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

We are grateful to Rodrigo Adão, David Autor, Moya Chin, Edward Glaeser, Paul Goldsmith-Pinkham, Larry Katz, Michal Kolesár, Gabriel Kreindler, Jack Liebersohn, Eduardo Morales, Jack Mountjoy, Jörn-Steffen Pischke, Isaac Sorkin, Jann Spiess, and Itzchak Tzachi Raz for helpful comments. We also thank David Autor, David Dorn, and Gordon Hanson for providing replication code and data. An earlier version of this draft was circulated in May 2017 under the title "Consistency and Inference in Bartik Research Designs." 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
JEL No. C18,C21,C26,F16,J21

ABSTRACT

Many empirical studies leverage shift-share (or "Bartik") instruments that combine a set of aggregate shocks with measures of shock exposure. We derive a necessary and sufficient shock-level orthogonality condition for these instruments to identify causal effects. We then show that orthogonality holds when observed shocks are as-good-as-randomly assigned and growing in number, with the average shock exposure sufficiently dispersed. We recommend that practitioners implement quasi-experimental shift-share designs with new shock-level regressions, which help visualize identifying shock variation, correct standard errors, choose appropriate specifications, test identifying assumptions, and optimally combine multiple sets of quasi-random shocks. We illustrate these points by revisiting Autor et al. (2013)'s analysis of the labor market effects of Chinese import competition.

Kirill Borusyak
Department of Economics
225 Julis Romo Rabinowitz Building
Princeton University
Princeton, NJ 08544
borusyak@princeton.edu

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

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

1 Introduction

A large and growing number of empirical studies use shift-share instruments, which average a set of observed shocks with unit-specific weights measuring shock exposure. In typical settings, such as Bartik (1991) and Blanchard and Katz (1992), a regional instrument is constructed from certain industry shocks with regional employment shares as weights, though other objects sometimes take the place of regions and industries.[1] Despite the popularity of these constructed instruments, it is not always clear what economic arguments and statistical conditions underlie their validity – that is, when shift-share instrumental variable (IV) regressions may plausibly reveal causal effects.

This paper develops a novel framework for understanding shift-share instruments, in which exogenous industry shocks provide identifying variation. Our approach starts with a numerical equivalence: we show that shift-share IV coefficients estimated in a sample of locations are identically obtained from a weighted IV regression in the industry dimension. In this regression the regional outcome and treatment are first averaged with exposure weights to obtain industry-level aggregates; the industry shocks then instrument for aggregate treatment. This result only relies on the structure of the shift-share instrument and thus applies to outcomes and treatments that are not typically computed for industries. The coefficient retains its location-level interpretation, but can be recovered using industry-level variation.

Motivated by this numerical equivalence, we show that the exclusion restriction for shift-share IV is equivalent to an industry-level orthogonality condition. For the instrument to be valid, industry shocks must be uncorrelated with an exposure-weighted average of untreated potential outcomes – that is, the average unobserved determinants of outcomes in the regions most exposed to each industry.

We then propose a set of industry-level assumptions sufficient for such orthogonality. The key requirement is that shocks are as-good-as-randomly assigned to industries, as if arising from a natural experiment. However, quasi-random shock assignment is not by itself enough. For the law of large numbers to apply in the industry dimension, the number of observed industries facing independent shocks must grow with the sample. Furthermore, even though our approach allows each region to be mostly exposed to only a small number of industries, shock exposure must be sufficiently dispersed on average (as measured by a Herfindahl index) such that no finite set of industries asymptotically drives variation in the shift-share instrument. Under these conditions, we show that the shift-share instrument is valid, even when shock exposure is endogenous.

Our quasi-experimental approach is readily extended to settings where shocks are as-good-as-

[1] In other shift-share designs, observations may represent regions impacted by immigration 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, 2017), or groups of individuals facing different income growth rates (Boustan et al., 2013). For simplicity we use the language of locations and industries throughout. Other influential and recent examples of shift-share IVs include Luttmer (2006), Saiz (2010), Kovak (2013), Autor et al. (2013), Nakamura and Steinsson (2014), Oberfield and Raval (2014), Greenstone et al. (2014), Diamond (2016), Suárez and Zidar (2016), and Hornbeck and Moretti (2018).

randomly assigned conditional on industry-level observables, and to panel data. For the former, we show that orthogonality is satisfied when one controls for exposure-weighted averages of the observables. For panels, we show that orthogonality can hold when either the number of observed shocks per period or the number of periods grows, and discuss different constructions of panel shift-share instruments and controls.

We then show how industry-level analyses may be used for shift-share estimation and testing. First, one can inspect the identifying industry-level variation, for example by plotting the outcomes and treatment (aggregated at the industry level) against the shocks. Second, by estimating appropriate industry-level regressions with standard statistical software, one can address deficiencies in traditional location-level inference. In particular conventional second-stage standard errors and measures of first-stage strength from these regressions are valid within the inferential framework of Adao et al. (2018), which builds on our identification approach. Third, the practitioner's view on the industry-level quasi-experiment can guide the choice of location-level controls in shift-share IV. For example when industries are grouped into larger sectors, controlling for the interaction of a region's employment share in each sector with period indicators isolates the within-period, within-sectoral variation in industry shocks. Fourth, one can validate the assumptions underlying the preferred specification with appropriate industry-level tests of balance, no pre-trends, and no auto- or intra-class correlation of shocks. Finally, we introduce new procedures to optimally combine multiple sets of quasi-random shocks and test overidentifying restrictions. For example a two-stage least squares regression of aggregate outcomes on aggregate treatment produces efficient shift-share IV estimates when shocks are homoskedastic. We illustrate each of these points in the setting of Autor, Dorn, and Hanson (2013, henceforth ADH), who use shift-share IV to estimate the effects of Chinese import competition on the regional growth of U.S. manufacturing employment.[2]

Our quasi-experimental assumptions are not the only way to satisfy the shift-share exclusion restriction. In related work Goldsmith-Pinkham et al. (2018) develop and advocate for a different framework based on the exogeneity of the industry exposure weights. Their approach is motivated by a different numerical equivalence: the shift-share IV coefficient also coincides with a generalized method of moments estimator, with exposure shares as excluded instruments. Though share exogeneity is indeed a sufficient condition for identification (and, as such, satisfies our industry-level orthogonality condition), we focus on plausible conditions under which it is not necessary.

In our view, identification through exogenous shocks may be better aligned with researchers' motivations and goals in many settings. Consider the ADH shift-share instrument, which combines industry-specific changes in Chinese import competition with local shock exposure given by the lagged industrial composition of U.S. labor markets. In such a setting, exogeneity of industry employment

[2]To help practitioners implement quasi-experimental shift-share research designs we have developed a Stata package, *ssaggregate*, which creates the industry-level aggregates from the equivalence result. Users can install this package with the command *ssc install ssaggregate*. See the associated help file and our replication archive at https://github.com/borusyak/shift-share for more details.

shares appears difficult to justify *a priori*; for instance it would fail if there are any unobserved industry shocks that affect regional outcomes (through the shares). In contrast, ADH attempt to purge their industry shocks from U.S.-specific factors by measuring Chinese import growth outside of the U.S. One may view this as an effort to obtain more plausible quasi-random variation across industries, consistent with our framework. 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 exposure weights). They argue these shocks are "idiosyncratic," which our approach formalizes as "independent from relevant country-by-product unobservables and from each other." Note, however, that our quasi-experimental framework may be less appropriate for some settings, particularly those involving a small number of shocks.[3]

Econometrically, our approach is related to Kolesar et al. (2015)'s analysis of k-class estimators with many invalid instruments. Consistency in that setting follows when violations of individual instrument exclusion restrictions are uncorrelated with their first-stage effects. For shift-share IV the exposure weights can be thought of as a set of invalid instruments (as they would be in the Goldsmith-Pinkham et al. (2018) interpretation), and our orthogonality condition requires their exclusion restriction violations to instead be uncorrelated with the observed industry shocks. Despite this formal similarity, we argue that shift-share identification is better understood through the quasi-random assignment of a single industry-level instrument (shocks), rather than through a large set of invalid location-level instruments (shares) that nevertheless produce a consistent estimate. This view is reinforced by the numerical equivalence, yields natural industry-level identification conditions, and suggests new validations and extensions of shift-share IV.

Our work also relates to other recent methodological studies of shift-share designs, including Jaeger et al. (2018) and Broxterman and Larson (2018). The former highlights biases of shift-share IV due to endogenous local labor market dynamics, while the latter studies the empirical performance of different shift-share instrument constructions. As discussed above we also draw on the inferential framework of Adao et al. (2018), 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 Borusyak and Jaravel (2017) and Goodman-Bacon (2018) for event study designs; Hudson et al. (2017) and de Chaisemartin and D'Haultfoeuille (2018) for instrumented difference-in-difference designs; and Hull (2018) for mover designs.

The rest of this paper is organized as follows. Section 2 introduces the framework, shows the equivalence between location-level shift-share IV and a particular industry-level IV estimator, and derives the key orthogonality condition. Section 3 then establishes the quasi-experimental assumptions under which orthogonality is satisfied, while Section 4 discusses various practical implications and results from our ADH application. Section 5 concludes.

[3] For example, Card (2001), Kovak (2013), and Acemoglu and Restrepo (2017) each construct a shift-share instrument with around twenty shocks.

2 Shocks as Instruments in Shift-Share Designs

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

2.1 The Shift-Share IV Estimator

Suppose we observe an outcome y_ℓ, a treatment x_ℓ, and a vector of controls w_ℓ (which includes a constant) in a random sample of size L. For concreteness, we refer to the sampled units as "locations," as they are in Bartik (1991), ADH, and many other shift-share applications. We wish to estimate the causal effect of the treatment, β, assuming a linear constant-effect model,

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

where by definition $\mathbb{E}\left[\varepsilon_\ell\right] = \mathbb{E}\left[w_\ell \varepsilon_\ell\right] = 0$.[4] Here ε_ℓ denotes the residual from projecting the untreated potential outcome of location ℓ – that is, the outcome we would observe there if treatment were set to zero – on the controls w_ℓ. For example in the ADH setting y_ℓ and x_ℓ denote the growth rates of manufacturing employment (or other outcomes) and Chinese import exposure in local labor market ℓ, while w_ℓ contains measures of labor force demographics from a previous period and other controls.[5] The residual ε_ℓ thus contains all factors that are uncorrelated with lagged demographics but which would drive local employment growth in the absence of rising Chinese imports.

In writing equation (1) we do not require the realized treatment x_ℓ to be uncorrelated with the potential outcomes ε_ℓ, in which case the causal parameter β could be consistently estimated by ordinary least squares (OLS). Instead, we assume that we also observe a set of N shocks g_n and weights $s_{\ell n} \geq 0$ which predict the exposure of location ℓ to each shock. Using these, we construct a shift-share instrument z_ℓ as the exposure-weighted average of shocks:

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

Here z_ℓ can be interpreted as a predicted local shock in ℓ. Again for concreteness, we refer to n as industries throughout: in ADH, g_n denotes industry n's growth of Chinese imports in eight non-U.S. countries, while $s_{\ell n}$ is the employment share of industry n in location ℓ. To start simply we first assume that the sum of exposure weights across industries is constant, i.e. $\sum_{n=1}^{N} s_{\ell n} = 1$, and treat

[4] It is straightforward to extend our framework to models with heterogeneous treatment effects. As shown in Appendix A.1, the shift-share IV estimator captures a convex average of location-specific linear effects β_ℓ under straightforward extensions of our orthogonality condition and an additional first stage monotonicity assumption, as with the local average treatment effects of Angrist and Imbens (1994). This follows similarly to the results on heterogeneous effects in reduced-form shift-share regressions in Adao et al. (2018).

[5] Because ADH do not observe imports by region, they proxy local import exposure in each industry by the national average to construct x_ℓ. We abstract from this issue in our discussion.

the shocks g_n as known. We relax the first assumption in Section 3 and discuss relevant issues for shift-share IV with estimated shocks in Appendix A.6.

The conventional shift-share IV estimator $\hat{\beta}$ uses z_ℓ to instrument for x_ℓ in equation (1). Letting $v_\ell^\perp$ denote the residual from a sample projection of variable v_ℓ on the controls w_ℓ, we have

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

where the numerator and denominator represent location-level covariances between the instrument and the residualized outcomes and treatment, respectively.

