 3 then establishes the quasi-experimental assumptions under which this condition is satisfied, and derives various extensions. Section 4 discusses shock-level procedures for valid SSIV inference and testing, while Section 5 illustrates the methodology in the ADH setting. Section 6 concludes. Additional results and proofs are included in the paper's appendix.

2 Shocks as Instruments in Shift-Share Designs

We begin by defining the SSIV estimator and showing that it coincides with a new IV procedure, estimated at the level of shocks. Motivated by this equivalence result, we derive a necessary and sufficient shock-level orthogonality condition for shift-share instrument validity.

2.1 The Shift-Share IV Estimator

We consider a sequence of data generating processes, indexed by the number of observations L. The data include an outcome variable y_ℓ, an endogenous (or "treatment") variable x_ℓ, a vector of controls w_ℓ (which includes a constant), and an observation importance weight $e_\ell > 0$ (with $\sum_{\ell=1}^{L} e_\ell = 1$; $e_\ell = \frac{1}{L}$ covers the unweighted case). For reasons we will discuss in detail in Section 3, we do not assume independent or identically-distributed draws of these variables across ℓ.

We are interested in estimation of the causal effect or structural parameter β in a linear model relating outcomes to treatment: $y_\ell = \beta x_\ell + \epsilon_\ell$.[4] To accommodate the control vector we assume that the large-sample auxiliary e_ℓ-weighted projection of unobserved untreated potential outcomes, ϵ_ℓ, on w_ℓ is well-defined; that is $\hat{\gamma} = \left(\sum_{\ell=1}^{L} e_\ell w_\ell w_\ell'\right)^{-1} \left(\sum_{\ell=1}^{L} e_\ell w_\ell \epsilon_\ell\right) \xrightarrow{p} \gamma$ for some γ as $L \to \infty$. We then

[4] We consider models with heterogeneous treatment effects in Appendix A.1; see footnote 8 for a summary.

study the expanded model

$$y_\ell = \beta x_\ell + w_\ell' \gamma + \varepsilon_\ell, \tag{1}$$

where $\varepsilon_\ell = \epsilon_\ell - w_\ell' \gamma$ is a structural residual.[5]

For example, we might be interested in estimating the inverse labor supply elasticity β from observations of log wage growth y_ℓ and log employment growth x_ℓ across local labor markets ℓ. The residual ε_ℓ in (1) would then contain all labor supply shocks, such as those arising from demographic, human capital, or migration changes, that are not asymptotically correlated with the control vector w_ℓ. In estimating labor supply we may weight observations by the overall lagged regional employment, e_ℓ. We return to this labor supply example throughout the following theoretic discussion.

To estimate β we construct a shift-share instrument from a set of shocks g_n, for $n = 1, \dots, N$, and shares $s_{\ell n} \geq 0$ which define the relative exposure of each observation ℓ to each shock n. Specifically, the instrument is given by an exposure-weighted average of the shocks:

$$z_\ell = \sum_{n=1}^{N} s_{\ell n} g_n. \tag{2}$$

In the labor supply example, where a local labor demand instrument is called for, g_n may for instance denote new government subsidies to the output of different industries n and $s_{\ell n}$ may be location ℓ's lagged shares of industry employment. For now we require that the sum of exposure weights is constant across observations, i.e. that $\sum_{n=1}^{N} s_{\ell n} = 1$; we relax this in Section 3.4. Although our focus is on shift-share IV, we also note that this setup nests shift-share reduced-form regressions of y_ℓ on z_ℓ, when $x_\ell = z_\ell$.

The SSIV estimator $\hat{\beta}$ uses z_ℓ to instrument for x_ℓ in equation (1), weighting by e_ℓ. By the Frisch-Waugh-Lovell Theorem this estimator can be represented as a bivariate IV regression of outcome and treatment residuals, i.e. as the ratio of e_ℓ-weighted sample covariances between the instrument and the residualized outcome and treatment:

$$\hat{\beta} = \frac{\sum_{\ell=1}^{L} e_\ell z_\ell y_\ell^\perp}{\sum_{\ell=1}^{L} e_\ell z_\ell x_\ell^\perp}, \tag{3}$$

where $v_\ell^\perp$ denotes the residual from an e_ℓ-weighted projection of variable v_ℓ on the control vector w_ℓ. Note that by the properties of such residualization, it is enough to residualize y_ℓ and x_ℓ without also residualizing the instrument z_ℓ.

[5] For simplicity we suppose here that w_ℓ is of fixed length, such that $\hat{\gamma}$ consistently estimates a fixed number of coefficients. We weaken this assumption in the appendix proofs to the following results, allowing for an increasing number of group fixed effects or other controls subject to regularity conditions.

2.2 An Equivalent Shock-Level IV Estimator

Cross-sectional variation in the shift-share instrument only arises from differences in the exposure shares $s_{\ell n}$, since the set of shocks g_n is the same in the construction of each z_ℓ. It may be natural to conclude that such share variation is central for the exogeneity of z_ℓ. The following result suggests that this view is incomplete:

Prop. 1 The SSIV estimator $\hat{\beta}$ equals the second-stage coefficient from a shock-level IV regression that uses the shocks g_n as the instrument in estimating

$$\bar{y}_n^\perp = \alpha + \beta \bar{x}_n^\perp + \bar{\varepsilon}_n^\perp, \tag{4}$$

where $\bar{v}_n = \frac{\sum_{\ell=1}^L e_\ell s_{\ell n} v_\ell}{\sum_{\ell=1}^L e_\ell s_{\ell n}}$ denotes an exposure-weighted average of variable v_ℓ and the IV estimation is weighted by average shock exposure $s_n = \sum_{\ell=1}^L e_\ell s_{\ell n}$.

***Proof*:** By definition of z_ℓ, and by exchanging the order of summation in (3),

$$\hat{\beta} = \frac{\sum_{\ell=1}^L e_\ell \left(\sum_{n=1}^N s_{\ell n} g_n\right) y_\ell^\perp}{\sum_{\ell=1}^L e_\ell \left(\sum_{n=1}^N s_{\ell n} g_n\right) x_\ell^\perp} = \frac{\sum_{n=1}^N g_n \left(\sum_{\ell=1}^L e_\ell s_{\ell n} y_\ell^\perp\right)}{\sum_{n=1}^N g_n \left(\sum_{\ell=1}^L e_\ell s_{\ell n} x_\ell^\perp\right)} = \frac{\sum_{n=1}^N s_n g_n \bar{y}_n^\perp}{\sum_{n=1}^N s_n g_n \bar{x}_n^\perp}. \tag{5}$$

Furthermore $\sum_{n=1}^N s_n \bar{y}_n^\perp = \sum_{\ell=1}^L e_\ell \left(\sum_{n=1}^N s_{\ell n}\right) y_\ell^\perp = \sum_{\ell=1}^L e_\ell y_\ell^\perp = 0$, since $y_\ell^\perp$ is an e_ℓ-weighted regression residual and $\sum_{n=1}^N s_{\ell n} = 1$. This and an analogous equality for $\bar{x}_n^\perp$ imply that $\hat{\beta}$ in (5) is a ratio of s_n-weighted covariances of $\bar{y}_n^\perp$ and $\bar{x}_n^\perp$ with g_n; hence it is obtained from the specified shock-level IV regression. □

Proposition 1 is our main equivalence result, which shows that SSIV estimates can also be thought to arise from variation across shocks, rather than across observations. The IV regression that leverages this variation uses shock-level aggregates of the original (residualized) outcome and treatment, $\bar{y}_n^\perp$ and $\bar{x}_n^\perp$. Specifically, $\bar{y}_n^\perp$ reflects the average residualized outcome of the observations most exposed to the nth shock, while $\bar{x}_n^\perp$ is the same weighted average of residualized treatment. Each shock in this regression is weighted by s_n, representing its average (e_ℓ-weighted) exposure across observations.[6]

The labor supply example is useful for unpacking this general result. One expects industries n which receive a higher output subsidy g_n to increase employment in the regions ℓ that they are most active in, such that the shock-level first-stage regression of $\bar{x}_n^\perp$ on g_n is positive. The shock-level reduced-form regression of $\bar{y}_n^\perp$ on g_n further reflects the extent to which industries with higher subsidies also tend to be concentrated in regions with higher wage growth. Proposition 1 shows that the resulting shock-level IV regression estimate is numerically the same as the original SSIV estimate

[6] Note that in the special case where $\hat{\beta}$ comes from a reduced-form shift-share regression, Proposition 1 shows that the equivalent shock-level procedure is still an IV regression, of $\bar{y}_n^\perp$ on the transformed shift-share instrument $\bar{z}_n^\perp$, again instrumented by g_n and weighted by s_n.

(3). The shock-level "importance weights" s_n are similarly intuitive in this setting: if, as is common in practice, the industry employment shares $s_{\ell n}$ are measured in the same pre-period as total regional employment e_ℓ, then s_n will be proportional to total lagged industry employment. Without location importance weights (i.e. $e_\ell = \frac{1}{L}$), s_n is the average employment share of industry n across locations.

It is worth emphasizing that $\bar{y}_n$ and $\bar{x}_n$ (and their residualized versions in Proposition 1) are unconventional shock-level objects. They can, for example, be computed for outcomes and treatments that are not typically observed at the level of the shocks, such as when n indexes industries and y_ℓ measures regional marriage rates (Autor et al. 2019). Furthermore, even if the outcome and treatment have natural measures at the shock level, $\bar{y}_n$ and $\bar{x}_n$ will generally not coincide with them. For example, in the labor supply setting $\bar{y}_n$ is not industry n's wage growth; rather, it measures the average wage growth in regions where industry n employs the most workers. Accordingly, the shift-share regression estimates the elasticity of regional, rather than industry, labor supply.[7]

Proposition 1 is an algebraic result which, by itself, does not speak to the consistency of $\hat{\beta}$. At the same time it suggests a shock-level orthogonality condition that is both necessary and sufficient for SSIV consistency, given the standard assumption of first-stage relevance. We derive and interpret this condition next, before developing a quasi-experimental framework that implies it.

2.3 Shock Orthogonality

As usual, SSIV consistency (that $\hat{\beta} \xrightarrow{p} \beta$ as $L \to \infty$) requires and implies instrument exclusion (that z_ℓ and $\varepsilon_\ell^\perp$ are asymptotically uncorrelated), given instrument relevance (that z_ℓ and $x_\ell^\perp$ are asymptotically correlated). We discuss relevance in the following section and focus here on exclusion. The following result shows how applying the equivalency logic of Proposition 1 translates the cross-sectional exclusion condition to a novel orthogonality condition at the shock level:

Prop. 2 Suppose $\sum_{\ell=1}^L e_\ell w_\ell z_\ell \xrightarrow{p} \Omega_{zw}$, $\sum_{\ell=1}^L e_\ell w_\ell \epsilon_\ell \xrightarrow{p} \Omega_{w\epsilon}$, and $\sum_{\ell=1}^L e_\ell w_\ell w_\ell' \xrightarrow{p} \Omega_{ww}$ with Ω_{ww} full rank. Suppose further than instrument relevance holds: $\sum_{\ell=1}^L e_\ell z_\ell x_\ell^\perp \xrightarrow{p} \pi$ with $\pi \neq 0$. Then $\hat{\beta}$ is consistent if and only if

$$\sum_{n=1}^N s_n g_n \bar{\varepsilon}_n \xrightarrow{p} 0. \tag{6}$$

Proof: By Proposition 1 and instrument relevance, $\hat{\beta} - \beta = \frac{\sum_{n=1}^N s_n g_n \bar{\varepsilon}_n^\perp}{\sum_{n=1}^N s_n g_n \bar{x}_n^\perp} = \frac{\sum_{n=1}^N s_n g_n \bar{\varepsilon}_n^\perp}{\sum_{\ell=1}^L e_\ell z_\ell x_\ell^\perp} = \frac{1}{\pi}\sum_{n=1}^N s_n g_n \bar{\varepsilon}_n^\perp + o_p(1)$, so the result is trivial absent a control vector (i.e. when $\bar{\varepsilon}_n = \bar{\varepsilon}_n^\perp$). Otherwise by the second two regularity conditions $\hat{\gamma} \xrightarrow{p} \Omega_{ww}^{-1}\Omega_{w\epsilon} = \gamma$ and $\sum_{n=1}^N s_n g_n \bar{\varepsilon}_n -$

[7] In Appendix A.2 we develop a stylized model to illustrate how the SSIV coefficient can differ from a "native" shock-level IV coefficient in the presence of local spillovers or treatment effect heterogeneity, though both parameters may be of interest. Intuitively, in the labor supply case one may estimate a low regional elasticity but a high elasticity of industry labor supply if, for example, migration is constrained but workers are mobile across industries within a region.

$\sum_{n=1}^{N} s_n g_n \bar{\varepsilon}_n^{\perp} = \sum_{\ell=1}^{L} e_\ell z_\ell \left(\varepsilon_\ell - \varepsilon_\ell^{\perp}\right) = \left(\sum_{\ell=1}^{L} e_\ell z_\ell w_\ell'\right)(\hat{\gamma} - \gamma) \xrightarrow{p} 0$ by the first regularity condition, so $\hat{\beta} - \beta = \frac{1}{\pi} \sum_{n=1}^{N} s_n g_n \bar{\varepsilon}_n + o_p(1)$. □

Equation (6) is our shock-level orthogonality condition for SSIV, restricting the s_n-weighted covariance of the shocks g_n and the aggregated structural residuals $\bar{\varepsilon}_n = \frac{\sum_{\ell=1}^{L} e_\ell s_{\ell n} \varepsilon_\ell}{\sum_{\ell=1}^{L} e_\ell s_{\ell n}}$. Proposition 2 shows that for SSIV to be consistent, given first-stage relevance and three mild regularity conditions, this covariance must be asymptotically zero. In our running labor supply example, such shock orthogonality holds when output subsidies g_n are not systematically higher or lower for industries with higher $\bar{\varepsilon}_n$; i.e. those that are most active (in terms of lagged employment $s_{\ell n} e_\ell$) in the regions that face high unobserved labor supply conditions ε_ℓ.

As a necessary condition, shock orthogonality is satisfied when the exposure shares are exogenous, as in the preferred interpretation of SSIV in Goldsmith-Pinkham et al. (2019). That is, if each $s_{\ell n}$ is uncorrelated with a mean-zero structural error ε_ℓ (given the importance weights e_ℓ), N is fixed, and a conventional law of large numbers applies to $\sum_{\ell=1}^{L} e_\ell s_{\ell n} \varepsilon_\ell$ for each n, then $(\bar{\varepsilon}_1, \ldots, \bar{\varepsilon}_N) \xrightarrow{p} 0$ and equation (6) is satisfied for any fixed set of shocks g_n. This share exogeneity generally rules out any unobserved shocks ν_n that affect the outcome via the exposure shares (i.e. when $\sum_{n=1}^{N} s_{\ell n} \nu_n$ is included in ε_ℓ), even if such ν_n are uncorrelated with the observed shocks g_n, the exposure shares are randomly assigned to observations, and N is allowed to grow (see Appendix A.3 for a formal argument). In the labor supply example this would mean there are no unobserved industry shocks impacting the local labor market besides government output subsidies – a strong restriction in practice.

When shares are endogenous, shock orthogonality (6) may instead be satisfied by certain properties of the SSIV shocks. We now develop a quasi-experimental framework that yields this result.

3 A Quasi-Experimental SSIV Framework

Our approach to satisfying the shock orthogonality condition specifies a quasi-experiment in which shocks are as-good-as-randomly assigned, mutually uncorrelated, large in number, and sufficiently dispersed in terms of their average exposure. Instrument relevance generally holds in such settings when the exposure of individual observations tends to be concentrated in a small number of shocks, and that those shocks affect treatment. We then show how this framework is naturally generalized to settings in which shocks are only conditionally quasi-randomly assigned or exhibit some forms of mutual dependence, such as clustering, and to settings with panel data. We also consider further extensions, to SSIV regressions with varying exposure share sums, estimated shocks, and multiple endogenous variables or instruments.

3.1 Quasi-Randomly Assigned and Mutually Uncorrelated Shocks

To establish SSIV consistency, we consider a two-step data-generating process in which the shocks g_n are drawn conditional on the shock-level unobservables $\bar{\varepsilon}_n$ and exposure weights s_n. Placing assumptions on this process, rather than on the sampling properties of observations, is akin to a standard analysis of randomized treatment assignment in experimental settings (Abadie et al. 2019) and has two key advantages. First, as explained below, conventional independent or clustered sampling processes are generally inconsistent with the shift-share data structure when the shocks are considered random variables. Second, in conditioning on $\bar{\varepsilon}_n$ and s_n we place no restrictions on the dependence between the $s_{\ell n}$ and ε_ℓ, allowing shock exposure to be endogenous (i.e. $\bar{\varepsilon}_n \overset{p}{\not\to} 0$). The following result shows that such endogeneity does not pose problems for SSIV exclusion in our framework:

Prop. 3 If $\mathrm{Var}\left[g_n \mid \bar{\varepsilon}_n, s_n\right]$ and $\mathbb{E}\left[\bar{\varepsilon}_n^2 \mid s_n\right]$ are uniformly bounded and the three regularity conditions in Proposition 2 hold, then shock orthogonality is satisfied by the following conditions:

Assumption 1 *(Quasi-random shock assignment)*: $\mathbb{E}\left[g_n \mid \bar{\varepsilon}_n, s_n\right] = \mu$, for all n;

Assumption 2 *(Many uncorrelated shocks)*: $\mathbb{E}\left[\sum_{n=1}^N s_n^2\right] \to 0$ and for all n and $n' \neq n$, $\mathrm{Cov}\left[g_n, g_{n'} \mid \bar{\varepsilon}_n, \bar{\varepsilon}_{n'}, s_n, s_{n'}\right] = 0.$

Proof: See Appendix B.1.

Proposition 3 shows that the SSIV exclusion restriction holds under two substantive assumptions and two additional regularity conditions. In words, Assumption 1 states that the shocks g_n are as-good-as-randomly assigned, in that the same mean shock μ is expected across n regardless of the realization of average exposure s_n or the relevant unobservable $\bar{\varepsilon}_n$. As shown in Appendix B.1 this is enough to make the left-hand side of the orthogonality condition (6) zero in large samples, in expectation over different shock draws.[8] For it to converge to this mean in probability, a shock-level law of large numbers must hold. This is ensured by Assumption 2, which states that shocks are mutually uncorrelated given the unobservables and that the expected Herfindahl index of average exposure, $\mathbb{E}\left[\sum_{n=1}^N s_n^2\right]$, converges to zero. The latter condition implies that the number of observed shocks grows with the sample (since $\sum_{n=1}^N s_n^2 \geq 1/N$). An equivalent condition is that the largest s_n becomes vanishingly small: that is, that the largest shock importance weight vanishes asymptotically.[9] Both of these conditions, while novel for SSIV, would be standard requirements for the consistency of a conventional shock-level IV estimator with s_n weights.

[8] Appendix A.1 shows how SSIV identifies a convex average of heterogeneous treatment effects (varying potentially across both ℓ and n) under a stronger notion of as-good-as-random shock assignment and a first-stage monotonicity condition. This can be seen as generalizing both the IV identification result of Angrist et al. (2000) to shift-share instruments, as well as the reduced-form shift-share identification result in Adão et al. (2019).

[9] Goldsmith-Pinkham et al. (2019) propose a different measure of the importance of a given n, termed "Rotemberg weights." In Appendix A.4 we show the formal connection between s_n and Rotemberg weights, and that the latter do not carry the sensitivity-to-misspecification interpretation as they do in the exogenous shares view of Goldsmith-Pinkham et al. (2019). Instead, the Rotemberg weight of shock n measures the leverage of n in the equivalent shock-level IV regression from Proposition 1. Shocks may have large leverage either because of large s_n, as would be captured by the Herfindahl index, or because the shocks have a heavy-tailed distribution, which is allowed by Proposition 3.

Two comments on the nature of our framework are warranted. First, by focusing on the quasi-experimental assignment of shocks, our approach to IV consistency contrasts with those based on independent or clustered sampling across observations ℓ which are commonly used in non-experimental settings. Independent sampling of the shift-share instrument is impossible when shocks are stochastic, since the shift-share instrument $z_\ell = \sum_{n=1}^{N} s_{\ell n} g_n$ is inherently correlated across observations. Unobserved shocks ν_n may further induce complex exposure-driven dependencies in the residual ε_ℓ, precluding independence of $(z_\ell, \varepsilon_\ell)$ even conditionally on the observed g_n. Independent sampling across n in the equivalent shock-level IV regression of Proposition 1 is also untenable, since the aggregated outcome and treatment residuals $\bar{y}_n^\perp$ and $\bar{x}_n^\perp$ are constructed from a common set of observations. These dependency issues also complicate SSIV inference, as we discuss in Section 4.

Second, and relatedly, the large-L and large-N asymptotic sequence of Proposition 3 is intended to approximate the finite-sample distribution of the SSIV estimator, rather than to define an actual process for realizations of shocks or observations. In practice shift-share instruments are typically constructed from a fixed population of relevant shocks (e.g. subsidies across all industries) and often from the full population of observations (e.g. all local labor markets). Our approach takes this data structure seriously, relying on the quasi-random assignment of shocks to a given set of industries.[10]

Per Proposition 2, Assumptions 1 and 2 and the additional regularity conditions imply consistency of $\hat{\beta}$ under the usual IV relevance condition, that $\sum_{\ell=1}^{L} e_\ell z_\ell x_\ell^\perp \xrightarrow{p} \pi \neq 0$. In practice, the existence of such a first stage can be inferred from the data. To see when instrument relevance might hold with quasi-experimental shocks, consider a stylized setting without controls w_ℓ and where treatment is a share-weighted average of shock-specific components: $x_\ell = \sum_{n=1}^{N} s_{\ell n} x_{\ell n}$, where $x_{\ell n} = \pi_{\ell n} g_n + \eta_{\ell n}$ with $\pi_{\ell n} \geq \bar{\pi}$ almost surely for some fixed $\bar{\pi} > 0$. In line with Assumptions 1 and 2, suppose further that the shocks are mean-zero and mutually independent given the exposure share and weight matrices s and e and the full set of $\pi_{\ell n}$ and $\eta_{\ell n}$, with variances $\sigma_n^2 \geq \overline{\sigma}_g^2$ for some fixed $\overline{\sigma}_g^2 > 0$. Then

$$\begin{aligned}\mathbb{E}\left[\sum_{\ell=1}^{L} e_\ell z_\ell x_\ell\right] &= \mathbb{E}\left[\sum_{\ell=1}^{L} e_\ell \left(\sum_{n=1}^{N} s_{\ell n} g_n\right)\left(\sum_{n=1}^{N} s_{\ell n}(\pi_{\ell n} g_n + \eta_{\ell n})\right)\right] \\ &\geq \bar{\pi}\bar{\sigma}_g^2 \mathbb{E}\left[\sum_{\ell=1}^{L} e_\ell \sum_{n=1}^{N} s_{\ell n}^2\right],\end{aligned} \tag{7}$$

and under appropriate regularity conditions $\sum_{\ell=1}^{L} e_\ell z_\ell^\perp x_\ell$ converges in probability to this mean. In this case, SSIV relevance holds when the e_ℓ-weighted average of local exposure Herfindahl indices $\sum_{n=1}^{N} s_{\ell n}^2$ across observations does not vanish in expectation. In our running labor supply example, where $x_{\ell n}$ is industry-by-region employment growth, SSIV relevance thus arises from individual regions ℓ tending

[10] This is similar to the way Bekker (1994) uses a non-standard asymptotic sequence to analyze IV estimators with many weak instruments: "The parameter sequence is designed to make the asymptotic distribution fit the finite sample distribution better. It is completely irrelevant whether or not further sampling will lead to samples conforming to this sequence" (p. 658).

to specialize in a small number of industries n, provided subsidies have a non-vanishing effect on local industry employment.[11] Compare this condition to the Herfindahl index condition in Assumption 2, which instead states that the *average* shares of industries across locations s_n grow small. Both conditions may simultaneously hold when most regions specialize in a small number of industries, differentially across a large number of industries.

3.2 Conditional Shock Assignment and Weak Shock Dependence

While Proposition 3 establishes SSIV consistency when shocks have the same expectation across n and are mutually uncorrelated, both requirements are straightforward to relax. Here we provide extensions of Assumptions 1 and 2 that allow the shock expectation to depend on observables and for weak mutual dependence (such as clustering or serial correlation) of the residual shock variation.

As with other research designs, one may wish to assume that Assumptions 1 and 2 only hold conditionally on a vector of shock-level observables q_n (that includes a constant). For example, it may be more plausible that shocks are as-good-as-randomly assigned within a set of observed clusters $c(n) \in \{1, \ldots, C\}$ with non-random cluster-average shocks, in which case q_n collects $C-1$ cluster dummies and a constant. In general we consider the following weakened version of Assumption 1:

Assumption 3 *(Conditional quasi-random shock assignment)*: $\mathbb{E}\left[g_n \mid \bar{\varepsilon}_n, q_n, s_n\right] = q_n'\mu$, for all n.

Similarly, one may prefer to impose the mutual uncorrelatedness condition of Assumption 2 on the residual $g_n^* = g_n - q_n'\mu$, in place of g_n:

Assumption 4 *(Many uncorrelated shock residuals)*: $\mathbb{E}\left[\sum_{n=1}^N s_n^2\right] \to 0$ and for all n and $n' \neq n$, $\operatorname{Cov}\left[g_n^*, g_{n'}^* \mid \bar{\varepsilon}_n, \bar{\varepsilon}_{n'}, s_n, s_{n'}\right] = 0$.

In the shock cluster example, Assumption 4 would allow the number of clusters to remain small, each with its own random effect, as in that case a law of large numbers may apply to the within-cluster residuals g_n^* but not the original shocks g_n.

By a simple extension of the proof to Proposition 3, the orthogonality condition (6) is satisfied when these conditions replace Assumptions 1 and 2 and the residual shift-share instrument $z_\ell^* = \sum_{n=1}^N s_{\ell n} g_n^*$ replaces z_ℓ. While this instrument is infeasible, the following result shows it implicitly drives variation in SSIV regressions that control for the exposure-weighted vector of shock-level controls, $w_\ell^* = \sum_{n=1}^N s_{\ell n} q_n$:

Prop. 4 Suppose Assumptions 3 and 4 hold, $\operatorname{Var}\left[g_n^* \mid \bar{\varepsilon}_n, q_n, s_n\right]$ and $\mathbb{E}\left[\bar{\varepsilon}_n^2 \mid q_n, s_n\right]$ are uniformly bounded, and the regularity conditions in Proposition 2 hold. Then shock orthogonality (6) is satisfied provided w_ℓ^* is included in w_ℓ.

[11] Note that this precludes consideration of an asymptotic sequence where L remains finite as N grows. With L (and also $e_1, \ldots, e_L$) fixed, Assumption 2 implies $\sum_{\ell=1}^L e_\ell^2 \mathbb{E}\left[\sum_{n=1}^N s_{\ell n}^2\right] \to 0$ and thus $\operatorname{Var}\left[z_\ell\right] = \operatorname{Var}\left[\sum_{n=1}^N s_{\ell n} g_n\right] \to 0$ for each ℓ if $\operatorname{Var}\left[g_n\right]$ is bounded. If the instrument has asymptotically no variation it cannot have a first stage, unless the $\pi_{\ell n}$ grow without bound.

Proof: See Appendix B.2.

In particular, Proposition 4 shows that controlling for each observation's individual exposure to each cluster, $\sum_{n=1}^{N} s_{\ell n}\mathbf{1}\left[c\left(n\right)=c\right]$, isolates the within-cluster variation in shocks.

Even conditional on observables, mutual shock uncorrelatedness may be undesirably strong. It is, however, straightforward to further relax this assumption to allow for shock assignment processes with weak mutual dependence, such as further clustering or autocorrelation. In Appendix B.1 we prove generalizations of Proposition 3 which replace Assumption 2 with one of the following alternatives, with generalizations of Proposition 4 following analogously:

Assumption 5 *(Many uncorrelated shock clusters)*: There exists a partition of industries into clusters $c(n)$ such that $\mathbb{E}\left[\sum_{c=1}^{C} s_c^2\right] \to 0$ for $s_c = \sum_{n\colon c(n)=c} s_n$ and for all n and n' such that $c(n) \neq c(n')$, $\text{Cov}\left[g_n, g_{n'} \mid \bar{\varepsilon}_n, \bar{\varepsilon}_{n'}, s_{c(n)}, s_{c(n')}\right] = 0$;

Assumption 6 *(Many weakly-correlated shocks)*: For some sequence of numbers $B_L \geq 0$ and a fixed function $f(\cdot) \geq 0$ with $\sum_{r=1}^{\infty} f(r) < \infty$, $B_L \mathbb{E}\left[\sum_{n=1}^{N} s_n^2\right] \to 0$ and for all n and n' $\left|\text{Cov}\left[g_n, g_{n'} \mid \bar{\varepsilon}_n, \bar{\varepsilon}_{n'}, s_n, s_{n'}\right]\right| \leq B_L \cdot f\left(|n'-n|\right)$.

Here Assumption 5 relaxes Assumption 2 by allowing shocks to be grouped within mutually mean-independent clusters $c(n)$, while placing no restriction on within-cluster shock correlation. At the same time, the Herfindahl index assumption of Assumption 2 is strengthened to hold for industry clusters, with s_c denoting the average exposure of cluster c. Assumption 6 takes a different approach, allowing all nearby shocks to be mutually correlated provided their covariance is bounded by a function $B_L \cdot f\left(|n'-n|\right)$. This accommodates, for example, the case of first-order autoregressive time series with the covariance bound declining at a geometric rate, i.e. $f(r) = \delta^r$ for $\delta \in [0,1)$ and constant B_L. With B_L growing, stronger dependence of nearby shocks is also allowed (see Appendix B.1).

3.3 Panel Data

In practice, SSIV regressions are often estimated with panel data, where the outcome $y_{\ell t}$, treatment $x_{\ell t}$, controls $w_{\ell t}$, importance weights $e_{\ell t}$, exposure shares $s_{\ell n t}$, and shocks g_{nt} are additionally indexed by time periods $t = 1, \ldots, T$.[12] In such settings a time-varying instrument,

$$z_{\ell t} = \sum_{n=1}^{N} s_{\ell n t} g_{nt}, \tag{8}$$

is used, and the controls $w_{\ell t}$ may include unit- or period-specific fixed effects.

[12] Exposure shares are typically lagged and sometimes fixed in a pre-period. Our subscript t notation indicates that these shares are used to construct the instrument for period t, not that they are measured in that period. We also note that, as in our running labor supply example, the SSIV outcome, treatment, and shocks may be already measured as changes or growth rates over time in the previous "cross-sectional" discussion.

