Quasi-Experimental Shift-Share Research Designs

Kirill Borusyak Peter Hull Xavier Jaravel

Princeton U Chicago LSE*

July 24, 2018

Abstract

Many empirical studies leverage shift-share (or "Bartik") instruments by combining a set of aggregate shocks with measures of differential shock exposure. We derive a necessary and sufficient shock-level orthogonality condition for such instruments to identify causal effects. We then show that this condition holds when shocks are as-good-as-randomly assigned, growing in number, and sufficiently dispersed in terms of average exposure. Our quasi-experimental framework suggests several tests of shift-share instrument validity, extends to settings with conditional random assignment or multiple sets of shocks, and highlights a possible inconsistency from estimating many shocks – similar to that of two-stage least squares with many instruments – which may be addressed by split-sample estimation.

*Contact: borusyak@gmail.com, hull@uchicago.edu, and x.jaravel@lse.ac.uk. We thank Rodrigo Adão, Moya Chin, Paul Goldsmith-Pinkham, Edward Glaeser, Larry Katz, Michal Kolesár, Jack Liebersohn, Eduardo Morales, Jörn-Steffen Pischke, Isaac Sorkin, and Itzchak Tzachi Raz for helpful comments. An earlier version of this draft was circulated in May 2017 under the title *Consistency and Inference in Bartik Research Designs*.

1 Introduction

A large and growing number of empirical studies exploit variation in shift-share (or "Bartik") instruments, which combine a set of aggregate shocks with measures of differential shock exposure. Bartik (1991) and Blanchard and Katz (1992) are widely considered to have pioneered this research design in studies of local labor market dynamics. Autor, Dorn, and Hanson (2013, henceforth ADH) is a more recent application. ADH combine industry-specific shocks to Chinese import competition, measured in non-U.S. countries, with the industrial composition of U.S. labor markets. They use this shift-share instrument to estimate the effects of import competition on regional employment growth.[1]

Despite the popularity of shift-share instrumental variable (IV) regressions, few papers have studied the formal conditions underlying their validity. A recent exception is Goldsmith-Pinkham et al. (2018), who argue that identification in shift-share designs hinges on the exogeneity of shock exposure. This interpretation of ADH – a paper which we will use as a main illustrative example – requires industrial composition to be as-good-as-randomly assigned to U.S. labor markets. The Goldsmith-Pinkham et al. (2018) reasoning is based on a numerical equivalence between the shift-share estimator and an overidentified generalized method of moments (GMM) procedure that uses the exposure profile as a set of instruments, with the aggregate shocks determining the weighting matrix. While providing a coherent econometric framework for shift-share IV, this shares-as-instruments interpretation has nevertheless proved controversial.[2]

This paper develops a novel framework for understanding shift-share research designs, based on the exogeneity of the aggregate shocks themselves. Our approach also starts with a numerical equivalence: shift-share estimates are identically obtained by a just-identified IV regression, in which shocks instrument for a particular weighted average of treatment. In contrast to conventional shift-share IV procedures, the equivalent regression is estimated in the space of shocks (the industry space, for ADH). Correspondingly, we show that the shift-share IV exclusion restriction is satisfied if and only if a simple shock-level orthogonality condition holds. Shocks must be uncorrelated with a relevant unobservable: an average of untreated potential outcomes of different observations, weighted by their exposure to a given shock. In the ADH example, this unobservable reflects the average growth of employment due to unobserved factors in U.S. regions specializing in each industry. Thus for the ADH shift-share instrument to be valid, the growth of Chinese import competition, measured outside the U.S., must not be systematically different for industries concentrated in regions where employment is falling for other reasons.

[1] While most shift-share designs study regional variation in outcomes and treatment, observations may also represent 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). Other influential and recent examples of shift-share IVs, spanning many settings and topics, include Blanchard and Katz (1992), Luttmer (2006), Card (2009), Saiz (2010), Kovak (2013), Nakamura and Steinsson (2014), Oberfield and Raval (2014), Greenstone et al. (2014), Diamond (2016), Suárez and Zidar (2016), and Hornbeck and Moretti (2018).