2.2 An Equivalent Industry-Level IV

To build intuition for how identification in shift-share designs may come from the industry shocks, we first show that $\hat{\beta}$ can also be expressed as an industry-level IV estimator that uses g_n as an instrument. By definition of z_ℓ, we have:

$$\begin{aligned}\hat{\beta} &= \frac{\frac{1}{L}\sum_{\ell=1}^{L}\left(\sum_{n=1}^{N} s_{\ell n} g_n\right) y_\ell^\perp}{\frac{1}{L}\sum_{\ell=1}^{L}\left(\sum_{n=1}^{N} s_{\ell n} g_n\right) x_\ell^\perp} \\ &= \frac{\sum_{n=1}^{N} g_n \left(\frac{1}{L}\sum_{\ell=1}^{L} s_{\ell n} y_\ell^\perp\right)}{\sum_{n=1}^{N} g_n \left(\frac{1}{L}\sum_{\ell=1}^{L} s_{\ell n} x_\ell^\perp\right)} \\ &= \frac{\sum_{n=1}^{N} \hat{s}_n g_n \bar{y}_n^\perp}{\sum_{n=1}^{N} \hat{s}_n g_n \bar{x}_n^\perp},\end{aligned} \tag{4}$$

where $\hat{s}_n = \frac{1}{L}\sum_{\ell=1}^{L} s_{\ell n}$ is the average exposure to industry n, and $\bar{v}_n = \sum_{\ell=1}^{L} s_{\ell n} v_\ell / \sum_{\ell=1}^{L} s_{\ell n}$ denotes a weighted average of variable v_ℓ, with larger weights given to locations more exposed to industry n. Specifically, $\bar{y}_n^\perp$ reflects the average residualized outcome of locations specializing in n, while $\bar{x}_n^\perp$ is the same weighted average of residualized treatment. Thus (4) represents $\hat{\beta}$ as a ratio of $\hat{s}_n$-weighted covariances of g_n with $\bar{y}_n^\perp$ and $\bar{x}_n^\perp$.[6]

This shows that the shift-share IV coefficient is numerically equivalent to the coefficient from an $\hat{s}_n$-weighted industry-level IV regression, with instrument g_n and second stage

$$\bar{y}_n^\perp = \alpha + \beta \bar{x}_n^\perp + \bar{\varepsilon}_n^\perp. \tag{5}$$

In the ADH example, it is expected that industries with a high non-U.S. China shock g_n account for a larger share of employment in U.S. regions facing increasing Chinese import competition. Thus the industry-level first-stage regression slope of $\bar{x}_n^\perp$ on g_n is likely to be positive. By estimating

[6]The covariance interpretation follows from observing that the weighted means of $\bar{y}_n^\perp$ and $\bar{x}_n^\perp$ are zero: e.g. $\sum_{n=1}^{N} \hat{s}_n \bar{y}_n^\perp = \frac{1}{L}\sum_{\ell=1}^{L}\left(\sum_{n=1}^{N} s_{\ell n}\right) y_\ell^\perp = \frac{1}{L}\sum_{\ell=1}^{L} y_\ell^\perp = 0$, since the $y_\ell^\perp$ are regression residuals.

the reduced form regression of $\bar{y}_n^{\perp}$ on g_n, one learns whether industries with higher shocks are also concentrated in areas with larger declines in employment. We emphasize that $\bar{y}_n^{\perp}$ is not simply the average outcome in industry n: the impact of import penetration shocks on industry employment is a distinct estimand studied by Acemoglu et al. (2016) via industry-level IV regressions that do not have location-level equivalents. In contrast, $\bar{y}_n^{\perp}$ can be formed for any location-level outcome, including those not typically computed for industries.

As usual a particular exclusion restriction must hold for IV estimates of equation (5) to reveal a causal effect. By the equivalence result (4), this same condition is required for causal interpretation of the shift-share IV coefficient $\hat{\beta}$. We derive and interpret this restriction next.

2.3 Shock Orthogonality

We seek conditions under which the shift-share IV estimator $\hat{\beta}$ is consistent for the causal parameter β; that is, $\hat{\beta} \xrightarrow{p} \beta$ as $L \to \infty$. As usual, IV consistency requires both instrument relevance (that z_ℓ and $x_\ell^{\perp}$ are asymptotically correlated) and validity (that z_ℓ and ε_ℓ are asymptotically uncorrelated). Since relevance can be inferred from the data, we assume it is satisfied and focus on validity. Applying the logic of (4) to the population covariance between the shift-share instrument and structural residual ε_ℓ, we have

$$\begin{aligned} \operatorname{Cov}\left[z_\ell, \varepsilon_\ell\right] &= \mathbb{E}\left[\sum_{n=1}^{N} s_{\ell n} g_n \varepsilon_\ell\right] \\ &= \sum_{n=1}^{N} s_n g_n \phi_n, \end{aligned} \tag{6}$$

where $s_n \equiv \mathbb{E}\left[s_{\ell n}\right]$ measures the expected exposure to industry n, and where $\phi_n \equiv \mathbb{E}\left[s_{\ell n}\varepsilon_\ell\right]/\mathbb{E}\left[s_{\ell n}\right]$ is an exposure-weighted expectation of untreated potential outcomes.[7] In terms of the notation above, these represent population analogs of $\hat{s}_n$ and $\bar{\varepsilon}_n^{\perp}$. Given a law of large numbers, i.e. that $\frac{1}{L}\sum_{\ell=1}^{L} z_\ell \varepsilon_\ell - \operatorname{Cov}\left[z_\ell, \varepsilon_\ell\right] \xrightarrow{p} 0$, the shift-share IV estimator is therefore consistent if and only if

$$\sum_{n=1}^{N} s_n g_n \phi_n \to 0. \tag{7}$$

Equation (7) is our key orthogonality condition. It shows that for shift-share IV to be valid, the large-sample covariance (measured in the set of N industries and weighted by s_n) between industry-level shocks g_n and unobservables ϕ_n must be zero.[8] Thus in ADH, equation (7) holds when the

[7] For initial simplicity in this section we derive the validity condition given fixed sequences of the set of (g_n, s_n, ϕ_n). In the quasi-experimental framing it is more natural to imagine a hierarchical sampling design, in which sets are drawn from a larger population. All expectations and covariances in this section should then be thought to condition on the industry-level draws, with validity satisfied by $\sum_{n=1}^{N} s_n g_n \phi_n \xrightarrow{p} 0$. We formalize this framework in Section 3.

[8] Although our focus is on shift-share IV, note that an analogous orthogonality restriction is necessary and sufficient for shift-share reduced-form regressions of y_ℓ on z_ℓ to have a causal interpretation, with ϕ_n constructed from the

industry-specific measure of Chinese import growth in non-U.S. countries (g_n) is asymptotically uncorrelated with the industry-specific average of unobserved factors affecting employment growth in locations specializing in each industry (ϕ_n), with weights given by national industry employment shares s_n.[9]

As a necessary condition, equation (7) is satisfied under the preferred exogenous-shares interpretation of shift-share IV in Goldsmith-Pinkham et al. (2018). That is, if $\mathbb{E}\left[s_{\ell n}\varepsilon_\ell\right] = 0$ for all n, then $\phi_n = 0$ and equation (7) would be satisfied for any shocks g_n. Share exogeneity may indeed be a better motivation for shift-share IV in certain applications. However, in settings like ADH which study a particular industry shock, it may prove difficult to argue that there are no unobserved industry factors that also impact locations through the industrial composition (and thus generate $\phi_n \neq 0$).[10]

Absent share exogeneity, industry-level orthogonality can still provide a natural formalization of the intuitive identification arguments found in many shift-share designs. In ADH, for example, one may be concerned that orthogonality fails due to reverse causality: China may gain market shares in certain industries (given by a high g_n) precisely because locations specializing in them face other adverse conditions, as reflected in ϕ_n. Measuring import growth outside of the U.S. addresses such a concern, to the extent the underlying performance of U.S. and non-U.S. industries are not correlated. In addition, equation (7) helps understand potential concerns about omitted variables bias, because of either unobserved industry shocks or regional unobservables that are correlated with industrial composition – again captured by ϕ_n. For example, industries with a higher growth of competition with China may also be more exposed to certain technological shocks (e.g., automation), or these industries could happen to have larger employment shares in regions that are affected by other employment shocks (e.g., low-skill immigration).

3 A Quasi-Experimental Framework

We now develop a set of assumptions that characterize a shock-level quasi-experiment satisfying the orthogonality condition (7). We then show how to extend this approach to settings where shocks are quasi-randomly assigned given a vector of industry-level observables, and to applications with panel data.

residuals of that equation. Interestingly the equivalent representation of the reduced form at the industry level is still an IV regression, of $\bar{y}_\ell^\perp$ on $\bar{z}_\ell^\perp$, instrumented by g_n and weighted by $\hat{s}_n$.

[9] In many applications with employment shares as the exposure measures, shift-share regressions are weighted by total location employment. In this case, s_n can be interpreted as the lagged employment of industry n; otherwise s_n would be interpreted as the expected share of industry n in a random location. A slight complication in the ADH case is that they use lagged employment shares as $s_{\ell n}$ but beginning-of-period employment as regression weights.

[10] This argument applies even to Bartik (1991), where shocks consist of broader industry-level variation. Unobserved local labor supply shifters may still be correlated with industry shares, for example if they both are affected by land prices in the region.

3.1 Baseline Assumptions

Our quasi-experimental framework assumes shocks are as-good-as-randomly assigned with respect to the relevant industry-level unobservable. To leverage such assignment, we further assume that researchers observe an increasing number of independent shocks as the sample grows, and that the exposure weights are such that no small set of shocks asymptotically drives variation in the shift-share instrument. Under these conditions we show that equation (7) is satisfied, and hence that the instrument is valid.

To formalize the shock quasi-experiment, we imagine a hierarchical data-generating process in which, given the set of industry exposure weights $s_1, \ldots, s_N$, the set of industry-level shocks g_n and unobservables ϕ_n are drawn from some unspecified distribution. We place no restrictions on the marginal distribution of the ϕ_n (which again reflect weighted averages of the structural residual ε_ℓ, with industry-specific exposure weights $s_{\ell n}$) and thus allow exposure to be endogenous (i.e. $\phi_n \neq 0$). Instead, we make two assumptions on the conditional shock data-generating process:

A1. **(Quasi-random shock assignment)** $\mathbb{E}\left[g_n \mid \phi_n\right] = \mu$, for all n;

A2. **(Many independent shocks)** $\mathbb{E}\left[(g_n - \mu)(g_m - \mu) \mid \phi_n, \phi_m\right] = 0$ for all n and $m \neq n$, and $\sum_{n=1}^{N} s_n^2 \to 0$.

We show in Appendix A.2 that under these assumptions and a weak regularity condition, the orthogonality condition holds.

In words, A1 states that the shocks are as-good-as-randomly assigned, such that each industry n is expected to face the same shock μ regardless of its ϕ_n. As shown in the proof this is enough to make the left-hand side of the orthogonality condition zero, in expectation over different shock realizations. For it to converge to this mean, a shock-level law of large numbers must hold. This is implied by the second assumption A2, which states that shocks are mutually uncorrelated given the unobservables and that the Herfindahl index of expected exposure converges to zero. Appendix A.2 shows that the mutual mean-independence condition can be relaxed to allow for shock assignment processes with weak mutual dependence, such as clustering or autocorrelation. The Herfindahl condition implies that the number of observed industries grows with the sample (since $\sum_{n=1}^{N} s_n^2 \geq 1/N$), and the average exposure is sufficiently dispersed across industries. This ensures the covariance in equation (6) is not asymptotically driven by any finite set of shocks.[11] An equivalent concentration condition is that the largest industry becomes vanishingly small as the sample grows.

[11] Note that one could also formulate analogs of A1 and A2 when L is finite but N grows. However, this asymptotic sequence would cause the shift-share instrument to have no variation asymptotically: with fixed L, $\sum_{n=1}^{N} s_n^2 \to 0$ implies $\sum_{\ell=1}^{L}\left(\sum_{n=1}^{N} s_{\ell n}^2\right) \to 0$ and thus $\text{Var}\left[z_\ell\right] = \text{Var}\left[\sum_{n=1}^{N} s_{\ell n} g_n\right] \to 0$.

3.2 Conditional Quasi-Random Assignment

As with other designs, one may wish to assume that A1 and A2 only hold conditionally on a set of observables. For example, when industries n are grouped in larger clusters $c(n) \in \{1, \ldots, C\}$, one may be more willing to assume shocks are exogenously assigned within clusters while allowing cluster-average shocks to be non-random. Letting the vector q_n collect $C-1$ cluster dummies (and a constant), this corresponds to a weakened mean-independence restriction A1:

$$\mathbb{E}\left[g_n \mid \phi_n, q_n\right] = q_n' \tau. \tag{8}$$

In general any industry-level observables may be included in such a q_n. Similarly, one may prefer to impose the mutual uncorrelatedness condition of A2 on the residual $g_n^* = g_n - q_n' \tau$ in place of g_n. In the industry cluster case, this would allow the number of clusters to remain small, as in that case the law of large numbers may apply to g_n^* but not to g_n.

Under these conditional versions of A1 and A2, the orthogonality condition is satisfied by the residual shift-share instrument $\hat{z}_\ell^* = \sum_{n=1}^N s_{\ell n}(g_n - q_n' \hat{\tau})$, given a consistent estimate of τ. This follows by a straightforward extension of the proof in Appendix A.2. In Appendix A.3, we further show that such an instrument is implicitly used when the exposure-weighted vector of industry controls $\sum_{n=1}^N s_{\ell n} q_n$ is included in the location-level control vector w_ℓ. Thus for industry clusters, controlling for the individual exposure to each cluster, $\sum_{n=1}^N s_{\ell n} \mathbf{1}\left[c\left(n\right) = c\right]$, isolates within-cluster shock variation.

An intuitive application of this result is found in shift-share designs where the sum of exposure measures $S_\ell = \sum_{n=1}^N s_{\ell n}$ varies across locations. For example, in ADH the set of $s_{\ell n}$ corresponds to lagged employment shares of manufacturing industries only, so that S_ℓ denotes the lagged share of manufacturing in local labor market ℓ. In such settings one can rewrite the shift-share instrument with the "missing" industry included:

$$z_\ell = s_{\ell 0} g_0 + \sum_{n=1}^N s_{\ell n} g_n, \tag{9}$$

where $g_0 = 0$ and $s_{\ell 0} = 1 - S_\ell$, so that $\sum_{n=0}^N s_{\ell n} = 1$ as before. It is then clear that A1 in this "incomplete shares" case requires $\mathbb{E}\left[g_n \mid \phi_n\right] = 0$ for $n > 0$; otherwise, even if shocks are random, the shift-share instrument will implicitly use variation in average shocks across the missing and non-missing industries, and violate orthogonality when the weighted mean ϕ_n of these sectors differ.[12] Including the indicator $\mathbf{1}[n = 0]$ in the industry control vector q_n allows the mean of observed shocks to be different than that of the missing industry (i.e., non-zero), and corresponds to controlling for S_ℓ in the shift-share IV regression, as $\sum_{n=0}^N s_{\ell n} \mathbf{1}\left[n = 0\right] = 1 - S_\ell$.