It is straightforward to map this panel case to the previous cross-sectional setting by a simple relabeling. Let $\tilde{\ell} \in \{(\ell, t) : \ell = 1, \dots, L; t = 1, \dots, T\}$ and $\tilde{n} \in \{(n, t) : n = 1, \dots, N; t = 1, \dots, T\}$, with the time-varying outcomes now indexed as $y_{\tilde{\ell}}$ and similarly for $x_{\tilde{\ell}}$, $w_{\tilde{\ell}}$, $e_{\tilde{\ell}}$, and $g_{\tilde{n}}$. Further let $\tilde{s}_{\tilde{\ell}\tilde{n}} = \tilde{s}_{(\ell,t),(n,p)} = s_{\ell n t}\mathbf{1}[t = p]$ denote the exposure of observation ℓ in period t to shock n in period p, which is by definition zero for $t \neq p$. The time-varying instrument (8) can then be rewritten $z_{\tilde{\ell}} = \sum_{\tilde{n}} \tilde{s}_{\tilde{\ell}\tilde{n}} g_{\tilde{n}}$, as in the cross-sectional case. Generalizations of the equivalency result (Proposition 1), orthogonality condition (Proposition 2), and quasi-experimental framework (Propositions 3–4) immediately follow.

Unpacking these results in the panel case delivers three new insights. First, consistency of the panel SSIV estimator is established as $LT \to \infty$, with the Herfindahl condition in Proposition 3 (that $\mathbb{E}\left[\sum_{\tilde{n}} s_{\tilde{n}}^2\right] = \mathbb{E}\left[\sum_{n=1}^{N} \sum_{t=1}^{T} s_{nt}^2\right] \to 0$) requiring $NT \to \infty$. This means that our quasi-experimental framework can be applied to settings, such as that of Nunn and Qian (2014), with relatively few shocks N (and perhaps few observations L) in the cross-section, but many time periods T. The asymptotic sequence may also well-approximate settings like those of Berman et al. (2017) and Imbert et al. (2019) where NT is large despite moderate N and T. Second, while shocks must be mutually uncorrelated across periods under the baseline Assumption 2, arbitrary clustering across periods can be accommodated by Assumption 5, provided $N \to \infty$. If N is finite but $T \to \infty$, weak serial dependence is accommodated by Assumption 6.[13]

A third set of insights concerns the role of fixed effect (FE) controls in panel SSIV regressions. As in any panel regression, unit fixed effects purge time-invariant unobservables from the residual ϵ_ℓ. Assumption 1 would thus hold if shocks are as-good-as-randomly assigned with respect to aggregated time-varying unobservables. However, when exposure shares are fixed across periods, i.e. $s_{\ell n t} \equiv s_{\ell n 0}$, unit FEs can also be understood as isolating the time-varying variation in *shocks*. This follows from Proposition 4: exposure-weighted averages of the shock FEs in q_{nt}, $w_{\ell t \bar{n}}^* = \sum_n s_{\ell n 0}\mathbf{1}[n = \bar{n}]$, are time-invariant and thus subsumed by the unit FEs. A similar argument applies to period FEs, which isolate within-period shock variation. This only relies on exposure shares adding up to one, since then one can represent period FEs in the original data as exposure-weighted averages of shock-level period FEs in the relabeled data, i.e. $\mathbf{1}[t = \bar{t}] = \sum_{(n,p)} \tilde{s}_{\tilde{\ell}(n,p)}\mathbf{1}[p = \bar{t}]$.[14]

To make these insights concrete, consider our labor supply example in a panel setting with subsidies allocated to industries in each period. Imagine that certain industries get permanent subsidy shocks that are not as-good-as-randomly assigned across industries. In that case, one may prefer to only use the changes in industry subsidies over time as identifying variation. With fixed exposure shares, one way to proceed is to include region FEs in the SSIV specification, which implicitly control for industry

[13] Proposition 4 also suggests an alternative way to handle serial correlation when the time series properties of shocks are known. For example given a first-order autoregressive process $g_{nt} = \rho_0 + \rho_1 g_{n,t-1} + g_{nt}^*$, controlling for the exposure-weighted average of past shocks $\sum_{n=1}^{N} s_{\ell n t} g_{n,t-1}$ extracts the idiosyncratic shock component g_{nt}^*.

[14] Technically these observations hold under the regularity conditions used to prove Propositions 2–4, which require the control coefficient vector γ to be consistently estimated as $LT \to \infty$, even if $w_{\ell t}$ contains an increasing number of FEs (see footnote 5). In Appendix A.5 we show how stronger notions of shock exogeneity address this incidental parameters problem.

FEs at the shock level while also removing any potential time-invariant regional characteristics that could yield omitted variables bias. Including time FEs further removes any non-random variation in the average level of subsidies in each period.

We lastly note that while fixing exposure shares in a pre-period is useful for isolating time-varying shocks, lagging shares by many periods is likely to make the shift-share instrument less predictive of treatment, thereby reducing panel SSIV power. Appendix A.6 formalizes this intuition with a panel extension to our first stage model used in Section 3.1. In such cases it may be preferable to employ other ways of isolating time-varying shocks, which allow for updated shares. For example, one may use the first-differenced specification

$$\Delta y_{\ell t} = \beta \Delta x_{\ell t} + \gamma' \Delta w_{\ell t} + \Delta \varepsilon_{\ell t}, \tag{9}$$

instrumenting $\Delta x_{\ell t}$ with $z_{\ell t,FD} = \sum_n s_{\ell n,t-1} \Delta g_{nt}$, where Δ is the first-differencing operator for both observations and shocks.[15]

3.4 Further Extensions

We conclude this section by presenting several further extensions of the quasi-experimental SSIV framework, accommodating features of shocks and shares that are often encountered in practice.

The incomplete shares problem

While we have previously assumed the sum of exposure shares is constant, in practice this $S_\ell = \sum_{n=1}^{N} s_{\ell n}$ may vary across ℓ. For example, in the labor supply setting, government output subsidies may be only introduced to manufacturing industries while the lagged manufacturing employment shares of $s_{\ell n}$ may be measured relative to total employment in region ℓ. In this case S_ℓ corresponds to the lagged total share of manufacturing employment in region ℓ.

Our framework highlights a problem with such "incomplete share" settings: even if Assumptions 1 and 2 hold, the SSIV estimator will in general leverage non-experimental variation in S_ℓ, in addition to quasi-experimental variation in shocks. To see this formally, note that one can always rewrite the shift-share instrument with the "missing" (e.g., non-manufacturing) shock included to return to the complete shares setting:

$$z_\ell = s_{\ell 0} g_0 + \sum_{n=1}^{N} s_{\ell n} g_n, \tag{10}$$

[15] There is another argument for fixing the shares in a pre-period that applies when the current shares are affected by lagged shocks in a way that is correlated with unobservables $\varepsilon_{\ell t}$. In the labor supply example suppose local labor markets vary in flexibility, with stronger reallocation of employment to industries with bigger subsidies in flexible markets. If subsides are random but persistent, more subsidized industries will be increasingly concentrated in regions with flexible labor markets and Assumption 1 will be violated if such flexibility is correlated with $\varepsilon_{\ell t}$. This concern is distinct from that in Jaeger et al. (2018) who focus on the endogeneity of shares to the lagged residuals, rather than shocks, in a setting closer to Goldsmith-Pinkham et al. (2019). Jaeger et al. also point out another issue relevant to panel SSIV, that the outcome may respond to both current and lagged shocks; we return to this issue in Section 3.4.

where $g_0 = 0$ and $s_{\ell 0} = 1 - S_\ell$, such that $\sum_{n=0}^{N} s_{\ell n} = 1$ for all ℓ. The previous quasi-experimental framework then applies to this expanded set of shocks $g_0, \ldots, g_N$. Since $g_0 = 0$, Proposition 3 requires in this case that $\mathbb{E}\left[g_n \mid s_n, \bar{\varepsilon}_n\right] = 0$ for $n > 0$ as well; that is, that the expected shock to each manufacturing industry is the same as the "missing" non-manufacturing shock of zero. Otherwise, even if the manufacturing shocks are random, regions with higher manufacturing shares S_ℓ will tend to have systematically different values of the instrument z_ℓ, leading to bias when these regions also have different unobservables.[16]

Cast in this way, the incomplete shares problem has a natural solution via Assumption 3. Namely, one can allow the missing and non-missing shocks to have endogenously different means by conditioning on the indicator $\mathbf{1}[n > 0]$ in the q_n vector. By Proposition 4, the SSIV estimator allows for such conditional quasi-random assignment when the control vector w_ℓ contains the exposure-weighted average of $\mathbf{1}[n > 0]$, which is here $\sum_{n=0}^{N} s_{\ell n} \mathbf{1}\left[n > 0\right] = S_\ell$. Thus, in the labor supply example, quasi-experimental variation in manufacturing shocks is isolated in regressions with incomplete shares provided one controls for a region's lagged manufacturing share S_ℓ.

Two further points on incomplete shares are worth highlighting. First, allowing the observed shock mean to depend on other observables will tend to involve controlling for share-weighted averages of these controls interacted with the indicator $\mathbf{1}[n > 0]$. For example with period indicators in q_{nt}, within-period shock variation is isolated by controlling for sums of period-specific exposure $S_{\ell t}$, interacted with period indicators (i.e. $\sum_{n=0}^{N} s_{\ell nt} \mathbf{1}[n > 0]\mathbf{1}[t = \bar{t}] = S_{\ell t}\mathbf{1}[t = \bar{t}]$). Second, by effectively "dummying out" the missing industry, SSIV regressions that control for S_ℓ require a weaker Herfindahl condition: $\mathbb{E}\left[\sum_{n=1}^{N} s_n^2\right] \to 0$, allowing the non-manufacturing industry share s_0 to stay large.

Shift-share designs with estimated shocks

In some shift-share designs, the shocks are equilibrium objects that can be difficult to view as being quasi-randomly assigned. For example, in the canonical Bartik (1991) estimation of the regional labor supply elasticity, the shocks are national industry employment growth rates. Such growth reflects labor demand shifters, which one may be willing to assume are as-good-as-randomly assigned across industries. However industry growth also aggregates regional labor supply shocks that directly enter the residual ε_ℓ. Here we show how the quasi-experimental SSIV framework can still apply in such cases, by viewing the g_n as noisy estimates of some latent true shocks g_n^* (labor demand shifters, in the Bartik (1991) example) which satisfy Assumption 1. We establish the conditions on estimation noise (aggregated labor supply shocks, in Bartik (1991)) such that a feasible shift-share instrument estimator, perhaps involving a leave-one-out correction as in Autor and Duggan (2003), is

[16] Formally, if Assumptions 1 and 2 hold for all $n > 0$ we have from the proof to Proposition 3 that $\sum_{n=1}^{N} s_n g_n \bar{\varepsilon}_n = \sum_{n=0}^{N} s_n g_n \bar{\varepsilon}_n = \mathbb{E}\left[\sum_{n=0}^{N} s_n (g_n - \mu)\bar{\varepsilon}_n\right] + o_p(1) = -\mu \mathbb{E}\left[s_0 \bar{\varepsilon}_0\right] + o_p(1)$. If $\mu \neq 0$ and the missing industry share is large ($s_0 \overset{p}{\not\to} 0$) this can only converge to zero when $\mathbb{E}\left[s_0 \bar{\varepsilon}_0\right] = \mathbb{E}\left[\sum_{\ell=1}^{L} e_\ell s_{\ell 0} \varepsilon_\ell\right]$ does, i.e. when S_ℓ is exogenous.

asymptotically valid.

We leave a more general treatment of this issue to Appendix A.7 and for concreteness focus on the Bartik (1991) example here. The industry growth rates g_n can be written as weighted averages of the growth of each industry in each region: $g_n = \sum_{\ell=1}^{L} \omega_{\ell n} g_{\ell n}$, where the weights $\omega_{\ell n}$ are the lagged share of industry employment located in region ℓ, with $\sum_{\ell=1}^{L} \omega_{\ell n} = 1$ for each n. In a standard model of regional labor markets, $g_{\ell n}$ includes (to first-order approximation) an industry labor demand shock g_n^* and a term that is proportional to the regional supply shock ε_ℓ.[17] We suppose that the demand shocks are as-good-as-randomly assigned across industries, such that the infeasible SSIV estimator which uses $z_\ell^* = \sum_{n=1}^{N} s_{\ell n} g_n^*$ as an instrument satisfies our quasi-experimental framework. The asymptotic bias of the feasible SSIV estimator which uses $z_\ell = \sum_{n=1}^{N} s_{\ell n} g_n$ then depends on the large-sample covariance between the labor supply shocks ε_ℓ and an aggregate of the supply shock "estimation error,"

$$\psi_\ell = z_\ell - z_\ell^* \propto \sum_{n=1}^{N} s_{\ell n} \sum_{\ell'=1}^{L} \omega_{\ell' n} \varepsilon_{\ell'}. \tag{11}$$

Two insights follow from considering the bias term $\sum_{\ell=1}^{L} e_\ell \psi_\ell \varepsilon_\ell$. First, part of the covariance between ψ_ℓ and ε_ℓ is mechanical, since ε_ℓ enters ψ_ℓ. In fact, if supply shocks are spatially uncorrelated this is the only source of bias from using z_ℓ rather than z_ℓ^* as an instrument. This motivates the use of a leave-one-out (LOO) shock estimator, $g_{n,-\ell} = \sum_{\ell' \neq \ell} \omega_{\ell' n} g_{\ell' n} / \sum_{\ell' \neq \ell} \omega_{\ell' n}$, and the feasible instrument $z_\ell^{LOO} = \sum_{n=1}^{N} s_{\ell n} g_{n,-\ell}$ to remove this mechanical covariance.[18] Conversely, if the regional supply shocks ε_ℓ are spatially correlated a LOO adjustment may not be sufficient to eliminate mechanical bias in the feasible SSIV instrument, though more restrictive split-sample methods (e.g. those estimating shocks from distant regions) may suffice.

Second, in settings where there are many regions contributing to each shock estimate even the mechanical part of $\sum_{\ell=1}^{L} e_\ell \psi_\ell \varepsilon_\ell$ may be ignorable, such that the conventional non-LOO shift-share instrument z_ℓ (which, unlike z_ℓ^{LOO}, has a convenient shock-level representation per Proposition 1) is asymptotically valid when z_ℓ^{LOO} is.[19] In Appendix A.7, we derive a heuristic for this case, under the assumption of spatially-independent supply shocks. In a special case when each region is specialized in a single industry and there are no importance weights, the key condition is $L/N \to \infty$, or that the average number of regions specializing in the typical industry is large. With incomplete specialization or weights, the corresponding condition requires the typical industry to be located in a much larger number of regions than the number of industries that a typical region specializes in.

[17] Appendix A.8 presents such a model, showing that $g_{\ell n}$ also depends on the regional average of g_n^* (via local general equilibrium effects) and on idiosyncratic region-specific demand shocks. Both of these are uncorrelated with the error term in the model and thus do not lead to violations of Assumption 1; we abstract away from this detail here.

[18] This problem of mechanical bias is similar to that of two-stage least squares with many instruments (Bound et al. 1995), and the solution is similar to the jackknife instrumental variable estimate approach of Angrist et al. (1999).

[19] Adão et al. (2019) derive the corrected standard errors for LOO SSIV and find that they are in practice very close to the non-LOO ones, in which case the SSIV standard errors we derive in the next section are approximately valid even when the LOO correction is used.

To illustrate the preceding points in the data, Appendix A.7 replicates the setting of Bartik (1991) with and without a LOO estimator, using data from Goldsmith-Pinkham et al. (2019). We find that in practice the LOO correction does not matter for the SSIV estimate, consistent with the findings of Goldsmith-Pinkham et al. (2019) and Adão et al. (2019), and especially so when the regression is estimated without regional employment weights. Our framework provides a explanation for this: the heuristic statistic we derive is much larger without importance weights. These findings imply that if, in the canonical Bartik (1991) setting, one is willing to assume quasi-random assignment of the underlying industry demand shocks and that the regional supply shocks are spatially-uncorrelated, one can interpret the uncorrected SSIV estimator as leveraging demand variation in large samples, as some of the literature has done (e.g. Suárez and Zidar (2016)).

Multiple shocks and treatments

In some shift-share designs one may have access to multiple sets of shocks satisfying Assumptions 1 and 2 or their extensions. For example while Autor et al. (2013) construct an instrument from average Chinese import growth across eight non-U.S. countries, in principle the industry shocks from each individual country may be each thought to be as-good-as-randomly assigned. In other settings multiple endogenous variables may be required for the SSIV exclusion restriction to plausibly hold. For example Jaeger et al. (2018) show that when local labor markets respond dynamically to immigrant inflows, it may be necessary to instrument both the current and lagged immigrant growth rate with current and lagged shift-share instruments. Another example is provided by Bombardini and Li (2019), who estimate the reduced-form effects of two shocks: the regional growth of all exports and the regional growth of exports in pollution-intensive sectors. Both are shift-share variables based on the same regional employment shares across industries but different shocks: an overall industry export shock and the overall shock interacted with industry pollution intensity.[20]

In Appendix A.9 we show how these extensions fit into our quasi-experimental framework. The key insight is that SSIV regressions with multiple instruments – with and without multiple endogenous variables – again have an equivalent representation as particular shock-level IV estimators provided the exposure shares used to construct the instruments are the same. This immediately implies extensions of the foregoing results that establish consistency for just-identified SSIV regressions with multiple instruments, such as the dynamic adjustment case of Jaeger et al. (2018). In overidentified settings, the appendix derives new shock-level IV estimators that optimally combine the quasi-experimental variation and permit omnibus tests of the identifying assumptions, via the generalized method of moments theory of Hansen (1982) and inference results discussed in the next section. For example

[20]The instruments here are $\sum_{n=1}^{N} s_{\ell n} g_n$ and $\sum_{n=1}^{N} s_{\ell n} g_n q_n$ where q_n denotes industry n's pollution intensity. Our framework applies in this case even if q_n is not randomly assigned: as long as the export shock g_n satisfies an appropriate version of Assumption 1, $\mathbb{E}\left[g_n \mid q_n, \bar{\varepsilon}_n, s_n\right] = \mu$, the interacted shock satisfies $\mathbb{E}\left[g_n q_n \mid q_n, \bar{\varepsilon}_n, s_n\right] = \mu q_n$, i.e. Assumption 3. The natural extension of Proposition 4 to multiple instruments applies as long as $\sum_{n=1}^{N} s_{\ell n} q_n$, a measure of pollution intensity of the region, is controlled for, as Bombardini and Li (2019) do in some specifications.

when shocks are homoskedastic, a two-stage least squares regression of $\bar{y}_n^\perp$ on $\bar{x}_n^\perp$, weighted by s_n and instrumented by multiple shocks $g_{1n}, \ldots, g_{Jn}$ yields an efficient estimate of β.

4 Shock-Level Inference and Testing

A shock-level view of SSIV also brings new insights to inference and testing. In this section we first show how a problem with conventional inference in SSIV settings, first studied by Adão et al. (2019), has a convenient solution based on the equivalence result in Proposition 1. In particular, we show that conventional standard error calculations are asymptotically valid when quasi-experimental SSIV coefficients are estimated at the level of identifying variation (shocks). We then discuss how other novel shock-level procedures can be used to assess first-stage relevance and to implement valid falsification tests of the SSIV exclusion restriction. Lastly, we summarize a variety of Monte-Carlo simulations illustrating the finite-sample properties of SSIV and relating them to conventional shock-level analyses.

4.1 Exposure-Robust Standard Errors

As with consistency, SSIV inference is complicated by the fact that the observed shocks g_n and any unobserved shocks ν_n induce dependencies in the instrument z_ℓ and residual ε_ℓ across observations with similar exposure shares, making it implausible to treat $(z_\ell, \varepsilon_\ell)$ as independent draws. This problem can be understood as an extension of the standard clustering concern (Moulton 1986), in which the instrument and structural residual are correlated across observations within predetermined clusters, with the additional complication that in SSIV every pair of observations with overlapping shares may have correlated $(z_\ell, \varepsilon_\ell)$. Adão et al. (2019) study and develop solutions to this problem in a setting that builds on our own quasi-experimental framework. Their analysis shows both that such "exposure clustering" is likely to arise in economic models motivating SSIV regressions and that it can lead to misleading conventional inference procedures in practice. In Monte-Carlo simulations, for example, Adão et al. (2019) show that tests based on conventional standard errors with nominal 5% significance can reject a true null in 55% of placebo shock realizations.

Our equivalence result motivates a novel and convenient solution to SSIV exposure-based clustering. Namely, we next show that by estimating SSIV coefficients with an equivalent shock-level IV regression one obtains standard errors that are *exposure-robust* – i.e. asymptotically valid in the primary framework of Adão et al. (2019). This framework assumes that (as in Proposition 4) the control vector can be partitioned as $w_\ell = [w_\ell^{*\prime}, u_\ell^\prime]^\prime$, where $w_\ell^* = \sum_{n=1}^N s_{\ell n} q_n$ for some q_n capturing all sources of shock-level confounding, while the other controls in u_ℓ are included only to potentially increase the efficiency of the estimator and not asymptotically correlated with z_ℓ. To formalize this we follow Adão et al. (2019) in writing the first stage as $x_\ell = \sum_n s_{\ell n} \pi_{\ell n} g_n + \eta_\ell$ and considering the following strengthened version of Assumption 3:

Assumption 7 *(Strong conditional quasi-random shock assignment)*: For all n,

$$\mathbb{E}\left[g_n \mid \{q_{n'}\}_{n'}, \{u_\ell, \epsilon_\ell, \eta_\ell, \{s_{\ell n'}, \pi_{\ell n'}\}_{n'}, e_\ell\}_\ell\right] = q_n' \mu.$$

We further adopt a strengthened version of Assumption 4 (Assumption A1 in Appendix B.3), which assumes fully-independent shock residuals and a stronger condition on concentration of average exposure s_n, as well as additional regularity conditions (Assumption A2 in Appendix B.3) consistent with Adão et al. (2019). We then have the following result:

Prop. 5 Consider s_n-weighted IV estimation of the second stage equation

$$\bar{y}_n^\perp = \alpha + \beta \bar{x}_n^\perp + q_n' \tau + \bar{\varepsilon}_n^\perp \tag{12}$$

where $w_\ell^* = \sum_{n=1}^N s_{\ell n} q_n$ is included in the control vector w_ℓ used to compute $\bar{y}_n^\perp$ and $\bar{x}_n^\perp$, and $\bar{x}_n^\perp$ is instrumented by g_n. The IV estimate of β is numerically equivalent to the SSIV estimate $\hat{\beta}$. Furthermore when Assumptions 7, A1, and A2 hold, and $\sum_{\ell=1}^L e_\ell x_\ell^\perp z_\ell \xrightarrow{p} \pi$ for $\pi \neq 0$, the conventional heteroskedasticity-robust standard error for $\hat{\beta}$ yields asymptotically-valid confidence intervals for β.

***Proof*:** See Appendix B.3.

Proposition 5 provides a straightforward way to compute standard errors that are valid regardless of the correlation structure of the error term ε_ℓ; rather, its validity derives from the properties of the shocks, in line with our quasi-experimental framework. Equation (12) extends the previous shock-level estimating equation (4) to include a vector of controls q_n which, as in Proposition 4, are included in the SSIV control vector w_ℓ as exposure-weighted averages. The first result in Proposition 5 is that the addition of these controls does not alter the equivalence of $\hat{\beta}$. The second result establishes conditions under which conventional shock-level standard error calculations from estimation of (12) yield valid asymptotic inference. Appendix B.3 further shows that absent any controls, i.e. with $w_\ell = q_n = 1$, the shock-level robust standard error for β is numerically equivalent to the standard error formula that Adão et al. (2019) propose. Outside this case, the appendix shows that standard errors estimated using our procedure are likely to be smaller than those of Adão et al. (2019) in finite samples.

To understand Proposition 5, it is useful to relate it to a conventional solution to the standard clustering problem. The clustering environment can be viewed as a special case of shift-share IV in which the exposure shares are binary: $s_{\ell,n(\ell)} = 1$ for some $n(\ell) \in \{1, \ldots, N\}$, with $s_{\ell n} = 0$ otherwise, for each ℓ. In this case $z_\ell = g_{n(\ell)}$, such that the shift-share instrument is constant among observations with the same $n(\ell)$. The usual clustering concern would then be that ε_ℓ is also correlated within the $n(\ell)$ groupings. One solution is to estimate a grouped-data regression at the level of identifying variation (Angrist and Pischke 2008, p. 313). Proposition 5 generalizes this solution by running a regression at the level of quasi-experimental shocks. Indeed, in the binary shares case all variables $\bar{v}_n$ in equation (12) correspond to importance-weighted group averages of the corresponding v_ℓ.

This shock-level approach to obtaining valid SSIV standard errors has three practical advantages. First, it can be performed with standard statistical software packages given an simple initial transformation of the data (i.e. to obtain $\bar{y}_n^\perp$, $\bar{x}_n^\perp$, and s_n), for which we have released a Stata package *ssaggregate* (see footnote 3). Second, it is readily extended to settings where shocks are clustered (a case also considered by Adão et al. (2019)) or autoregressive, implying Assumptions 5 and 6, respectively; conventional cluster-robust or heteroskedastic-and-autocorrelation-consistent (HAC) standard error calculations applied to equation (12) are then valid. Third, Appendix B.3 shows that the shock-level inference approach continues to work in some cases where the original standard error calculation procedure of Adão et al. (2019) (which involves regressing $z_\ell^\perp$ on the vector of shares) fails: when $N > L$ or when some exposure shares are collinear.[21]

Our shock-level equivalence also provides a convenient implementation for alternative inference procedures which may have superior finite-sample performance. Adão et al. (2019) show, in particular, how standard errors that impose a given null hypothesis $\beta = \beta_0$ in estimating the residual ε_ℓ can generate confidence intervals with better coverage in situations with few shocks (and a similar argument can be made in the case of shocks with a heavy-tailed distribution).[22] Building on Proposition 5, such confidence intervals can be constructed in the same way as in any regular shock-level IV regression. To test $\beta = \beta_0$, one regresses $\bar{y}_n^\perp - \beta_0 \bar{x}_n^\perp$ on the shocks g_n (weighting by s_n and including any relevant shock-level controls q_n) and uses a null-imposed residual variance estimate. This procedure corresponds to the standard shock-level Lagrange multiplier test for $\beta = \beta_0$ that can be implemented by standard statistical software.[23] The confidence interval for β is constructed by collecting all candidate β_0 that are not rejected.

4.2 Falsification and Relevance Tests

Our Proposition 5 also provides a practical way to perform valid regression-based tests of the SSIV exclusion restriction (i.e. falsification tests) and first-stage relevance. In the quasi-experimental SSIV framework, such tests also require computing exposure-robust standard errors, and shock-level inference procedures allow for convenient implementation via equivalent shock-level IV regressions.

As usual, the SSIV exclusion restriction cannot be tested directly. However, indirect falsification

[21] This is an empirically relevant issue: for instance, employment shares of some industries are collinear in the Autor et al. (2013) setting. To give one example, SIC code 2068 "Salted and roasted nuts and seeds" was part of code 2065 "Candy and other confectionery products" until the 1987 revision of the classification; the rest of code 2065 we reassigned to code 2064. Therefore, when using 1980 employment shares to construct the shift-share instrument for the 1990s, Autor et al. (2013) allocate employment between 2064 and 2068 codes proportionately.

[22] As explained by Adão et al. (2019), the problem that this "AKM0" confidence interval addresses generalizes the standard finite-sample bias of cluster-robust standard errors with few clusters (Cameron and Miller 2015). With few or heavy-tailed shocks, estimates of the residual variance will tend to be biased downwards, leading to undercoverage of confidence intervals based on standard errors that do not impose the null.

[23] For example in Stata one can use the *ivreg2* overidentification test statistic from regressing $\bar{y}_n^\perp - \beta_0 \bar{x}_n^\perp$ on q_n with no endogenous variables and with g_n specified as the instrument (again with s_n weights).

tests may be conducted given an observed variable r_ℓ thought to proxy for the structural error ε_ℓ. Namely one may test whether r_ℓ is uncorrelated with the shift-share instrument z_ℓ, while controlling for w_ℓ. Examples of such an r_ℓ may include a baseline characteristic realized prior to the shocks or a lagged observation of the outcome y_ℓ (resulting in what is often called a "pre-trend" test). To interpret the magnitude of the reduced form falsification regression coefficient, researchers may also scale it by the first stage regression of x_ℓ, yielding a placebo SSIV coefficient.

As with any SSIV regression, inference for such falsification tests must account for the exposure-induced correlation of z_ℓ across observations with similar exposure profiles; the insights of Adão et al. (2019) and the previous section apply directly to this case. For exposure-robust inference on a placebo e_ℓ-weighted SSIV regression of r_ℓ on x_ℓ, instrumented by z_ℓ, one may use the conventional standard errors from an s_n-weighted regression of $\bar{r}_n^\perp$ on $\bar{x}_n^\perp$, instrumenting by g_n and controlling for any shock-level covariates q_n. Similarly, for valid inference from a reduced form regression of r_ℓ on z_ℓ, one may use the conventional standard errors from an IV regression of $\bar{r}_n^\perp$ on $\bar{z}_n^\perp$, with the same instrument, weights, and controls (see footnote 6).[24] If a researcher observes a shock-level confounder r_n, they can construct its observation-level average $r_\ell = \sum_{n=1}^N s_{\ell n} r_n$ and perform a similar test.

Unlike exclusion, the SSIV relevance condition can be evaluated directly, via OLS regressions of x_ℓ on z_ℓ that control for w_ℓ. For exposure-robust inference on this OLS coefficient, one may again use an equivalent shock-level IV regression: of $\bar{x}_n^\perp$ on $\bar{z}_n^\perp$, instrumenting by g_n, weighting by s_n, and controlling for q_n. The F-statistic, which is a common heuristic for the strength of the first-stage relationship, is then obtained as a squared coefficient t-statistic. We generalize this result to the case of multiple shift-share instruments in Appendix A.9 by detailing the appropriate construction of the "effective" first-stage F-statistic of Montiel Olea and Pflueger (2013), again based on the equivalent shock-level IV regression.

4.3 Monte-Carlo Simulations

Though the exposure-robust standard errors obtained from estimating equation (12) are asymptotically valid, it is useful to verify that they offer appropriate coverage with a finite number of observations and shocks. Of interest especially is whether the finite-sample performance of the equivalent regression (12) is comparable to that of more conventional shock-level IV regressions, in which the outcome and instrument are not aggregated from a common set of y_ℓ and x_ℓ.

In Appendix A.10 we provide Monte-Carlo evidence suggesting that the finite-sample properties of

[24] One might also consider a simpler shock-level OLS regression of $\bar{r}_n^\perp$ on g_n weighted by s_n and controlling for q_n (i.e. the reduced form of the proposed IV regression). This produces a coefficient that typically cannot be generated from the original observations of (r_ℓ, z_ℓ, w_ℓ). Moreover, the power of the shock-level OLS balance test is likely to be lower than the proposed IV: the robust Wald statistic for both tests has the same form, $\left(\sum_{n=1}^N s_n \bar{r}_n^\perp \hat{g}_n\right)^2 / \sum s_n^2 \hat{\kappa}_n^2 \hat{g}_n^2$, where $\hat{g}_n$ is the residual from an auxiliary s_n-weighted projection of g_n on q_n and the only difference is in the $\bar{r}_n^\perp$ residuals, $\hat{\kappa}_n$. Under the alternative model of $r_\ell = \alpha_0 + \alpha_1 z_\ell + \kappa_\ell$, with $\alpha_1 \neq 0$, the variance of these $\hat{\kappa}_n$ is likely to be smaller in the correctly specified IV balance test than the OLS regression, leading to a likely higher value of the test statistic.