[12] To see this formally, note that if A1 and A2 hold for $n > 0$ the orthogonality condition converges in probability to its expectation, $\mathbb{E}\left[\sum_{n=0}^N s_n g_n \phi_n\right] = \sum_{n=1}^N s_n \mathbb{E}\left[\mathbb{E}\left[g_n \mid \phi_n\right] \phi_n\right] = \mu \sum_{n=1}^N s_n \mathbb{E}\left[\phi_n\right]$. Since $s_0 \mathbb{E}\left[\phi_0\right] + \sum_{n=1}^N s_n \mathbb{E}\left[\phi_n\right] = \mathbb{E}\left[\varepsilon_\ell\right] = 0$, when $\mu \neq 0$ this can only be zero if $\sum_{n=1}^N s_n \mathbb{E}\left[\phi_n\right] = \mathbb{E}\left[\phi_0\right] = 0$.

3.3 Panel Data

We conclude this section by extending the quasi-experimental framework to settings with panel data. Letting $y_{\ell t}$, $x_{\ell t}$, $s_{\ell nt}$, and g_{nt} denote observations of the outcome, treatment, exposure weights, and shocks in periods $t = 1, \ldots, T$, we consider validity of the time-varying shift-share instrument

$$z_{\ell t} = \sum_{n=1}^{N} s_{\ell nt} g_{nt}. \tag{10}$$

Here note that (10) can be rewritten

$$z_{\ell t} = \sum_{n=1}^{N} \sum_{p=1}^{T} s_{\ell ntp} g_{np}, \tag{11}$$

where $s_{\ell ntp} = s_{\ell nt}\mathbf{1}[t = p]$ denotes the exposure of location ℓ in time t to industry shock n in period p, which is by definition zero for $t \neq p$. This formulation of $z_{\ell t}$ nests the panel shift-share instrument in the preceding framework; the key orthogonality condition is

$$\text{Cov}\left[z_{\ell t}, \varepsilon_{\ell t}\right] = \sum_{n=1}^{N} \sum_{p=1}^{T} s_{np} g_{np} \phi_{np} \to 0, \tag{12}$$

where $s_{np} = \mathbb{E}\left[s_{\ell ntp}\right]$ and $\phi_{np} = \mathbb{E}\left[s_{\ell ntp}\varepsilon_{\ell t}\right] / \mathbb{E}\left[s_{\ell ntp}\right]$, with $\varepsilon_{\ell t}$ denoting the time-varying structural residual. Analogs of assumptions A1 and A2 applied to $((s_{np}, g_{np}, \phi_{np})_{p=1}^{T})_{n=1}^{N}$ moreover provide sufficient quasi-experimental conditions for the validity of a panel shift-share instrument.

Unpacking these assumptions in the panel setting delivers three insights. First, when the analogs of assumptions A1 and A2 hold (implying that shocks are mutually mean-independent across both locations and time periods), shift-share orthogonality may be satisfied by either observing many industries N or from observing a finite number of industries over many periods T.[13] Second, one can relax the two quasi-experimental assumptions by including location and time fixed effects in the shift-share IV regression. With location fixed effects included in w_{ℓ}, only time-varying unobservables enter ϕ_{np}; shocks are then allowed to be correlated with the exposure-weighted averages of all time-invariant unobservables. Including period fixed effects, moreover, is analogous to allowing shocks to be "clustered" within periods, and thus allows for period-specific shock means. In the incomplete shares case discussed above, this is accomplished by controlling for the sums of exposure weights S_{ℓ}, interacted with period indicators. Third, serial correlation in industry shocks may be accommodated either by a modification of A2 which allows for weak mutual shock dependence (see Appendix A.2 for an

[13] Goldsmith-Pinkham et al. (2018) also consider shift-share identification with many time periods, leveraging a stronger identification condition. In particular they assume exposure shares are constant over time and shocks to each industry are uncorrelated with structural residuals in each location, over time. Our asymptotics are also slightly different: L is fixed in the Goldsmith-Pinkham et al. (2018) panel extension but large in ours (though the latter can be relaxed, since the problem from footnote 11 does not arise in long panels).

example) or by controlling for certain share-weighted lags of shocks when the time series properties are known. For example when shocks follow a first-order autoregressive process, $g_{nt} = \rho_0 + \rho_1 g_{n,t-1} + g^*_{nt}$, controlling for $\sum_{n=1}^N s_{\ell nt} g_{n,t-1}$ extracts the idiosyncratic period-specific shock variation in g^*_{nt}.

Lastly, we note that our framework implies a tradeoff between letting the exposure shares used to construct the instrument vary over time or, as in some applications, holding them fixed in some pre-period (i.e., $s_{\ell nt} = s_{\ell n0}$ for all t). On the one hand, when shocks exhibit no serial correlation conditional on the controls, updating shares may improve the IV first stage by making $z_{\ell t} = \sum_{n=1}^N s_{\ell nt} g_{nt}$ more predictive of $x_{\ell t}$. On the other hand, when shocks are serially correlated and shares are affected by lagged shocks, orthogonality is unlikely to hold with time-varying $s_{\ell nt}$; that is, g_{nt} is likely to be correlated with $\mathbb{E}\left[s_{\ell nt}\varepsilon_{\ell t}\right]/\mathbb{E}\left[s_{\ell nt}\right]$ via lagged $g_{n,t-1}$. For example, if the high-ε_ℓ regions in ADH are relatively more resilient to increased Chinese import competition in their industries, the industries repeatedly affected by adverse shocks will be increasingly concentrated in locations where manufacturing would fall for other reasons. In this case A1 can be satisfied if shares are measured in a period prior to the determination of any shocks correlated with g_{nt}, e.g. before the serially correlated natural experiment has begun.[14]

4 Quasi-Experimental Shift-Share Designs in Practice

We now present five practical implications of our theoretic framework, illustrated in the Autor et al. (2013) setting. First, following the equivalence result in Section 2, practitioners can perform industry-level analyses to inspect the identifying shock variation. Second, they can correct deficiencies in location-level standard errors by estimating a slightly modified version of the industry-level regression (4), with standard statistical software. Third, they can leverage the theory in Section 3 to guide their choice of shift-share specification, with the implicit natural experiment in mind. Fourth, they can test the assumptions underlying the quasi-experimental designs at the industry level. Lastly, they can combine multiple observed shocks to form more efficient shift-share IV estimates and test overidentifying restrictions of the quasi-experimental framework, with new industry-level regressions. We implement these ideas in the ADH application with the help of a new Stata command *ssaggregate* that constructs the necessary industry-level aggregates (see footnote 2).

4.1 Visualizing Shocks as Instruments

Column 1 of Table 1 replicates the headline shift-share IV estimate of ADH, using a repeated cross section of 722 locations (U.S. commuting zones) and 397 four-digit SIC manufacturing industries in two decades: 1990-1999 and 2000-2009. Here again y_ℓ is the growth rate of manufacturing employment;

[14]This reason for lagging the shares 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 the preferred interpretation of Goldsmith-Pinkham et al. (2018).

x_ℓ measures the local growth of imports from China (in \$1,000 per worker); w_ℓ is the vector of controls used in the preferred specification of ADH, containing beginning-of-period measures of labor force demographics, period fixed effects, and census region fixed effects; $s_{\ell n}$ is the employment share of manufacturing industry n, measured a decade before each period begins; and g_n is industry n's growth of Chinese exports to eight non-U.S. economies (again in \$1,000 per U.S. worker). For brevity, we omit the period subscript t except where it is necessary. The coefficient of -0.596 coincides with Table 3, column 6 in ADH.[15]

Our equivalence result (4) shows that this IV coefficient can be equivalently obtained from an industry-level regression. In particular, estimating equation (5) by regressing exposure-weighted outcome residuals $\bar{y}_n^\perp$ on exposure-weighted treatment residuals $\bar{x}_n^\perp$, weighting by $\hat{s}_n$ and instrumenting by shocks, also produces a coefficient of -0.596. Since ADH is an example of a shift-share design with manufacturing industry exposure weights that do not sum to one (recall the discussion in Section 3.2), we include the "missing" non-manufacturing industry group in the regression, with a shock of zero.

Panel A of Figure 1 illustrates the equivalent industry-level IV coefficient via binned scatter plots of $\bar{x}_n^\perp$ and $\bar{y}_n^\perp$ against g_n. The former reveals a strong positive first-stage relationship, with a $\hat{s}_n$-weighted slope of 0.0065. The latter reduced form scatter trends downward, with a weighted slope of -0.0039. Scaling this by the first-stage slope again yields the IV coefficient of -0.596.

Practitioners can use industry-level plots to inspect the identifying variation in quasi-experimental shift-share designs. For example in Figure 1A we see that the distribution of realized shocks is highly skewed, with a small number occupying the far-right of the x-axis. In these situations, one may wish to check robustness of the shift-share IV coefficient by excluding these shocks, which may be viewed as outliers. Notably, the industries generating these shocks are also likely to be of interest in the alternative shares-as-instruments framework, as they may generate skewness in the "Rotemberg weights" that Goldsmith-Pinkham et al. (2018) use to decompose the location-level regression.

4.2 Using Industry-Level Standard Errors

Industry-level regressions may also be used to overcome shortcomings of conventional shift-share inference. When shares are not valid instruments (i.e., when orthogonality is satisfied despite $\phi_n \neq 0$), location-level standard errors may fail to capture the asymptotic variance of the shift-share IV estimator. This issue has been recently studied by Adao et al. (2018); intuitively, the quasi-random assignment of g_n guarantees that the orthogonality condition holds in expectation but not for any given realization of N shocks. When N is large, the covariance of g_n and ϕ_n may be as important than the location sampling variation targeted by conventional methods. Indeed, Adao et al. (2018) show in simulations that this covariance can lead to large distortions in the coverage probabilities of

[15] Appendix Table A1 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.

standard location-level confidence intervals.

Adao et al. (2018) derive corrected shift-share IV standard errors in a shocks-as-instruments framework that builds on our own. They work under the additional assumption that each control $w_{\ell j}$ either has a shift-share structure, i.e. $w_{\ell j} = \sum_{n=1}^{N} s_{\ell n} q_{nj}$ for some q_{nj}, or is uncorrelated with the shift-share instrument conditional on all such q_{nj}. This captures two main reasons for including controls in shift-share designs: either because shocks satisfy A1 only conditionally on the industry-level vector q_n (as in equation (8)) or because the $w_{\ell j}$ absorb some of the residual variance in outcomes, thereby increasing the efficiency of the estimator.

Our equivalence result (4) suggests a convenient implementation of the Adao et al. (2018) standard errors with industry-level regressions. As shown in Appendix A.4, a heteroskedasticity-robust standard error applied to the estimate of β in the $\hat{s}_n$-weighted industry-level IV regression

$$\bar{y}_n^{\perp} = \beta \bar{x}_n^{\perp} + q_n' \theta + \bar{\varepsilon}_n, \tag{13}$$

with $\bar{x}_n^{\perp}$ instrumented by g_n, coincides asymptotically with the corresponding formula of Adao et al. (2018). This equivalence extends to cluster-robust standard errors: shocks are then allowed to have a random cluster component, as with the extension of A2 in Appendix A.2. The Adao et al. (2018) inferential framework, and thus our industry-level implementation, moreover apply when the structural errors are themselves clustered in the location space, as with the state-level clustering scheme of ADH.

Practitioners may thus obtain shift-share IV coefficients and valid standard errors by estimating equation (13) with standard statistical software. Note that this IV regression differs from (5) by the inclusion of the control vector q_{nt}, but this will not change the coefficient estimate $\hat{\beta}$.[16] Thus $\hat{\beta}$ retains its equivalence to the conventional shift-share IV coefficient, and continues to estimate a location-level causal effect under the appropriate quasi-experimental assumptions.

By estimating (13) in column 1 of Table 1 we obtain a standard error, clustered by three-digit SIC (SIC3) codes, of 0.114. This is indeed quite similar to the corresponding standard error in Table 5 of Adao et al. (2018) in their reanalysis of ADH. Conveniently, the first-stage F-statistic also becomes valid when estimated at the industry level with standard software. This is because the conventional standard error on the estimated industry-level first stage coefficient π is also asymptotically valid in the Adao et al. (2018) framework, by an argument analogous to that of Appendix A.4, and because tests of $\pi = 0$ asymptotically coincide with tests of shift-share instrument relevance.[17] In column 1 of Table 1 we report an F-statistic of 39.7, indicating a strong ADH first stage.

[16] This is because both $\bar{y}_n^{\perp}$ and $\bar{x}_n^{\perp}$ are uncorrelated with q_n when weighted by $\hat{s}_n$: e.g. $\sum_{n=1}^{N} \hat{s}_n q_n \bar{y}_n^{\perp} = \frac{1}{L} \sum_{\ell=1}^{L} \left(\sum_{n=1}^{N} s_{\ell n} q_n \right) y_\ell^{\perp} = \frac{1}{L} \sum_{\ell=1}^{L} w_\ell y_\ell^{\perp} = 0$. Also note that, as in Section 3, equivalence between shift-share IV and (5) in the incomplete shares case requires including the missing industry as an observation, with $g_0 = 0$.

[17] That is, both test the null hypothesis that $\mathrm{Cov}\left[z_\ell, x_\ell^{\perp}\right] = \sum_{n=1}^{N} s_n g_n \bar{x}_n^{\perp}$ is asymptotically zero.

4.3 Choosing a Specification to Fit the Quasi-Experiment

We now demonstrate how the quasi-experimental framework may guide the choice of the shift-share IV specification itself. Since a comprehensive analysis of possible specifications in ADH is outside the scope of this paper, our discussion is meant only to illustrate how the framework can help think through potential omitted variables and select appropriate controls in quasi-experimental shift-share designs.

First, as discussed above, ADH is a setting in which the sum of exposure weights (the 10-year lagged manufacturing share) varies across locations. Thus the replicated specification in column 1 of Table 1 implicitly compares non-zero Chinese import growth in the manufacturing sector with the zero shock in the "missing" non-manufacturing sector, and its estimate may be affected by unobserved shocks that systematically differ across these sectors. To address this concern and ensure the estimate is only driven by variation in shocks within manufacturing, one may include the sum of shares as an IV control; in the industry regression, this amounts to "dummying out" the non-manufacturing sector by including $\mathbf{1}[n = 0]$ in the control vector q_n, or by simply dropping the $n = 0$ observation. In fact, ADH follow a similar strategy: they control for the beginning-of-period manufacturing employment share, rather than for the lagged manufacturing employment share that corresponds to the instrument exposure weights. In column 2 of Table 1 we substitute the former with the latter; as expected, this results in a similar, though slightly smaller, IV coefficient of -0.489.

Other industry-level controls can be included to further reduce the possibility of omitted variables bias, after excluding the non-manufacturing sector. For the ADH repeated cross section, a natural candidate is the set of period indicators, as the manufacturing sector may have faced different unobserved shocks across periods (e.g., due to automation, or changes in union regulation). Including period indicators in q_n allows mean Chinese import growth to differ between 1990-1999 and 2000-2009, and isolates within-period within-manufacturing shock variation.[18] The corresponding location-level control set w_ℓ contains the exposure-weighted vector $\sum_{n=1}^{N} s_{\ell n} q_n$. As the exposure weights in ADH correspond to manufacturing industries only, this amounts to controlling for sums of lagged manufacturing shares, interacted with period indicators. Column 3 of Table 1 shows that adding these controls in ADH reduces the IV coefficient to -0.267. This change reflects the fact that 2000-2009 saw both a faster increase in Chinese import competition and a faster decline in U.S. manufacturing.[19]

We take the specification in column 3 as our baseline for illustrating other practical considerations. Panel B of Figure 1 visualizes this specification by plotting its $\bar{y}_n^{\perp}$ and $\bar{x}_n^{\perp}$ against g_n, showing that

[18] Likewise, studies of trade shocks at the industry level typically rely on within-period within-manufacturing shock variation, by running specifications with manufacturing-by-period fixed effects in a sample of all industries or with period fixed effects in a sample restricted to manufacturing industries (e.g., Acemoglu et al. (2016)).

[19] 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 A2 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). We again compute standard errors at the industry level, exploiting a straightforward extension of the numerical equivalence in Section 2.2 and the argument in Section 4.2. 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.

it still involves a skewed distribution of shocks, as with the replicated ADH specification in Panel A. In Column 4 of Table 1, we examine the sensitivity of the IV coefficient to the exclusion of potential outlier shocks, with $g_n > 47.7$ thousand dollars per worker. Dummying out these shocks in q_n (and controlling for the associated period-specific industry shares in w_ℓ) isolates within-period variation in the other industries: the IV coefficient grows in magnitude, from -0.267 in column 3 to -0.349 in column 4. Appendix Figure A1 plots the remaining industry-level variation in the restricted specification.

Finally, we include in the last two columns of Table 1 examples of estimates which exploit more narrow shock variation, from within industrial sectors. These specifications add the lagged shares of sectors, interacted with period dummies, to the baseline control vector w_ℓ and correspondingly add sector-by-period fixed effects to q_n. Specifically, column 5 uses 10 broad manufacturing sectors from Autor et al. (2014), while column 6 controls for 20 two-digit SIC codes nested in these sectors. The estimate is largely unchanged in column 5 (at -0.252) and shrinks to -0.118 in column 6, though the column 3 estimate falls within the 95% confidence interval of both specifications. Notably, these specifications are more demanding of the data, as some of the variation in exposure to rising Chinese import competition is absorbed by the fixed effects; indeed, the first stage F-statistic drops from 34.3 to in column 3 to 18.6 in column 6.

4.4 Testing Assumptions A1 and A2

With our baseline quasi-experimental specification in place, we now show how its identifying assumptions may be tested. As usual, the key shift-share orthogonality condition, and its sufficient conditions A1 and A2, can not be checked directly. However, given an observed variable r_ℓ that is plausibly correlated with the structural error ε_ℓ, one may indirectly verify the validity of a shift-share instrument by testing $\text{Cov}\left[z_\ell, r_\ell^\perp\right] = 0$. For example in the preferred ADH specification, predetermined regional characteristics (such as the average demographics of residents, or the average "offshorability" of local business) may be regressed on the instrument and controls, while lagged (pre-1990) manufacturing employment growth may be used to check instrument pre-trends.

In the quasi-experimental shift-share framework, inference for such tests must account for the randomness of shocks and is thus again non-standard. As with second- and first-stage inference in shift-share IV, however, valid balance and pre-trend tests can be easily conducted at the industry level. Specifically, researchers may construct the exposure-weighted residual $\bar{r}_n^\perp$ (which here proxies for ϕ_n). They may then regress $\bar{r}_n^\perp$ on shocks g_n and any industry-level controls q_n, weighting by $\hat{s}_n$, and validate their preferred specification with conventional heteroskedastic-robust or clustered standard errors. Of course, shock balance with respect to other observed industry characteristics thought to be correlated with ϕ_n may also be tested this way. Table 2 reports the results of these regressions in the ADH data, using various predetermined regional characteristics as well as manufacturing employment pre-trends (1970-1979 and 1980-1989) as r_ℓ. Six out of seven of the validation tests pass at conventional levels of statistical significance, with the seventh suggesting that industries facing higher Chinese import growth are concentrated in locations with more immigrants. The p-value for the joint test of significance across all seven industry-level regressions is 0.179.

One can also use industry-level regressions to validate the mutual uncorrelatedness of shocks in A2, conditional on observables q_n. For example, one can check that the residuals of the auxiliary regression of g_n on q_n are not systematically clustered within observed industry groups, nor over time. When such tests fail, the quasi-experimental interpretation of the shift-share design may be retained by a modified A2 that permits weak mutual dependence in shocks (see Appendix A.2 for examples), but the preferred industry-level standard errors should be consistent with this dependence.

We test for mutual correlations in the ADH setting by exploiting the available industry clustering structure and time dimension of the pooled cross section. Specifically, we decompose the residual variance of China import shocks (given the industry controls from our baseline shift-share specification) via a hierarchical linear model:

$$g_{nt} = q'_{nt}\tau + a_{\text{ten}(n),t} + b_{\text{sic2}(n),t} + c_{\text{sic3}(n),t} + d_n + e_{nt}, \tag{14}$$

where the control vector q_{nt} contains period indicators, as in column 3 of Table 1; $a_{\text{ten}(n),t}$, $b_{\text{sic2}(n),t}$, and $c_{\text{sic3}(n),t}$ denote time-varying (and possibly auto-correlated) random effects generated by the ten industry groups in Autor et al. (2014), SIC2 codes, and SIC3 codes, respectively; and d_n is a time-invariant industry random effect. Following convention for hierarchical linear models, we estimate equation (14) by maximum likelihood, assuming Gaussian residual components.[20]

Table 3 reports intra-class correlations (ICCs) from this analysis, which reflect 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), consistent with using the SIC3-clustered standard errors in the industry-level regressions. 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, which underlies our baseline specification.

4.5 Leveraging Multiple Sets of Shocks

In some shift-share designs practitioners may have access to multiple industry shocks satisfying A1 and A2. For example, while ADH measure Chinese import growth by its sum across eight non-U.S. countries, one may think that each set of country-specific import growth measures g_{nk} for $k =$

[20] In particular we estimate an unweighted mixed-effects regression using Stata's *mixed* command, imposing an exchangeable variance matrix for $(a_{\text{ten}(n),0}, a_{\text{ten}(n),1})$, $(b_{\text{ten}(n),0}, b_{\text{ten}(n),1})$, and $(c_{\text{ten}(n),0}, c_{\text{ten}(n),1})$ and reporting robust standard errors. See the replication code for the exact syntax.

$1, \ldots, 8$ (still in \$1,000 per U.S. worker) is itself as-good-as-randomly assigned with respect to U.S. industry-level unobservables ϕ_n, and thus each may be used to consistently estimate β. In other settings additional shocks may be generated via non-linear transformations of the original g_{nt}, if stronger independence assumptions hold. As usual, such overidentification of β raises the possibility of an efficient generalized method of moments (GMM) estimator that optimally combines the quasi-experimental variation, as well as a Hansen (1982) omnibus test of the identifying assumptions. Here we show how one may obtain these estimator and test from certain novel industry-level regressions.

Given a set of K shift-share instruments $z_{\ell k} = \sum_{n=1}^{N} s_{\ell n} g_{nk}$, a usual overidentified IV procedure is the two stage least squares (2SLS) regression of y_ℓ on x_ℓ, controlling for w_ℓ and instrumenting by the set of $z_{\ell k}$. The resulting estimate can be interpreted as a weighted average of instrument-specific shift-share IV coefficients and is thus consistent when the orthogonality condition (7) holds for each set of shocks. However, it will only be the efficient estimate of β when the vector of structural errors ε is spherical conditional on the instrument matrix z, i.e. when $\mathbb{E}\left[\varepsilon\varepsilon' \mid z\right] = \sigma_\varepsilon^2 I$ (Wooldridge, 2002, p. 96). This condition is unlikely to hold in the shift-share setting, even when shocks are independently assigned, as locations with similar exposure profiles to observed shocks may also be exposed to similar unobserved disturbances (Adao et al., 2018).[21]

To characterize the optimal GMM estimator of β, we again establish and leverage a numerical equivalence. Namely, the location-level moment function based on the validity of a shift-share instrument vector $\mathbf{z}_\ell = (z_{\ell 1}, \ldots z_{\ell K})'$ can be rewritten

$$\begin{aligned} m(b) &= \sum_{\ell=1}^{L} \mathbf{z}_\ell \left(y_\ell^\perp - b x_\ell^\perp\right) \\ &= \sum_{n=1}^{N} \hat{s}_n \mathbf{g}_n \left(\bar{y}_n^\perp - b\bar{x}_n^\perp\right), \end{aligned} \tag{15}$$

where $\mathbf{g}_n$ is a $K \times 1$ vector collecting the set of shocks g_{nk}. This corresponds to a weighted industry-level moment function, exploiting the asymptotic orthogonality of shocks $\mathbf{g}_n$ and industry-level unobservables ϕ_n. The optimal GMM estimator using $m(b)$ is then given by

$$\hat{\beta}^* = \arg\min_b m(b)' W^* m(b), \tag{16}$$

where W^* is a consistent estimate of the inverse asymptotic variance of $m(\beta)$'s limiting distribution.

As before, the asymptotic theory of Adao et al. (2018) can be used to derive such a W^*. For example in the case of conditionally homoskedastic shocks, i.e. $\mathrm{Var}\left[\mathbf{g}_n \mid \phi_n, q_n\right] = G$, the optimal

[21] Note that this 2SLS estimator also differs from the optimal IV estimator in the Goldsmith-Pinkham et al. (2018) setting of exogenous shares, even under homoskedasticity. The 2SLS regression of $y_{\ell t}$ on $x_{\ell t}$ controlling for $w_{\ell t}$ and instrumenting by $s_{\ell 1t}, \ldots, s_{\ell Nt}$ will in general not involve growth rates at all. For example, when treatment has a shift-share structure, the first stage of 2SLS with shares as instruments produces perfect fit, and so shares-as-instruments 2SLS is the same as OLS. This almost corresponds to the ADH case (see footnote 5), except that they use employment shares from different years in constructing the instrument and treatment.

moment-weighting matrix is proportional to a consistent estimate of G^{-1}. As shown in Appendix A.5, the efficient estimate (16) is then equivalent to the coefficient from a $\hat{s}_n$-weighted industry-level 2SLS regression of $\bar{y}_n^{\perp}$ on $\bar{x}_n^{\perp}$, instrumenting with $\mathbf{g}_n$ and controlling for q_n. Natural extensions of the Adao et al. (2018) theory and our equivalence result moreover imply the conventional standard error and first stage F-statistic from this regression are valid. Column 2 of Table 4 presents the 2SLS estimate using our baseline ADH specification (column 3 of Table 1). At -0.238, it is very similar to the just-identified estimate, although the industry-level first-stage F-statistic (computed via the procedure in Montiel Olea and Pflueger (2013)) falls to 7.05.[22]

The industry-level regression interpretation of $\hat{\beta}^*$ extends to other choices of the weighting matrix as well. Column 2 of Table 4, for example, shows the $\hat{s}_n$-weighted industry-level limited information maximum likelihood (LIML) coefficient from regressing $\bar{y}_n^{\perp}$ on $\bar{x}_n^{\perp}$, while column 3 reports two-step GMM-IV estimates that are optimal given heteroskedasticity and clustering of shocks at the three-digit SIC code level (i.e. using a weighting matrix W^* that is efficient under the preferred inference scheme). Overidentified shift-share IV coefficients from these procedures are -0.247 and -0.158, respectively, with the latter seeing a modest decline in its standard error. Reassuringly, the coefficient does not change much with LIML, despite a potential for weak identification indicated by the low F-statistic. Given the efficient estimate in column 4, an omnibus chi-squared specification test, with $K-1$ degrees of freedom, is obtained from the minimized criteria function in (16). This test statistic, again computed with standard statistical software, is 10.92 in the ADH application (with 7 degrees of freedom), and the p-value for joint orthogonality of the ADH shocks is 0.142.