SSIV and conventional shock-level IV regressions are similar, and that the variance of SSIV estimators is well-approximated by our exposure-robust standard errors. Specifically, we compare the finite-sample behavior of two estimators which both measure the effects of import competition with China on U.S. employment and use the same industry variation to construct the instruments. The first estimator is an SSIV and takes regional growth of manufacturing employment as the outcome (following Autor et al. (2013)) while the second is a conventional IV where the outcome is industry employment growth (similar to Acemoglu et al. (2016)). While finite sample properties of SSIV are not always perfect, e.g. there is moderate overrejection with heavy-tailed shocks, traditional industry-level IV suffers from the same problems. Our results further indicate that the conventional rule of thumb to detect weak instruments – the Montiel Olea and Pflueger (2013) first-state F-statistic – applies equally well to both the SSIV and the traditional IV estimators, when computed for SSIV as describe in Appendix A.9. Finally, our results also show that with the Herfindahl concentration index $\sum_{n=1}^{N} s_n^2$ as high as 1/20 (i.e. with the "effective sample size" of the shock-level analyses as low as 20), the asymptotic approximation for the SSIV estimator is still reasonable – with rejection rates in the vicinity of 7% for tests with a nominal size of 5%, despite heavy tails of the shock distribution. Together, these results indicate that a researcher who is comfortable with the finite-sample performance of a shock-level analysis with some set of g_n should also be comfortable using such shocks in SSIV, provided there is sufficient variation in exposure shares to yield a strong SSIV first stage.[25]

5 Application

We next apply our theoretical framework to the setting of Autor et al. (2013; hereafter abbreviated as ADH). We first describe this setting and then use it to illustrate the tools and lessons that emerge from viewing it as leveraging a quasi-experimental shift-share research design.

5.1 Setting

ADH use a shift-share IV to estimate the causal effect of rising import penetration from China on U.S. local labor markets. They do so with a repeated cross section of 722 commuting zones ℓ and 397 four-digit SIC manufacturing industries n over two periods t, 1990-2000 and 2000-2007. In these years U.S. commuting zones were exposed to a dramatic rise in import penetration from China, a historic change in trade patterns commonly referred to as the "China shock." Variation in exposure to this change across commuting zones results from the fact that different areas were initially specialized in different industries which saw different changes in the aggregate U.S. growth of Chinese imports.

[25] Naturally, these simulation results may be specific to the data-generating process we consider here, modeled after the "China shock" setting of Autor et al. (2013). In practice, we recommend that researchers perform similar simulations based on their data if they are concerned with the quality of asymptotic approximation—a suggestion that of course applies to conventional shock-level IV analyses as well.

ADH combine import changes across industries in eight comparable developed economies (as shocks) with lagged industry employment (as exposure shares) to construct their shift-share instrument.

To apply our framework to this setting we focus on ADH's primary outcome of the change in total manufacturing employment as a percentage of working-age population during period t in location ℓ, which we write as $y_{\ell t}$. The treatment variable $x_{\ell t}$ measures local exposure to the growth of imports from China in \$1,000 per worker. The vector of controls $w_{\ell t}$, which comes from the preferred specification of ADH (Column 6 of their Table 3), contains start-of-period measures of labor force demographics, period fixed effects, Census region fixed effects, and the start-of-period total manufacturing share to which we return below. The shift-share instrument is $z_{\ell t} = \sum_{n=1}^{397} s_{\ell nt} g_{nt}$, where $s_{\ell nt}$ is the share of manufacturing industry n in total employment in location ℓ, measured a decade before each period t begins, and g_{nt} is industry n's growth of imports from China in the eight comparable economies over period t (also expressed in \$1,000 per U.S. worker).[26] Importantly, the sum of lagged manufacturing shares across industries ($S_{\ell t} = \sum_{n=1}^{397} s_{\ell nt}$) is not constant across locations and periods, placing the ADH instrument in the "incomplete shares" class discussed in Section 3.4. All regressions are weighted by $e_{\ell t}$, which measures the start-of-period population of the commuting zone, and all variables are measured in ten-year equivalents.

To see how the ADH instrument can be viewed as leveraging quasi-experimental shocks, it is useful to imagine an idealized experiment generating random variation in the growth of imports from China across industries. One could imagine, for example, random variation in industry-specific productivities in China affecting import growth in both the U.S. and in comparable economies. If we observed and used these productivity changes directly as g_{nt}, the resulting SSIV exclusion restriction may be satisfied by our Assumptions 1 and 2. This would require idiosyncratic productivity shocks across many industries, with small average exposure to each shock across commuting zones. Weaker versions of this experimental ideal, in which productivity shocks can be partly predicted by industry observables and are only weakly dependent across industries, are accommodated by the extensions discussed in Section 3.

In practice, industry-specific productivity shocks in China are not directly observed, and the observed changes in trade patterns between the U.S. and China also depend on changes in U.S. supply and demand conditions across industries. This raises potential concerns over omitted variables bias, since U.S. supply and demand shocks may have direct effects on employment dynamics across

[26] To be precise, local exposure to the growth of imports from China is constructed for period t as $x_{\ell t} = \sum_{n=1}^{397} s_{\ell nt}^{\text{current}} g_{nt}^{\text{US}}$. Here $g_{nt}^{\text{US}} = \frac{\Delta M_{nt}^{\text{US}}}{E_{nt}^{\text{current}}}$ is the growth of U.S. imports from China in thousands of dollars ($\Delta M_{nt}^{\text{US}}$) divided by the industry employment in the U.S. at the beginning of the current period (E_{nt}^{current}) and $s_{\ell nt}^{\text{current}}$ are local employment shares, also measured at the beginning of the period. The instrument, in contrast, is constructed as $z_{\ell t} = \sum_{n=1}^{N} s_{\ell nt} g_{nt}$ with $g_{nt} = \frac{\Delta M_{nt}^{8 \text{ countries}}}{E_{nt}}$, where $\Delta M_{nt}^{8 \text{ countries}}$ measures the growth of imports from China in eight comparable economies (in thousands of U.S. dollars) and both local employment shares $s_{\ell nt}$ and U.S. employment E_{nt} are lagged by 10 years. The eight countries are Australia, Denmark, Finland, Germany, Japan, New Zealand, Spain, and Switzerland. Note that Autor et al. (2013) express the same instrument differently, based on employment shares relative to the industry total, rather than the regional total. Our way of writing $z_{\ell t}$ aims to clearly separate the exposure shares from the industry shocks, highlighting the shift-share structure of the instrument.

U.S. regions. To address this challenge, ADH define g_{nt} as the observed changes in trade patterns between China and a group of developed countries excluding the United States. Such variation reflects Chinese productivity shocks as well as supply and demand shocks in these other countries; in this way, the ADH strategy eliminates bias from shocks that are specific to the United States.

5.2 Properties of Industry Shocks and Exposure Shares

Our quasi-experimental view of the ADH research design places particular emphasis on the variation in Chinese import growth rates g_{nt} and their average exposure s_{nt} across U.S. commuting zones. With few or insufficiently-variable shocks, or highly concentrated shocks exposure, the large-N asymptotic approximation developed in Section 3.1 is unlikely to be a useful tool for characterizing the finite-sample behavior of the SSIV estimator. Moreover if the mean of shocks clearly varies by observables, such as time periods or industry groups, controlling for these variables may be useful to avoid omitted variables bias. We thus first summarize the distribution of g_{nt}, as well as the industry-level weights from our equivalence result, $s_{nt} \propto \sum_{\ell=1}^{722} e_{\ell t} s_{\ell n t}$ (normalized to add up to one in the entire sample), to gauge the plausibility of this framework. We additionally summarize the shift-share instrument $z_{\ell t}$ itself, to show how variation at the industry level translates to variation in predicted Chinese import growth across commuting zones, a necessary requirement of SSIV relevance.

In summarizing industry-level variation it is instructive to recall that the ADH instrument is constructed with "incomplete" manufacturing shares. Per the discussion in Section 3.4, this means that absent any regression controls the SSIV estimator uses variation not only in manufacturing industry shocks but also implicitly the variation in the 10-year lagged total manufacturing share $S_{\ell t}$ across commuting zones and periods. In practice, ADH control for the start-of-period manufacturing share, which is highly – though not perfectly – correlated with $S_{\ell t}$. We thus summarize the ADH shocks both with and without the "missing industry" shock $g_{0t} = 0$, which here represents the lack of a "China shock" in service (i.e. non-manufacturing) industries.

Table 1 reports summary statistics for the ADH shocks g_{nt} computed with importance weights s_{nt} and characterizes these weights.[27] Column 1 includes the "missing" service industry shock of zero in each period. It is evident that with this shock the distribution of g_{nt} is unconventional: for example, its interquartile range is zero. This is because the service industry accounts for a large fraction of total employment (s_{0t} is 71.9% of the period total in the 1990s and 79.5% in the 2000s). As a result we see a high concentration of industry exposure, as measured by the inverse of its Herfindahl index (HHI), $1/\sum_{n,t} s_{nt}^2$, which intuitively corresponds to the "effective sample size" and plays a key role in our Assumption 2. With the "missing" service industry included, the effective sample size is only 3.5. For an HHI computed at the level of three-digit industry codes $\sum_c s_c^2$, where s_c aggregates

[27] Note that s_{nt} would be proportional to lagged industry employment if the ADH regression weights $e_{\ell t}$ were lagged regional employment. ADH however use a slightly different $e_{\ell t}$: the start-of-period commuting zone population.

exposure across the two periods and industries within the same 3-digit group c, it is even lower, at 1.7. This suggests even less industry-level variation is available when shocks are allowed to be serially correlated or clustered by groups. Furthermore, the mean of manufacturing shocks is significantly different from the zero shock of the missing service industry.[28] Together, these analyses suggest that the service industry should be excluded from the identifying variation, because it is likely to violate both Assumption 1 ($\mathbb{E}\left[g_{nt} \mid \bar{\varepsilon}_{nt}, s_{nt}\right] \neq g_{0t} = 0$) and Assumption 2 ($\sum_{n,t} s_{nt}^2$ is not close to zero).

Column 2 of Table 1 therefore summarizes the sample with the service industry excluded. Within manufacturing, the average shock is 7.37, with a standard deviation of 20.93 and an interquartile range of 6.61. The inverse HHI of exposure shares is now relatively high, 191.6 across industry-by-period cells and 58.4 when exposure is aggregated by SIC3 group. The largest shock weights in this column are only 3.4% across industry-by-periods and 6.5% across SIC3 groups. This suggests a sizable degree of variation at the industry level, consistent with Assumption 2.

With Chinese import growth shocks measured in two sequential periods, differences in expected shocks across periods present an obvious threat to the simplest assumption of quasi-experimental shock assignment (Assumption 1) that would typically be addressed by including period fixed effects in the industry-level analysis. Column 3 of Table 1 confirms that even within periods there is sizable residual shock variation to implement the SSIV version of this strategy. The standard deviation and interquartile range of shock residuals (obtained from regressing shocks on period fixed effects with s_{nt} weights) are only somewhat smaller than in Column 2, despite the higher mean shock in the later period, at 12.6 versus 3.6. Motivated by the controls used in an industry-level analysis of similar shocks in Acemoglu et al. (2016), in Column 4 we further explore residual variation from regressing shocks on 10 broad industrial sector fixed effects interacted with period fixed effects.[29] These more stringent controls again reduce the variation in shocks only slightly, indicating that there is substantial residual shock variation within periods and industrial sectors.

To see how this variation translates to the commuting zone level, Table 2 summarizes the shift-share instrument $z_{\ell t}$, weighting by $e_{\ell t}$. Column 1 presents the raw variation in the data, with a standard deviation of 1.55 and an interquartile range of 1.74. Columns 2–4 add controls that mirror the specifications from Table 1. Specifically, Column 2 residualizes the instrument on the lagged manufacturing share, effectively excluding the service industry. Column 3 similarly controls for the lagged manufacturing share interacted with period dummies, while Column 4 controls for lagged shares of each of the 10 manufacturing sectors, again interacted with periods indicators. These isolate variation in $z_{\ell t}$ that is due to variation in shocks within manufacturing industries within periods (Column 3) and within sector (Column 4), per the discussion in Section 3.4. The residual variation in the instrument falls with richer controls but remains substantial, with a standard deviation of

[28] The weighted mean of manufacturing shocks is 7.4, with a standard error clustered at the 3-digit SIC level (as in our analysis below) of 1.3.

[29] See Autor et al. (2014), Figure II, for the definitions of these sectors, which aggregate two-digit SIC codes.

0.87 and an interquartile range of 0.51 in Column 4. These results indicate that heterogeneity in exposure shares across locations and the underlying residual variation in China import shocks creates meaningful variation in the instrument, which is necessary for SSIV relevance, even with a relatively stringent set of sectoral controls.

To more formally assess the plausibility of Assumption 2 (shock independence) and choose the appropriate level of clustering for exposure-robust standard errors, we next analyze the correlation patterns of shocks across manufacturing industries using available industry classifications and the time dimension of the pooled cross section. In particular, we compute intraclass-correlation coefficients (ICCs) of shocks within different industry groups, as one might do to correct for conventional clustering parametrically (e.g. Angrist and Pischke (2008), p. 312).[30] These ICCs come from a random effects model, providing a hierarchical decomposition of residual within-period shock variation:

$$g_{nt} = \mu_t + a_{\text{ten}(n),t} + b_{\text{sic2}(n),t} + c_{\text{sic3}(n),t} + d_n + e_{nt}, \tag{13}$$

where μ_t are period fixed effects; $a_{\text{ten}(n),t}$, $b_{\text{sic2}(n),t}$, and $c_{\text{sic3}(n),t}$ denote time-varying (and possibly auto-correlated) random effects generated by the ten industry groups in Acemoglu et al. (2016), 20 groups identified by SIC2 codes, and 136 groups corresponding to SIC3 codes, respectively; and d_n is a time-invariant industry random effect (across our 397 four-digit SIC industries). Following convention, we estimate equation (13) as a hierarchical linear model by maximum likelihood, assuming Gaussian residual components.[31]

Table 3 reports estimated ICCs from equation (13), reflecting the share of the overall shock residual variance due to each random effect. These reveal moderate clustering of shock residuals at the industry and SIC3 level (with ICCs of 0.169 and 0.073, respectively). At the same time, there is less evidence for clustering of shocks at a higher SIC2 level and particularly by ten cluster groups (ICCs of 0.047 and 0.016, respectively, with standard errors of comparable magnitude). This supports the assumption that shocks are mean-independent across SIC3 clusters, so it will be sufficient to cluster standard errors at the level of SIC3 groups, as Acemoglu et al. (2016) do in their conventional industry-level IV regressions. The inverse HHI estimates in Table 1 indicate that at this level of shock clustering there is still an adequate effective sample size.

[30] Note that similar ICC calculations could be implemented in a setting that directly regresses industry outcomes on industry shocks, such as Acemoglu et al. (2016). Mutual correlation in the instrument is a generic concern that is not specific to shift-share designs, although one that is rarely tested for. Getting the correlation structure in shocks right is especially important for inference in our framework, since the outcome and treatment in the industry-level regression ($\bar{y}^{\perp}_{nt}$ and $\bar{x}^{\perp}_{nt}$) are by construction correlated across industries.

[31] In particular we estimate an unweighted mixed-effects regression using Stata's *mixed* command, imposing an exchangeable variance matrix for $(a_{\text{ten}(n),1}, a_{\text{ten}(n),2})$, $(b_{\text{sic2}(n),1}, b_{\text{sic2}(n),2})$, and $(c_{\text{sic3}(n),1}, c_{\text{sic3}(n),2})$.

5.3 Estimates from Shock-Level Regressions

Table 4 reports SSIV coefficients from regressing regional manufacturing employment growth in the U.S. on the growth of import competition from China, instrumented by predicted Chinese import growth.[32] Per the results in Section 4.1, we estimate these coefficients with equivalent industry-level regressions in order to obtain valid exposure-robust standard errors. Consistent with the above analysis of shock ICCs, we cluster standard errors at the SIC3 level when estimating these regressions. We also report first-stage F-statistics with corresponding exposure-robust inference. As discussed in Section 4.2, these come from industry-level IV regressions of the aggregated treatment and instrument (i.e. $\bar{x}_{nt}^{\perp}$ on $\bar{z}_{nt}^{\perp}$), instrumented with shocks and weighting by s_{nt}. The F-statistics are well above the conventional threshold of ten in all columns of the table.

Columns 1 through 4 of Table 4 document the sensitivity of SSIV estimates to the inclusion of various controls, designed to isolate the conditional random assignment of shocks under alternative versions of Assumption 3. Column 1 first replicates column 6 of Table 3 in Autor et al. (2013) by including in $w_{\ell t}$ period fixed effects, Census division fixed effects, start-of-period conditions (% college educated, % foreign-born, % employment among women, % employment in routine occupations, and the average offshorability index), and the start-of-period manufacturing share. The point estimate is -0.596, with a corrected standard error of 0.114.[33]

As noted, the ADH specification in column 1 does not include the lagged manufacturing share control $S_{\ell t}$, which is necessary to solve the "incomplete shares" problem in Section 3.4, though it does include a highly correlated control (start-of-period manufacturing share). In column 2 of Table 4 we isolate within-manufacturing variation in shocks by replacing the latter sum-of-share control with the former. The SSIV point estimate remains almost unchanged, at -0.489 (with a standard error of 0.100). Here exposure-robust standard errors are obtained from an industry-level regression that drops the implicit service sector shock of $g_{0t} = 0$.

Isolating the within-period variation in manufacturing shocks requires further controls in the incomplete shares case, as discussed in Section 3.4. Specifically, Column 3 controls for lagged manufacturing shares interacted with period indicators, which are the share-weighted sums of period effects in q_{nt}. With these controls the SSIV point estimate is to -0.267 with an exposure-robust standard error of 0.099.[34] While the coefficient remains statistically and economically significant, it is smaller

[32] Appendix Table C1 reports estimates for other outcomes in ADH: growth rates of unemployment, labor force non-participation, and average wages, corresponding to columns 3 and 4 of Table 5 and column 1 of Table 6 in ADH.

[33] Appendix Table C2 implements three alternative methods for conducting inference in Table 4, reporting conventional state-clustered standard errors as in ADH (which are not exposure-robust), the Adão et al. (2019) standard errors (which are asymptotically equivalent to ours but differ in finite samples), and null-imposed confidence intervals obtained from shock-level Lagrange multiplier tests (which may have better finite-sample properties). Consistent with the theoretical discussion in Appendix B.3, the conventional standard errors are generally too low, while the Adão et al. (2019) standard errors are slightly larger than those from Table 4, by 10–15% for the main specifications. Imposing the null widens the confidence interval more substantially, by 30–50%, although more so on the left end, suggesting that much larger effects are not rejected by the data. This last finding is consistent with Adão et al. (2019), except that we use the equivalent industry-level regression to compute the null-imposed confidence interval.

[34] Appendix Figure C1 reports binned scatter plots that illustrate the first-stage and reduced-form relationships

in magnitude than the estimates in Columns 1 and 2. The difference stems from the fact that 2000–07 saw both a faster growth in imports from China (e.g., due to its entry to the WTO) and a faster decline in U.S. manufacturing. The earlier columns attribute the faster manufacturing decline to increased trade with China, while the specification in Column 3 controls for any unobserved shocks specific to the manufacturing sector overall in the 2000s (e.g., faster automation), as would a conventional industry-level IV regression with period fixed effects.

Column 4 illustrates how our framework makes it straightforward to introduce more detailed industry-level controls in SSIV, which are commonly used in industry-level studies of the China shock. In particular, Acemoglu et al. (2016) examine the impact of trade with China on U.S. employment across industries and control for fixed effects of 10 broad industry groups. By Proposition 4, we can also exploit shock variation within these industry groups (and periods) in the SSIV design, weakening the assumption of quasi-random shock assignment. This entails controlling for the shares of exposure to the 10 industry groups defined by Acemoglu et al. (2016), interacted with period dummies. The resulting point estimate in column 4 of Table 4 remains very similar to that of column 3, at -0.252 with a standard error of 0.136. Here exposure-robust standard errors are obtained by including the 10-groups-by-period fixed effects in the equivalent industry-level regression, per Section 4.1.

Finally, columns 5 and 6 of Table 4 report falsification tests of the SSIV specifications in columns 3 and 4 by lagging the manufacturing employment growth outcome by two decades. These specifications thus provide "pre-trend" tests which are, as usual, informative about the plausibility of Assumption 1 if employment pre-trends correlate with the contemporaneous trend residual $\varepsilon_{\ell t}$. Column 5 shows no significant relationship with lagged manufacturing employment when within-period manufacturing shocks are leveraged. The point estimate of -0.028 is a full order of magnitude smaller than the comparable effect estimate in Column 3, with an exposure-robust standard error of 0.099. When adding the more stringent set of 10-sectoral controls in column 6, the coefficient flips signs while remaining statistically insignificant at conventional levels. These tests lend support to the plausibility of the identification assumption. Further support is found in Appendix Table C4, which shows we obtain similar estimates from different overidentified SSIV procedures (leveraging variation in country-specific Chinese import growth, instead of the ADH total), and a p-value for the the shock-level overidentification test derived in Appendix A.9 of 0.142.

Taken together, these sensitivity, falsification, and overidentification exercises suggest that the ADH approach can be reasonably viewed as leveraging exogenous shock variation via our framework. This is notably in contrast to the analysis of Goldsmith-Pinkham et al. (2019), who find the ADH exposure shares to be implausible instruments via different balance and overidentification tests.

corresponding to the Column 3 IV specification. Also note that the column 3 estimate can be interpreted as a weighted average of two period-specific shift-share IV coefficients. Column 1 of Appendix Table C3 shows the underlying estimates, from a just-identified IV regression where both treatment and the instrument are interacted with period indicators (as well as the manufacturing share control, as in column 3), with exposure-robust standard errors obtained by the equivalent industry-level regression discussed in Section 4. The estimated effect of increased Chinese import competition is negative in both periods (–0.491 and –0.225). Other columns repeat the analysis for other outcomes.

6 Conclusion

Shift-share instruments combine a common set of observed shocks with variation in shock exposure. In this paper, we provide a quasi-experimental framework for the validity of such instruments based on identifying variation in the shocks, allowing the exposure shares to be endogenous. Our framework revolves around a novel equivalence result: shift-share IV estimates can be reframed as coefficients from weighted shock-level IV regressions, in which the shocks instrument directly for an exposure-weighted average of the original endogenous variable. Shift-share instruments are therefore valid when shocks are idiosyncratic with respect to an exposure-weighted average of the unobserved factors determining the outcome variable, and yield consistent IV estimates when the number of shocks is large and they are sufficiently dispersed in terms of their average exposure.

Through various extensions and illustrations, we show how our quasi-experimental SSIV framework can guide empirical work in practice. By controlling for exposure-weighted averages of shock-level confounders, researchers can isolate more plausibly exogenous variation in shocks, such as over time or within narrow industry groups. By estimating SSIV coefficients, placebo regressions, and first stage F-statistics at the level of shocks, researchers can easily perform exposure-robust inference that accounts for the inherent non-standard clustering of observations with common shock exposure. Our shock-level analysis also raises new concerns: SSIV designs with few or insufficiently dispersed shocks may have effectively small samples, despite there being many underlying observations, and instruments constructed from exposure shares that do not add up to a constant require appropriate controls in order to isolate quasi-random shock variation.

Ultimately, the plausibility of our exogenous shocks framework, as with the alternative framework of Goldsmith-Pinkham et al. (2019) based on exogenous shares, depends on the SSIV application. We encourage practitioners to use shift-share instruments based on an *a priori* argument supporting the plausibility of either one of these approaches; various diagnostics and tests of the framework that is most suitable for the setting may then be applied.[35] While our paper develops such procedures for the "shocks" view, Goldsmith-Pinkham et al. (2019) provide different tools for the "shares" view.

Examples from the vast and expanding set of SSIV applications help to illustrate *a priori* arguments for or against the two SSIV frameworks. In some settings, the exposure shares underlying the instrument are tailored to a specific economic question, and to the particular endogenous variable included in the model. In this case, the scenario considered in Section 2.3 – that there are unobserved shocks ν_n which enter ε_ℓ through the shares – may be less of a concern, making shares more plausible instruments. Mohnen (2019), for example, uses the age profile of older workers in local labor markets as the exposure shares in a shift-share instrument for the change in the local elderly employment rate the following decade. He argues, based on economic intuition, that these tailored shares are uncorre-

[35] In principle, shares and shocks may simultaneously provide valid identifying variation, but in practice it would seem unlikely for both sources of variation to be *a priori* plausible in the same setting, because these approaches to identification are substantively different.

lated with unobserved trends in youth employment rates. This argument notably does not require one to specify the age-specific shocks g_n, which only affect power of the instrument; in fact, the shocks are dispensed with altogether in robustness checks that directly instrument with the shares and report 2SLS and LIML estimates. Similarly, Algan et al. (2017) use the lagged share of the construction sector in the regional economy as an instrument for unemployment growth during the Great Recession, arguing that it does not predict changes in voting outcomes in other ways. With a single industry considered, the identification assumption reduces to that of conventional difference-in-differences with continuous treatment intensity, and our approach cannot be applied.

In contrast, our framework is more appropriate in settings where shocks are tailored to a specific question while the shares are "generic," in that they could conceivably measure an observation's exposure to multiple shocks (both observed and unobserved). Both Autor et al. (2013) and Acemoglu and Restrepo (Forthcoming), for example, build shift-share instruments with similar lagged employment shares but different shocks – rising trade with China and the adoption of industrial robots, respectively. According to the shares view, these papers use essentially the same instruments (lagged employment shares) for different endogenous variables (growth of import competition and growth of robot usage), and are therefore mutually inconsistent. Our framework helps reconcile these identification strategies, provided the variation in each set of shocks can be described as arising from a natural experiment.

In sum, our analysis formalizes the claim that SSIV identification may "come from" the exogeneity of shocks, while providing new guidance for SSIV estimation and inference that may be applied across a number of economic fields, including international trade, labor economics, urban economics, macroeconomics, and public finance. Our shock-level assumptions connect SSIV in these settings to conventional shock-level IV estimation, bringing shift-share instruments to more familiar econometric territory and facilitating the assessment of SSIV credibility in practice.

Table 1: Shock Summary Statistics in the Autor et al. (2013) Setting

	(1)	(2)	(3)	(4)
Mean	1.79	7.37	0	0
Standard deviation	10.79	20.92	20.44	19.39
Interquartile range	0	6.61	6.11	5.50
Specification				
Excluding service industries		✓	✓	✓
Residualized on manufacturing-by-period FE			✓	✓
Residualized on 10-sectors-by-period FE				✓
Effective sample size ($1/HHI$ of s_{nt} weights)				
Across industries and periods	3.5	191.6	191.6	191.6
Across SIC3 groups	1.7	58.4	58.4	58.4
Largest s_{nt} weight				
Across industries and periods	0.398	0.035	0.035	0.035
Across SIC3 groups	0.757	0.066	0.066	0.066
Observation counts				
# of industry-by-period shocks	796	794	794	794
# of industries	398	397	397	397
# of SIC3 groups	137	136	136	136
# of periods	2	2	2	2

Notes: This table summarizes the distribution of China import shocks g_{nt} across industries n and periods t in the Autor et al. (2013) replication. Shocks are measured using flows of imports from China in eight developed economics outside of the United States. All statistics are weighted by the average industry exposure shares s_{nt}, computed as in Proposition 1; shares are measured from lagged manufacturing employment, as described in the text. Column 1 includes the non-manufacturing industry aggregate in each period with a shock of 0, while columns 2-4 restrict to manufacturing industries. The following columns residualize the shocks on period indicators (column 3) or the indicators for each of the 10 sectors defined in Acemoglu et al. (2016) interacted with period indicators (column 4). We report the effective sample size (the inverse renormalized Herfindahl index of the share weights, as described in the text) with and without the non-manufacturing industry, at both the industry-by-period level and at the level of aggregate SIC3 groups (across periods), along with the largest share weights.

Table 2: Shift-Share Instrument Summary Statistics in the Autor et al. (2013) Setting

	(1)	(2)	(3)	(4)
Mean	1.79	0	0	0
Standard deviation	1.55	1.51	1.06	0.87
Interquartile range	1.74	1.750	0.69	0.51
Controls				
Lagged mfg. share		✓	✓	✓
Period-specific lagged mfg. share			✓	✓
Period-specific lagged 10-sector shares				✓
Observation counts				
# of CZs-by-periods	1,444	1,444	1,444	1,444
# of commuting zones (CZs)	722	722	722	722
# of periods	2	2	2	2

Notes: This table summarizes the distribution of the shift-share instrument $z_{\ell t}$ across commuting zones ℓ and periods t in the Autor et al. (2013) replication. The shocks used in the instrument are measured using flows of imports from China in eight developed economies outside of the United States and the exposure shares are measured from lagged manufacturing employment, as described in the text. All statistics are weighted by the start-of-period commuting zone population, as in Autor et al. (2013). Columns 2–4 residualize the instrument on the lagged commuting zone manufacturing share, the lagged manufacturing share interacted with period indicators, and the lagged manufacturing share for each of the 10 sectors defined in Acemoglu et al. (2016), interacted with period indicators. These correspond to the share-weighted averages of industry-level controls included in each column of Table 1.

Table 3: Shock Intra-Class Correlations in the Autor et al. (2013) Setting

	Estimate	SE
	(1)	(2)
Shock ICCs		
10 sectors	0.016	(0.022)
SIC2	0.047	(0.052)
SIC3	0.073	(0.057)
Industry	0.169	(0.047)
Period means		
1990s	4.65	(1.38)
2000s	16.87	(3.34)
# of industry-by-periods	794	

Notes: This table reports shock intra-class correlation coefficients in the Autor et al. (2013) replication, estimated from the hierarchical model described in the text. Estimates come from a maximum likelihood procedure with an exchangeable covariance structure for each industry random effect and with period fixed effects. Robust standard errors are reported in parentheses.