5 Conclusion

Shift-share instruments combine a set of observed shocks with variation in local shock exposure. In this paper we provide a general condition for the validity of these instruments while focusing on the role of shock-level variation. Our quasi-experimental framework is motivated by an equivalence result: shift-share IV estimates can be reframed as coefficients from weighted industry-level regressions, in which shocks instrument for an exposure-weighted average of treatment. Shift-share instruments are therefore valid when shocks are idiosyncratic with respect to an exposure-weighted average of the unobserved factors determining outcomes. While this orthogonality condition can technically be satisfied when either the exposure measures or the shocks are as-good-as-randomly assigned, the latter may be more plausible and better aligned with researchers' motivations in many settings, such as that of Autor et al. (2013). In such settings, the quasi-experimental approach assumes shocks are drawn as-good-as-randomly and independently across industries, perhaps conditional on observables or over time, with the average exposure to any single observed industry becoming small as the sample grows.

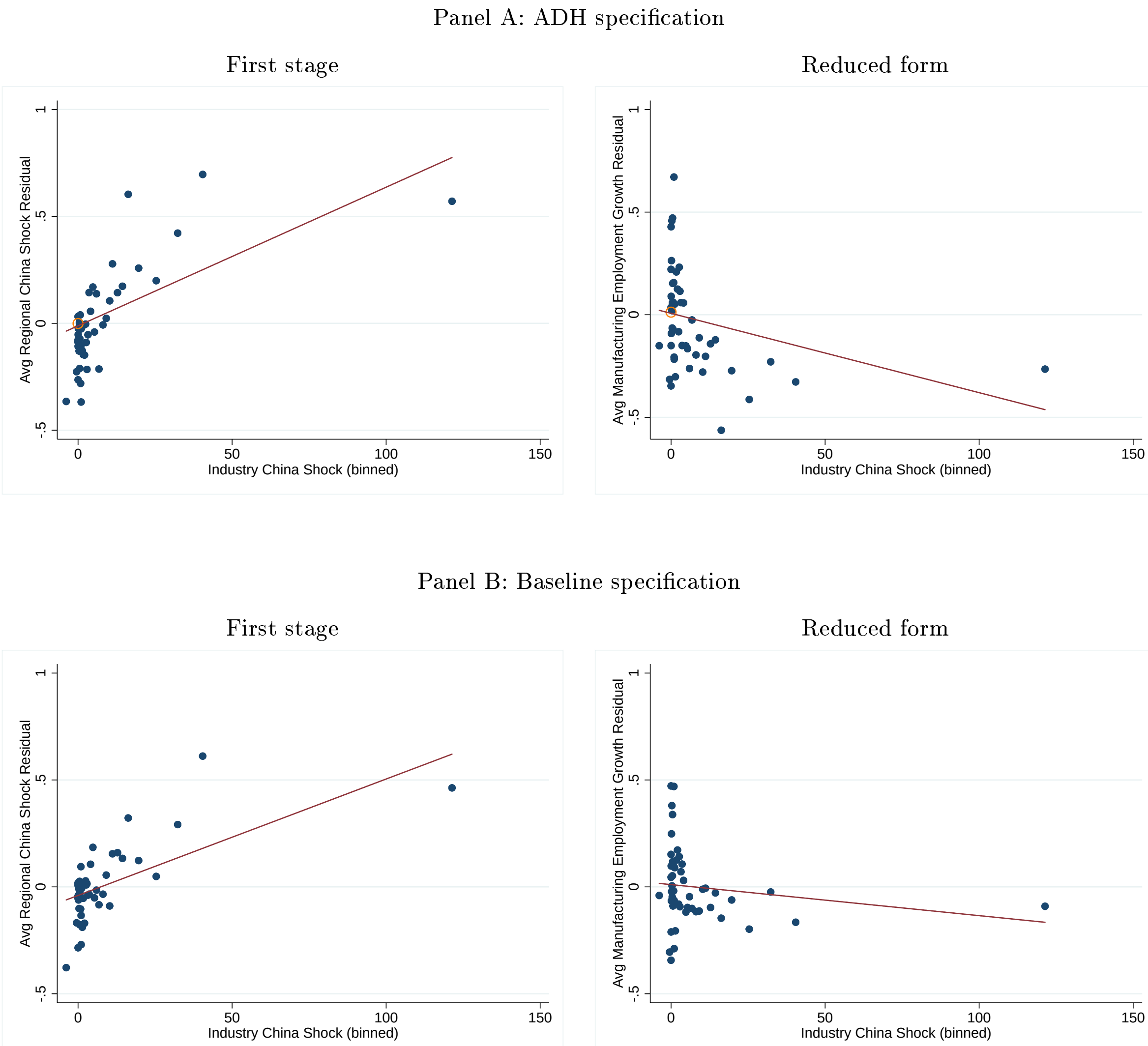
[22] Appendix Table A3 reports analogous estimates for other outcome variables.

We then discuss quasi-experimental shift-share IV in practice. As with our theoretic discussion, each of our five practical observations makes use of the equivalence result. Researchers can use industry-level analyses to inspect the identifying variation in shift-share designs, to correct conventional standard errors using the theory of Adao et al. (2018), to guide their choice of a location-level IV specification, to test the assumptions of the quasi-experimental framework, and to optimally combine multiple natural experiments. In this way our framework allows researchers to draw on intuitions from other quasi-experimental settings, bringing shift-share IV estimators to familiar econometric territory.

We conclude by noting one important limitation of our main analysis. In some shift-share applications, such as Bartik (1991) and in contrast to ADH, the industry shocks are estimated within the same sample of locations used for IV estimation. Since Autor and Duggan (2003), this is often done with leave-one-out (or other split-sample) averaging of location-specific shock measures. In Appendix A.6 we discuss shift-share IV with estimated shocks, and show that our quasi-experimental interpretation applies to these cases provided the split-sample shock estimates are unbiased. Moreover, we demonstrate that split-sample estimation may be necessary in some settings, as other shock estimators may cause the shift-share IV estimate to be inconsistent in the presence of many shocks – a problem closely related to the bias of 2SLS with many instruments. However, split-sample estimators may not have exact industry-level equivalents, and their inferential properties are not well understood. We leave further study of these issues to future research.

Figures and Tables

Figure 1: Industry-level variation in Autor et al. (2013)



Notes: This figure shows binned scatterplots of industry-level outcome and treatment residuals for fifty weighted bins of industry shocks in the ADH application, with each bin representing around 2% of total employment. The OLS lines of best fit, reported in red, use the same average exposure weights. The residuals in panel A come from the ADH replication in Table 1 column 1, while panel B illustrates the industry-level first stage and reduced form from our baseline specification in column 3. Hollow circles in panel A correspond to the “missing” non-manufacturing industry (in both periods), as explained in the text.

Table 1: Shift-share IV estimates of the effect of Chinese imports on manufacturing employment

	(1)	(2)	(3)	(4)	(5)	(6)
Coefficient	-0.596	-0.489	-0.267	-0.349	-0.252	-0.118
	(0.114)	(0.100)	(0.099)	(0.154)	(0.136)	(0.134)
Controls:						
ADH	✓	✓	✓	✓	✓	✓
Beginning-of-period mfg. share	✓					
Lagged mfg. share		✓	✓	✓	✓	✓
(Lagged mfg. share)×period			✓	✓	✓	✓
Lagged exposure to outlier industries				✓		
(10 sectoral shares)×period					✓	✓
(SIC2 shares)×period						✓
First stage F-stat.	39.65	43.90	34.26	35.25	28.86	18.59
Industry×period observations	796	794	794	760	794	794

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, instrumenting by predicted Chinese 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 beginning-of-period manufacturing shares. Column 2 replaces the beginning-of-period manufacturing shares control with lagged manufacturing shares, while column 3 interacts this control with period indicators. Column 4 adds controls for the regional exposure to each of the 34 “outlier” industry-by-period shocks detected in Figure 1. Columns 5 and 6 instead add regional shares of the 10 industry sectors defined in Autor et al. (2014) and of 20 two-digit SIC sectors, respectively, all interacted with period indicators. Standard errors and Kleibergen-Paap first-stage F-statistics are obtained from equivalent industry-level regressions, allow for clustering at the level of three-digit SIC codes, and are valid in the framework of Adao et al. (2018). The sample includes 722 locations (commuting zones) and 397 industries, each observed in two periods. The last row reflects the number of observations in the equivalent industry-level regressions, with the non-manufacturing industry aggregate included in column 1 (with a shock of zero).

Table 2: ADH balance checks

	(1)	(2)
	Coef.	SE
% college educated	0.656	(0.935)
% foreign born	2.092	(0.959)
% employment among women	-0.114	(0.372)
% employment in routine occupations	-0.216	(0.210)
Average offshorability index	0.062	(0.061)
Growth of mfg. employment in 1970s	0.389	(0.225)
Growth of mfg. employment in 1980s	0.040	(0.137)
Joint $\chi^2(7)$ significance test stat: 10.18 [p-val: 0.179]		

Notes: This table reports coefficients from regressing industry-specific averages of beginning-of-period regional characteristics (residualized by the period-specific sum of share controls and period indicators) on industry shocks and period fixed effects. Standard errors are clustered by SIC3 codes, regressions are weighted by average industry exposure, and coefficients are multiplied by 100 for readability. The first five characteristics are measured at the start of each 10-year period (1990 and 2000), while the final two pre-trend variables measure the growth in manufacturing employment in 1970-1979 and 1980-1989. The sample includes 794 industry$\times$period observations.

Table 3: Intra-class correlations of ADH shocks

	(1) ICC	(2) SE
10 sectors	0.016	(0.022)
SIC2	0.047	(0.052)
SIC3	0.073	(0.057)
Industry	0.169	(0.047)
Industry×period obs.	794	

Notes: This table reports estimated intra-class correlation coefficients from the industry-level hierarchical model described in the text. Estimates come from a maximum likelihood mixed-effects procedure with an exchangeable covariance structure for each industry cluster and period indicator controls. Robust standard errors are reported in parentheses.

Table 4: 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)
Estimator	2SLS	LIML	GMM
First stage F-stat.		7.05	
$\chi^2(7)$ overid. test stat. [p-value]		10.92 [0.142]	

Notes: Column 2 of this table reports an overidentified estimate from the baseline ADH specification in Table 1, column 3, via a two-stage least squares regression of industry-level average manufacturing employment growth residuals on industry-level average Chinese import competition growth residuals, instrumenting by growth of imports (per U.S. worker) in eight non-U.S. countries from ADH separately, controlling for period fixed effects, and weighting by average industry exposure. 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, Montiel Olea and Pflueger (2013) first-stage F-statistics, and the Hansen (1982) χ^2 test of overidentifying restrictions all allow for clustering of shocks at the SIC3 industry group level. The sample includes 794 industry$\times$period observations.

References

Acemoglu, D., D. H. Autor, D. Dorn, G. H. Hanson, and B. Price (2016): "Import Competition and the Great U.S . Employment Sag of the 2000s," *Journal of Labor Economics*, 34, S141–S198.

Acemoglu, D. and P. Restrepo (2017): "Robots and Jobs: Evidence from US Labor Markets," NBER Working Paper 23285.

Adao, R., M. Kolesar, and E. Morales (2018): "Shift-Share Designs: Theory and Inference," NBER Working Paper 24944.

Angrist, J., G. Imbens, and A. Krueger (1999): "Jackknife Instrumental Variables Estimation," *Journal of Applied Econometrics*, 14, 57–67.

Angrist, J. D. and G. W. Imbens (1994): "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 62, 467–475.

Angrist, J. D. and A. B. Krueger (1995): "Split-Sample Instrumental Variables Estimates of the Returns to Schooling," *Journal of Business and Economic Statistics*, 13, 225–235.

Autor, D. H., D. Dorn, and G. H. Hanson (2013): "The China Syndrome: Local Labor Market Impacts of Import Competition in the United States," *American Economic Review*, 103, 2121–2168.

Autor, D. H., D. Dorn, G. H. Hanson, and J. Song (2014): "Trade Adjustment: Worker-Level Evidence," *Quarterly Journal of Economics*, 129, 1799–1860.

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

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

Blanchard, O. J. and F. Katz (1992): "Regional Evolutions," *Brookings Papers on Economic Activity*, 1–75.

Borusyak, K. and X. Jaravel (2017): "Revisiting Event Study Designs, with an Application to the Estimation of the Marginal Propensity to Consume," Working Paper.

Boustan, L., F. Ferreira, H. Winkler, and E. M. Zolt (2013): "The effect of rising income inequality on taxation and public expenditures: Evidence from U.S. Municipalities and school districts, 1970-2000," *Review of Economics and Statistics*, 95, 1291–1302.

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

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

Chetty, R., J. N. Friedman, N. Hilger, E. Saez, W. Schanzenbach, and D. Yagan (2011): "How Does Your Kindergarten Classroom Affect Your Earnings? Evidence from Project STAR," *Quarterly Journal of Economics*, 126, 1593–1660.

de Chaisemartin, C. and X. D'Haultfoeuille (2018): "Two-way Fixed Effects Estimators with Heterogeneous Treatment Effects," Working Paper.