Table 4: Shift-Share IV Estimates of the Effect of Chinese Imports on Manufacturing Employment

	Effects				Pre-trends	
	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	-0.596	-0.489	-0.267	-0.252	-0.028	0.142
	(0.114)	(0.100)	(0.099)	(0.136)	(0.092)	(0.090)
CZ-level controls ($w_{\ell t}$)						
Autor et al. (2013) baseline	✓	✓	✓	✓	✓	✓
Start-of-period mfg. share	✓					
Lagged mfg. share		✓	✓	✓	✓	✓
Period-specific lagged mfg. share			✓	✓	✓	✓
Period-specific lagged 10-sector shares				✓		✓
Industry-level controls (q_{nt})						
Period indicators			✓	✓	✓	✓
Sector-by-period indicators				✓		✓
SSIV first stage F-stat.	185.59	166.73	123.64	46.50	123.64	46.50
# of industries-by-periods	796	794	794	794	794	794
# of CZs-by-periods	1,444	1,444	1,444	1,444	1,444	1,444

Notes: This table reports shift-share IV coefficients from regressions of regional manufacturing employment growth in the U.S. on the growth of import competition from China, instrumented with predicted China import growth as described in the text. Column 1 replicates column 6 of Table 3 in Autor et al. (2013) by controlling for period fixed effects, Census division fixed effects, beginning-of-period conditions (% college educated, % foreign-born, % employment among women, % employment in routine occupations, and the average offshorability index), and the start-of-period manufacturing share. Column 2 replaces the start-of-period manufacturing shares control with the lagged manufacturing shares underlying the instrument, while column 3 interacts this control with period indicators. Column 4 adds lagged shares of the 10 industry sectors defined in Acemoglu et al. (2016), again interacted with period indicators. Columns 5 and 6 report falsification tests using manufacturing employment growth lagged by two decades as an outcome, in specifications that parallel those in columns 3 and 4. Exposure-robust standard errors and first-stage F-statistics are obtained from equivalent industry-level IV regressions, as described in the text, with the indicated industry-level controls and allowing for clustering of shocks at the level of three-digit SIC codes. The sample in columns 2–6 includes 722 locations (commuting zones) and 397 industries, each observed in two periods; the estimate in column 1 implicitly includes an additional two periods of the non-manufacturing industry with a shock of zero.

References

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey M. Wooldridge. 2019. "Sampling-based vs. Design-based Uncertainty in Regression Analysis." *Working Paper.*

Acemoglu, Daron, David H. Autor, David Dorn, Gordon H. Hanson, and Brendard Price. 2016. "Import Competition and the Great U.S . Employment Sag of the 2000s." *Journal of Labor Economics* 34:S141–S198.

Acemoglu, Daron, and Pascual Restrepo. Forthcoming. "Robots and Jobs: Evidence from US Labor Markets." *Journal of Political Economy.*

Adão, Rodrigo, Michal Kolesár, and Eduardo Morales. 2019. "Shift-Share Designs: Theory and Inference." *The Quarterly Journal of Economics* 134:1949–2010.

Algan, Yann, Sergei Guriev, Elias Papaioannou, and Evgenia Passari. 2017. "The European Trust Crisis and the Rise of Populism." *Brookings Papers on Economic Activity* 2017:309–400.

Angrist, Joshua D., Kate Graddy, and Guido W. Imbens. 2000. "The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish." *The Review of Economic Studies* 67:499–527.

Angrist, Joshua D., Guido W. Imbens, and Alan B. Krueger. 1999. "Jackknife Instrumental Variables Estimation." *Journal of Applied Econometrics* 14:57–67.

Angrist, Joshua D., and JS Pischke. 2008. *Mostly Harmless Econometrics: An Empiricist's Companion.* Princeton University Press.

Autor, David H., David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review* 103:2121–2168.

———. 2019. "When Work Disappears: Manufacturing Decline and the Falling Marriage-Market Value of Young Men." *American Economic Review: Insights* 1:161–178.

Autor, David H., David Dorn, Gordon H. Hanson, and Jae Song. 2014. "Trade Adjustment: Worker-Level Evidence." *The Quarterly Journal of Economics* 129:1799–1860.

Autor, David H., and Mark G. Duggan. 2003. "The Rise in the Disability Rolls and the Decline in Unemployment." *The Quarterly Journal of Economics* 118:157–205.

Bartik, Timothy J. 1991. *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Bekker, Paul A. 1994. "Alternative Approximations to the Distributions of Instrumental Variable Estimators." *Econometrica* 62:657.

Berman, Nicolas, Mathieu Couttenier, Dominic Rohner, and Mathias Thoenig. 2017. "This mine is mine! How minerals fuel conflicts in Africa." *American Economic Review* 107:1564–1610.

Blanchard, Olivier Jean, Lawrence F. Katz, Robert E. Hall, and Barry Eichengreen. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity* 1992:1–75.

Bombardini, Matilde, and Bingjing Li. 2019. "Trade , Pollution and Mortality in China." *Working Paper.*

Borusyak, Kirill, and Xavier Jaravel. 2017. "Revisiting Event Study Designs, with an Application to the Estimation of the Marginal Propensity to Consume." Working Paper.

Bound, John, David A. Jaeger, and Regina M. Baker. 1995. "Problems with instrumental variables estimation when the correlation between the instruments and the endogenous explanatory variable is weak." *Journal of the American Statistical Association* 90:443–450.

Boustan, Leah, Fernando Ferreira, Hernan Winkler, and Eric M. Zolt. 2013. "The Effect of Rising Income Inequality on Taxation and Public Expenditures: Evidence from U.S. Municipalities and School Districts, 1970-2000." *The Review of Economics and Statistics* 95:1291–1302.

Broxterman, Daniel A, and William D Larson. 2018. "An Examination of Industry Mix Demand Indicators : The Bartik Instrument at Twenty-Five." Working Paper.

Cameron, A. Colin, and Douglas L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *The Journal of Human Resources* 50:317–372.

Card, David. 2001. "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics* 19:22–64.

Chaisemartin, Clément de, and Xavier D'Haultfoeuille. 2019. "Two-way Fixed Effects Estimators with Heterogeneous Treatment Effects." NBER Working Paper No. 25904.

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000." *American Economic Review* 106:479–524.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2019. "Bartik Instruments: What, When, Why, and How." NBER Working Paper 24408.

Goodman-Bacon, Andrew. 2018. "Difference-in-Differences with Variation in Treatment Timing." NBER Working Paper 25018.

Greenstone, Michael, Alexandre Mas, and Hoai-Luu Nguyen. 2014. "Do Credit Market Shocks Affect the Real Economy? Quasi-Experimental Evidence from the Great Recession and 'Normal' Economic Times." NBER Working Paper 20704.

Hansen, Lars Peter. 1982. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 50:1029–1054.

Hornbeck, Richard, and Enrico Moretti. 2019. "Who Benefits From Productivity Growth? Direct and Indirect Effects of Local TFP Growth on Wages, Rents, and Inequality." NBER Working Paper 24661.

Hudson, Sally, Peter Hull, and Jack Liebersohn. 2017. "Interpreting Instrumented Difference-in-Differences." Mimeo.

Hull, Peter. 2018. "Estimating Treatment Effects in Mover Designs." Working Paper.

Hummels, David, Rasmus Jorgensen, Jakob Munch, and Chong Xiang. 2011. "The Wage Effects of Offshoring: Evidence from Danish Matched Worker-Firm Data." NBER Working Paper 17496.

Hummels, David, Rasmus Jorgensen, Jakob Munch, and Chong Xiang. 2014. "The Wage Effects of Offshoring: Evidence From Danish Matched Worker-Firm Data." *American Economic Review* 104:1597–1629.

Imbert, Clement, Marlon Seror, Yifan Zhang, and Yanos Zylberberg. 2019. "Migrants and Firms: Evidence from China." Working Paper.

Jaeger, David A, Joakim Ruist, and Jan Stuhler. 2018. "Shift-Share Instruments and Dynamic Adjustments: The Case of Immigration." Working Paper.

Jaravel, Xavier. 2019. "The Unequal Gains from Product Innovations: Evidence from the U.S. Retail Sector." *The Quarterly Journal of Economics* 134:715–783.

Kolesar, Michal, Raj Chetty, John Friedman, Edward Glaeser, and Guido W. Imbens. 2015. "Identification and Inference With Many Invalid Instruments." *Journal of Business and Economic Statistics* 33:474–484.

Kovak, Brian K. 2013. "Regional Effects of Trade Reform: What is the Correct Measure of Liberalization?" *American Economic Review* 103:1960–1976.

Luttmer, Erzo F.P. 2005. "Neighbors as Negatives: Relative Earnings and Well-being." *The Quarterly Journal of Economics* 120:963–1002.

Mohnen, Paul. 2019. "The Impact of the Retirement Slowdown on the U.S. Youth Labor Market." *Working Paper.*

Montiel Olea, Jose Luis, and Carolin Pflueger. 2013. "A robust test for weak instruments." *Journal of Business and Economic Statistics* 31:358–369.

Moulton, Brent R. 1986. "Random group effects and the precision of regression estimates." *Journal of Econometrics* 32:385–397.

Nakamura, Emi, and Jón Steinsson. 2014. "Fiscal Stimulus in a Monetary Union: Evidence from US Regions." *American Economic Review* 104:753–792.

Nunn, Nathan, and Nancy Qian. 2014. "US Food Aid and Civil Conflict." *American Economic Review* 104:1630–1666.

Oberfield, Ezra, and Devesh Raval. 2014. "Micro Data and Macro Technology." NBER Working Paper No. 20452.

Saiz, Albert. 2010. "The Geographic Determinants of Housing Supply." *The Quarterly Journal of Economics* 125:1253–1296.

Stuen, Eric T., Ahmed Mushfiq Mobarak, and Keith E. Maskus. 2012. "Skilled Immigration and Innovation: Evidence from Enrolment Fluctuations in U.S. Doctoral Programmes." *The Economic Journal* 122:1143–1176.

Suárez, Juan Carlos, and Owen Zidar. 2016. "Who Benefits from State Corporate Tax Cuts? A Local Labor Markets Approach with Heterogeneous Firms." *American Economic Review* 106:2582–2624.

Xu, Chenzi. 2019. "Reshaping Global Trade: The Immediate and Long-Run Effects of Bank Failures." Working Paper.

Appendix to "Quasi-Experimental Shift-Share Research Designs"

Kirill Borusyak, UCL

Peter Hull, University of Chicago

Xavier Jaravel, London School of Economics

December 2019

Contents

A Appendix Results

A.1 Heterogeneous Treatment Effects

In this appendix we consider what a linear SSIV identifies when the structural relationship between y_ℓ and x_ℓ is nonlinear. We show that under a first-stage monotonicity condition the large-sample SSIV coefficient estimates a convexly weighted average of heterogeneous treatment effects. This holds even when the instrument has different effects on the outcome depending on the underlying realization of shocks, for example when $y_\ell = \sum_n s_{\ell n}\tilde{\beta}_{\ell n} x_{\ell n} + \varepsilon_\ell$ with $\tilde{\beta}_{\ell n}$ capturing the effects of (possibly unobserved) observation- and shock-specific treatments $x_{\ell n}$ making up the observed $x_\ell = \sum_n s_{\ell n} x_{\ell n}$.

Consider a general structural outcome model of

$$y_\ell = y(x_{\ell 1}, \dots, x_{\ell R}, \varepsilon_\ell), \tag{14}$$

where the R treatments are similarly defined as $x_{\ell r} = x_r(g, \eta_{\ell r})$ with g collecting the vector of shocks g_n and $\eta_\ell = (\eta_{\ell 1}, \dots, \eta_{\ell R})$ capturing first-stage heterogeneity. We consider an IV regression of y_ℓ on some aggregated treatment $x_\ell = \sum_r \alpha_{\ell r} x_{\ell r}$ with $\alpha_{\ell r} \geq 0$. Note that this nests the case of a single aggregate treatment ($R = 1$ and $\alpha_{\ell 1} = 1$) with arbitrary effect heterogeneity, as well as the special case above ($R = N$ and $\alpha_{\ell r} = s_{\ell n}$). We abstract away from controls w_ℓ and assume each shock is as-good-as-randomly assigned (mean-zero and mutually independent) conditional on the vector of second-stage unobservables ε and the matrices of first-stage unobservables η, exposure shares s, importance weights e, and aggregation weights α, collected in $\mathcal{I} = (\varepsilon, \eta, s, e, \alpha)$. This assumption is stronger than A1 but generally necessary in a non-linear setting while still allowing for the endogeneity of exposure shares. For further notational simplicity we assume that $y(\cdot, \varepsilon_\ell)$ and each $x_r(\cdot, \eta_{\ell r})$ are almost surely continuously differentiable, such that $\beta_{\ell r}(\cdot) = \frac{\partial}{\partial x_r} y(\cdot, \varepsilon_\ell)$ captures the effect, for observation ℓ, of marginally increasing treatment r on the outcome and $\pi_{\ell n r}(\cdot) = \frac{\partial}{\partial g_n} x_r(\cdot, \eta_{\ell r})$ captures the effect of marginally increasing the nth shock on the rth treatment at ℓ.

Under an appropriate law of large numbers, the shift-share IV estimator approximates a ratio of sums of "reduced-form" and "first-stage" expectations:

$$\hat{\beta} = \frac{\mathbb{E}\left[\sum_\ell e_\ell z_\ell y_\ell\right]}{\mathbb{E}\left[\sum_\ell e_\ell z_\ell x_\ell\right]} + o_p(1) = \frac{\sum_\ell \sum_n \mathbb{E}\left[s_{\ell n} e_\ell g_n y_\ell\right]}{\sum_\ell \sum_n \sum_r \mathbb{E}\left[s_{\ell n} e_\ell g_n \alpha_{\ell r} x_{\ell r}\right]} + o_p(1). \tag{15}$$

Given this, we then have the following result:

Prop. A1 When $\pi_{\ell n r}([\gamma; g_{-n}]) \geq 0$ almost surely, equation (15) can be written

$$\hat{\beta} = \frac{\sum_\ell \sum_n \sum_r \mathbb{E}\left[\int_{-\infty}^{\infty} \tilde{\beta}_{\ell n r}(\gamma)\omega_{\ell n r}(\gamma)\right] d\gamma}{\sum_\ell \sum_n \sum_r \mathbb{E}\left[\int_{-\infty}^{\infty} \omega_{\ell n r}(\gamma)\right] d\gamma} + o_p(1), \tag{16}$$

where $\omega_{\ell nr}(\gamma) \geq 0$ almost surely and

$$\tilde{\beta}_{\ell nr}(\gamma) = \frac{\beta_{\ell r}(x_1([\gamma; g_{-n}], \eta_{\ell 1}), \ldots x_R([\gamma; g_{-n}], \eta_{\ell R}))}{\alpha_{\ell r}} \tag{17}$$

is a rescaled treatment effect, evaluated at $(x_1([\gamma; g_{-n}], \eta_{\ell 1}), \ldots x_R([\gamma; g_{-n}], \eta_{\ell R})$ for $[g_n; g_{-n}] = (g_1, \ldots, g_{n-1}, g_n, g_{n+1}, \ldots g_N)'$.

Proof: See Appendix B.4.

This shows that in large samples $\hat{\beta}$ estimates a convex average of rescaled treatment effects, $\tilde{\beta}_{\ell nr}(\cdot)$, when each $x_r(\cdot, \eta_{\ell r})$ is almost surely monotone in each shock. Appendix B.4 shows that the weights $\omega_{\ell nr}(\gamma)$ are proportional to the first-stage effects $\pi_{\ell nr}([\gamma; g_{-n}])$, exposure shares $s_{\ell n}$, regression weights e_ℓ, treatment aggregation weights $\alpha_{\ell r}$, and a function of the shock distribution. In the case without aggregation, i.e. $R = \alpha_{\ell r} = 1$, there is no rescaling in the $\tilde{\beta}_{\ell nr}(\gamma)$. Equation (16) then can be seen as generalizing the result of Angrist et al. (2000), on the identification of heterogeneous effects of continuous treatments, to the continuous shift-share instrument case. Intuition for the $\omega_{\ell nr}(\gamma)$ weights follows similarly from this connection. With aggregation – that is, when the realization of shocks may have heterogeneous effects on y_ℓ holding the aggregated x_ℓ fixed – equation (16) shows that SSIV captures a convex average of treatment effects per aggregated unit. Thus in the leading example of $y_\ell = \sum_n s_{\ell n} \tilde{\beta}_{\ell n} x_{\ell n} + \varepsilon_\ell$ and $x_\ell = \sum_n s_{\ell n} x_{\ell n}$, this result establishes identification of a convex average of the $\tilde{\beta}_{\ell n}$. In this way the result generalizes Adão et al. (2019b), who establish the identification of convex averages of rescaled treatment effects in reduced form shift-share regressions.

A.2 Comparing SSIV and Native Shock-Level Regression Estimands

In this appendix we illustrate economic differences between the estimands of two regressions that researchers may consider: SSIV using outcome and treatment observations y_ℓ and x_ℓ (which we show in Proposition 1 are equivalent to certain shock-level IV regressions), and more conventional shock-level IV regressions using "native" y_n and x_n. These outcomes and treatments capture the same economic concepts as the original y_ℓ and x_ℓ, in contrast to the constructed $\bar{y}_n$ and $\bar{x}_n$ discussed in Section 2.2. In line with the labor supply and other key SSIV examples, we will for concreteness refer to the ℓ and n as indexing regions and industries, respectively. We consider the case where both the outcome and treatment can be naturally defined at the level of region-by-industry cells (henceforth, cells) – $y_{\ell n}$ and $x_{\ell n}$, respectively – and thus suitable for aggregation across either dimension with some weights $E_{\ell n}$ (e.g., cell employment growth rates aggregated with lagged cell employment weights): $y_\ell = \sum_n s_{\ell n} y_{\ell n}$ for $s_{\ell n} = \frac{E_{\ell n}}{\sum_{n'} E_{\ell n'}}$ and $y_n = \sum_\ell \omega_{\ell n} y_{\ell n}$ for $\omega_{\ell n} = \frac{E_{\ell n}}{\sum_{\ell'} E_{\ell' n}}$, with analogous expressions for x_ℓ and x_n. We further define $E_\ell = \sum_n E_{\ell n}$ and $E_n = \sum_\ell E_{\ell n}$ for conciseness.[36]

[36]This formulation nests reduced-form shift-share regressions when $x_{\ell n} = g_n$ for each ℓ. The labor supply example of Section 2.1 fits only partially in this formal setup because the industry or regional wage growth y_n is not equal to a

We consider the estimands of two regression specifications: β from the regional level model (1), instrumented by z_ℓ and weighted by $e_\ell = E_\ell/E$ for $E = \sum_\ell E_\ell$, and β_{ind} from a simpler industry-level IV regression of

$$y_n = \beta_{\text{ind}} x_n + \varepsilon_n, \tag{18}$$

instrumented by the industry shock g_n and weighted by $s_n = E_n/E$. For simplicity we do not include any controls in either specification and implicitly condition on $\{E_{\ell n}\}_{\ell,n}$ (and some other variables as described below), viewing them as non-stochastic.[37]

We show that β and β_{ind} generally differ when there are within-region spillover effects or when treatment effects are heterogenous. We study these cases in turn, maintaining several assumptions: (i) a first stage relationship analogous to the one considered in Section 3.1:

$$x_{\ell n} = \pi_{\ell n} g_n + \eta_{\ell n}, \tag{19}$$

for non-stochastic $\pi_{\ell n} \geq \bar{\pi} > 0$, (ii) a stronger version of our Assumption 1 that imposes $\mathbb{E}[g_n] = \mathbb{E}[g_n \varepsilon_{\ell n'}] = \mathbb{E}[g_n \eta_{\ell n'}] = 0$ for all ℓ, n, and n', with $\varepsilon_{\ell n'}$ denoting the structural cell-level residual of each model, (iii) the assumption that g_n is uncorrelated with $g_{n'}$ for all n and n', and (iv) that all appropriate laws of large numbers hold.

Within-Region Spillover Effects Suppose the structural model at the cell level is given by

$$y_{\ell n} = \beta_0 x_{\ell n} - \beta_1 \sum_{n'} s_{\ell n'} x_{\ell n'} + \varepsilon_{\ell n}. \tag{20}$$

Here β_0 captures the direct effect of the shock on the cell outcome, and β_1 captures a within-region spillover effect. The local employment effects of industry demand shocks from the model in Appendix A.8 fit in this framework, see equation (47).[38] The following proposition shows that the SSIV estimand β captures the effect of treatment net of spillovers (i.e. $\beta_0 - \beta_1$), whereas β_{ind} subtracts the spillover only partially; this is intuitive since the spillover effect is fully contained within regions but not within industries.

Prop. A2 Suppose equation (20) holds and the average local concentration index $H_L = \sum_{\ell,n} e_\ell s_{\ell n}^2$ is bounded from below by a constant $\bar{H}_L > 0$. Further assume $\pi_{\ell n} = \bar{\pi}$ and $\text{Var}[g_n] = \sigma_g^2$ for all ℓ and n. Then the SSIV estimator satisfies

$$\hat{\beta} = \beta_0 - \beta_1 + o_p(1) \tag{21}$$

weighted average of wage growth across cells: reallocation of employment affects the average wage growth even in the absence of wage changes in any given cell.

[37] Note that we thereby condition on the shares $s_{\ell n}$ and importance weights e_ℓ. Yet we still allow for share endogeneity by not restricting $\mathbb{E}[\varepsilon_{\ell n}]$ to be zero.

[38] In the labor supply example from the main text $y_{\ell n}$ is the cell wage, which is equalized within the region, and $x_{\ell n}$ is cell employment. Equation 20 therefore holds for $\beta_0 = 0$ and $-\beta_1$ being the inverse labor supply elasticity.

while the native industry-level IV estimator satisfies

$$\hat{\beta}_{\text{ind}} = \beta_0 - \beta_1 H_L + o_p(1), \tag{22}$$

If $\beta_1 \neq 0$ (i.e. in presence of within-region spillovers), $\hat{\beta}$ and $\hat{\beta}_{\text{ind}}$ asymptotically coincide if and only if $H_L \xrightarrow{p} 1$, which corresponds to the case where the average region is asymptotically concentrated in one industry.

***Proof*:** See Appendix B.5.

Treatment Effect Heterogeneity Now consider a different structural model which allows for heterogeneity in treatment effects:

$$y_{\ell n} = \beta_{\ell n} x_{\ell n} + \varepsilon_{\ell n}. \tag{23}$$

We also allow the first-stage coefficients $\pi_{\ell n}$ and shock variance σ_n^2 to vary. The following proposition shows that β and β_{ind} differ in how they average effect $\beta_{\ell n}$ (here treated as non-stochastic) across the (ℓ, n) cells. The weights corresponding to the SSIV estimand β are relatively higher for cells that represent a larger fraction of the regional economy. This follows because in the regional regression $s_{\ell n}$ determines the cell's weight in both the outcome and the shift-share instrument, while in the industry regression only the former argument applies. Heterogeneity in the $\pi_{\ell n}$ and σ_n^2 has equivalent effects on the weighting scheme of both estimands.

Prop. A3 In the casual model (23),

$$\hat{\beta} = \frac{\sum_{\ell,n} E_{\ell n} s_{\ell n} \pi_{\ell n} \sigma_n^2 \cdot \beta_{\ell n}}{\sum_{\ell,n} E_{\ell n} s_{\ell n} \pi_{\ell n} \sigma_n^2} + o_p(1) \tag{24}$$

and

$$\hat{\beta}_{\text{ind}} = \frac{\sum_{\ell,n} E_{\ell n} \pi_{\ell n} \sigma_n^2 \cdot \beta_{\ell n}}{\sum_{\ell,n} E_{\ell n} \pi_{\ell n} \sigma_n^2} + o_p(1), \tag{25}$$

***Proof*:** See Appendix B.6.

A.3 Unobserved n-level Shocks Violate Share Exogeneity

In this appendix, we show that the assumption of SSIV share exogeneity from Goldsmith-Pinkham et al. (2019) is violated when there are unobserved shocks ν_n that affect outcomes via the exposure shares $s_{\ell n}$, i.e. when the residual has the structure

$$\varepsilon_\ell = \sum_n s_{\ell n} \nu_n + \check{\varepsilon}_\ell. \tag{26}$$

This is clear in the simple case of fixed N: even when shares are randomly assigned across observations (i.e. each $s_{\ell n}$ is independent of each ν_n and $\check{\varepsilon}_\ell$), the structure of (26) will ensure $\mathbb{E}\left[s_{\ell n}\varepsilon_\ell\right] \neq 0$ and thus $\bar{\varepsilon}_n \overset{p}{\nrightarrow} 0$ for each n.

We next show that this result generalizes to the case of increasing N, where the contribution of each ν_n to the variation in ε_ℓ becomes small. The intuition is that the SSIV relevance condition typically requires individual observations to be sufficiently concentrated in a small number of shocks (see Section 3.1), and under this condition the share exclusion restriction violations remain asymptotically non-ignorable even as $N \to \infty$.

In this case we define share endogeneity as non-vanishing $\text{Var}\left[\bar{\varepsilon}_n\right]$ at least for some n. This will tend to violate SSIV exclusion, unless shocks are as-good-as-randomly assigned (Assumption 1), even if the importance weights of individual shocks, s_n, converge to zero (Assumption 2). As in the previous section, we continue to treat e_ℓ and $s_{\ell n}$ as non-stochastic to show this result with simple notation.

Prop. A4 Suppose condition (26) holds with the ν_n mean-zero and uncorrelated with the $\check{\varepsilon}_\ell$ and with each other, and with $\text{Var}\left[\nu_n\right] = \sigma_n^2 \geq \sigma_\nu^2$ for a fixed $\sigma_\nu^2 > 0$. Also assume $H_L = \sum_\ell e_\ell \sum_n s_{\ell n}^2 \to \bar{H} > 0$ such that first-stage relevance can be satisfied. Then there exists a constant $\delta > 0$ such that $\max_n \text{Var}\left[\bar{\varepsilon}_n\right] > \delta$ for sufficiently large L.

***Proof*:** See Appendix B.7.

A.4 Connection to Rotemberg Weights

In this appendix we rewrite the decomposition of the SSIV coefficient $\hat{\beta}$ from Goldsmith-Pinkham et al. (2019) that gives rise to their "Rotemberg" weight interpretation, and show that these weights measure the leverage of shocks in our equivalent shock-level IV regression. We then show that, in our framework, skewed Rotemberg weights do not measure sensitivity to misspecification (of share exogeneity) and do not pose a problem for consistency. We finally discuss the implications of high-leverage observations for SSIV inference.

Proposition 1 implies the following decomposition:

$$\hat{\beta} = \frac{\sum_n s_n g_n \bar{y}_n^\perp}{\sum_n s_n g_n \bar{x}_n^\perp} = \sum_n \alpha_n \hat{\beta}_n, \tag{27}$$

where

$$\hat{\beta}_n = \frac{\bar{y}_n^\perp}{\bar{x}_n^\perp} = \frac{\sum_\ell e_\ell s_{\ell n} y_\ell^\perp}{\sum_\ell e_\ell s_{\ell n} x_\ell^\perp} \tag{28}$$

and

$$\alpha_n = \frac{s_n g_n \bar{x}_n^\perp}{\sum_{n'} s_{n'} g_{n'} \bar{x}_{n'}^\perp}. \tag{29}$$

This is a shock-level version of the decomposition discussed in Goldsmith-Pinkham et al. (2019): $\hat{\beta}_n$ is the IV estimate of β that uses share $s_{\ell n}$ as the instrument, and α_n is the so-called Rotemberg weight.

To see the connection with leverage (defined, typically in the context of OLS, as the derivative of each observation's fitted value with respect to its outcome) in our equivalent IV regression, note that

$$\frac{\partial\left(\bar{x}_n^{\perp}\hat{\beta}\right)}{\partial\bar{y}_n^{\perp}} = \bar{x}_n^{\perp}\frac{s_n g_n}{\sum_{n'} s_{n'} g_{n'} \bar{x}_{n'}^{\perp}} = \alpha_n. \tag{30}$$

In this way, α_n measures the sensitivity of $\hat{\beta}$ to $\hat{\beta}_n$. For Goldsmith-Pinkham et al. (2019), exposure to each shock should be a valid instrument and thus $\hat{\beta}_n \xrightarrow{p} \beta$ for each n. However, in our framework deviations of $\hat{\beta}_n$ from β reflect nonzero $\bar{\varepsilon}_n$ in large samples, and such share endogeneity is not ruled out; thus α_n does not have the same sensitivity-to-misspecification interpretation. Moreover, a high leverage of certain shocks ("skewed Rotemberg weights," in the language of Goldsmith-Pinkham et al. (2019)) is not a problem for consistency in our framework, provided it results from a heavy-tailed and high-variance distribution of shocks (that still satisfies our regularity conditions, such as finite shock variance), and each s_n is small as required by Assumption 2.

Nevertheless, skewed α_n may cause issues with SSIV inference, as may high leverage observations in any regression. In general, the estimated residuals $\hat{\bar{\varepsilon}}_n^{\perp}$ of high-leverage observations will tend to be biased toward zero, which may lead to underestimation of the residual variance and too small standard errors (e.g., Cameron and Miller 2015). This issue can be addressed, for instance, by computing confidence intervals with the null imposed, as Adão et al. (2019b) recommend and as we discuss in Section 4.1. In practice our Monte-Carlo simulations in Appendix A.10 find that the coverage of conventional exposure-robust confidence intervals to be satisfactory even with Rotemberg weights as skewed as those reported in the applications of Goldsmith-Pinkham et al. (2019) analysis.

A.5 SSIV Consistency in Short Panels

This appendix shows how alternative shock exogeneity assumptions imply the consistency of panel SSIV regressions with many inconsistently-estimated fixed effects (FEs). We consider this incidental parameters problem in "short" panels, with fixed T and $L \to \infty$ and with unit FE. Similar arguments apply with L fixed and $T \to \infty$ and with period FEs.

Suppose for the linear causal model $y_{\ell t} = \beta x_{\ell t} + \epsilon_{\ell t}$ and control vector $w_{\ell t}$ (which includes the FEs), $\sum_\ell e_{\ell t} w_{\ell t}^{\Delta} z_{\ell t} \xrightarrow{p} \Omega_{zw}$, $\sum_\ell e_{\ell t} w_{\ell t}^{\Delta} \epsilon_{\ell t}^{\Delta} \xrightarrow{p} \Omega_{w\epsilon}$, and $\sum_\ell e_{\ell t} w_{\ell t}^{\Delta} w_{\ell t}^{\Delta\prime} \xrightarrow{p} \Omega_{ww}$ for full-rank Ω_{ww}, where $v_{\ell t}^{\Delta}$ is a subvector of the (weighted) unit-demeaned observation of variable $v_{\ell t}$, $v_{\ell t} - \frac{\sum_{t'} e_{\ell t'} v_{\ell t'}}{\sum_{t'} e_{\ell t'}}$, that drops any elements that are identically zero (e.g. those corresponding to the unit FEs in $w_{\ell t}$). Then defining $\gamma = \Omega_{ww}^{-1}\Omega_{w\epsilon}$ we can write $y_{\ell t}^{\Delta} = \beta x_{\ell t}^{\Delta} + w_{\ell t}^{\Delta\prime}\gamma + \varepsilon_{\ell t}^{\Delta}$. Suppose also that $\sum_\ell e_{\ell t} z_{\ell t} x_{\ell t}^{\perp} \xrightarrow{p} \pi$ for some $\pi \neq 0$. Then, following the proof to Proposition 2, $\hat{\beta}$ is consistent if and only if

$$\sum_n \sum_t s_{nt} g_{nt} \bar{\varepsilon}^{\Delta}_{nt} \xrightarrow{p} 0, \tag{31}$$

where $s_{nt} = \sum_\ell e_{\ell t} s_{\ell nt}$ and $\bar{\varepsilon}^{\Delta}_{nt} = \frac{\sum_\ell e_{\ell t} s_{\ell nt} \varepsilon^{\Delta}_{\ell t}}{\sum_\ell e_{\ell t} s_{\ell nt}}$. This condition is satisfied when analogs of Assumptions 1 and 2 and the regularity conditions in Proposition 3 hold, as are the various extensions discussed in Section 3. In particular when $w_{\ell t}$ contains t-specific FE the key assumption of quasi-experimental shock assignment is $\mathbb{E}\left[g_{nt} \mid \bar{\varepsilon}^{\Delta}_{nt}, s_{nt}\right] = \mu_t$, $\forall n, t$, allowing endogenous period-specific shock means μ_t via Proposition 4. This assumption avoids the incidental parameters problem by considering shocks as-good-as-randomly assigned given an unobserved $\bar{\varepsilon}^{\Delta}_{nt}$, which is a function of the time-varying $\varepsilon_{\ell p}$ across all periods p.

An intuitive special case of this approach is when the exposure shares and importance weights are fixed: $s_{\ell nt} = s_{\ell n0}$ and $e_{\ell t} = e_{\ell 0}$. Then the weights in (31) are time-invariant, $s_{nt} = s_{n0}$, and

$$\begin{aligned}
\bar{\varepsilon}^{\Delta}_{nt} &= \frac{\sum_\ell e_{\ell 0} s_{\ell n0} \varepsilon^{\Delta}_{\ell t}}{\sum_\ell e_\ell s_{\ell n0}} \\
&= \frac{\sum_\ell e_{\ell 0} s_{\ell n0} (\varepsilon_{\ell t} - \frac{1}{T} \sum_{t'} \varepsilon_{\ell t'})}{\sum_\ell e_{\ell 0} s_{\ell n0}} \\
&= \bar{\varepsilon}_{nt} - \frac{1}{T} \sum_{t'} \bar{\varepsilon}_{nt'},
\end{aligned} \tag{32}$$

where $\bar{\varepsilon}_{nt} = \frac{\sum_\ell e_{\ell 0} s_{\ell n0} \varepsilon_{\ell t}}{\sum_\ell e_{\ell 0} s_{\ell n0}}$ is an aggregate of period-specific unobservables $\varepsilon_{\ell t}$. It is then straightforward to extend Propositions 3 and 4 under a shock-level assumption of strong exogeneity, i.e. that $\mathbb{E}\left[g_{nt} \mid \bar{\varepsilon}_{n1}, \ldots \bar{\varepsilon}_{nT}, s_{n0}\right] = \mu_n + \tau_t$ for all n and t. Here endogenous n-specific shock means are permitted by the observation in Section 3.3, that share-weighted n-specific FE at the shock level are subsumed by ℓ-specific FE in the SSIV regression when shares and weights are fixed.

A.6 SSIV Relevance with Panel Data

This appendix shows that holding the exposure shares fixed in a pre-period is likely to weaken the SSIV first-stage in panel regressions. Consider a panel extension of the first stage model used in Section 3.1, where $x_{\ell t} = \sum_n s_{\ell nt} x_{\ell nt}$ with $x_{\ell nt} = \pi_{\ell nt} g_{nt} + \eta_{\ell nt}$, $\pi_{\ell n} \geq \bar{\pi}$ for some fixed $\bar{\pi} > 0$, and the g_{nt} are mutually independent and mean-zero with variance $\sigma^2_{nt} \geq \bar{\sigma}^2_g$ for fixed $\sigma^2_g > 0$, independently of $\{\eta_{\ell nt}\}_{\ell,n,t}$. As in other appendices, we here treat $s_{\ell nt}$, $e_{\ell t}$, and $\pi_{\ell nt}$ as non-stochastic. Then again omitting controls for simplicity, an SSIV regression with $z_{\ell t} = \sum_{n=1}^{N} s^*_{\ell nt} g_{nt}$ as an instrument, where

$s^*_{\ell nt}$ is either $s_{\ell nt}$ (updated shares) or $s_{\ell n0}$ (fixed shares), yields an expected first-stage covariance of

$$\mathbb{E}\left[\sum_\ell \sum_t e_{\ell t} z_{\ell t} x^\perp_{\ell t}\right] \geq \bar{\sigma}^2_g \bar{\pi} \sum_\ell \sum_t e_{\ell t} \sum_n s^*_{\ell nt} s_{\ell tn}. \tag{33}$$

For panel SSIV relevance we require the $e_{\ell t}$-weighted average of $\sum_n s^*_{\ell nt} s_{\ell nt}$ to not vanish asymptotically. With updated shares this is satisfied when the Herfindahl index of an average observation-period (across shocks) is non-vanishing, while in the fixed shares case the overlap of shares in periods 0 and t, $\sum_n s_{\ell n0} s_{\ell nt}$, may become weak or even vanish as $T \to \infty$, on average across observations.

A.7 Estimated Shocks

This appendix establishes the formal conditions for the SSIV estimator, with or without a leave-one-out correction, to be consistent when shocks g_n are noisy estimates of some latent g^*_n satisfying Assumptions 1 and 2. We also propose a heuristic measure that indicates whether the leave-one-out correction is likely to be important and compute it for the Bartik (1991) setting. Straightforward extensions to other split-sample estimators follow.

Suppose a researcher estimates shocks via a weighted average of variables $g_{\ell n}$. That is, given weights $\omega_{\ell n} \geq 0$ such that $\sum_\ell \omega_{\ell n} = 1$ for all n, she computes

$$g_n = \sum_\ell \omega_{\ell n} g_{\ell n}. \tag{34}$$

A leave-one-out (LOO) version of the shock estimator is instead

$$g_{n,-\ell} = \frac{\sum_{\ell' \neq \ell} \omega_{\ell' n} g_{\ell' n}}{\sum_{\ell' \neq \ell} \omega_{\ell' n}}. \tag{35}$$

We assume that each $g_{\ell n}$ is a noisy version of the same latent shock g^*_n:

$$g_{\ell n} = g^*_n + \psi_{\ell n}, \tag{36}$$

where g^*_n satisfies Assumptions 1 and 2 and $\psi_{\ell n}$ is estimation error (in Section 3.4 we considered the special case of $\psi_{\ell n} \propto \varepsilon_\ell$). This implies a feasible shift-share instrument of $z_\ell = z^*_\ell + \psi_\ell$ and its LOO version $z^{LOO}_\ell = z^*_\ell + \psi^{LOO}_\ell$, where $z^*_\ell = \sum_n s_{\ell n} g^*_n$, $\psi_\ell = \sum_n s_{\ell n} \sum_{\ell'} \omega_{\ell' n} \psi_{\ell' n}$, and $\psi^{LOO}_\ell = \sum_n s_{\ell n} \frac{\sum_{\ell' \neq \ell} \omega_{\ell' n} \psi_{\ell' n}}{\sum_{\ell' \neq \ell} \omega_{\ell' n}}$. The orthogonality condition for these instruments requires that $\sum_\ell e_\ell \varepsilon_\ell \psi_\ell \xrightarrow{p} 0$ and $\sum_\ell e_\ell \varepsilon_\ell \psi^{LOO}_\ell \xrightarrow{p} 0$, respectively.

We now present three sets of results. First, we establish a simple sufficient condition under which the LOO instrument satisfies exclusion. We also propose stronger conditions that guarantee consistency of LOO-SSIV. Second, we explore the conditions under which the covariance between ε_ℓ and $\psi_{\ell n}$

is ignorable, i.e. asymptotically does not lead to a "mechanical" bias of the conventional leave-one-out estimator. We propose a heuristic measure that is large when the bias is likely to be small. Lastly, we apply these ideas to the setting of Bartik (1991) using the data from Goldsmith-Pinkham et al. (2019). In line with previous appendices, we condition on $s_{\ell n}$, $\omega_{\ell n}$, and e_ℓ and treat them as non-stochastic for notational convenience. We also assume the SSIV regressions are estimated without controls w_ℓ.

LOO Exclusion and Consistency The following proposition establishes three results. The first is the most important one, providing the condition for exclusion to hold in expectation, which we discuss below. The second strengthens this condition so that the estimator converges, which naturally requires that most shocks are estimated with sufficient amount of data. A tractable case of complete specialization is considered in last part, where there should be many more observations than shocks.

Prop. A5

1. If $\mathbb{E}\left[\varepsilon_\ell \psi_{\ell' n}\right] = 0$ for all $\ell \neq \ell'$ and n, then $\mathbb{E}\left[\sum_\ell e_\ell \varepsilon_\ell \psi_{\ell,LOO}\right] = 0$.

2. If $\mathbb{E}\left[(\varepsilon_\ell, \psi_{\ell n}) \mid \{(\varepsilon_{\ell'}, \psi_{\ell' n'})\}_{\ell' \neq \ell, n'}\right] = 0$ for all ℓ and n, then the LOO estimator is consistent, provided it has a first stage and two regularity conditions hold: $\mathbb{E}\left[\left|\varepsilon_{\ell_1}\varepsilon_{\ell_2}\psi_{\ell_1' n_1}\psi_{\ell_2' n_2}\right|\right] \leq B$ for a constant B and all $(\ell_1, \ell_2, \ell_1', \ell_2', n_1, n_2)$ and

$$\sum_{\substack{(\ell_1,\ell_2,\ell_1',\ell_2')\in\mathcal{J},\\ n_1,n_2}} e_{\ell_1} e_{\ell_2} s_{\ell_1 n_1} s_{\ell_2 n_2} \frac{\omega_{\ell_1' n_1}}{\sum_{\ell \neq \ell_1} \omega_{\ell n_1}} \frac{\omega_{\ell_2' n_2}}{\sum_{\ell \neq \ell_2} \omega_{\ell n_2}} \to 0, \tag{37}$$

   with $\mathcal{J}$ denoting the set of tuples $(\ell_1, \ell_2, \ell_1', \ell_2')$ for which one of the two conditions hold: (i) $\ell_1 = \ell_2$ and $\ell_1' = \ell_2' \neq \ell_1$, (ii) $\ell_1 = \ell_2'$ and $\ell_2 = \ell_1' \neq \ell_1$.

3. Condition (37) is satisfied if $\frac{N}{L} \to 0$ in the special case where each region is specialized in one industry, i.e. $s_{\ell n} = \mathbf{1}\left[n = n(\ell)\right]$ for some $n(\cdot)$, there are no importance weights ($e_\ell = \frac{1}{L}$), and shocks estimated by simple LOO averaging among observations exposed to a given shock ($\omega_{\ell n} = \frac{1}{L_n}$ for $L_n = \sum_\ell \mathbf{1}\left[n(\ell) = n\right]$), assuming further that $L_n \geq 2$ for each n so that the LOO estimator is well-defined.

***Proof*:** See Appendix B.8.

The condition of (1) would be quite innocuous in random samples of ℓ – the environment in which leave-one-out adjustments are often considered (e.g. Angrist et al. (1999)) – but is strong without random sampling. It requires ε_ℓ and $\psi_{\ell' n}$ to be uncorrelated for $\ell' \neq \ell$, which may easily be violated when both ℓ and ℓ' are exposed to the same shocks—a situation in which excluding own observation is not sufficient. Moreover, since we have conditioned on the exposure shares throughout, $\mathbb{E}\left[\varepsilon_\ell \psi_{\ell' n}\right] = 0$ generally requires either ε_ℓ or $\psi_{\ell' n}$ to have a zero *conditional* mean—the share exogeneity assumption applied to either the residuals or the estimation error. At the same time, this condition does not require $\mathbb{E}\left[\varepsilon_\ell \psi_{\ell' n}\right]=0$ for $\ell=\ell'$, which reflects the benefit of LOO: eliminating the mechanical bias from the residual directly entering shock estimates.

Heuristic for Importance of LOO Correction We now return to the non-LOO SSIV estimator. As in Proposition A4, we assume that $\mathbb{E}\left[\varepsilon_\ell \psi_{\ell' n}\right]=0$ for $\ell' \neq \ell$ and all n, so the LOO estimator is consistent under the additional regularity conditions. Then the "mechanical bias" mentioned in Section 3.4 is the only potential problem: under appropriate regularity conditions (similar to those in part 2 of Proposition A4),

$$\begin{aligned}
\hat{\beta}-\beta &= \frac{\mathbb{E}\left[\sum_\ell e_\ell \varepsilon_\ell \psi_\ell\right]}{\mathbb{E}\left[\sum_\ell e_\ell z_\ell^{\perp} x_\ell\right]}+o_p(1) \\
&= \frac{\sum_{\ell, n} e_\ell s_{\ell n} \omega_{\ell n} \mathbb{E}\left[\varepsilon_\ell \psi_{\ell n}\right]}{\mathbb{E}\left[\sum_\ell e_\ell z_\ell^{\perp} x_\ell\right]}+o_p(1). \qquad (38)
\end{aligned}$$

With $\left|\mathbb{E}\left[\varepsilon_\ell \psi_{\ell n}\right]\right|$ bounded by some $B_1>0$ for all ℓ and n, the numerator of (38) is bounded by $H_N B_1$, for an observable composite of the relevant shares $H_N=\sum_{\ell, n} e_\ell s_{\ell n} \omega_{\ell n}$. The structure of the shares also influences the strength of the first stage in the denominator. We assume, without loss of generality, that z_ℓ is mean-zero and impose our standard model of the first stage from Section 3.1 (but specified based on the latent shock g_n^*): $x_\ell=\sum_n s_{\ell n} x_{\ell n}$ for $x_{\ell n}=\pi_{\ell n} g_n^*+\eta_{\ell n}$, $\eta_{\ell n}$ mean-zero and uncorrelated with $g_{n'}^*$ for all ℓ, n, n', $\operatorname{Var}\left[g_n^*\right] \geq \bar{\sigma}_g^2>0$ and $\pi_{\ell n} \geq \bar{\pi}>0$:

$$\begin{aligned}
\mathbb{E}\left[\sum_\ell e_\ell z_\ell^{\perp} x_\ell\right] &= \sum_\ell e_\ell \mathbb{E}\left[\left(\sum_n s_{\ell n}\left(g_n^*+\psi_{\ell n}\right)\right)\left(\sum_{n'} s_{\ell n'}\left(\pi_{\ell n} g_{n'}^*+\eta_{\ell n'}\right)\right)\right] \\
&= \sum_{\ell, n} e_\ell s_{\ell n}^2 \cdot \pi_{\ell n} \operatorname{Var}\left[g_n^*\right]+\sum_\ell e_\ell \sum_{n, n'} s_{\ell n} s_{\ell n'} \mathbb{E}\left[\psi_{\ell n}\left(\pi_{\ell n} g_{n'}^*+\eta_{\ell n'}\right)\right]. \qquad (39)
\end{aligned}$$

Excepting knife-edge cases where the two terms in (39) cancel out, $\mathbb{E}\left[\sum_\ell e_\ell z_\ell x_\ell\right] \not\rightarrow 0$ provided $H_L=\sum_{\ell, n} e_\ell s_{\ell n}^2 \geq \bar{H}$ for some fixed $\bar{H}>0$.

We thus define the following heuristic:

$$H=\frac{H_L}{H_N}=\frac{\sum_{\ell, n} e_\ell s_{\ell n}^2}{\sum_{\ell, n} e_\ell s_{\ell n} \omega_{\ell n}}. \qquad (40)$$

When H is large, we expect the non-LOO SSIV estimator to be relatively insensitive to the mechanical bias generated by the average covariance between $\psi_{\ell n}$ and ε_ℓ, and thus similar to the LOO estimator.

We note an important special case. Suppose all weights are derived from variable $E_{\ell n}$ (e.g. lagged employment level in region ℓ and industry n) as $s_{\ell n}=\frac{E_{\ell n}}{E_\ell}$, $\omega_{\ell n}=\frac{E_{\ell n}}{E_n}$, and $e_\ell=\frac{E_\ell}{E}$, for $E_\ell=\sum_n E_{\ell n}$,

$E_n = \sum_\ell E_{\ell n}$, and $E = \sum_\ell E_\ell$. Then

$$H_N = \sum_{\ell,n} \frac{E_\ell}{E} \frac{E_{\ell n}}{E_\ell} \frac{E_{\ell n}}{E_n} = \sum_{\ell,n} \frac{E_n}{E} \left(\frac{E_{\ell n}}{E_n}\right)^2 = \sum_n s_n \sum_\ell \omega_{\ell n}^2, \tag{41}$$

where $s_n = \frac{E_n}{E}$ is the weight in our equivalent shock-level regression. Therefore, H_N is the weighted average across n of n-specific Herfindahl concentration indices, while H_L is the weighted average across ℓ of the ℓ-specific Herfindahl indices. With $E_{\ell n}$ denoting lagged employment, H is high (and thus we expect the LOO correction to be unnecessary) when employment is much more concentrated across industries in a typical region than it is concentrated across regions for a typical industry.

The formula simplifies further with $E_{\ell n} = \mathbf{1}\left[n = n(\ell)\right]$ for all ℓ, n, corresponding to the case of complete specialization of observations in shocks with no regression or shock estimation weights, as in part 3 of Proposition A4. In that case,

$$H = \frac{1}{\sum_\ell \frac{1}{L} \frac{1}{L_{n(\ell)}}} = \frac{1}{\frac{1}{L} \sum_n \sum_{\ell:\, n(\ell)=n} \frac{1}{L_n}} = \frac{L}{N}. \tag{42}$$

Our heuristic is therefore large when there are many observations per estimated shock.[39]

Application to Bartik (1991) We finally apply our insights to the Bartik (1991) setting, using the Goldsmith-Pinkham et al. (2019) replication code and data. Table C5 reports the results. Column 1 shows the estimates of the inverse local labor supply elasticity using SSIV estimators with and without the LOO correction and using population weights, replicating Table 3, column 2, of Goldsmith-Pinkham et al. (2019). Column 2 repeats the analysis without the population weights.[40] We find all estimates to range between 1.2 and 1.3, showing that in practice for Bartik (1991) the LOO correction does not play a substantial role.

This is however especially true without weights, where the LOO and conventional SSIV estimators are 1.30 and 1.29, respectively. Our heuristic provides an explanation: H is almost 8 times bigger when computed without weights. The intuition is that large commuting zones, such as Los Angeles and New York, may constitute a substantial fraction of employment in industries of their comparative advantage. This generates a potential for the mechanical bias: labor supply shocks in those regions affect shock estimates; this bias is avoided by LOO estimators. However, the role of the largest commuting zones is only significant in weighted regressions (by employment or, as in Goldsmith-Pinkham et al. (2019), population).

[39] Here $1/H = N/L$ is proportional to the "bias" of the non-LOO estimator, which is similar to how the finite-sample bias of conventional 2SLS is proportional to the number of instruments over the sample size (Nagar 1959).

[40] Industry growth shocks in this column are the same as in Column 1, again estimated with employment weights.

A.8 Equilibrium Industry Growth in a Model of Local Labor Markets

This appendix develops a simple model of regional labor supply and demand, similar to the model in Adão et al. (2019a). Our goal is to show how the national growth rate of industry employment can be viewed as a noisy version of the national industry-specific labor demand shocks, and how regional labor supply shocks (along with some other terms) generate the "estimation error."

Consider an economy that consists of a set of L regions. In each region ℓ there is a prevailing wage W_ℓ, and labor supply has constant elasticity ϕ:

$$E_\ell = M_\ell W_\ell^\phi, \tag{43}$$

where E_ℓ is total regional employment and M_ℓ is the supply shifter that depends on the working-age population, the outside option, and other factors. Labor demand in each industry n is given by a constant-elasticity function

$$E_{\ell n} = A_n \xi_{\ell n} W_\ell^{-\sigma}, \tag{44}$$

where $E_{\ell n}$ is employment, A_n is the national industry demand shifter, $\xi_{\ell n}$ is its idiosyncratic component, and σ is the elasticity of labor demand. The equilibrium is given by

$$\sum_n E_{\ell n} = E_\ell. \tag{45}$$

Now consider small changes in fundamentals A_n, $\xi_{\ell n}$ and M_ℓ. We use log-linearization around the observed equilibrium and employ the Jones 1965 hat algebra notation, with $\hat{v}$ denoting the relative change in v between the equilibria. We then establish:

Prop. A6 After a set of small changes to fundamentals, the national industry employment growth is characterized by

$$g_n = \sum_\ell \omega_{\ell n} g_{\ell n}, \tag{46}$$

for $\omega_{\ell n} = E_{\ell n} / \sum_{\ell'} E_{\ell' n}$ denoting the share of region ℓ in industry employment, and the change in region-by-industry employment $g_{\ell n}$ is characterized by

$$g_{\ell n} = g_n^* + \frac{\sigma}{\sigma + \phi} \varepsilon_\ell + \hat{\xi}_{\ell n} - \frac{\sigma}{\sigma + \phi} \sum_n s_{\ell n} \left(g_n^* + \hat{\xi}_{\ell n} \right), \tag{47}$$

where $g_n^* = \hat{A}_n$ is the national industry labor demand shock, $\varepsilon_\ell = \hat{M}_\ell$ is the regional labor supply shock, and $s_{\ell n} = E_{\ell n} / \sum_{n'} E_{\ell n'}$.

Proof: See Appendix B.9.

The first term in (47) justifies our interpretation of the observed industry employment growth as a noisy estimate of the latent labor demand shock g_n^*. The other terms constitute the "estimation error." The first of them is proportional to the structural residual of the labor supply equation, ε_ℓ; we have previously established the conditions under which it may or may not confound SSIV estimation. The other terms, that we abstracted away from in Section 3.4, include the idiosyncratic demand shock $\hat{\xi}_{\ell n}$ and shift-share averages of both national and idiosyncratic demand shocks. If the model is correct, all of these are uncorrelated with ε_ℓ, thus not affecting Assumption 1.

A.9 SSIV with Multiple Endogenous Variables or Instruments

This appendix first generalizes our equivalence result for SSIV regressions with multiple endogenous variables and instruments, and discusses corresponding extensions of our quasi-experimental framework via the setting of Jaeger et al. (2018). We also describe how to construct the effective first-stage F-statistic of Montiel Olea and Pflueger (2013) for SSIV with one endogenous variable but multiple instruments. We then consider new shock-level IV procedures in this framework, which can be used for efficient estimation and specification testing. Finally, we illustrate these new procedures in the Autor et al. (2013) setting.

Generalized Equivalence and SSIV Consistency We consider a class of SSIV estimators of an outcome model with multiple treatment channels,

$$y_\ell = \beta' x_\ell + \gamma' w_\ell + \varepsilon_\ell, \tag{48}$$

where $x_\ell = (x_{1\ell}, \ldots, x_{K\ell})'$ is instrumented by $z_\ell = (z_{1\ell}, \ldots, z_{J\ell})'$, for $z_{j\ell} = \sum_n s_{\ell n} g_{jn}$ and $J \geq K$, and observations are weighted by e_ℓ. Members of this class are parameterized by a (possibly stochastic) full-rank $K \times J$ matrix $\boldsymbol{c}$, which is used to combine the instruments into a vector of length J, $\boldsymbol{c} z_\ell$. For example the two-stage least squares (2SLS) estimator sets $\boldsymbol{c} = \boldsymbol{x}^{\perp\prime} \boldsymbol{e} \boldsymbol{z} (\boldsymbol{z}^{\perp\prime} \boldsymbol{e} \boldsymbol{z}^\perp)^{-1}$, where $\boldsymbol{z}^\perp$ stacks observations of the residualized $z_\ell^{\perp\prime}$. IV estimates using a given combination are written as

$$\hat{\beta} = (\boldsymbol{c}\boldsymbol{z}'\boldsymbol{e}\boldsymbol{x}^\perp)^{-1} \boldsymbol{c}\boldsymbol{z}'\boldsymbol{e}\boldsymbol{y}^\perp, \tag{49}$$

where $\boldsymbol{y}^\perp$ and $\boldsymbol{x}^\perp$ stack observations of the residualized $y_\ell^\perp$ and $x_\ell^{\perp\prime}$, $\boldsymbol{z}$ stacks observations of z_ℓ', and $\boldsymbol{e}$ is an $L \times L$ diagonal matrix of e_ℓ weights. In just-identified IV models (i.e. $J = K$) the two $\boldsymbol{c}$'s cancel in this expression and all IV estimators are equivalent. Note that while the shocks g_{jn} are different across the multiple instruments, we assume here that the exposure shares $s_{\ell n}$ are all the same.

As in Proposition 1, $\hat{\beta}$ can be equivalently obtained by a particular shock-level IV regression. Intuitively, when the shares are the same, $\boldsymbol{c} z_\ell$ also has a shift-share structure based on a linear combination of shocks $\boldsymbol{c} g_n$, and thus Proposition 1 extends. Formally, write $\boldsymbol{z} = \boldsymbol{s}\boldsymbol{g}$ where $\boldsymbol{s}$ is an

$L \times N$ matrix of exposure shares and $\boldsymbol{g}$ stacks observations of the shock vector g'_n; then,

$$
\begin{aligned}
\hat{\beta} &= (\boldsymbol{c}\boldsymbol{g}'\boldsymbol{s}'\boldsymbol{e}\boldsymbol{x}^{\perp})^{-1}\boldsymbol{c}\boldsymbol{g}'\boldsymbol{s}'\boldsymbol{e}\boldsymbol{y}^{\perp} \\
&= (\boldsymbol{c}\boldsymbol{g}'\boldsymbol{S}\bar{\boldsymbol{x}}^{\perp})^{-1}(\boldsymbol{c}\boldsymbol{g}'\boldsymbol{S}\bar{\boldsymbol{y}}^{\perp}),
\end{aligned} \tag{50}
$$

where $\boldsymbol{S}$ is an $N \times N$ diagonal matrix with elements s_n, $\bar{\boldsymbol{x}}^{\perp}$ is an $N \times K$ matrix with elements $\bar{x}^{\perp}_{kn}$, and $\bar{\boldsymbol{y}}^{\perp}$ is an $N \times 1$ vector of $\bar{y}^{\perp}_n$. This is the formula for an s_n-weighted IV regression of $\bar{y}^{\perp}_n$ on $\bar{x}^{\perp}_{1n}, \ldots, \bar{x}^{\perp}_{Kn}$ with shocks as instruments, no constant, and the same $\boldsymbol{c}$ matrix. Furthermore, as in Proposition 1,

$$
\iota'\boldsymbol{S}\bar{\boldsymbol{y}}^{\perp} = \sum_n s_n \bar{y}^{\perp}_n = \sum_{\ell} e_{\ell}\left(\sum_n s_{\ell n}\right) y^{\perp}_{\ell} = \sum_{\ell} e_{\ell} y^{\perp}_{\ell} = 0, \tag{51}
$$

and similarly for $\iota'\boldsymbol{S}\bar{\boldsymbol{x}}'$, where ι is a $N \times 1$ vector of ones. Therefore, the same estimate is obtained by including a constant in this IV procedure (and the same result holds including a shock-level control vectorq_n provided $\sum_n s_{\ell n}$ has been included in w_{ℓ}, as in Proposition 5). The $\boldsymbol{c}$ matrix is again redundant in the just-identified case.

A natural generalization of the orthogonality condition from Section 2.3 and the quasi-experimental framework of Section 3 follows. Rather than rederiving all of these results, we discuss them intuitively in the setting of Jaeger et al. (2018). Here y_{ℓ} denotes the growth rate of wages in region ℓ in a given period (residualized on Mincerian controls), $x_{1\ell}$ is the immigrant inflow rate in that period, and $x_{2\ell}$ is the previous period's immigration rate. The residual ε_{ℓ} captures changes to local productivity and other regional unobservables. Jaeger et al. (2018, Table 5) estimate this model with two "past settlement" instruments $z_{1\ell} = \sum_n s_{\ell n} g_{1n}$ and $z_{2\ell} = \sum_n s_{\ell n} g_{2n}$, where $s_{\ell n}$ is the share of immigrants from country of origin n in location ℓ at a previous reference date and $\boldsymbol{g}_n = (g_{1n}, g_{2n})'$ gives the current and previous period's national immigration rate from n. When this path of immigration shocks is as-good-as-randomly assigned with respect to the aggregated productivity shocks $\bar{\varepsilon}_n$ (satisfying a generalized Assumption 1), the $\boldsymbol{g}_n$ are uncorrelated across countries and $\mathbb{E}\left[\sum_n s_n^2\right] \to 0$ (satisfying a generalized Assumption 2), and appropriately generalized regularity conditions from Proposition 3 hold, the multiple-treatment shock orthogonality condition is satisfied: $\sum_n s_n g_{kn} \bar{\varepsilon}_n \xrightarrow{p} 0$ for each k. Then under the relevance and regularity conditions from Proposition 2, again appropriately generalized, the SSIV estimates are consistent: $\hat{\beta} \xrightarrow{p} \beta$.

Effective First-Stage $\boldsymbol{F}$-statistics With one endogenous variable and multiple instruments, the Montiel Olea and Pflueger (2013) effective first-stage F-statistic provides a state-of-art heuristic for detecting a weak first-stage. Here we describe a correction to it for SSIV that generalizes the F-statistic in the single instrument case discussed in Section 4.2. The Stata command *weakssivtest*, provided with our replication archive, implements this correction.[41]

Consider a structural first stage with multiple instruments and one endogenous variable:

$$x_\ell = \pi' z_\ell + \rho w_\ell + \eta_\ell. \tag{52}$$

Suppose each of the shocks satisfies Assumption 3, i.e. $\mathbb{E}\left[g_{jn} \mid q_n, \bar{\varepsilon}_n, s_n\right] = \mu_j' q_n$, where $\sum_n s_{\ell n} q_n$ is included in w_ℓ, and the residual shocks $g_{jn}^* = g_{jn} - \mu_j' q_n$ are independent from $\{\eta_\ell\}_{\ell=1}^L$. The Montiel Olea and Pflueger (2013) effective F-statistic for the 2SLS regression of y_ℓ on x_ℓ, instrumenting with $z_{1\ell}, \ldots, z_{J\ell}$, controlling for w_ℓ, and weighting by e_ℓ, is given by

$$F_{\text{eff}} = \frac{\left(\sum_\ell e_\ell x_\ell^\perp z_\ell^\perp\right)' \left(\sum_\ell e_\ell x_\ell^\perp z_\ell^\perp\right)}{\operatorname{tr} \hat{V}}, \tag{53}$$

where $\hat{V}$ estimates $V = \text{Var}\left[\sum_\ell e_\ell z_\ell^\perp \eta_\ell\right]$. Note that, as before, the first-stage covariance of the original SSIV regression equals that of the equivalent shock-level one from Proposition 5:

$$\sum_\ell e_\ell x_\ell^\perp z_\ell^\perp = \sum_\ell e_\ell x_\ell^\perp z_\ell = \sum_n s_n g_n \bar{x}_n^\perp = \sum_n s_n g_{n\perp} \bar{x}_n^\perp, \tag{54}$$

where $g_{n\perp}$ is an s_n-weighted projection of g_n on q_n, which consistently estimates g_n^*. A natural extension of Proposition 5 to many mutually-uncorrelated shocks further implies that V is well-approximated by

$$\hat{V} = \sum_n s_n^2 g_{n\perp}^2 \hat{\bar{\eta}}_n^2, \tag{55}$$

where, per the discussion in Section 4.2, $\hat{\bar{\eta}}_n$ denotes the residuals from an IV regression of $\bar{x}_n^\perp$ on $\bar{z}_{1n}^\perp, \ldots, \bar{z}_{Jn}^\perp$, instrumented with $g_{1n}, \ldots, g_{Jn}$, weighted by s_n and controlling for q_n. Plugging this $\hat{V}$ into (53) yields the corrected effective first-stage F-statistic.