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

Dobbie, W., J. Goldin, and C. Yang (2018): "The Effects of Pre-Trial Detention on Conviction, Future Crime, and Employment: Evidence from Randomly Assigned Judges," *American Economic Review*, 108(2), 201–240.

Doyle, J., J. Graves, J. Gruber, and S. Kleiner (2015): "Measuring Returns to Hospital Care: Evidence from Ambulance Referral Patterns," *Journal of Political Economy*, 123, 170–214.

Feng, J. and X. Jaravel (2017): "Crafting Intellectual Property Rights: Implications for Patient Assertion Entities, Litigation, and Innovation," Working Paper.

Goldsmith-Pinkham, P., I. Sorkin, and H. Swift (2018): "Bartik Instruments: What, When, Why, and How," NBER Working Paper 24408.

Goodman-Bacon, A. (2018): "Difference-in-Differences with Variation in Treatment Timing," NBER Working Paper 25018.

Greenstone, M., A. Mas, and H.-L. 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, L. P. (1982): "Large Sample Properties of Generalized Method of Moments Estimators," *Econometrica*, 50, 1029–1054.

Hornbeck, R. and E. Moretti (2018): "Who Benefits From Productivity Growth? Direct and Indirect Effects of Local TFP Growth on Wages, Rents, and Inequality," NBER Working Paper 24661.

Hudson, S., P. Hull, and J. Liebersohn (2017): "Interpreting Instrumented Difference-in-Differences," Mimeo.

Hull, P. (2018): "Estimating Treatment Effects in Mover Designs," Working Paper.

Hummels, D., R. Jorgensen, J. Munch, and C. Xiang (2014): "The Wage Effects of Offshoring: Evidence From Danish Matched Worker-Firm Data," *American Economic Review*, 104, 1597–1629.

Jaeger, D. A., J. Ruist, and J. Stuhler (2018): "Shift-Share Instruments and Dynamic Adjustments: The Case of Immigration," Working Paper.

Jaravel, X. (2017): "The Unequal Gains from Product Innovations: Evidence from the US Retail Sector," Working Paper.

Kolesar, M., R. Chetty, J. Friedman, E. Glaeser, and G. W. Imbens (2015): "Identification and Inference With Many Invalid Instruments," *Journal of Business and Economic Statistics*, 33, 474–484.

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

Luttmer, E. F. (2006): "Neighbors as Negatives: Relative Earnings and Well-being," *Quarterly Journal of Economics*, 120, 963–1002.

Maestas, N., K. J. Mullen, and A. Strand (2013): "Does Disability Insurance Receipt Discourage Work? Using Examiner Assignment to Estimate Causal Effects of SSDI Receipt," *American Economic Review*, 103, 1797–1829.

Montiel Olea, J. L. and C. Pflueger (2013): "A robust test for weak instruments in Stata," *Journal of Business and Economic Statistics*, 31, 358–369.

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

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

Oberfield, E. and D. Raval (2014): "Micro Data and Macro Technology," Working Paper.

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

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

Wooldridge, J. M. (2002): *Econometric Analysis of Cross Section and Panel Data*, MIT Press.

A Appendix

A.1 Heterogeneous Treatment Effects

In this section we extend our shift-share IV identification result to allow for location-specific treatment effects. Maintaining linearity, suppose the structural outcome model is

$$y_\ell = \alpha + \beta_\ell x_\ell + \varepsilon_\ell, \tag{17}$$

where now β_ℓ denotes the treatment effect for location ℓ, and where we abstract away from other controls for simplicity. The shift-share IV estimator can then be written

$$\begin{aligned}\hat{\beta} &= \frac{\widehat{Cov}(z_\ell, y_\ell^\perp)}{\widehat{Cov}(z_\ell, x_\ell^\perp)} \\ &= \frac{\frac{1}{L}\sum_{\ell=1}^L \beta_\ell z_\ell x_\ell^\perp}{\frac{1}{L}\sum_{\ell=1}^L z_\ell x_\ell^\perp} + \frac{\frac{1}{L}\sum_{\ell=1}^L z_\ell \varepsilon_\ell^\perp}{\frac{1}{L}\sum_{\ell=1}^L z_\ell x_\ell^\perp},\end{aligned} \tag{18}$$

where $\widehat{Cov}(\cdot,\cdot)$ denotes a sample covariance. When our orthogonality condition holds $\frac{1}{L}\sum_{\ell=1}^L z_\ell \varepsilon_\ell^\perp \xrightarrow{p} 0$. Given instrument relevance ($p\lim \frac{1}{L}\sum_{\ell=1}^L z_\ell x_\ell^\perp \neq 0$), we thus have

$$\hat{\beta} = \sum_{\ell=1}^L \beta_\ell \frac{z_\ell x_\ell^\perp}{\sum_{k=1}^L z_k x_k^\perp} + o_p(1). \tag{19}$$

This shows that the shift-share IV coefficient approximates a weighted average of heterogeneous treatment effects β_ℓ, with weights $z_\ell x_\ell^\perp / \sum_{k=1}^L z_k x_k^\perp$ that sum to one. As with the classic result of Angrist and Imbens (1994), a further monotonicity assumption ensures that the weighted average IV captures is convex. Suppose treatment is generated from a linear, heterogeneous-effects first stage model

$$x_\ell = \kappa + \sum_{n=1}^N \pi_{\ell n} g_n + \upsilon_\ell, \tag{20}$$

where the $\pi_{\ell n}$ denote industry- and location-specific effects of the industry shocks on treatment. Suppose further that the shift-share orthogonality condition holds not only for the second-stage industry-level unobservable $\phi_n = \mathbb{E}[s_{\ell n}\varepsilon_\ell]/\mathbb{E}[s_{\ell n}]$, but that g_n is also uncorrelated with unobserved $\mathbb{E}[s_{\ell n}\upsilon_\ell]/\mathbb{E}[s_{\ell n}]$ and $\mathbb{E}[s_{\ell n}\beta_\ell\upsilon_\ell]/\mathbb{E}[s_{\ell n}]$, when weighted by s_n. Finally, suppose shocks are mutually mean independent conditional on $\mathbb{E}[s_{\ell n}\pi_{\ell n}]$, similar to A2. Then combining equations (20) and (19) and simplifying gives

$$\hat{\beta} = \sum_{\ell=1}^L \beta_\ell \frac{\omega_\ell}{\sum_{k=1}^L \omega_k} + o_p(1), \tag{21}$$

where

$$\omega_\ell = \sum_{n=1}^{N} \pi_{\ell n} s_{\ell n} \mathrm{Var}\left[g_n\right]. \tag{22}$$

This shows that the shift-share IV coefficient approximates a convex average of heterogenous treatment effects when the first-stage effects of shocks on treatment satisfy $\pi_{\ell n} \geq 0$ almost-surely.

A.2 Proving Quasi-Experimental Consistency and Weakening A2

This section proves that under A1, A2, and a weak regularity condition the shift-share orthogonality condition is satisfied. We then extend the proof to accommodate certain forms of the mutual shock dependence, relaxing A2.

Under A1, $\mathbb{E}\left[\sum_{n=1}^{N} s_n g_n \phi_n\right] = \sum_{n=1}^{N} \mathbb{E}\left[s_n \phi_n\right] \mu = \mathbb{E}\left[\varepsilon_\ell\right] \mu = 0$. Suppose the fourth moments of each ϕ_n and $g_n - \mu$ are bounded by finite B_ϕ^4 and B_g^4, respectively. Then by A2 and the Cauchy-Schwartz inequality

$$\begin{aligned}
\mathrm{Var}\left[\sum_{n=1}^{N} s_n g_n \phi_n\right] &= \mathbb{E}\left[\left(\sum_{n=1}^{N} s_n \phi_n \left(g_n - \mu\right)\right)^2\right] \\
&= \sum_{n=1}^{N} s_n^2 \mathbb{E}\left[\phi_n^2 \left(g_n - \mu\right)^2\right] \\
&\leq \sum_{n=1}^{N} s_n^2 \sqrt{\mathbb{E}\left[\phi_n^4\right] \mathbb{E}\left[\left(g_n - \mu\right)^4\right]} \\
&\leq B_\phi^2 B_g^2 \sum_{n=1}^{N} s_n^2 \\
&\to 0.
\end{aligned} \tag{23}$$

This implies L^2-convergence and thus weak convergence: $\sum_{n=1}^{N} s_n g_n \phi_n \xrightarrow{p} 0$.

A similar proof applies when A2 is replaced with one of two alternatives:

A2′. **(Many independent clusters)** $g_n = g_{c(n)} + g_n^*$ with $\mathbb{E}\left[\left(g_c - \mu\right)\left(g_d - \mu\right) \mid \{\phi_n\}_{n\colon c(n) \in \{c,d\}}\right] = 0$ for all $c \neq d$ and all n, and $\sum_{c=1}^{N} s_c^2 \to 0$ for $s_c = \sum_{n\colon c(n)=c} s_n$;

A2″. **(Many strongly-mixing shocks)** $\left|\mathbb{E}\left[\left(g_n - \mu\right)\left(g_m - \mu\right) \mid \phi_n, \phi_m\right]\right| \leq B_L \cdot f\left(|m-n|\right)$ for all n and m and $B_L \sum_{n=1}^{N} s_n^2 \to 0$, where $B_L \geq 0$ and $f(\cdot) \geq 0$ is a fixed function with $\sum_{r=1}^{\infty} f(r) < \infty$.

The first condition of A2′ relaxes that of A2 by allowing shocks to be grouped within mutually mean-independent clusters $c(n)$, while placing no restriction on within-cluster shock correlation. The Herfindahl index assumption of A2 is then strengthened to hold for industry clusters, with s_c denoting the expected exposure of cluster c. Assumption A2″ takes a different approach, allowing all nearby shocks to be mutually correlated provided their covariance is bounded by a function $B_L \cdot f(|m-n|)$, with the series $f(1), f(2), \ldots$ converging. For example if $f(r) = \delta^r$ for $\delta \in [0,1)$, the covariance bound for nearby shocks declines at a geometric rate δ, as when shocks are first-order autoregressive.

We now show that equation (23) holds when either A2′ or A2″ replace A2. Under A2′,

$$\begin{aligned}
\operatorname{Var}\left[\sum_{n=1}^{N} s_n g_n \phi_n\right] &= \mathbb{E}\left[\left(\sum_{c=1}^{C}\sum_{n\in N(c)} s_n\left(g_n-\mu\right)\phi_n\right)^2\right] \\
&= \sum_{c=1}^{C} s_c^2 \mathbb{E}\left[\left(\sum_{n\in N(c)} \frac{s_n}{s_c}(g_n-\mu)\phi_n\right)^2\right] \\
&= \sum_{c=1}^{C} s_c^2 \sum_{n,m\in N(c)} \frac{s_n}{s_c}\frac{s_m}{s_c}\mathbb{E}\left[(g_n-\mu)(g_m-\mu)\phi_n\phi_m\right] \\
&\leq B_\phi^2 B_g^2 \sum_{c=1}^{C} s_c^2 \to 0,
\end{aligned}$$

where $N(c) = \{n \colon c(n) = c\}$ and we again use the Cauchy-Schwartz inequality. Similarly, under A2″

$$\begin{aligned}
\operatorname{Var}\left[\sum_{n=1}^{N} s_n g_n \phi_n\right] &= \mathbb{E}\left[\left(\sum_{n=1}^{N} s_n\left(g_n-\mu\right)\phi_n\right)^2\right] \\
&= \sum_{n=1}^{N}\sum_{m=1}^{N} s_n s_m \mathbb{E}\left[\left(g_n-\mu\right)\left(g_m-\mu\right)\phi_n\phi_m\right] \\
&\leq B_L \sum_{n=1}^{N}\sum_{m=1}^{N} s_n s_m f\left(|m-n|\right)\mathbb{E}\left[\phi_n\phi_m\right] \\
&\leq B_\phi^2 B_L \sum_{n=1}^{N}\sum_{m=1}^{N} s_n s_m f\left(|m-n|\right) \\
&= B_\phi^2 B_L \left(\sum_{n=1}^{N} s_n^2 f(0) + 2\sum_{r=1}^{N-1}\sum_{m=1}^{N-r} s_{m+r} s_m f(r)\right) \\
&\leq B_\phi^2 \left(B_L \sum_{n=1}^{N} s_n^2\right)\left(f(0) + 2\sum_{r=1}^{N-1} f(r)\right) \to 0.
\end{aligned} \tag{24}$$

Here we used the fact that $\sum_{m=1}^{N-r} s_m s_{m+r} \leq \sum_{n=1}^{N} s_n^2$ for each r, since

$$\begin{aligned}
\sum_{n=1}^{N} s_n^2 &\geq \frac{1}{2}\left(\sum_{n=1}^{N-r} s_n^2 + \sum_{n=r+1}^{N} s_n^2\right) \\
&= \frac{1}{2}\left(\sum_{n=1}^{N-r} \left(s_n - s_{n+r}\right)^2 + 2 s_n s_{n+r}\right) \\
&\geq \sum_{n=1}^{N-r} s_n s_{n+r}.
\end{aligned} \tag{25}$$

A.3 Controlling for Industry Observables

This section shows that when $\mathbb{E}\left[g_n \mid \phi_n, q_n\right] = q_n'\tau$, as in Section 3.2, including an exposure-weighted vector of industry-level controls $\sum_{n=1}^{N} s_{\ell n} q_n$ in the location-level control vector w_ℓ is asymptotically equivalent to using a shift-share instrument $z_\ell^* = \sum_n s_{\ell n} g_n^*$, where $g_n^* = g_n - q_n'\tau$. For expositional simplicity, we abstract from other location-level controls and treat the matrix $\mathbb{E}\left[s_{\ell n} s_{\ell n}'\right]$ as fixed.

By the Frisch-Waugh-Lovell theorem, the shift-share IV estimator is

$$\hat{\beta} = \frac{\sum_{\ell=1}^{L} \hat{z}_\ell^* y_\ell}{\sum_{\ell=1}^{L} \hat{z}_\ell^* x_\ell}, \tag{26}$$

where $\hat{z}_\ell^*$ is the sample residual from regressing the instrument on the controls:

$$z_\ell = \sum_{n=1}^{N} s_{\ell n} q_n' \tau + z_\ell^*. \tag{27}$$

With $\sum_{n=1}^{N} s_{\ell n} = 1$, this regression can also be written

$$\sum_{n=1}^{N} s_{\ell n} g_n = \sum_{n=1}^{N} s_{\ell n} q_n' \tau + \sum_{n=1}^{N} s_{\ell n} z_\ell^*. \tag{28}$$

Let s be the $L \times N$ matrix collecting observations of $s_{\ell n}$ and let g be the $N \times 1$ vector stacking the g_n. Then we can write OLS estimates of the parameters of (28) as

$$\begin{aligned} \hat{\tau} &= \left((sq)'sq\right)^{-1} (sq)' sg \\ &= \left(q' \left(s's\right) q\right)^{-1} q' \left(s's\right) g, \end{aligned} \tag{29}$$

which is a matrix-weighted regression of g_n on q_n, with weight matrix $s's$. Under the assumptions, $\hat{\tau}$ is a consistent estimate of τ, and thus using

$$\begin{aligned} \hat{z}_\ell^* &= z_\ell - \sum_{n=1}^{N} s_{\ell n} q_n' \hat{\tau} \\ &= \sum_{n=1}^{N} s_{\ell n} \left(g_n - q_n' \hat{\tau}\right). \end{aligned} \tag{30}$$

as an instrument in (26) is asymptotically equivalent to using z_ℓ^*.

A.4 Industry-Level Standard Errors

In this section we show that conventional standard errors from certain industry-level IV regressions coincide asymptotically with the formulas in Adao et al. (2018). We prove this for heteroskedasticity-robust standard errors, but it can be similarly shown for homoskedastic or clustered standard errors.

Conventional standard errors for the $\hat{s}_n$-weighted regression of $\bar{y}_n^{\perp}$ on $\bar{x}_n^{\perp}$ and a constant, instrumented by g_n, are given by

$$\widehat{se}_{\text{reg}} = \frac{\sqrt{\sum_{n=1}^{N} \hat{s}_n^2 \hat{\varepsilon}_n^2 \left(g_n - \bar{g}\right)^2}}{\left|\sum_{n=1}^{N} \hat{s}_n \bar{x}_n^{\perp} g_n\right|}, \tag{31}$$

where $\hat{\varepsilon}_n = \bar{y}_n^{\perp} - \hat{\beta}\bar{x}_n^{\perp}$ is the estimated industry-level regression residual and $\bar{g} = \sum_{n=1}^{N} \hat{s}_n g_n$ is the $\hat{s}_n$-weighted average of shocks. Note that

$$\begin{aligned}\hat{\varepsilon}_n &= \frac{\sum_{\ell=1}^{L} s_{\ell n}\left(y_\ell^{\perp} - \hat{\beta}x_\ell^{\perp}\right)}{\hat{s}_n} \\ &= \frac{\sum_{\ell=1}^{L} s_{\ell n}\hat{\varepsilon}_\ell}{\hat{s}_n},\end{aligned} \tag{32}$$

where $\hat{\varepsilon}_\ell$ is the estimated residual from the location-level shift-share IV regression by the equivalence (4). Therefore, the numerator of (31) can be rewritten

$$\sqrt{\sum_{n=1}^{N} \hat{s}_n^2 \hat{\varepsilon}_n^2 \left(g_n - \bar{g}\right)^2} = \sqrt{\sum_{n=1}^{N} \left(\sum_{\ell=1}^{L} s_{\ell n}\hat{\varepsilon}_\ell\right)^2 \left(g_n - \bar{g}\right)^2}. \tag{33}$$

The expression in the denominator of (31) estimates the magnitude of the industry-level first-stage covariance, which matches the covariance at the location level:

$$\begin{aligned}\sum_{n=1}^{N} \hat{s}_n \bar{x}_n^{\perp} g_n &= \sum_{n=1}^{N} \left(\sum_{\ell=1}^{L} s_{\ell n} x_\ell^{\perp}\right) g_n \\ &= \sum_{\ell=1}^{L} x_\ell^{\perp} z_\ell.\end{aligned} \tag{34}$$

Thus

$$\widehat{se}_{reg} = \frac{\sqrt{\sum_{n=1}^{N} \left(g_n - \bar{g}\right)^2 \left(\sum_{\ell=1}^{L} s_{\ell n}\hat{\varepsilon}_\ell\right)^2}}{\left|\sum_{\ell=1}^{L} x_\ell^{\perp} z_\ell\right|}. \tag{35}$$

We now compare this expression to the corresponding standard errors in Adao et al. (2018). Absent location-level controls, equation (44) in that paper derives a valid IV standard error estimate as

$$\widehat{se}_{AKM} = \frac{\sqrt{\sum_{n=1}^{N} \ddot{g}_n^2 \left(\sum_{\ell=1}^{L} s_{\ell n}\hat{\varepsilon}_\ell\right)^2}}{\left|\sum_{\ell=1}^{L} x_\ell^{\perp} z_\ell\right|}, \tag{36}$$

where $\ddot{g}_n$ denote coefficients from regressing $z_\ell - \frac{1}{L}\sum_{\ell=1}^{L} z_\ell$ on all shares $s_{\ell n}$, without a constant. To understand these coefficients, note that

$$
\begin{aligned}
\frac{1}{L}\sum_{\ell=1}^{L} z_\ell &= \frac{1}{L}\sum_{\ell=1}^{L}\sum_{n=1}^{N} s_{\ell n} g_n \\
&= \sum_{n=1}^{N} \hat{s}_n g_n \\
&= \bar{g},
\end{aligned} \tag{37}
$$

and, with $\sum_{n=1}^{N} s_{\ell n} = 1$, we can rewrite

$$
\begin{aligned}
z_\ell - \frac{1}{L}\sum_{\ell=1}^{L} z_\ell &= \sum_{\ell=1}^{L} s_{\ell n} g_n - \bar{g} \\
&= \sum_{\ell=1}^{L} s_{\ell n}\left(g_n - \bar{g}\right).
\end{aligned} \tag{38}
$$

Therefore the auxiliary regression in Adao et al. (2018) has exact fit and produces $\ddot{g}_n = g_n - \bar{g}$, making $\widehat{se}_{\text{reg}}$ and $\widehat{se}_{AKM}$ identical. We emphasize that for this equivalence to hold a constant must be included in the industry-level regression; otherwise, standard statistical software packages would use g_n^2 in the numerator of (35) instead of $\left(g_n - \bar{g}\right)^2$. Including a constant does not change the shift-share IV coefficient estimate, as the weighted means of both $\bar{y}_n^\perp$ and $\bar{x}_n^\perp$ are already zero.

Next, consider the case with controls. In the Adao et al. (2018) framework, the location-level control vector can be partitioned into two components, $w_\ell = \left(w_\ell^1, w_\ell^2\right)$. The first subvector has a shift-share structure, $w_\ell^1 = \sum_n s_{\ell n} q_n$ where q_n is a vector of observed industry controls, while the second is uncorrelated with z_ℓ, controlling for w_ℓ^1. Note that q_n includes a constant; when shares sum to one this element corresponds to the constant in w_ℓ, while in the incomplete shares case discussed in Section 3.2 it corresponds to the sum of exposure measure S_ℓ. Adao et al. (2018) further assume $\mathbb{E}\left[g_n \mid \phi_n, q_n\right] = \mathbb{E}\left[g_n \mid q_n\right] = q_n'\tau$, as in equation (8).

In this case, the residuals $\ddot{g}_n$ in equation (36) satisfy $\ddot{g}_n = g_n - q_n'\hat{\tau}_{\text{AKM}}$, where $\hat{\tau}_{\text{AKM}}$ consistently estimates τ. Replacing $\ddot{g}_n$ with de-meaned $\breve{g}_n = g_n - q_n'\hat{\tau}$ for some other consistent estimate $\hat{\tau}$ thus yields asymptotically equivalent standard errors. These standard errors, along with $\hat{\beta}$, can be obtained from the $\hat{s}_n$-weighted industry-level regression of $\bar{y}_n^\perp$ on $\bar{x}_n^\perp$ and a constant, instrumented by de-meaned $\breve{g}_n$ instead of g_n.[23] Finally, note that the industry-level regression of $\bar{y}_n^\perp$ on $\bar{x}_n^\perp$, instrumented by the original g_n and controlling for q_n (which, again, includes a constant) is equivalent to using such a $\breve{g}_n$. This follows from the Frisch-Waugh-Lovell theorem: $\hat{\tau}$ here is given by the $\hat{s}_n$-weighted projection of g_n on q_n, which is consistent under the Adao et al. (2018) assumptions. Note that while this IV does

[23] Note that the weighted covariances of q_n with both $\bar{x}_n^\perp$ and $\bar{y}_n^\perp$ are zero when w_ℓ^1 is included in the first-step regression producing $y_\ell^\perp$ and $x_\ell^\perp$ (see footnote 16),. Therefore, $\sum_{n=1}^{N} \hat{s}_n \bar{x}_n^\perp \breve{g}_n = \frac{1}{L}\sum_{\ell=1}^{L} x_\ell^\perp z_\ell$ and $\sum_{n=1}^{N} \hat{s}_n \bar{y}_n^\perp \breve{g}_n = \frac{1}{L}\sum_{\ell=1}^{L} y_\ell^\perp z_\ell$. This in turn implies that the coefficient from the modified industry-level regression is $\hat{\beta}$ and the residuals are $\hat{\varepsilon}_\ell$, and standard errors from (35) apply with de-meaned $\breve{g}_n$ replacing $g_n - \bar{g}$.

not directly involve w_ℓ^2, controls without a shift-share structure may still affect standard errors via the calculation of $\bar{y}_n^\perp$ and $\bar{x}_n^\perp$. In particular they may decrease standard errors by reducing variation in the regression error $\hat{\varepsilon}_n = \bar{y}_n^\perp - \hat{\beta}\bar{x}_n^\perp$ without a commensurate reduction in the first-stage covariance.

A.5 Industry-Level Two-Stage Least Squares

This section shows that with multiple sets of shocks satisfying conditional homoskedasticity, i.e. $\operatorname{Var}[\mathbf{g}_n \mid \phi_n, q_n] = G$ for a constant matrix G, the industry-level $\hat{s}_n$-weighted 2SLS regression of $\bar{y}_n^\perp$on $\bar{x}_n^\perp$, instrumented by $\mathbf{g}_n$ and controlling for q_n, produces is the efficient GMM estimate of β.

Let $\mathbf{g}_n^*$ denote the $K \times 1$ vector of shock residuals $g_{nk}^* = g_{nk} - q_n'\hat{\tau}_k$ given consistent estimates of the τ_k in $\mathbb{E}[g_{nk} \mid \phi_n, q_n] = q_n'\tau_k$, and let $\mathbf{g}^*$ be the $N \times K$ matrix stacking observations of $\mathbf{g}_n^{*\prime}$. Next, let

$$P_{g^*} \quad = \quad \mathbf{g}^*(\mathbf{g}^{*\prime} S \mathbf{g}^*)^{-1}\mathbf{g}^{*\prime} S \tag{39}$$

be a $N \times N$ weighted projection matrix, where S is a diagonal matrix with elements $\hat{s}_n$. Finally, let $\hat{G} = \mathbf{g}^{*\prime}\boldsymbol{S}\mathbf{g}^*$ be the $\hat{s}_n$-weighted sample variance of $\mathbf{g}_n^*$ (or proportional to it when the sum of exposure weights varies), and let $\bar{x}^\perp$ and $\bar{y}^\perp$ be $N \times 1$ vectors stacking observations of $\bar{x}_n^\perp$ and $\bar{y}_n^\perp$. Under conditional homoskedasticity, $\hat{G}^{-1}$ is an efficient weight matrix W^*, so that equation (16) can be rewritten

$$\begin{aligned}
\hat{\beta}^* &= \arg\min_b \left(\mathbf{g}' S \left(\bar{y}^\perp - b\bar{x}^\perp\right)\right)' \hat{G}^{-1}\mathbf{g}' S \left(\bar{y}^\perp - b\bar{x}^\perp\right) \\
&= \arg\min_b \left(\mathbf{g}^{*\prime} S \left(\bar{y}^\perp - b\bar{x}^\perp\right)\right)' \hat{G}^{-1}\mathbf{g}^{*\prime} S \left(\bar{y}^\perp - b\bar{x}^\perp\right) \\
&= \arg\min_b \left(\bar{y}^\perp - b\bar{x}^\perp\right)' P_{g^*}' S \left(\bar{y}^\perp - b\bar{x}^\perp\right), \qquad (40)
\end{aligned}$$

where the first equality follows from the fact that both $\bar{y}_n^\perp$ and $\bar{x}_n^\perp$ are by construction orthogonal to q_n when weighted by $\hat{s}_n$ (see footnote 23), and the second follows by definition of P_{g^*}. Thus

$$\begin{aligned}
\hat{\beta}^* &= \left(\bar{x}^{\perp\prime} P_{g^*}' S \bar{x}^\perp\right)^{-1} \bar{x}^{\perp\prime} P_{g^*}' S \bar{y}^\perp \\
&= \left((P_{g^*}\bar{x}^\perp)' S \bar{x}^\perp\right)^{-1} (P_{g^*}\bar{x}^\perp)' S \bar{y}^\perp, \qquad (41)
\end{aligned}$$

which is the formula for a $\hat{s}_n$-weighted regression of $\bar{y}_n^\perp$ on the first stage weighted projection of $\bar{x}_n^\perp$ on g_n^*. By the Frisch-Waugh-Lovell theorem this is equivalent to the target 2SLS procedure, where the $\hat{\tau}_k$ vectors come from $\hat{s}_n$-weighted regressions of g_{nk} on q_n.