Efficient Shift-Share GMM In overidentified settings ($J > K$), it is natural to consider which estimators are most efficient; for quasi-experimental SSIV, this can be answered by combining the asymptotic results of Adão et al. (2019b) with the classic generalized methods of moments (GMM) theory of Hansen (1982). Here we show how standard shock-level IV procedures – such as 2SLS – may yield efficient coefficient estimates $\hat{\beta}^*$, depending on the variance structure of multiple quasi-randomly assigned shocks.

We first note that the equivalence result (50) applies to SSIV-GMM estimators as well:

$$\begin{aligned}\hat{\beta} &= \arg\min_b \left(\boldsymbol{y}^\perp - \boldsymbol{x}^\perp b\right)' \boldsymbol{ezWz'e} \left(\boldsymbol{y}^\perp - \boldsymbol{x}^\perp b\right) \\ &= \arg\min_b \left(\bar{\boldsymbol{y}}^\perp - \bar{\boldsymbol{x}}^\perp b\right)' \boldsymbol{SgWg'S} \left(\bar{\boldsymbol{y}}^\perp - \bar{\boldsymbol{x}}^\perp b\right),\end{aligned} \tag{56}$$

[41] Our package extends the *weakivtest* command developed by Pflueger and Wang (2015).

where $\boldsymbol{W}$ is an $J \times J$ moment-weighting matrix. This leads to an IV estimator with $\boldsymbol{c} = \bar{\boldsymbol{x}}^{\perp\prime}\boldsymbol{SgW}$. For 2SLS estimation, for example, $\boldsymbol{W} = (\boldsymbol{z}^{\perp\prime}\boldsymbol{e}\boldsymbol{z}^{\perp})^{-1}$. Under appropriate regularity conditions, the efficient choice of $\boldsymbol{W}^*$ consistently estimates the inverse asymptotic variance of $\boldsymbol{z}'\boldsymbol{e}\left(\boldsymbol{y}^{\perp} - \boldsymbol{x}^{\perp}\beta\right) = \boldsymbol{g}'S\bar{\varepsilon} + o_p(1)$. Generalizations of results in Adão et al. (2019b) can then be used to characterize this $\boldsymbol{W}^*$ when shocks are as-good-as-randomly assigned with respect to $\bar{\varepsilon}$. Given an estimate $\hat{\boldsymbol{W}}^*$, an efficient coefficient estimate $\hat{\beta}^*$ is given by shock-level IV regressions (50) that set $\boldsymbol{c}^* = \bar{\boldsymbol{x}}^{\perp\prime}\boldsymbol{Sg}\hat{\boldsymbol{W}}^*$. A χ^2_{J-K} test statistic based on the minimized objective in (56) can be used for specification testing.

As an example, suppose shocks are conditionally homoskedastic with the same variance-covariance matrix across n, $\mathrm{Var}\left[\boldsymbol{g}_n \mid \bar{\varepsilon}_n, s_n\right] = \boldsymbol{G}$ for a constant $J \times J$ matrix $\boldsymbol{G}$. Then the optimal $\hat{\beta}^*$ is obtained by a shock-level 2SLS regression of $\bar{y}_n^{\perp}$ on all $\bar{x}_{kn}^{\perp}$ (instrumented by g_{jn} and weighted by s_n). We show this in the case of no controls (and mean-zero shocks) for notational simplicity. Then,

$$
\begin{aligned}
\mathrm{Var}\left[\boldsymbol{g}'S\left(\bar{\boldsymbol{y}}^{\perp} - \bar{\boldsymbol{x}}^{\perp}\beta\right)\right] &= \mathbb{E}\left[\bar{\varepsilon}'\boldsymbol{Sgg}'\boldsymbol{S}\bar{\varepsilon}\right] \\
&= \mathrm{tr}\left(\mathbb{E}\left[\bar{\varepsilon}'\boldsymbol{SGS}\bar{\varepsilon}\right]\right) \\
&= k\boldsymbol{G} \qquad (57)
\end{aligned}
$$

for $k = tr\left(\mathbb{E}\left[\boldsymbol{S}\bar{\varepsilon}\bar{\varepsilon}'\boldsymbol{S}\right]\right)$. The optimal weighting matrix thus should consistently estimate $\boldsymbol{G}$, which is satisfied by $\hat{\boldsymbol{G}} = \boldsymbol{g}'\boldsymbol{Sg}$. Under appropriate regularity conditions, a feasible optimal GMM estimate is thus given by

$$
\begin{aligned}
\hat{\beta}^* &= (\bar{\boldsymbol{x}}^{\perp\prime}\boldsymbol{Sg}\hat{\boldsymbol{G}}^{-1}\boldsymbol{g}'\boldsymbol{S}\bar{\boldsymbol{x}}^{\perp})^{-1}(\bar{\boldsymbol{x}}^{\perp\prime}\boldsymbol{Sg}\hat{\boldsymbol{G}}^{-1}\boldsymbol{g}'\boldsymbol{S}\bar{\boldsymbol{y}}^{\perp}) \\
&= \left(\left(P_{\boldsymbol{g}}\bar{\boldsymbol{x}}^{\perp}\right)' S\bar{\boldsymbol{x}}^{\perp}\right)^{-1}\left(P_{\boldsymbol{g}}\bar{\boldsymbol{x}}^{\perp}\right)' S\bar{y}^{\perp}, \qquad (58)
\end{aligned}
$$

where $P_{\boldsymbol{g}} = \boldsymbol{g}(\boldsymbol{g}'\boldsymbol{Sg})^{-1}\boldsymbol{g}'\boldsymbol{S}$ is an s_n-weighted shock projection matrix. This is the formula for an s_n-weighted IV regression of $\bar{y}_n^{\perp}$ on the fitted values from projecting the $\bar{x}_{kn}^{\perp}$ on the shocks, corresponding to the 2SLS regression above. Straightforward extensions of this equivalence between optimally-weighted estimates of β and shock-level overidentified IV procedures follow in the case of heteroskedastic or clustered shocks, in which case the 2SLS estimator (58) is replaced by the estimator of White (1982). We emphasize that these shock-level estimators are generally different than 2SLS or White (1982) estimators at the level of original observations, which are optimal under conditional homoskedasticity and independence assumptions placed on the residual ε_ℓ – assumptions which are generally violated in our quasi-experimental framework.

Many Shocks in Autor et al. (2013) Appendix Table C4 illustrates different shock-level overidentified IV estimators in the setting of Autor et al. (2013), introduced in Section 5.1. ADH construct their shift-share instrument based on the growth of Chinese imports in eight economies comparable to the U.S., together. We separate them to produce eight sets of industry shocks g_{jn}, each reflecting the growth of Chinese imports in one of those countries. As in Section 5, the outcome of interest is a commuting zone's growth in total manufacturing employment with the single treatment variable measuring a commuting zone's local exposure to the growth of imports from China (see footnote 26 for precise variable definitions). The vector of controls coincides with that of column 3 of Table 4, isolating within-period variation in manufacturing shocks. Per Section 4, exposure-robust standard errors are obtained by controlling for period main effects in the shock-level IV procedures, and we report corrected first stage F-statistics constructed as detailed above.

Column 1 reports estimates of the ADH coefficient β using the industry-level two-stage least squares procedure (58). At -0.238, this estimate it is very similar to the just-identified estimate in column 3 of Table 4. Column 2 shows that we also obtain a very similar coefficient of -0.247 with an industry-level limited information maximum likelihood (LIML) estimator. Finally, in column 3 we report a two-step optimal IV estimate of β using an industry-level implementation of the White (1982) estimator. Both the coefficient and standard error fall somewhat, with the latter consistent with the theoretical improvement in efficiency relative to columns 1 and 2. From this efficient estimate we obtain an omnibus overidentification test statistic of 10.92, distributed as chi-squared with seven degrees of freedom under the null of correct specification. This yields a p-value for the test of joint orthogonality of all eight ADH shocks of 0.142. Table C4 also reports the corrected effective first-stage F-statistic which measures the strength of the relationship between the endogenous variable and the eight shift-share instruments across regions. At 15.10 it is substantially lower than with one instrument in column 3 of Table 4 but still above the conventional heuristic threshold of 10.

A.10 Finite-Sample Performance of SSIV: Monte-Carlo Evidence

In this appendix we study the finite-sample performance of the SSIV estimator via Monte-Carlo simulation. We base this simulation on the data of Autor et al. (2013), as described in Section 5. For comparison, we also simulate more conventional shock-level IV estimators, similar to those used in Acemoglu et al. (2016), which also estimate the effects of import competition with China on U.S. employment. We begin by describing the design of these simulations and the benchmark Monte-Carlo results. We then explore how the simulation results change with various deviations from the benchmark: with different levels of industry concentration, different numbers of industries and regions, and with many shock instruments. Besides showing the general robustness of our framework, these extensions allow us to see how informative some conventional rules of thumb are on the finite-sample performance of shift-share estimators.

Simulation design We base our benchmark data-generating process for SSIV on the specification in column 3 of Table 4. The outcome variable $y_{\ell t}$ corresponds to the change in manufacturing employment as a fraction of working-age population of region ℓ in period t, treatment $x_{\ell t}$ is a measure of regional import competition with China, and the shift-share instrument is constructed by combining the industry-level growth of China imports in eight developed economies, g_{nt}, with lagged regional employment weights of different industries $s_{\ell nt}$. We also include pre-treatment controls $w_{\ell t}$ as in column 3 of Table 4 and and estimate regressions with regional employment weights $e_{\ell t}$; see Section 5 for more detail on the Autor et al. (2013) setting.

In a first step we obtain an estimated SSIV second and first stage of

$$y_{\ell t} = \hat{\beta} x_{\ell t} + \hat{\gamma}' w_{\ell t} + \hat{\varepsilon}_{\ell t}, \tag{59}$$

$$x_{\ell t} = \hat{\pi} z_{\ell t} + \hat{\rho}' w_{\ell t} + \hat{u}_{\ell t}. \tag{60}$$

We then generate 10,000 simulated samples by drawing shocks g^*_{nt}, as detailed below, and constructing the simulated shift-share instrument $z^*_{\ell t} = \sum_n s_{\ell nt} g^*_{nt}$ and treatment $x^*_{\ell t} = \hat{\pi} z^*_{\ell t} + \hat{u}_{\ell t}$. Imposing a true causal effect of $\beta^* = 0$, we use the same $y^*_{\ell t} \equiv \hat{\varepsilon}_{\ell t}$ as the outcome in each simulation (note that it is immaterial whether we include $\hat{\pi}' w_{\ell t}$ and $\hat{\rho}' w_{\ell t}$, since all our specifications control for $w_{\ell t}$). By keeping $\hat{\varepsilon}_{\ell t}$ and $\hat{u}_{\ell t}$ fixed, we study the finite sample properties of the estimator that arises from the randomness of shocks, which is the basis of the inferential framework of Adão et al. (2019b); we also avoid having to take a stand on the joint data generating process of $(\varepsilon_{\ell t}, u_{\ell t})$, which this inference framework does not restrict.

We estimate SSIV specifications that parallel (59)-60) from the simulated data

$$y^*_{\ell t} = \beta^* x^*_{\ell t} + \gamma^* w_{\ell t} + \varepsilon^*_{\ell t}, \tag{61}$$

$$x^*_{\ell t} = \pi^* z^*_{\ell t} + \rho^* w_{\ell t} + u^*_{\ell t}. \tag{62}$$

using the original weights $e_{\ell t}$ and controls $w_{\ell t}$. We then test the (true) hypothesis $\beta^* = 0$ using either the heteroskedasticity-robust standard errors from the equivalent industry-level regression or their version with the null imposed, as in Section 4.1.[42] As in column 3 of Table 4, we control for period indicators as q_{nt} in the industry-level regression.

Our comparison estimator is a conventional industry-level IV inspired by Acemoglu et al. (2016). However, we try to keep the IV regression as similar to the SSIV as possible, thus diverging from Acemoglu et al. (2016) in some details. Specifically, the outcome y_{nt} is the industry employment growth as measured by these authors. It is defined for 392 out of the 397 industries in Autor et al. (2013), so we drop the remaining five industries in each period. The endogenous regressor $x_{nt} \equiv g^{US}_{nt}$ (growth of U.S. imports from China per worker) and the instrument g_{nt} (growth of China imports

[42] Note that there is no need for clustering since we generate the shocks independently across industries in all simulations. We have verified, however, that allowing for correlation in shocks within industry groups and using clustered standard errors yields similar results.

into eight developed economies) are those from which we built the shift-share endogenous regressor and treatment, respectively (see footnote 26). Construction of those variables differ from Acemoglu et al. (2016) who measure imports relative to domestic absorption rather than employment. We also follow our SSIV analysis in using period indicators as the only industry-level control variables q_{nt} and taking identical regression importance weights s_{nt}.

The Monte-Carlo strategy for the conventional shock-level IV parallels the one for SSIV; we obtain an estimated industry-level second and first stage of

$$y_{nt} = \hat{\beta}_{\text{ind}} x_{nt} + \hat{\gamma}' q_{nt} + \hat{\varepsilon}_{nt}, \tag{63}$$

$$x_{nt} = \hat{\pi}_{\text{ind}} g_{nt} + \hat{\rho}' q_{nt} + \hat{u}_{nt}. \tag{64}$$

using the s_{nt} importance weights. We then perform 10,000 simulations where we regenerate shocks g^*_{nt} and regress $y^*_{nt} = \hat{\varepsilon}_{nt}$ (consistent with a true causal effect of $\beta^{\text{ind}} = 0$, given that we control for q_{nt}) on $x^*_{nt} = \hat{\pi}_{\text{ind}} g^*_{nt} + \hat{u}_{nt}$, instrumenting by g^*_{nt}, controlling for q_{nt}, and weighting by s_{nt}. We test $\beta_{\text{ind}} = 0$ by using robust standard errors in this IV regression or the version with the null imposed, which corresponds to a standard Lagrange Multiplier test for this true null hypothesis.

In both simulations we report the rejection rate of nominal 5% level tests for $\beta = 0$ and $\beta_{\text{ind}} = 0$ to gauge the quality of each asymptotic approximation. We do not report the bias of the estimators because they are all approximately unbiased (more precisely, the simulated median bias is at most 1% of the estimator's standard deviation). However we return to the question of bias at the end of the section, where we extend the analysis to having many instruments with a weak first stage.

Main results Table C6 reports the rejection rates for shift-share IV (columns 1 and 2) and conventional industry-level IV (columns 3 and 4) in various simulations. Specifically, column 1 corresponds to using exposure-robust standard errors from the equivalent industry-level IV, and column 2 implements the version with the null hypothesis imposed. Columns 3 and 4 parallel columns 1 and 2 when applied to conventional IV: the former uses heteroskedasticity-robust standard errors and the latter tests $\beta_{\text{ind}} = 0$ with the null imposed, which amounts to using the Lagrange multiplier test.

The simulations in Panel A vary the data-generating process of the shocks. Following Adão et al. (2019b) in row (a) we draw the shocks *iid* from a normal distribution with the variance matched to the sample variance of the shocks in the data after de-meaning by year. The rejection rate is close to the nominal rate of 5% for both SSIV and conventional IV (7.6% and 6.8%, respectively), and in both cases it becomes even closer when the null is imposed (5.2% and 5.0%).

This simulation may not approximate the data-generating process well because of heteroskedasticity: smaller industries have more volatile shocks.[43] To match unrestricted heteroskedasticity, in row

[43]This is established by unreported regressions of $|g_{nt}|$ on s_{nt}, for year-demeaned g_{nt} from ADH, with or without weights. The negative relationship is significant at conventional levels.

(b) we use wild bootstrap, generating $g^*_{nt} = g_{nt}\nu^*_{nt}$ by multiplying the year-demeaned observed shocks g_{nt} by $\nu^*_{nt} \overset{iid}{\sim} \mathcal{N}(0,1)$ (Liu 1988). This approach also provides a better approximation for the marginal distribution of shocks than the normality assumption. Here the relative performance of SSIV is even better: the rejection rate is 8.0% vs. 14.2% for conventional IV.

We now depart from the row (b) simulation in several directions, as a case study for the sensitivity of the asymptotic approximation to different features of the SSIV setup. Specifically, we study the role of the Herfindahl concentration index across industries, the number of regions and industries, and the many weak instrument bias. We uniformly find that the performance of the SSIV estimator is similar to that of industry-level IV. Our results also suggest that the Herfindahl index is a useful statistic for measuring the effective number of industries in SSIV, and the first-stage F-statistic is informative about the weak instrument bias, as usual.

The Role of Industry Concentration Since Assumption 2 requires small concentration of industry importance weights, measured using the Herfindahl index $\sum_{n,t} s^2_{nt} / \left(\sum_{n,t} s_{nt}\right)^2$, Panel B of Table C6 studies how increasing the skewness of s_{nt} towards the bigger industries affects coverage of the tests.[44] For conventional IV this simply amounts to reweighting the regression. Specifically, for a parameter $\alpha > 1$, we use weights

$$\tilde{s}_{nt} = s^{\alpha}_{nt} \cdot \frac{\sum_{n',t'} s_{n't'}}{\sum_{n',t'} s^{\alpha}_{n't'}}.$$

We choose the unique α to match the target level of $\widetilde{HHI}$ by solving, numerically,

$$\frac{\sum_{n,t} \left(\tilde{s}_{nt}\right)^2}{\left(\sum_{n,t} \tilde{s}_{nt}\right)^2} = \widetilde{HHI}. \tag{65}$$

Matching the Herfindahl index in SSIV is more complicated since we need to choose how exactly to amend shares $\tilde{s}_{\ell nt}$ and regional weights $\tilde{e}_{\ell t}$ that would yield $\tilde{s}_{nt}$ from (65). We proceed as follows: we consider the lagged level of manufacturing employment by industry $E_{\ell nt} = e_{\ell t} s_{\ell nt}$ and the total regional non-manufacturing employment $E_{\ell 0t} = e_{\ell t}\left(1 - \sum_n s_{\ell nt}\right)$.[45] We then define $\tilde{E}_{\ell nt} = E_{\ell nt} \cdot \frac{\tilde{s}_{nt}}{s_{nt}}$ for manufacturing industries (and leave non-manufacturing employment unchanged, $\tilde{E}_{\ell 0t} = E_{\ell 0t}$). This increases employment in large manufacturing industries proportionately in all regions, while

[44] Note that in ADH $\sum_{n=1}^{N} s_{\ell nt}$ equals the lagged share of regional manufacturing employment, which is below one. We thus renormalize the shares when computing the Herfindahl.

[45] The interpretation of $E_{\ell nt}$ as the lagged level is approximate since $e_{\ell t}$ is measured at the beginning of period in ADH, while $s_{\ell nt}$ is lagged.

reducing it in smaller ones. We then recompute shares $\tilde{s}_{\ell nt}$ and weights $\tilde{e}_{\ell t}$ accordingly:

$$\tilde{e}_{\ell t} = \sum_{n=0}^{N} \sum_{t} \tilde{E}_{\ell nt},$$
$$\tilde{s}_{\ell nt} = \frac{\tilde{E}_{\ell nt}}{\tilde{e}_{\ell t}}.$$

Rows (c)–(e) of Table C6 Panel B implement this procedure for target Herfindahl levels of 1/50, 1/20, and 1/10, respectively. For comparison, the Herfindahl in the actual ADH data is 1/191.6 (Table 1, column 2). The table finds that even with the Herfindahl index of 1/20 (corresponding to the "effective" number of shocks of 20 in both periods total) the rejection rate is still around 7%, a level that may be considered satisfactory. It also shows that the rejection rate grows when the Herfindahl is even higher, at 1/10, suggesting that the Herfindahl can be used as an indicative rule of thumb. More importantly, the rejection rates are similar for SSIV and conventional industry-level IV, as before.

Varying the Number of Industries and Regions The asymptotic sequence we consider in Section 3.1 relies on both N and L growing. Here we study how the quality of the asymptotic approximation depends on these parameters

First, to consider the case of small N, we aggregate industries in a natural way: from 397 four-digit manufacturing SIC industries into 136 three-digit ones and further into 20 two-digit ones and reconstruct the endogenous right-hand side variable and the instrument using aggregated data.[46] Rows (f) and (g) of Table C6 Panel C report simulation results based on the aggregated data. They show that rejection rates are similar to the case of detailed industries, and between SSIV and conventional IV. This does not mean that disaggregated data are not useful: the dispersion of the simulated distribution (not reported) increases with industry aggregation, reducing test power. However, standard errors correctly reflect this variability, resulting in largely unchanged test coverage rates.

Second, to study the implications of having fewer regions L, we select a random subset of them in each simulation. The results are presented in Rows (h) and (i) of Panel C for $L = 100$ and 25, compared to the original $L = 722$, respectively.[47] They show once again that rejection rates are not significantly affected (even though unreported standard errors expectedly increase).

Many Weak Instruments In this final simulation we return to the question of SSIV bias. Since our previous simulations confirm that just-identified SSIV is median-unbiased, we turn to the case of

[46] Specifically, we aggregate imports from China to the U.S. and either developed economies as well as the number of U.S. workers by manufacturing industry to construct the new g_{nt} and g_{nt}^{US}. We then aggregate the shares $s_{\ell nt}$ and $s_{\ell nt}^{\text{current}}$ to construct $x_{\ell t}$ and $z_{\ell t}$ (see footnote 26 for formulas). We do not change the regional outcome, controls, or importance weights. For conventional IV, we additionally reconstruct the outcome (industry employment growth) by aggregating employment levels by year in the Acemoglu et al. (2016) data and measuring growth according to their formulas.

[47] When we select regions, we always keep observations from both periods for each selected region. We keep the second- and first-stage coefficients from the full sample to focus on the noise that arises from shock randomness.

multiple instruments. We show that the problem of many weak instruments is similar between SSIV and conventional IV, and that first-stage F-statistics, when properly constructed, can serve as useful heuristics.

For clarity, we begin by describing the procedure for the conventional shock-level IV that is a small departure from Column 3 of Table C6. For a given number of instruments $J \geq 1$, in each simulation we generate g^*_{jnt}, $j = 1, \ldots, J$, independently across j using wild bootstrap (as in Table C6 Row (b)).[48] We make only the first instrument relevant by setting $x^*_{nt} = \hat{\pi}_{\text{ind}} g^*_{1nt} + \sum_{j=2}^{J} 0 \cdot g^*_{jnt} + \hat{u}_{nt}$. We then estimate the IV regression of $y^*_{nt} \equiv \hat{\varepsilon}_{nt}$ on x^*_{nt}, instrumenting with $g^*_{1nt}, \ldots, g^*_{Jnt}$, controlling for q_{nt}, and weighting by s_{nt}. We use robust standard errors and compute the effective first-stage F-statistic using the Montiel Olea and Pflueger (2013) method.

The procedure for SSIV is more complex but as usual parallels the one for the conventional shock-level IV as much as possible. Given simulated shocks g^*_{jnt}, we construct shift-share instruments $z^*_{j\ell t} = \sum_{\ell} s_{\ell nt} g^*_{jnt}$ and make only the first of them relevant, $x^*_{\ell t} = \hat{\pi} z^*_{1\ell t} + \sum_{j=2}^{J} 0 \cdot z^*_{jnt} + \hat{u}_{\ell t}$. Since the equivalence result from Section 2.2 need not hold for overidentified SSIV, we rely on the results in Appendix A.9: we estimate β^* from the industry-level regression of $\bar{y}^{*\perp}_{nt}$ (based on $y^*_{\ell t} = \hat{\varepsilon}_{\ell t}$ as before) on $\bar{x}^{*\perp}_{nt}$ by 2SLS, instrumenting by $g^*_{1nt}, \ldots, g^*_{Jnt}$, controlling for q_{nt} and weighting by s_{nt}. We compute robust standard errors from this regression to test $\beta^* = 0$. For effective first-stage F-statistics, we follow the procedure described in Appendix A.9 and implemented via our *weakssivtest* command in Stata.

Table C7 reports the result for $J = 1, 5, 10, 25$, and 50, presenting the rejection rate corresponding to the 5% nominal, the median bias as a percentage of the simulated standard deviation, and the median first-stage F-statistic. Panel A corresponds to SSIV and Panel B to the conventional shock-level IV. For higher comparability, we adjust the first-stage coefficient $\hat{\pi}_{\text{ind}}$ in the latter in order to make the F-statistics approximately match between the two panels. We find that the median bias is now non-trivial and grows with J, at the same time as the F-statistic declines. However, the level of bias is similar for the two estimators. The rejection rates tend to be higher for conventional IV than SSIV, although they converge as J grows.

[48] For computational reasons we perform only $15,000/J$ simulations when $J > 1$ (but 10,000 for $J = 1$ as before).

B Appendix Proofs

B.1 Proposition 3 and Extensions

This section proves Proposition 3 and its extensions which allow for certain forms of mutual shock dependence. As previewed in footnote 5, we do so under weaker regularity conditions than in the main text: we assume that there is a sequence of vectors γ_L such that $\|\hat{\gamma} - \gamma_L\|_1 = o_p(1)$, that $\max_m \left|\sum_\ell e_\ell w_{\ell m}\right| = O_p(1)$ and that $\max_m \left|\sum_\ell e_\ell w_{\ell m} z_\ell\right| = O_p(1)$. These are implied by the regularity conditions of Proposition 3, while they also allow the dimension of $\hat{\gamma}$ to grow with L.

We first verify that Proposition 2 holds under these weaker regularity conditions, such that if shock orthogonality holds and the SSIV relevance condition is satisfied then $\hat{\beta}$ is consistent. The relevant step from the main text proof is

$$\sum_n s_n g_n \bar{\varepsilon}_n - \sum_n s_n g_n \bar{\varepsilon}_n^{\perp} = \sum_\ell e_\ell z_\ell \left(\varepsilon_\ell - \varepsilon_\ell^{\perp}\right) = \left(\sum_\ell e_\ell z_\ell w_\ell'\right)(\hat{\gamma} - \gamma_L) \xrightarrow{p} 0, \tag{66}$$

since $\|\hat{\gamma} - \gamma_L\|_1 \xrightarrow{p} 0$ and $\max_m \left|\sum_\ell e_\ell w_{\ell m} z_\ell\right| = O_p(1)$.

We next prove that $\sum_n s_n g_n \bar{\varepsilon}_n$ is asymptotically mean-zero. Since $\sum_n s_{\ell n} = 1$ and $\sum_\ell e_\ell \varepsilon_\ell^{\perp} = 0$,

$$\begin{aligned}\sum_n s_n \bar{\varepsilon}_n &= \sum_\ell e_\ell \varepsilon_\ell \\ &= \sum_\ell e_\ell (\varepsilon_\ell - \varepsilon_\ell^{\perp}) \\ &= \left(\sum_\ell e_\ell w_\ell'\right)(\hat{\gamma} - \gamma_L) \xrightarrow{p} 0,\end{aligned} \tag{67}$$

when $\|\hat{\gamma} - \gamma_L\|_1 \xrightarrow{p} 0$ and $\max_m \left|\sum_\ell e_\ell w_{\ell m}\right| = O_p(1)$. Thus

$$\sum_n s_n g_n \bar{\varepsilon}_n = \sum_n s_n (g_n - \mu)\bar{\varepsilon}_n + o_p(1), \tag{68}$$

with

$$\mathbb{E}\left[\sum_n s_n (g_n - \mu)\bar{\varepsilon}_n\right] = 0 \tag{69}$$

under Assumption 1.

Finally, note that when $\mathbb{E}\left[(g_n - \mu)^2 \mid s_n, \bar{\varepsilon}_n\right]$ and $\mathbb{E}\left[\bar{\varepsilon}_n^2 \mid s_n\right]$ are uniformly bounded by finite B_g

and B_ε, respectively, we have by Assumption 2 and the Cauchy-Schwartz inequality

$$\begin{aligned}
\mathrm{Var}\left[\sum_n s_n\left(g_n-\mu\right)\bar{\varepsilon}_n\right] &= \mathbb{E}\left[\left(\sum_n s_n\left(g_n-\mu\right)\bar{\varepsilon}_n\right)^2\right] \\
&= \sum_n \mathbb{E}\left[s_n^2\mathbb{E}\left[\mathbb{E}\left[\left(g_n-\mu\right)^2 \mid s_n, \bar{\varepsilon}_n\right]\bar{\varepsilon}_n^2 \mid s_n\right]\right] \\
&\leq B_g B_\varepsilon \mathbb{E}\left[\sum_n s_n^2\right] \to 0. \qquad (70)
\end{aligned}$$

This implies l_2-convergence and so weak convergence: $\sum_n s_n(g_n-\mu)\bar{\varepsilon}_n \xrightarrow{p} 0$, proving Proposition 3.

Similar last steps apply when Assumption 2 is replaced by either Assumptions 5 or 6, maintaining analogous regularity conditions. Under Assumption 5 we have, for $N\left(c\right)=\{n\colon c(n)=c\}$,

$$\begin{aligned}
\mathrm{Var}\left[\sum_n s_n\left(g_n-\mu\right)\bar{\varepsilon}_n\right] &= \mathbb{E}\left[\left(\sum_c\sum_{n\in N(c)} s_n\left(g_n-\mu\right)\bar{\varepsilon}_n\right)^2\right] \\
&= \mathbb{E}\left[\sum_c s_c^2\mathbb{E}\left[\left(\sum_{n\in N(c)}\frac{s_n}{s_c}\left(g_n-\mu\right)\bar{\varepsilon}_n\right)^2 \mid s_c\right]\right] \\
&= \mathbb{E}\left[\sum_c s_c^2\sum_{n,n'\in N(c)}\frac{s_n}{s_c}\frac{s_{n'}}{s_c}\mathbb{E}\left[\left(g_n-\mu\right)\left(g_{n'}-\mu\right)\bar{\varepsilon}_n\bar{\varepsilon}_{n'} \mid s_c\right]\right] \\
&\leq B_g B_\varepsilon \mathbb{E}\left[\sum_c s_c^2\right] \to 0, \qquad (71)
\end{aligned}$$

when $\mathbb{E}\left[(g_n-\mu)^2 \mid \{\bar{\varepsilon}_{n'}\}_{n'\in N(c(n))}, s_c\right] \leq B_g$ and $\mathbb{E}\left[\bar{\varepsilon}_n^2 \mid s_c\right] \leq B_\varepsilon$ uniformly. Here the last line used the Cauchy-Schwartz inequality twice: to establish, for $n, n' \in N(c)$,

$$\begin{aligned}
\mathbb{E}\left[\left(g_n-\mu\right)\left(g_{n'}-\mu\right) \mid \bar{\varepsilon}_n, \bar{\varepsilon}_{n'}, s_c\right] &\leq \sqrt{\mathbb{E}\left[(g_n-\mu)^2 \mid \bar{\varepsilon}_n, \bar{\varepsilon}_{n'}, s_c\right]\mathbb{E}\left[(g_{n'}-\mu)^2 \mid \bar{\varepsilon}_n, \bar{\varepsilon}_{n'}, s_c\right]} \\
&\leq B_g \qquad (72)
\end{aligned}$$

and

$$\begin{aligned}
\mathbb{E}\left[\left|\bar{\varepsilon}_n\right|\left|\bar{\varepsilon}_{n'}\right| \mid s_c\right] &\leq \sqrt{\mathbb{E}\left[\bar{\varepsilon}_n^2 \mid s_c\right]\mathbb{E}\left[\bar{\varepsilon}_{n'}^2 \mid s_c\right]} \\
&\leq B_\varepsilon. \qquad (73)
\end{aligned}$$

If we instead replace Assumption 2 with Assumption 6, we have

$$
\begin{aligned}
\operatorname{Var}\left[\sum_n s_n\left(g_n-\mu\right)\bar{\varepsilon}_n\right] &= \mathbb{E}\left[\left(\sum_n s_n\left(g_n-\mu\right)\bar{\varepsilon}_n\right)^2\right] \\
&= \sum_n\sum_{n'}\mathbb{E}\left[s_n s_{n'}\mathbb{E}\left[(g_n-\mu)(g_{n'}-\mu)\mid s_n,\bar{\varepsilon}_n,s_{n'},\bar{\varepsilon}_{n'}\right]\bar{\varepsilon}_n\bar{\varepsilon}_{n'}\right] \\
&\leq B_L\sum_n\sum_{n'} f\left(|n'-n|\right)\mathbb{E}\left[|s_n\bar{\varepsilon}_n s_{n'}\bar{\varepsilon}_{n'}|\right] \\
&= B_L\left(\sum_n\mathbb{E}\left[\left(s_n\bar{\varepsilon}_n\right)^2\right]f(0)+2\sum_{r=1}^{N-1}\sum_{n=1}^{N-r}\mathbb{E}\left[|s_{n+r}\bar{\varepsilon}_{n+r}|\cdot|s_n\bar{\varepsilon}_n|\right]f(r)\right) \\
&\leq \left(B_L\sum_n\mathbb{E}\left[s_n^2\mathbb{E}\left[\bar{\varepsilon}_n^2\mid s_n\right]\right]\right)\left(f(0)+2\sum_{r=1}^{N-1}f(r)\right) \\
&\leq B_\varepsilon\left(f(0)+2\sum_{r=1}^{N-1}f(r)\right)\left(B_L\mathbb{E}\left[\sum_n s_n^2\right]\right)\to 0,
\end{aligned} \tag{74}
$$

provided $\mathbb{E}\left[\bar{\varepsilon}_n^2\mid s_n\right]<B_\varepsilon$ uniformly. Here the second-to-last line follows because for any sequence of numbers $a_1,\ldots,a_N$ and any $r>0$,

$$
\begin{aligned}
\sum_n a_n^2 &\geq \frac{1}{2}\left(\sum_{n=1}^{N-r}a_n^2+\sum_{n=1}^{N-r}a_{n+r}^2\right) \\
&= \frac{1}{2}\sum_{n=1}^{N-r}\left(a_n-a_{n+r}\right)^2+\sum_{n=1}^{N-r}a_n a_{n+r} \\
&\geq \sum_{n=1}^{N-r}a_n a_{n+r},
\end{aligned} \tag{75}
$$

and the same is true in expectation if $a_n=|s_n\bar{\varepsilon}_n|$ are random variables. We note that allowing B_L to grow in the asymptotic sequence imposes much weaker conditions on the correlation structure of shocks. For example, with shock importance weights s_n approximately equal, i.e. $\sum_n s_n^2=O_p\left(1/N\right)$, it is enough to have $|\operatorname{Cov}\left[g_n,g_{n'}\mid\bar{\varepsilon}_n,\bar{\varepsilon}_{n'},s_n,s_{n'}\right]|\leq B_1N^{-\alpha}$ for any $\alpha>0$: in this case one can satisfy Assumption 6 by setting $B_L=B_1N^{1-\alpha/2}$ and $f(r)=r^{-1-\alpha/2}$.

B.2 Proposition 4

This section proves Proposition 4, again allowing for a growing number of controls (as in the above proof) under the following regularity conditions: that $\|\hat{\gamma}-\gamma_L\|_1=o_p(1)$, $\max_m|\sum_\ell e_\ell w_{\ell m}w_\ell^{*\prime}\mu|=O_p(1)$, and $\max_m|\sum_\ell e_\ell w_{\ell m}z_\ell|=O_p(1)$. Note that these are the same as in the proof of Proposition 3 when q_n just includes a constant (so $w_\ell^*=1$), and Proposition 2 can be similarly shown to hold.

As in the proof to Proposition 3, we first note that

$$\sum_n s_n g_n \bar{\varepsilon}_n = \sum_n s_n \left(g_n - q_n' \mu\right) \bar{\varepsilon}_n + \left(\sum_n s_n q_n' \bar{\varepsilon}_n\right) \mu. \tag{76}$$

By a straightforward extension of the proof to Proposition 2, the first term of this expression is $o_p(1)$ when $\text{Var}\left[g_n^* \mid \bar{\varepsilon}_n, q_n, s_n\right]$ and $\mathbb{E}\left[\bar{\varepsilon}_n^2 \mid q_n, s_n\right]$ are uniformly bounded and Assumption 4 holds. Moreover, when w_ℓ^* is an included control $\sum_\ell e_\ell w_\ell^{*\prime} \bar{\varepsilon}_\ell^\perp = 0$, such that

$$\begin{aligned}
\left(\sum_n s_n q_n' \bar{\varepsilon}_n\right) \mu &= \left(\sum_\ell e_\ell w_\ell^{*\prime} \varepsilon_\ell\right) \mu \\
&= \left(\sum_\ell e_\ell w_\ell^{*\prime} \left(\varepsilon_\ell - \bar{\varepsilon}_\ell^\perp\right)\right) \mu \\
&= \left(\hat{\gamma} - \gamma_L\right)' \left(\sum_\ell e_\ell w_\ell w_\ell^{*\prime}\right) \mu \xrightarrow{p} 0,
\end{aligned} \tag{77}$$

by the assumed regularity conditions.

B.3 Proposition 5

This section first shows that shock-level IV coefficients obtained from estimating (12) are numerically equivalent to the SSIV estimate $\hat{\beta}$, and that its heteroskedasticity-robust standard error is numerically equivalent to the baseline IV standard error of Adão et al. (2019b) when there are no controls in w_ℓ. We then prove that these two standard errors asymptotically coincide more generally under the assumptions that Adão et al. (2019b) use for valid conditional SSIV inference, and that the shock-level heteroskedasticity-robust standard error remains valid without the conditions of $L > N$ and non-collinear exposure shares which Adão et al. (2019b) require. We conclude with a discussion of the likely finite-sample difference between the two standard error formulas.

To establish the equivalence of IV coefficients, note that when $\sum_n s_{\ell n} q_n$ is included in w_ℓ

$$\sum_n s_n q_n \bar{y}_n^\perp = \sum_\ell e_\ell y_\ell^\perp \left(\sum_n s_{\ell n} q_n\right) = 0 \tag{78}$$

and similarly for $\sum_n s_n q_n \bar{x}_n^\perp$. The s_n-weighted regression of $\bar{y}_n^\perp$ and $\bar{x}_n^\perp$ on q_n thus produces a coefficient vector that is numerically zero, implying the s_n-weighted and g_n-instrumented regression of $\bar{y}_n^\perp$ on $\bar{x}_n^\perp$ is unchanged with the addition of q_n controls. Proposition 1 shows that the IV coefficient from this regression is equivalent to the SSIV estimate $\hat{\beta}$.

To establish standard error equivalence in the case without controls, note that the conventional heteroskedasticity-robust standard error from for the s_n-weighted shock-level IV regression of $\bar{y}_n^\perp$ on

$\bar{x}_n^\perp$ and a constant, instrumented by g_n, is given by

$$\widehat{se}_{\text{equiv}} = \frac{\sqrt{\sum_n s_n^2 \hat{\varepsilon}_n^2 \hat{g}_n^2}}{\left|\sum_n s_n \bar{x}_n^\perp g_n\right|}, \tag{79}$$

where $\hat{\varepsilon}_n = \bar{y}_n^\perp - \hat{\beta}\bar{x}_n^\perp$ is the estimated shock-level regression residual (where we used the fact that the estimated constant in that regression is numerically zero) and $\hat{g}_n = g_n - \sum_n s_n g_n$ is the s_n-weighted demeaned shock. By Proposition 1 $\hat{\varepsilon}_n$ coincides with the share-weighted aggregate of the SSIV estimated residuals $\hat{\varepsilon}_\ell = y_\ell^\perp - \hat{\beta} x_\ell^\perp$:

$$\hat{\varepsilon}_n = \frac{\sum_\ell e_\ell s_{\ell n} y_\ell^\perp}{\sum_\ell e_\ell s_{\ell n}} - \hat{\beta} \cdot \frac{\sum_\ell e_\ell s_{\ell n} x_\ell^\perp}{\sum_\ell e_\ell s_{\ell n}} = \frac{\sum_\ell e_\ell s_{\ell n} \hat{\varepsilon}_\ell}{\sum_\ell e_\ell s_{\ell n}}. \tag{80}$$

The squared numerator of (79) can thus be rewritten

$$\sum_n s_n^2 \hat{\varepsilon}_n^2 \hat{g}_n^2 = \sum_n \left(\sum_\ell e_\ell s_{\ell n} \hat{\varepsilon}_\ell\right)^2 \hat{g}_n^2. \tag{81}$$

The expression in the denominator of (79) estimates the magnitude of the shock-level first-stage covariance, which matches the e_ℓ-weighted sample covariance of x_ℓ and z_ℓ:

$$\sum_n s_n \bar{x}_n^\perp g_n = \sum_n \left(\sum_\ell e_\ell s_{\ell n} x_\ell^\perp\right) g_n = \sum_\ell e_\ell x_\ell^\perp z_\ell. \tag{82}$$

Thus

$$\widehat{se}_{\text{equiv}} = \frac{\sqrt{\sum_n \left(\sum_\ell e_\ell s_{\ell n} \hat{\varepsilon}_\ell\right)^2 \hat{g}_n^2}}{\left|\sum_\ell e_\ell x_\ell^\perp z_\ell\right|}. \tag{83}$$

We now compare this expression to the standard error formula from Adão et al. (2019b). Absent controls and incorporating the e_ℓ importance weights, equation (39) in their paper yields

$$\widehat{se}_{\text{AKM}} = \frac{\sqrt{\sum_n \left(\sum_\ell e_\ell s_{\ell n} \hat{\varepsilon}_\ell\right)^2 \ddot{g}_n^2}}{\left|\sum_\ell e_\ell x_\ell^\perp z_\ell\right|}, \tag{84}$$

where $\ddot{g}_n$ denotes the coefficients from regressing the demeaned instrument $z_\ell - \sum_\ell e_\ell z_\ell$ on all shares $s_{\ell n}$, without a constant. It thus remains to show that $\ddot{g}_n = \hat{g}_n$. Note that

$$\sum_\ell e_\ell z_\ell = \sum_\ell e_\ell \sum_n s_{\ell n} g_n = \sum_n s_n g_n, \tag{85}$$

so that, with $\sum_n s_{\ell n} = 1$,

$$z_\ell - \sum_\ell e_\ell z_\ell = \sum_\ell s_{\ell n} g_n - \sum_n s_n g_n = \sum_\ell s_{\ell n} \hat{g}_n. \tag{86}$$

This means that the auxiliary regression in Adão et al. (2019b) has exact fit and produces $\ddot{g}_n = \hat{g}_n$.

With controls, $\widehat{se}_{\text{equiv}}$ and $\widehat{se}_{\text{AKM}}$ differ only in respect to the construction of shock residuals. The former sets $\hat{g}_n = g_n - q_n'\hat{\mu}$ where $\hat{\mu} = \left(\sum_n s_n q_n q_n'\right)^{-1} \sum_n s_n q_n g_n$ is the coefficient vector from the auxiliary projection of the instrument in equation (12) on the control vector q_n. The Adão et al. (2019b) standard error formula sets $\ddot{g}_n$ to the N coefficients from regressing the residualized instrument $z_\ell^\perp$ on all shares $s_{\ell n}$, without a constant; note that to compute this requires $L > N$ and that the matrix of exposure shares $s_{\ell n}$ is full rank.

We establish the general asymptotic equivalence of $\widehat{se}_{\text{equiv}}$ and $\widehat{se}_{\text{AKM}}$ under assumptions which mirror those developed by Adão et al. (2019b) to show that $\widehat{se}^2_{\text{AKM}}$ captures the conditional asymptotic variance of $\hat{\beta}$. This variance is conditioned on $\mathcal{I}_L = \left\{\{q_n\}_n, \{u_\ell, \epsilon_\ell, \eta_\ell, \{s_{\ell n}, \pi_{\ell n}\}_n, e_\ell\}_\ell\right\}$, where $w_\ell = \left[\sum_n s_{\ell n} q_n', u_\ell'\right]'$ and $x_\ell = \sum_n s_{\ell n}\pi_{\ell n} g_n + \eta_\ell$; we note that the resulting confidence intervals are asymptotically valid unconditionally, since if $Pr(\beta \in \widehat{CI} \mid \mathcal{I}_L) = \alpha$ then $Pr(\beta \in \widehat{CI}) = \mathbb{E}\left[\mathbb{E}\left[\mathbf{1}[\beta \in \widehat{CI} \mid \mathcal{I}_{\mathcal{L}}\right]\right] = \alpha$ by the law of iterated expectations. We follow Adão et al. (2019b) in implicitly treating the set of shares $s_{\ell n}$ (and, for us, importance weights e_ℓ) as non-stochastic. Along with Assumption 7, we consider two additional conditions:

Assumption A1: The g_n are mutually independent given $\mathcal{I}_L$, $\max_n s_n \to 0$, and $\max_n \frac{s_n^2}{\sum_{n'} s_{n'}^2} \to 0$;

Assumption A2: $\mathbb{E}\left[|g_n|^{4+v} \mid \mathcal{I}_L\right]$ is uniformly bounded for some $v > 0$ and $\sum_\ell e_\ell \sum_n s_{\ell n}^2 \text{Var}\left[g_n \mid \mathcal{I}_L\right] \pi_{\ell n} \neq 0$. The support of $\pi_{\ell n}$ is bounded, the fourth moments of ϵ_ℓ, η_ℓ, $\tilde{u}_\ell$, and $\tilde{q}_n$ exist and are uniformly bounded, and $\sum_\ell e_\ell w_\ell w_\ell' \xrightarrow{p} \Omega_{ww}$ for positive definite Ω_{ww}. For $\gamma = \mathbb{E}\left[\sum_\ell e_\ell w_\ell w_\ell'\right]^{-1} \mathbb{E}\left[\sum_\ell e_\ell w_\ell \epsilon_\ell\right]$ and $\hat{\gamma} = \left(\sum_\ell e_\ell w_\ell w_\ell'\right)^{-1} \sum_\ell e_\ell w_\ell \epsilon_\ell$, $\hat{\gamma} \xrightarrow{p} \gamma$.

Here Assumption A1 corresponds to Assumption 2 in Adão et al. (2019b), and Assumption A2 includes the relevant conditions from their Assumptions 4 and A.3. Together, Assumptions 7, A1, and A2 imply our Assumptions 3 and 4. Under these assumptions, Proposition A.1 of Adão et al. (2019b) shows that

$$\sqrt{r_L}(\hat{\beta} - \beta) = \mathcal{N}\left(0, r_L \frac{\mathcal{V}_L}{\pi^2}\right) + o_p(1) \tag{87}$$

where $r_L = 1/\left(\sum_n s_n^2\right)$ and $\mathcal{V}_L = \sum_n \left(\sum_\ell e_\ell s_{\ell n} \varepsilon_\ell\right)^2 \text{Var}\left[g_n \mid \mathcal{I}_L\right]$, provided $r_L \mathcal{V}_L$ converges to a non-random limit. To establish the asymptotic validity of $\widehat{se}_{\text{AKM}}$, i.e. that $r_L \left(\sum_n \left(\sum_\ell e_\ell s_{\ell n} \hat{\varepsilon}_\ell\right)^2 \ddot{g}_n^2 - \mathcal{V}_L\right) \xrightarrow{p} 0$, Adão et al. (2019b) further assume that $L \geq N$, the matrix of $s_{\ell n}$ is always full rank, and additional regularity conditions (see their Proposition 5).

We establish $r_L \left(\sum_n \left(\sum_\ell e_\ell s_{\ell n} \hat{\varepsilon}_\ell\right)^2 \hat{g}_n^2 - \mathcal{V}_L\right) \xrightarrow{p} 0$ for $\hat{g}_n = g_n - q_n'\hat{\mu}$, and thus the asymptotic validity of $\widehat{se}_{\text{equiv}}$ under Assumptions 7, A1, and A2, without imposing the additional regularity conditions in Adão et al. (2019b)'s Proposition 5 and thereby allowing for $N > L$ or for collinear exposure shares. To start, write $g_n^* = g_n - q_n'\mu$ and decompose

$$\begin{aligned} r_L\left(\sum_n\left(\sum_\ell e_\ell s_{\ell n}\hat{\varepsilon}_\ell\right)^2\hat{g}_n^2 - \mathcal{V}_L\right) =& r_L\left(\sum_n\left(\sum_\ell e_\ell s_{\ell n}\varepsilon_\ell\right)^2 g_n^{*2} - \mathcal{V}_L\right) \\ &+ r_L\sum_n\left(\left(\sum_\ell e_\ell s_{\ell n}\hat{\varepsilon}_\ell\right)^2 - \left(\sum_\ell e_\ell s_{\ell n}\varepsilon_\ell\right)^2\right)g_n^{*2} \\ &+ r_L\sum_n\left(\sum_\ell e_\ell s_{\ell n}\hat{\varepsilon}_\ell\right)^2\left(\hat{g}_n^2 - g_n^{*2}\right). \end{aligned} \tag{88}$$

Adão et al. (2019b) show that the second term of this expression is $o_p(1)$ under our assumptions, using the fact (their Lemma A.3, again generalized to include importance weights) that for a triangular array $\{A_{L1},\ldots,A_{LL},B_{L1},\ldots,B_{LL},C_{L1},\ldots,C_{LN_L}\}_{L=1}^\infty$ with $\mathbb{E}\left[A_{L\ell}^4 \mid \{\{s_{\ell' n}\}_n, e_{\ell'}\}_{\ell'}\right]$, $\mathbb{E}\left[B_{L\ell}^4 \mid \{\{s_{\ell' n}\}_n, e_{\ell'}\}_{\ell'}\right]$, and $\mathbb{E}\left[C_{Ln}^4 \mid \{\{s_{\ell' n}\}_n, e_{\ell'}\}_{\ell'}\right]$ uniformly bounded,

$$r_L\sum_\ell\sum_{\ell'}\sum_n e_\ell e_{\ell'} s_{\ell n}s_{\ell' n}A_{L\ell}B_{L\ell'}C_{Ln} = O_p(1). \tag{89}$$

Here with $D_\ell = (z_\ell, w_\ell')'$, $\theta = (\beta,\gamma')'$, and $\hat{\theta} = (\hat{\beta},\hat{\gamma}')'$ we can write

$$\begin{aligned} \left(\sum_\ell e_\ell s_{\ell n}\hat{\varepsilon}_\ell\right)^2 =& \left(\sum_\ell e_\ell s_{\ell n}\varepsilon_\ell\right)^2 + 2\sum_\ell\sum_{\ell'} e_\ell e_{\ell'} s_{\ell n}s_{\ell' n}D_\ell'\left(\theta - \hat{\theta}\right)\varepsilon_{\ell'} \\ &+ \sum_\ell\sum_{\ell'} e_\ell e_{\ell'} D_\ell'(\theta-\hat{\theta})D_{\ell'}'(\theta-\hat{\theta}), \end{aligned} \tag{90}$$

and both D_ℓ and ε_ℓ have bounded fourth moments by the assumption of bounded fourth moments of $\epsilon_\ell, \eta_\ell, u_\ell$, and p_n in Assumption A2. Thus by the lemma

$$\begin{aligned} r_L\sum_n\left(\left(\sum_\ell e_\ell s_{\ell n}\hat{\varepsilon}_\ell\right)^2 - \left(\sum_\ell e_\ell s_{\ell n}\varepsilon_\ell\right)^2\right)g_n^{*2} =& 2\left(\theta-\hat{\theta}\right)'\left(r_L\sum_\ell\sum_{\ell'}\sum_n e_\ell e_{\ell'} s_{\ell n}s_{\ell' n}g_n^{*2}D_\ell\varepsilon_{\ell'}\right) \\ &+ \left(\theta-\hat{\theta}\right)'\left(r_L\sum_\ell\sum_{\ell'}\sum_n e_\ell e_{\ell'} s_{\ell n}s_{\ell' n}g_n^{*2}D_\ell D_{\ell'}'\right)\left(\theta-\hat{\theta}\right) \\ =& \left(\theta-\hat{\theta}\right)'O_p(1) + \left(\theta-\hat{\theta}\right)'O_p(1)\left(\theta-\hat{\theta}\right), \end{aligned} \tag{91}$$

which is $o_p(1)$ by the consistency of $\hat{\theta}$ (implied by Assumptions 7, A1, and A2). Adão et al. (2019b) further show the first term of equation (88) is $o_p(1)$, without using the additional regularity conditions of Proposition 5.

It thus remains for us to show the third term of (88) is also $o_p(1)$. Note that

$$\hat{g}_n^2 = \left(g_n - q_n'\hat{\mu}\right)^2 = g_n^{*2} + \left(q_n'\left(\hat{\mu}-\mu\right)\right)^2 - 2g_n^*q_n'\left(\hat{\mu}-\mu\right), \tag{92}$$

so that

$$
\begin{aligned}
& r_L \sum_n \left(\sum_\ell e_\ell s_{\ell n} \hat{\varepsilon}_\ell \right)^2 \left(\hat{g}_n^2 - g_n^{*2} \right) \\
=& r_L \sum_n \left(\sum_\ell e_\ell s_{\ell n} \hat{\varepsilon}_\ell \right)^2 \left(q_n' \left(\hat{\mu} - \mu \right) - 2 g_n^* \right) q_n' \left(\hat{\mu} - \mu \right) \\
=& r_L \sum_n \left(\sum_\ell e_\ell s_{\ell n} \varepsilon_\ell \right)^2 \left(q_n' \left(\hat{\mu} - \mu \right) - 2 g_n^* \right) q_n' \left(\hat{\mu} - \mu \right) \\
& + r_L \sum_n \left(\left(\sum_\ell e_\ell s_{\ell n} \hat{\varepsilon}_\ell \right)^2 - \left(\sum_\ell e_\ell s_{\ell n} \varepsilon_\ell \right)^2 \right) \left(q_n' \left(\hat{\mu} - \mu \right) - 2 g_n^* \right) q_n' \left(\hat{\mu} - \mu \right).
\end{aligned} \tag{93}
$$

Using the previous lemma, the first term of this expression is $O_p(1)(\hat{\mu} - \mu)$ since ε_ℓ, p_n, and g_n^* have bounded fourth moments under Assumption A2. The second term is similarly $O_p(1)(\hat{\mu} - \mu)$ by the lemma and the decomposition used in equation (91). Noting that $\hat{\mu} - \mu = \left(\sum_n s_n q_n q_n' \right)^{-1} \sum_n s_n q_n g_n^* \xrightarrow{p} 0$ under the assumptions completes the proof.

This characterization also suggests that in finite samples it is likely that $\widehat{se}_{\text{equiv}} < \widehat{se}_{\text{AKM}}$. To see this, consider versions of the two standard error formulas obtained under shock homoskedasticity (i.e. $\text{Var}\left[g_n \mid \mathcal{I}_L \right] = \sigma_g^2$):

$$
\hat{se}_{\text{equiv}} = \frac{\sqrt{\left(\sum_n s_n^2 \hat{\varepsilon}_n^2 \right) \left(\sum_n s_n \hat{g}_n^2 \right)}}{\left| \sum_n s_n \bar{x}_n^\perp g_n \right|} \tag{94}
$$

$$
\hat{se}_{\text{AKM}} = \frac{\sqrt{\left(\sum_n s_n^2 \hat{\varepsilon}_n^2 \right) \left(\sum_n s_n \ddot{g}_n^2 \right)}}{\left| \sum_\ell e_\ell x_\ell^\perp z_\ell \right|}, \tag{95}
$$

which differ by a factor of $\sqrt{\sum_n s_n \hat{g}_n^2 / \sum_n s_n \ddot{g}_n^2}$. When the SSIV controls have an exact shift-share structure, $w_\ell = \sum_n s_{\ell n} q_n$, the share projection producing $\ddot{g}_n$ has exact fit such that one can represent $\ddot{g}_n = g_n - q_n' \hat{\mu}_{AKM}$ for some $\hat{\mu}_{AKM}$. In this case the s_n-weighted sum of squares of shock residuals is lower in our equivalent regression by construction of $\hat{\mu}$: $\sum_n s_n \hat{g}_n^2 \leq \sum_n s_n \ddot{g}_n^2$ (with strict inequality when $\hat{\mu}_{\text{AKM}} \neq \hat{\mu}$). Similarly, when w_ℓ instead contains controls that are included for efficiency only and are independent of the shocks, projection of z_ℓ on the shares produces a noisy estimate of $g_n - \bar{g}$, which again has a higher sum of squares. In practice, we find that the heteroskedastic standard errors of Adão et al. (2019b) are also larger in the application in Section 5.

B.4 Proposition A1

We consider each expectation in equation (15) in turn. For each n, write

$$
\kappa_n(g_{-n}, \varepsilon_\ell, \eta_\ell) = \lim_{g_n \to -\infty} y(x_1([g_n; g_{-n}], \eta_{\ell 1}), \ldots, x_M([g_n; g_{-n}], \eta_{\ell M}), \varepsilon_\ell) \tag{96}
$$

such that

$$
\begin{aligned}
s_{\ell n} e_{\ell} g_n y_{\ell} = & s_{\ell n} e_{\ell} g_n \kappa_n(g_{-n}, \varepsilon_{\ell}, \eta_{\ell}) \\
& + s_{\ell n} e_{\ell} g_n \int_{-\infty}^{g_n} \frac{\partial}{\partial g_n} y(x_1([\gamma; g_{-n}], \eta_{\ell 1}), \ldots x_M([\gamma; g_{-n}], \eta_{\ell M}), \varepsilon_{\ell}) d\gamma.
\end{aligned} \tag{97}
$$

By as-good-as-random shock assignment, the expectation of the first term is

$$
\mathbb{E}\left[s_{\ell n} e_{\ell} g_n \kappa_n(g_{-n}, \varepsilon_{\ell}, \eta_{\ell})\right] = \mathbb{E}\left[s_{\ell n} e_{\ell} \mathbb{E}\left[g_n \mid s, e, g_{-n}, \varepsilon, \eta_{\ell}\right] \kappa_n(g_{-n}, \varepsilon_{\ell}, \eta_{\ell})\right] = 0, \tag{98}
$$

while the expectation of the second is

$$
\begin{aligned}
& \mathbb{E}\left[s_{\ell n} e_{\ell} g_n \int_{-\infty}^{g_n} \frac{\partial}{\partial g_n} y(x_1([\gamma; g_{-n}], \eta_{\ell 1}), \ldots x_M([\gamma; g_{-n}], \eta_{\ell M}), \varepsilon_{\ell}) d\gamma\right] \\
& = \mathbb{E}\left[s_{\ell n} e_{\ell} \int_{-\infty}^{\infty} \int_{-\infty}^{g_n} g_n \frac{\partial}{\partial g_n} y(x_1([\gamma; g_{-n}], \eta_{\ell 1}), \ldots x_M([\gamma; g_{-n}], \eta_{\ell M}), \varepsilon_{\ell}) d\gamma dF_n(g_n \mid \mathcal{I})\right] \\
& = \mathbb{E}\left[s_{\ell n} e_{\ell} \int_{-\infty}^{\infty} \frac{\partial}{\partial g_n} y(x_1([\gamma; g_{-n}], \eta_{\ell 1}), \ldots x_M([\gamma; g_{-n}], \eta_{\ell M}), \varepsilon_{\ell}) \int_{\gamma}^{\infty} g_n dF_n(g_n \mid \mathcal{I}) d\gamma\right]
\end{aligned} \tag{99}
$$

where $F_n(\cdot \mid \mathcal{I})$ denotes the conditional distribution of g_n. Thus

$$
\begin{aligned}
\mathbb{E}\left[s_{\ell n} e_{\ell} g_n y_{\ell}\right] & = \mathbb{E}\left[s_{\ell n} e_{\ell} \int_{-\infty}^{\infty} \frac{\partial}{\partial g_n} y(x_1([\gamma; g_{-n}], \eta_{\ell}), \ldots x_M([\gamma; g_{-n}], \eta_{\ell}), \varepsilon_{\ell}) \mu_n(\gamma \mid \mathcal{I}) d\gamma\right] \\
& = \sum_m \mathbb{E}\left[\int_{-\infty}^{\infty} s_{\ell n} e_{\ell} \alpha_{\ell m} \pi_{\ell m n}([\gamma; g_{-n}]) \mu_n(\gamma \mid \mathcal{I}) \tilde{\beta}_{\ell m n}(\gamma) d\gamma\right]
\end{aligned} \tag{100}
$$

where

$$
\begin{aligned}
\mu_n(\gamma \mid \mathcal{I}) \equiv & \int_{\gamma}^{\infty} g_n dF_n(g_n \mid \mathcal{I}). \\
= & \left(\mathbb{E}\left[g_n \mid g_n \geq \gamma, \mathcal{I}\right] - \mathbb{E}\left[g_n \mid g_n < \gamma, \mathcal{I}\right]\right) Pr\left(g_n \geq \gamma \mid \mathcal{I}\right)\left(1 - Pr\left(g_n \geq \gamma \mid \mathcal{I}\right)\right) \geq 0 \ a.s.
\end{aligned} \tag{101}
$$

Similarly

$$
\mathbb{E}\left[s_{\ell n} e_{\ell} g_n \alpha_{\ell m} x_{\ell m}\right] = \sum_m \mathbb{E}\left[\int_{-\infty}^{\infty} s_{\ell n} e_{\ell} \alpha_{\ell m} \pi_{\ell m n}([\gamma; g_{-n}]) \mu_n(\gamma \mid \mathcal{I}) d\gamma\right]. \tag{102}
$$

Combining equations (100) and (102) completes the proof, with

$$
\omega_{\ell m n}(\gamma) = s_{\ell n} e_{\ell} \alpha_{\ell m} \mu_n(\gamma \mid \mathcal{I}) \pi_{\ell m n}([\gamma; g_{-n}]) \geq 0 \ a.s. \tag{103}
$$

B.5 Proposition A2

To prove (21), we aggregate (20) across industries within a region using $E_{\ell n}$ weights:

$$y_\ell = (\beta_0 - \beta_1)\, x_\ell + \varepsilon_\ell, \tag{104}$$

where $\varepsilon_\ell = \sum_n s_{\ell n}\varepsilon_{\ell n}$. The shift-share instrument z_ℓ is relevant because

$$\begin{aligned}\mathbb{E}\left[\sum_\ell e_\ell x_\ell z_\ell\right] &= \sum_\ell e_\ell \mathbb{E}\left[\sum_n s_{\ell n}\left(\bar{\pi} g_n + \eta_{\ell n}\right)\cdot \sum_{n'} s_{\ell n'} g_{n'}\right] \\ &= \sum_{\ell,n} e_\ell s_{\ell n}^2 \bar{\pi}\sigma_g^2 \\ &\geq \bar{H}_L \bar{\pi}\sigma_g^2,\end{aligned} \tag{105}$$

while exclusion holds because

$$\begin{aligned}\mathbb{E}\left[\sum_\ell e_\ell z_\ell \varepsilon_\ell\right] &= \sum_\ell e_\ell \mathbb{E}\left[\sum_n s_{\ell n}\varepsilon_{\ell n} \cdot \sum_{n'} s_{\ell n'} g_{n'}\right] \\ &= 0.\end{aligned} \tag{106}$$

Thus by an appropriate law of large numbers, $\hat{\beta} = \beta_0 - \beta_1 + o_p(1)$.