A.6 Estimated Shocks and Many-Shock Bias

This section extends the main analysis to settings where, in contrast to ADH, industry shocks are estimated within the IV sample of locations. For example, Bartik (1991) estimates national industry employment growth rate shocks by the sample average growth of industry employment across observed locations. He then uses these estimated shocks as instruments for local employment growth in a regression of wage growth estimated in the same location sample. Since Autor and Duggan (2003), in-sample shock estimation typically involves split-sample techniques which exclude locations from their own shock estimate (e.g. leave-one-out averages). This approach is often motivated informally by concerns over mechanical correlations between the instrument and the structural second-stage residual.

Here we establish two results. First, our quasi-experimental assumptions A1 and A2 guarantee consistency of split-sample shift-share estimators, as long as the split-sample shock estimates are unbiased and the sampling of locations is partially independent. Second, shift-share IV estimates that do not exclude the own observation from shock estimation may be inconsistent when the number of estimated shocks grows with the sample size. Intuitively, this bias arises from the fact that shock estimation error need not vanish asymptotically and may be systematically correlated with untreated potential outcomes in large samples. This result provides a strong justification for split-sample shock estimation.[24]

Suppose a researcher estimates location-specific shocks via a weighted split-sample average of variables $g_{\ell n}$. That is, given weights $\omega_{\ell n}$, she computes for each ℓ

$$\hat{g}_{\ell n} = \frac{\sum_{m \notin A_\ell} \omega_{mn} g_{mn}}{\sum_{m \notin A_\ell} \omega_{mn}}, \tag{42}$$

where A_ℓ denotes a set of observation indices that includes ℓ. For example, in Bartik (1991) $g_{\ell n}$ is the local employment growth rate of industry n in location ℓ and $\omega_{\ell n}$ is the local lagged level of industry employment; a leave-one-out estimate of shocks would then set $A_\ell = \{\ell\}$. The researcher uses $\hat{g}_{\ell n}$ to form a split-sample shift-share instrument $\hat{z}_\ell = \sum_{n=1}^{N} s_{\ell n} \hat{g}_{\ell n}$.

Suppose also that the sampling of locations is such that each ℓ is drawn independently of the locations whose indices are contained in A_ℓ. This would be satisfied by simple random sampling, but also accommodates clustered sampling provided A_ℓ includes the cluster containing ℓ. Following the

[24] For notational simplicity, we return to the convention of treating the shocks g_n as fixed, as in Section 2

logic of Section 2, validity of the split-sample shift-share instrument is then based on

$$\begin{aligned}\mathrm{Cov}\left[\tilde{z}_\ell, \varepsilon_\ell\right] &= \sum_{n=1}^{N} \mathbb{E}\left[s_{\ell n}\varepsilon_\ell \mathbb{E}\left[\tilde{g}_{\ell n} \mid s_{\ell n}, \varepsilon_\ell\right]\right] \\ &= \sum_{n=1}^{N} \hat{s}_n \mathbb{E}\left[\tilde{g}_{\ell n}\right]\phi_n, \end{aligned} \tag{43}$$

where (with slight abuse of notation in the clustered sampling case) $\mathbb{E}\left[\tilde{g}_{\ell n}\right] - g_n$ captures the bias of the split-sample shock estimator. When the estimator is unbiased, we thus have the same orthogonality condition as before, and the theory of Section 3 similarly applies.

Now consider the case without split-sample estimation. The researcher forms the shift-share instrument $\hat{z}_\ell = \sum_{n=1}^{N} s_{\ell n}\hat{g}_n$ where

$$\hat{g}_n = \frac{\sum_{\ell=1}^{L} \omega_{\ell n} g_{\ell n}}{\sum_{\ell=1}^{L} \omega_{\ell n}} \tag{44}$$

is a full-sample weighted average of the $g_{\ell n}$. Note that we can write $\hat{z}_\ell = \sum_{n=1}^{N} s_{\ell n}\hat{g}_n = z_\ell + \psi_\ell$, where

$$\psi_\ell = \sum_{n=1}^{N} s_{\ell n}(\hat{g}_n - g_n) \tag{45}$$

is a weighted average of industry-level estimation error, $\hat{g}_n - g_n$. When the orthogonality condition in Section 2 holds, i.e. when $\mathrm{Cov}\left[z_\ell, \varepsilon_\ell\right] \to 0$, validity of $\hat{z}_\ell$ requires an additional condition:

$$\mathrm{Cov}\left[\psi_\ell, \varepsilon_\ell\right] \to 0, \tag{46}$$

or that the measurement error ψ_ℓ is asymptotically uncorrelated with the structural residual ε_ℓ.

To see why this condition may fail, note that we can rewrite (45) as the sum of two terms,

$$\psi_\ell = \sum_{n=1}^{N} s_{\ell n}\frac{\omega_{\ell n}(g_{\ell n} - g_n)}{\sum_{m=1}^{L} \omega_{mn}} + \sum_{n=1}^{N} s_{\ell n}\frac{\sum_{k\neq\ell} \omega_{kn}(g_{kn} - g_n)}{\sum_{m=1}^{L} \omega_{mn}}, \tag{47}$$

where the first term captures location ℓ's own contribution to estimation error $\tilde{g}_n - g_n$ and the second is the contribution of all other locations. For $\mathrm{Cov}\left[\psi_\ell, \varepsilon_\ell\right] \not\to 0$, it is sufficient for the first term to be systematically correlated with the structural error, though in principle both may be. Intuitively the first term may be asymptotically non-ignorable if, as the number of industries grows large, the number of observations determining the $\tilde{g}_n$ estimate through the weights $\omega_{\ell n}$ remains small. If these observations also have a large exposure to g_n (as measured by $s_{\ell n}$) with deviations $g_{\ell n} - g_n$ that are systematically correlated with ε_ℓ, the feasible shift-share IV estimator will be inconsistent. In the case of Bartik (1991), where both $\omega_{\ell n}$ and $s_{\ell n}$ reflect the size of industry n in location ℓ, condition

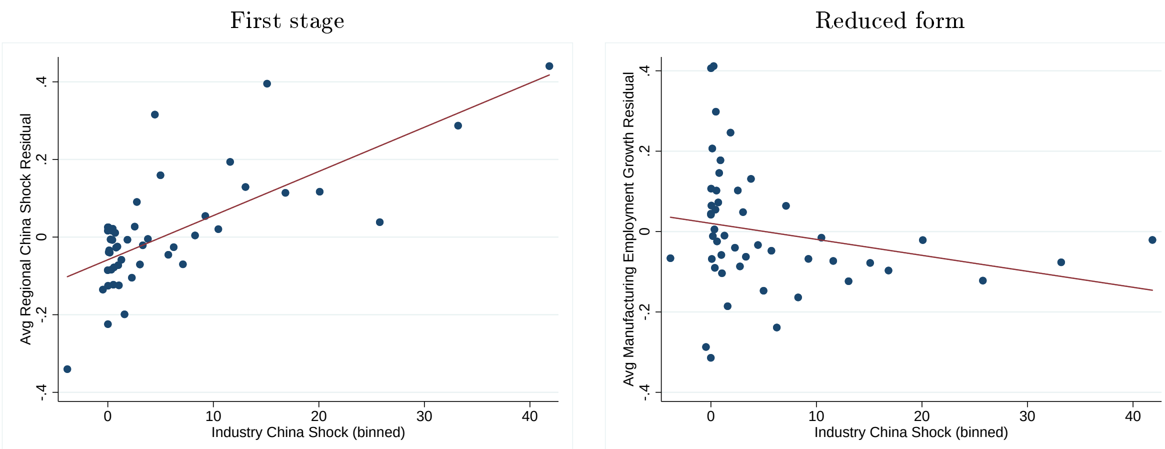
(46) may thus fail if regions with faster untreated potential wage growth also see faster employment growth in their dominant industries. This is exactly the sort of endogeneity motivating the use of shift-share IV in Bartik (1991).

This potential for inconsistency from many estimated shocks is analogous to the bias of conventional 2SLS estimation with many instruments. To make this link explicit, consider the special case of an "examiner" or "judge" design, in which each location is exposed to only one shock and the shares (e.g., examiner dummies) are used as instruments for treatment. First-stage fitted values in this setting are examiner group-specific averages of treatment, hence 2SLS with examiner dummy instruments can be thought of as a shift-share IV estimator in which shocks are given by group-specific expectations of treatment.[25] Here under homoskedasticity and independent sampling $|\text{Cov}\left[\psi_\ell, \varepsilon_\ell\right]|$ is proportional to N/L, aligning with the original many-instrument 2SLS bias term of Nagar (1959). This bias persists when the number of industries (or examiner groups) is non-negligible relative to the sample size. While stark, the examiner example may be a reasonable approximation to many shift-share designs where locations tend specialize in a few industries, so that the exposure shares resemble "fuzzy" industry group assignment. The split-sample solution to inconsistency with many estimated shocks is similarly analogous to the split-sample or jackknife instrumental variables estimators proposed by Angrist and Krueger (1995) and Angrist et al. (1999).

[25] Recent examples of examiner designs include Chetty et al. (2011), Maestas et al. (2013), Doyle et al. (2015), Feng and Jaravel (2017), and Dobbie et al. (2018). Notably, Kolesar et al. (2015) study the Chetty et al. (2011) design under the assumption that examiner groups are invalid instruments for treatment, leveraging an orthogonality condition similar to ours. Our results in this section can thus be thought to generalize this analysis to settings where shocks are not necessarily given by the instrument first stage.

Appendix Figures and Tables

Figure A1: Industry-level variation in ADH baseline specification, excluding outliers



Notes: This figure plots binned scatterplots of industry-level outcome and treatment residuals against industry shocks in the baseline ADH specification without outliers, corresponding to Table 1 column 4. Binscatters and lines of best fit use average industry exposure weights.

Table A1: Shift-share IV estimates of the effect of Chinese imports on other regional outcomes

	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Growth of unemployment						
Coefficient	0.221	0.217	0.063	0.001	0.100	0.116
	(0.049)	(0.046)	(0.060)	(0.101)	(0.083)	(0.086)
Panel B: Growth of NILF						
Coefficient	0.553	0.534	0.098	0.475	0.126	-0.067
	(0.185)	(0.183)	(0.133)	(0.218)	(0.134)	(0.149)
Panel C: Growth of log weekly wages						
Coefficient	-0.759	-0.607	0.227	0.360	0.133	0.095
	(0.258)	(0.226)	(0.242)	(0.303)	(0.281)	(0.308)
Controls:						
ADH	✓	✓	✓	✓	✓	✓
Beginning-of-period mfg. share	✓					
Lagged mfg. share		✓	✓	✓	✓	✓
(Lagged mfg. share)×period			✓	✓	✓	✓
Lagged exposure to outlier industries				✓		
(10 sectoral shares)×period					✓	✓
(SIC2 shares)×period						✓
Industry observations	796	794	794	760	794	794

Notes: This table extends the analysis of Table 1 to different outcomes in ADH: regional growth of unemployment, labor force non-participation (NILF), and log average weekly wages. The specifications in columns 1–6 otherwise mirror those of Table 1. Standard errors allow for clustering at the level of three-digit SIC codes.

Table A2: Period-specific effects in ADH

	(1) Mfg. employment	(2) Unemployment	(3) NILF	(4) Weekly 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 Tables 1 and A1, 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 beginning-of-period conditions (% college educated, % foreign-born, % employment among women, % employment in routine occupations, and the average offshorability index). Standard errors allow for clustering at the level of three-digit SIC codes. The sample includes 794 industry$\times$period observations.

Table A3: Overidentified estimates for other ADH outcomes

	(1)	(2)	(3)
Panel A: Growth of unemployment			
Coefficient	0.061	0.079	0.082
	(0.065)	(0.073)	(0.041)
$\chi^2(7)$ overid. test stat. [p-value]		8.37 [0.301]	
Panel B: Growth of NILF			
Coefficient	0.061	0.140	0.001
	(0.135)	(0.182)	(0.108)
$\chi^2(7)$ overid. test stat. [p-value]		12.59 [0.083]	
Panel C: Growth of log weekly wages			
Coefficient	0.265	0.227	-0.019
	(0.254)	(0.305)	(0.167)
$\chi^2(7)$ overid. test stat. [p-value]		13.96 [0.052]	
Estimator	2SLS	LIML	GMM

Notes: This table extends the analysis of Table 4 to different outcomes in ADH: regional growth of unemployment, labor force non-participation (NILF), and log average weekly wages. The specifications in columns 1–3 otherwise mirror those of Table 4. 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 sample includes 794 industry$\times$period observations.