To study $\hat{\beta}_{\text{ind}}$, we aggregate (20) across regions (again with $E_{\ell n}$ weights):

$$y_n = \beta_0 x_n - \beta_1 \sum_\ell \omega_{\ell n} \sum_{n'} s_{\ell n'} x_{\ell n'} + \varepsilon_n, \tag{107}$$

for $\varepsilon_n = \sum_\ell \omega_{\ell n}\varepsilon_{\ell n}$. The resulting IV estimate yields

$$\begin{aligned}\hat{\beta}_{\text{ind}} - \beta_0 &= \frac{\sum_n s_n y_n g_n}{\sum_n s_n x_n g_n} - \beta_0 \\ &= \frac{\sum_n s_n \left(-\beta_1 \sum_\ell \omega_{\ell n} \sum_{n'} s_{\ell n'} x_{\ell n'} + \varepsilon_n\right) g_n}{\sum_n s_n x_n g_n}.\end{aligned} \tag{108}$$

The expected denominator of $\hat{\beta}_{\text{ind}}$ is non-zero:

$$\begin{aligned}\mathbb{E}\left[\sum_n s_n x_n g_n\right] &= \sum_n s_n \mathbb{E}\left[\sum_\ell \omega_{\ell n}\left(\bar{\pi} g_n + \eta_{\ell n}\right) g_n\right] \\ &= \sum_n s_n \omega_{\ell n} \bar{\pi}\sigma^2 \\ &= \sum_n \frac{E_n}{E}\cdot\frac{E_{\ell n}}{E}\bar{\pi}\sigma^2 \\ &= \bar{\pi}\sigma^2,\end{aligned} \tag{109}$$

while the expected numerator is

$$\mathbb{E}\left[\sum_n s_n \left(-\beta_1 \sum_\ell \omega_{\ell n} \sum_{n'} s_{\ell n'} x_{\ell n'} + \varepsilon_n\right) g_n\right] = -\beta_1 \sum_{n,\ell} s_n \omega_{\ell n} s_{\ell n} \bar{\pi} \sigma^2$$
$$= -\beta_1 H_L \bar{\pi} \sigma^2, \tag{110}$$

where the last equality follows because

$$\begin{aligned}\sum_{n,\ell} s_n \omega_{\ell n} s_{\ell n} &= \sum_{n,\ell} \frac{E_n}{E} \frac{E_{\ell n}}{E_n} \frac{E_{\ell n}}{E_\ell} \\ &= \sum_{n,\ell} \frac{E_\ell}{E} \left(\frac{E_{\ell n}}{E_\ell}\right)^2 \\ &= \sum_{n,\ell} e_\ell s_{\ell n}^2 \\ &= H_L. \end{aligned} \tag{111}$$

Thus by an appropriate law of large numbers,

$$\hat{\beta}_{\text{ind}} = \beta_0 - \beta_1 H_L + o_p(1). \tag{112}$$

B.6 Proposition A3

By the appropriate law of large numbers,

$$\begin{aligned}\hat{\beta} &= \frac{\mathbb{E}\left[\sum_\ell E_\ell \left(\sum_n s_{\ell n} y_{\ell n}\right)\left(\sum_{n'} s_{\ell n'} g_{n'}\right)\right]}{\mathbb{E}\left[\sum_\ell E_\ell \left(\sum_n s_{\ell n} x_{\ell n}\right)\left(\sum_{n'} s_{\ell n'} g_{n'}\right)\right]} + o_p(1) \\ &= \frac{\sum_{\ell,n} E_\ell s_{\ell n}^2 \pi_{\ell n} \sigma_n^2 \beta_{\ell n}}{\sum_{\ell,n} E_\ell s_{\ell n}^2 \pi_{\ell n} \sigma_n^2} + o_p(1) \\ &= \frac{\sum_{\ell,n} E_{\ell n} s_{\ell n} \pi_{\ell n} \sigma_n^2 \beta_{\ell n}}{\sum_{\ell,n} E_{\ell n} s_{\ell n} \pi_{\ell n} \sigma_n^2} + o_p(1) \end{aligned} \tag{113}$$

while

$$\begin{aligned}\hat{\beta}_{\text{ind}} &= \frac{\sum_n E_n y_n g_n}{\sum_n E_n x_n g_n} \\ &= \frac{\mathbb{E}\left[\sum_n E_n \left(\sum_\ell \omega_{\ell n} y_{\ell n}\right) g_n\right]}{\mathbb{E}\left[\sum_n E_n \left(\sum_\ell \omega_{\ell n} x_{\ell n}\right) g_n\right]} + o_p(1) \\ &= \frac{\sum_{\ell,n} E_n \omega_{\ell n} \pi_{\ell n} \sigma_n^2 \beta_{\ell n}}{\sum_{\ell,n} E_n \omega_{\ell n} \pi_{\ell n} \sigma_n^2} + o_p(1) \\ &= \frac{\sum_{\ell,n} E_{\ell n} \pi_{\ell n} \sigma_n^2 \beta_{\ell n}}{\sum_{\ell,n} E_{\ell n} \pi_{\ell n} \sigma_n^2} + o_p(1). \end{aligned} \tag{114}$$

B.7 Proposition A4

By definition of $\bar{\varepsilon}_n$,

$$\begin{aligned}\bar{\varepsilon}_n &= \frac{\sum_\ell e_\ell s_{\ell n}\left(\sum_{n'} s_{\ell n'}\nu_{n'} + \check{\varepsilon}_\ell\right)}{\sum_\ell e_\ell s_{\ell n}} \\ &\equiv \sum_{n'} \alpha_{nn'}\nu_{n'} + \bar{\check{\varepsilon}}_n, \end{aligned} \tag{115}$$

for $\alpha_{nn'} = \frac{\sum_\ell e_\ell s_{\ell n} s_{\ell n'}}{\sum_\ell e_\ell s_{\ell n}}$ and $\bar{\check{\varepsilon}}_n = \frac{\sum_\ell e_\ell s_{\ell n}\check{\varepsilon}_\ell}{\sum_\ell e_\ell s_{\ell n}}$. Therefore,

$$\begin{aligned}\mathrm{Var}\left[\bar{\varepsilon}_n\right] &= \sum_{n'} \sigma^2_{n'}\alpha^2_{nn'} + \mathrm{Var}\left[\bar{\check{\varepsilon}}_n\right] \\ &\geq \sigma^2_\nu \alpha^2_{nn}, \end{aligned} \tag{116}$$

and

$$\max_n \mathrm{Var}\left[\bar{\varepsilon}_n\right] \geq \sigma^2_\nu \max_n \alpha^2_{nn}. \tag{117}$$

To establish a lower bound on this quantity, observe that the s_n-weighted average of α_{nn} satisfies:

$$\begin{aligned}\sum_n s_n \alpha_{nn} &= \sum_n s_n \frac{\sum_\ell e_\ell s^2_{\ell n}}{s_n} \\ &= H_L. \end{aligned} \tag{118}$$

Since $\sum_n s_n = 1$, it follows that $\max_n \alpha_{nn} \geq H_L$ and therefore $\max_n \mathrm{Var}\left[\bar{\varepsilon}_n\right] \geq \sigma^2_\nu H^2_L$. Since $H_L \to \bar{H} > 0$, we conclude that, for sufficiently large L, $\max_n \mathrm{Var}\left[\bar{\varepsilon}_n\right]$ is bounded from below by any positive $\delta < \sigma^2_\nu \bar{H}^2$.

B.8 Proposition A5

We prove each part of this proposition in turn.

1. Expanding the exclusion condition yields:

$$\begin{aligned}\mathbb{E}\left[\sum_\ell e_\ell \varepsilon_\ell \psi_{\ell,LOO}\right] &= \sum_\ell \mathbb{E}\left[e_\ell \varepsilon_\ell \sum_n s_{\ell n} \frac{\sum_{\ell' \neq \ell} \omega_{\ell' n}\psi_{\ell' n}}{\sum_{\ell' \neq \ell}\omega_{\ell' n}}\right] \\ &= \sum_\ell e_\ell \sum_n s_{\ell n} \frac{\sum_{\ell' \neq \ell} \omega_{\ell' n}\mathbb{E}\left[\varepsilon_\ell \psi_{\ell' n}\right]}{\sum_{\ell' \neq \ell}\omega_{\ell' n}} \\ &= 0. \end{aligned} \tag{119}$$

2. The assumption of part (1) is satisfied here, so $\mathbb{E}\left[\sum_\ell e_\ell \varepsilon_\ell \psi_{\ell,LOO}\right] = 0$. We now establish that $\mathbb{E}\left[\left(\sum_\ell e_\ell \varepsilon_\ell \psi_{\ell,LOO}\right)^2\right] \to 0$, which implies the exclusion condition $\sum_\ell e_\ell \varepsilon_\ell \psi_{\ell,LOO} \xrightarrow{p} 0$ and thus

consistency of the LOO SSIV estimator provided it has a first stage:

$$\begin{aligned}\mathbb{E}\left[\left(\sum_{\ell} e_\ell \varepsilon_\ell \psi_{\ell,LOO}\right)^2\right] &= \sum_{\substack{\ell_1,\ell_2,n_1,n_2,\\ \ell_1'\neq\ell_1,\ell_2'\neq\ell_2}} e_{\ell_1} e_{\ell_2} s_{\ell_1 n_1} s_{\ell_2 n_2} \frac{\omega_{\ell_1' n_1}}{\sum_{\ell\neq\ell_1}\omega_{\ell n_1}} \frac{\omega_{\ell_2' n_2}}{\sum_{\ell\neq\ell_2}\omega_{\ell n_2}} \mathbb{E}\left[\varepsilon_{\ell_1}\varepsilon_{\ell_2}\psi_{\ell_1' n_1}\psi_{\ell_2' n_2}\right] \\ &\leq \sum_{\substack{(\ell_1,\ell_2,\ell_1',\ell_2')\in\mathcal{J},\\ n_1,n_2}} e_{\ell_1} e_{\ell_2} s_{\ell_1 n_1} s_{\ell_2 n_2} \frac{\omega_{\ell_1' n_1}}{\sum_{\ell\neq\ell_1}\omega_{\ell n_1}} \frac{\omega_{\ell_2' n_2}}{\sum_{\ell\neq\ell_2}\omega_{\ell n_2}} \cdot B \to 0.\end{aligned} \tag{120}$$

Here the second line used the first regularity condition, which implies that $\mathbb{E}\left[\varepsilon_{\ell_1}\varepsilon_{\ell_2}\psi_{\ell_1' n_1}\psi_{\ell_2' n_2}\right] = 0$ whenever there is at least one index among $\{\ell_1,\ell_2,\ell_1',\ell_2'\}$ which is not equal to any of the others, i.e. for all $(\ell_1,\ell_2,\ell_1',\ell_2')\notin\mathcal{J}$.

3. We show that under the given assumptions on $s_{\ell n}$, e_ℓ, and $\omega_{\ell n}$, the expression in (37) is bounded by $4N/L$:

$$\begin{aligned}&\sum_{\substack{(\ell_1,\ell_2,\ell_1',\ell_2')\in\mathcal{J},\\ n_1,n_2}} e_{\ell_1} e_{\ell_2} s_{\ell_1 n_1} s_{\ell_2 n_2} \frac{\omega_{\ell_1' n_1}}{\sum_{\ell\neq\ell_1}\omega_{\ell n_1}} \frac{\omega_{\ell_2' n_2}}{\sum_{\ell\neq\ell_2}\omega_{\ell n_2}} \\ &= \sum_{(\ell_1,\ell_2,\ell_1',\ell_2')\in\mathcal{J}} \frac{1}{L^2} \frac{\omega_{\ell_1' n(\ell_1)}}{\sum_{\ell\neq\ell_1}\omega_{\ell n(\ell_1)}} \frac{\omega_{\ell_2' n(\ell_2)}}{\sum_{\ell\neq\ell_2}\omega_{\ell n(\ell_2)}} \\ &= \frac{1}{L^2} \sum_{\substack{(\ell_1,\ell_2,\ell_1',\ell_2')\in\mathcal{J}\\ n(\ell_1')=n(\ell_1),\\ n(\ell_2')=n(\ell_2)}} \frac{1}{L_{n(\ell_1)}-1} \frac{1}{L_{n(\ell_2)}-1} \\ &= \frac{1}{L^2}\sum_n \frac{2L_n\left(L_n-1\right)}{\left(L_n-1\right)^2} \leq 4\frac{N}{L}.\end{aligned} \tag{121}$$

Here the second line plugs in the expressions for $s_{\ell n}$ and e_ℓ, and the third line plugs in $\omega_{\ell n}$. The last line uses the fact that any tuple $(\ell_1,\ell_2,\ell_1',\ell_2')\in\mathcal{J}$ such that $n(\ell_1')=n(\ell_1)$ and $n(\ell_2')=n(\ell_2)$ has all four elements exposed to the same shock n. Moreover, it is easily verified that all of these tuples have a structure $(\ell_A,\ell_B,\ell_A,\ell_B)$ or $(\ell_A,\ell_B,\ell_B,\ell_A)$ for any $\ell_A\neq\ell_B$ exposed to the same shock. Therefore, there are $2L_n\left(L_n-1\right)$ of them for each n. Finally, $\frac{L_n}{L_n-1}\leq 2$ as $L_n\geq 2$.

B.9 Proposition A6

National industry employment satisfies $E_n = \sum_\ell E_{\ell n}$; log-linearizing this immediately implies (46). To solve for $g_{\ell n}$, log-linearize (43), (44), and (45):

$$\hat{E}_\ell = \phi \hat{W}_\ell + \varepsilon_\ell, \tag{122}$$

$$g_{\ell n} = g_n^* + \hat{\xi}_{\ell n} - \sigma \hat{W}_\ell, \tag{123}$$

$$\hat{E}_\ell = \sum_n s_{\ell n} g_{\ell n}. \tag{124}$$

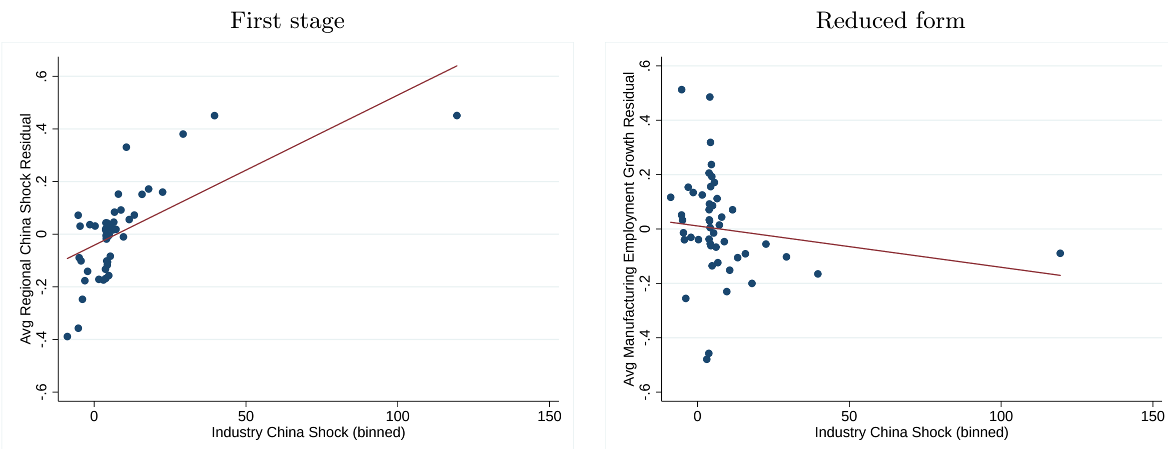
Solving this system of equations yields

$$\hat{W}_\ell = \frac{1}{\sigma + \phi} \left(\sum_n s_{\ell n} \left(g_n^* + \hat{\xi}_{\ell n} \right) - \varepsilon_\ell \right) \tag{125}$$

and expression (47).

C Appendix Figures and Tables

Figure C1: Industry-level variation in ADH



Notes: The figure shows binned scatterplots of shock-level outcome and treatment residuals, $\bar{y}_{nt}^{\perp}$ and $\bar{x}_{nt}^{\perp}$, corresponding to the SSIV specification in column 3 of Table 4. The shocks, g_{nt}, are residualized on year indicators (with the full-sample mean added back) and grouped into fifty weighted bins, with each bin representing around 2% of total share weight s_{nt}. Lines of best fit, indicated in red, are weighted by the same s_{nt}. The slope coefficients equal 5.71×10^{-3} and -1.52×10^{-3}, respectively, with the ratio (-0.267) equalling the SSIV coefficient in column 3 of Table 4.

Table C1: Shift-Share IV Estimates of the Effect of Chinese Imports on Other Outcomes

	(1)	(2)	(3)	(4)
Unemployment growth	0.221	0.217	0.063	0.100
	(0.049)	(0.046)	(0.060)	(0.083)
NILF growth	0.553	0.534	0.098	0.126
	(0.185)	(0.183)	(0.133)	(0.134)
Log weekly wage growth	-0.759	-0.607	0.227	0.133
	(0.258)	(0.226)	(0.242)	(0.281)
CZ-level controls ($w_{\ell t}$)				
Autor et al. (2013) baseline	✓	✓	✓	✓
Start-of-period mfg. share	✓			
Lagged mfg. share		✓	✓	✓
Period-specific lagged mfg. share			✓	✓
Period-specific lagged 10-sector shares				✓
# of industries-by-periods	796	794	794	794

Notes: This table extends the analysis of Table 4 to different regional outcomes in Autor et al. (2013): unemployment growth, labor force non-participation (NILF) growth, and log average weekly wage growth. The specifications in columns 1–4 are otherwise the same as those in Table 4. Exposure-robust standard errors are computed using equivalent industry-level IV regressions and allow for clustering of shocks at the level of three-digit SIC codes.

Table C2: Alternative Exposure-Robust Standard Errors

	Effects				Pre-trends	
	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	-0.596	-0.489	-0.267	-0.252	-0.028	0.142
Table 4 standard error (SE)	(0.114)	(0.100)	(0.099)	(0.136)	(0.092)	(0.090)
State-clustered SE	(0.099)	(0.086)	(0.086)	(0.095)	(0.105)	(0.111)
Adão et al. (2019b) SE	(0.126)	(0.116)	(0.113)	(0.152)	(0.125)	(0.121)
Confidence interval	[-1.018,	[-0.840,	[-0.584,	[-0.762,	[-0.349,	[-0.124,
with the null imposed	-0.362]	-0.273]	-0.009]	0.028]	0.287]	0.480]

Notes: This table extends the analysis of Table 4 by reporting conventional state-clustered SE, the Adão et al. (2019b) SIC3-clustered standard errors, and confidence intervals based on the equivalent industry-level IV regression with the null imposed, as discussed in Section 4.1. For comparison we repeat the coefficient estimates and exposure-robust standard errors from Table 4, also obtained by the equivalent shock-level regression.

Table C3: Period-Specific Effects in the Autor et al. (2013) Setting

	(1) Mfg. emp.	(2) Unemp.	(3) NILF	(4) Wages
Coefficient (1990s)	-0.491	0.329	1.209	-0.649
	(0.266)	(0.155)	(0.347)	(0.571)
Coefficient (2000s)	-0.225	0.014	-0.109	0.391
	(0.103)	(0.083)	(0.123)	(0.288)

Notes: This table estimates the shift-share IV regression from column 3 of Table 4 and Table C1, for different outcomes, allowing the treatment coefficient to vary by period. This specification uses two endogenous treatment variables (treatment interacted with period indicators) and two corresponding shift-share instruments. Controls include lagged manufacturing shares interacted with period dummies, period fixed effects, Census division fixed effects, and start-of-period conditions (% college educated, % foreign-born, % employment among women, % employment in routine occupations, and the average offshorability index). Exposure-robust standard errors, obtained by the equivalent shock-level regressions, allow for clustering at the level of three-digit SIC codes.

Table C4: Overidentified Shift-Share IV Estimates of the Effect of Chinese Imports on Manufacturing Employment

	(1)	(2)	(3)
Coefficient	-0.238	-0.247	-0.158
	(0.099)	(0.105)	(0.078)
Shock-level estimator	2SLS	LIML	GMM
Effective first stage F-statistic		15.10	
$\chi^2(7)$ overid. test stat. [p-value]		10.92 [0.142]	

Notes: Column 1 of this table reports an overidentified estimate of the coefficient corresponding to column 3 of Table 4, obtained with a two-stage least squares regression of shock-level average manufacturing employment growth residuals $\bar{y}^{\perp}_{nt}$ on shock-level average Chinese import competition growth residuals $\bar{x}^{\perp}_{nt}$, instrumenting by the growth of imports (per U.S. worker) in eight non-U.S. countries from ADH g_{nk}, controlling for period fixed effects q_{nt}, and weighting by average industry exposure s_{nt}. Column 2 reports the corresponding limited information maximum likelihood estimate, while column 3 reports a two-step optimal GMM estimate. Standard errors, the optimal GMM weight matrix, and the Hansen (1982) χ^2 test of overidentifying restrictions all allow for clustering of shocks at the SIC3 industry group level. The first-stage F-statistic is computed by a shift-share version of the Montiel Olea and Pflueger (2013) method described in Appendix A.9.

Table C5: Bartik (1991) Application

	(1)	(2)
Leave-one-out estimator	1.277	1.300
	(0.150)	(0.124)
Conventional estimator	1.215	1.286
	(0.139)	(0.121)
H heuristic	1.32	10.50
Population weights	✓	
Region-by-period observations	2,166	

Notes: Column 1 replicates column 2 of Table 3 from Goldsmith-Pinkham et al. (2019), reporting two SSIV estimators of the inverse labor supply elasticity, with and without the LOO adjustment. Regions are U.S. commuting zones; periods are 1980s, 1990s, and 2000s; all specifications include controls for 1980 regional characteristics interacted with period indicators (see Goldsmith-Pinkham et al. (2019) for more details). Standard errors allow for clustering by commuting zones. Column 1 uses 1980 population weights, while column 2 repeats the same analysis without population weights. The table also reports the H heuristic for the importance of the leave-one-out adjustment proposed in Appendix A.7 (equation (40)).

Table C6: Simulated 5% Rejection Rates for Shift-Share and Conventional Shock-Level IV

		SSIV Exposure-Robust SE		Shock-level IV Robust SE	
		Null not Imposed	Null Imposed	Null not Imposed	Null Imposed
		(1)	(2)	(3)	(4)
	Panel A: Benchmark Monte-Carlo Simulation				
(a)	Normal shocks	7.6%	5.2%	6.8%	5.0%
(b)	Wild bootstrap (benchmark)	8.0%	4.9%	14.2%	4.0%
	Panel B: Higher Industry Concentration				
(c)	$1/HHI = 50$	5.6%	4.9%	8.4%	6.1%
(d)	$1/HHI = 20$	7.3%	5.5%	7.0%	10.7%
(e)	$1/HHI = 10$	9.0%	8.2%	14.8%	23.8%
	Panel C: Smaller Numbers of Industries or Regions				
(f)	$N = 136$ (SIC3 industries)	5.4%	4.5%	7.7%	4.3%
(g)	$N = 20$ (SIC2 industries)	7.7%	3.7%	7.9%	3.2%
(h)	$L = 100$ (random regions)	9.7%	4.5%	N/A	
(i)	$L = 25$ (random regions)	10.4%	4.3%	N/A	

Notes: This table summarizes the result of the Monte-Carlo analysis described in Appendix A.10, reporting the rejection rates for a nominal 5% level test of the true null that $\beta^* = 0$. In all panels, columns 1 and 2 are simulated from the SSIV design based on Autor et al. (2013), as in column 3 of Table 4, while columns 3 and 4 are based on the conventional industry-level IV in Acemoglu et al. (2016). Column 1 uses exposure-robust standard errors from the equivalent industry-level IV and column 2 implements the version with the null hypothesis imposed. Columns 3 and 4 parallel columns 1 and 2 when applied to conventional IV. In Panel A, the simulations approximate the data-generating process using a normal distribution in row (a), with the variance matched to the sample variance of the shocks in the data after de-meaning by year, while wild bootstrap is used in row (b), following Liu (1988). Panel B documents the role of the Herfindahl concentration index across industries, varying 1/HHI from 50 to 10 in rows (c) to (e), compared with 191.6 for shift-share IV and 189.7 for conventional IV. Panel C documents the role of the number of regions and industries. We aggregate industries from 397 four-digit manufacturing SIC industries into 136 three-digit industries in row (f) and further into 20 two-digit industries in row (g). In rows (h) and (i), we select a random subset of region in each simulation. See Appendix A.10 for a complete discussion.

Table C7: First Stage F-statistics as a Rule of Thumb: Monte-Carlo Evidence

	Number of Instruments				
	1	5	10	25	50
	(1)	(2)	(3)	(4)	(5)
Panel A: SSIV					
5% rejection rate	8.0%	8.9%	11.5%	15.0%	23.0%
Median bias, % of std. dev.	0.3%	14.6%	28.3%	43.2%	72.1%
Median first-stage F	54.3	14.8	9.1	6.4	7.7
Panel B: Conventional Shock-Level IV					
5% rejection rate	13.6%	13.9%	14.9%	17.7%	22.0%
Median bias, % of std. dev.	-0.3%	10.1%	27.1%	57.0%	80.2%
Median first-stage F	59.4	19.4	13.2	10.0	11.2
Number of simulations	10,000	3,000	1,500	500	300

Notes: This table reports the result of the Monte-Carlo analysis with many weak instruments, described in Appendix A.10. Panel A is simulated from the SSIV design based on Autor et al. (2013), as in column 3 of Table 4, while Panel B is based on the conventional industry-level IV in Acemoglu et al. (2016). The five columns gradually increase the number of shocks $J = 1, 5, 10, 25$, and 50, with only one shock relevant to treatment. The table reports the rejection rates corresponding to a nominal 5% level test of the true null that $\beta^* = 0$, the median bias of the estimator as a percentage of the simulated standard deviation, and the median first-stage F-statistic obtained via the Montiel Olea and Pflueger (2013) method (extended to shift-share IV in Panel A, following Appendix A.9). See Appendix A.10 for a complete discussion.

References

Acemoglu, Daron, David H. Autor, David Dorn, Gordon H. Hanson, and Brendard Price. 2016. "Import Competition and the Great U.S . Employment Sag of the 2000s." *Journal of Labor Economics* 34:S141–S198.

Adão, Rodrigo, Costas Arkolakis, and Federico Esposito. 2019a. "Spatial Linkages, Global Shocks, and Local Labor Markets: Theory and Evidence." *NBER Working Paper 25544.*

Adão, Rodrigo, Michal Kolesár, and Eduardo Morales. 2019b. "Shift-Share Designs: Theory and Inference." *The Quarterly Journal of Economics* 134:1949–2010.

Angrist, Joshua D., Kate Graddy, and Guido W. Imbens. 2000. "The Interpretation of Instrumental Variables Estimators in Simultaneous Equations Models with an Application to the Demand for Fish." *The Review of Economic Studies* 67:499–527.

Angrist, Joshua D., Guido W. Imbens, and Alan B. Krueger. 1999. "Jackknife Instrumental Variables Estimation." *Journal of Applied Econometrics* 14:57–67.

Autor, David H., David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Effects of Import Competition in the United States." *American Economic Review* 103:2121–2168.

Bartik, Timothy J. 1991. *Who Benefits from State and Local Economic Development Policies?* Kalamazoo, MI: W. E. Upjohn Institute for Employment Research.

Cameron, A. Colin, and Douglas L. Miller. 2015. "A Practitioner's Guide to Cluster-Robust Inference." *The Journal of Human Resources* 50:317–372.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2019. "Bartik Instruments: What, When, Why, and How." NBER Working Paper 24408.

Hansen, Lars Peter. 1982. "Large Sample Properties of Generalized Method of Moments Estimators." *Econometrica* 50:1029–1054.

Jaeger, David A, Joakim Ruist, and Jan Stuhler. 2018. "Shift-Share Instruments and Dynamic Adjustments: The Case of Immigration." Working Paper.

Jones, Ronald W. 1965. "The Structure of Simple General Equilibrium Models." *Journal of Political Economy* 73:557–572.

Liu, Regina Y. 1988. "Bootstrap Procedures under Some Non-I.I.D. Models." *The Annals of Statistics* 16:1696–1708.

Montiel Olea, Jose Luis, and Carolin Pflueger. 2013. "A robust test for weak instruments." *Journal of Business and Economic Statistics* 31:358–369.

Nagar, Anirudh L. 1959. "The Bias and Moment Matrix of the General k-class Estimators of the Parameters in Simultaneous Equations." *Econometrica* 27:575–595.

Pflueger, Carolin E., and Su Wang. 2015. "A robust test for weak instruments in Stata." *Stata Journal* 15:216–225.

White, Halbert. 1982. "Instrumental Variables Regression with Independent Observations." *Econometrica* 50:483–499.