[2] See, for example, Tim Bartik's comment on a recent online discussion: http://blogs.worldbank.org/impactevaluations/rethinking-identification-under-bartik-shift-share-instrument.

Using this characterization, we next propose a set of intuitive shock-level restrictions that satisfy the shift-share orthogonality condition. The key requirement of our approach is that shocks are as-good-as-randomly assigned, as if arising from a natural experiment. However, quasi-random shock assignment is not, by itself, enough. The equivalence result shows that for the law of large numbers to apply to shift-share quasi-experiments, the number of observed independent shocks must grow with the sample. Moreover, even though our approach allows each observation to be mostly exposed to only a small number of shocks, on average shock exposure must be sufficiently dispersed such that no finite set of shocks 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 shocks-as-instruments interpretation bears several new insights for shift-share estimation and inference in practice. First, we outline different validations of the quasi-experimental framework, such as shock-level balance tests, pre-trend checks, and tests for auto- and intra-class correlation of shocks. Second, we show how researchers can weaken the key quasi-experimental assumptions by controlling for exposure-weighted averages of shock-level observables. For example when shocks are naturally grouped into larger clusters (such as industry sectors in ADH), controlling for measures of cluster exposure allows for endogenous cluster shocks and avoids inconsistency from observing only a small number of clusters. Third, we show that for the orthogonality condition to hold when – as in ADH – the exposure weights do not sum to one, researchers must either control for the sum or de-mean the aggregate shocks. Fourth, we derive optimal shift-share IV estimators when multiple sets of shocks (such as country-specific China import shocks in the ADH setting) satisfy the orthogonality condition, along with the corresponding omnibus specification test of overidentifying restrictions. Finally, we note a potential for inconsistency when the aggregate shocks are estimated within the IV estimation sample, as in Bartik (1991). This issue is closely related to the classic bias of two-stage least squares with many instruments, which Angrist et al. (1999) show can be overcome with split-sample IV estimation. Correspondingly, we recommend that researchers use split-sample estimates of aggregate shocks when constructing shift-share instruments.

Regarding inference, we follow Adao et al. (2018) in arguing that researchers should account for the variance of the quasi-experimental shocks, in addition to usual observation sampling variation. In particular we show that conventional standard error formulas applied to a modified version of the shock-level IV regression from our equivalence result coincide with the formulas in Adao et al. (2018), and can thus be used for valid inference under their assumptions. This result extends to conventional diagnostics of first-stage instrument strength and instrument balance, suggesting researchers may wish to use shock-level regressions in practice.

A quasi-experimental shock framework may in many applications be better aligned with researchers' motivations and goals, relative to the shares-as-instruments interpretation. For example Hummels et al. (2014), who combine country-by-product supply shocks with lagged firm-specific exposure to sources of intermediate inputs, motivate their approach by writing:

> "While these shocks are exogenous to Danish firms, their impact varies markedly across firms [...]. That is, if only one Danish firm buys titanium hinges from Japan, *idiosyncratic shocks* to the supply or transport costs of those hinges affects just that one firm." (p. 1598; emphasis added)

Similarly, as noted ADH construct an instrument for Chinese import penetration in U.S. regions using the industry growth of Chinese imports in other developed economies. This research design attempts to purge Chinese industry import penetration from U.S.-specific factors, which one may view as an attempt to obtain quasi-random variation in the industry space. We emphasize, however, that the shocks-as-instruments interpretation may be less appropriate for some applications, particularly those involving only a small number of shocks.[3]

Econometrically, our approach is related to the Kolesar et al. (2015) study of consistency in IV designs with many invalid instruments. Identification in that setting follows when violations of individual instrument exclusion restrictions are uncorrelated with their first-stage effects. Here, per Goldsmith-Pinkham et al. (2018), the shock exposure measures can be thought of as a set of invalid instruments, while our key orthogonality condition requires their exclusion restriction violations to be uncorrelated with the aggregate shocks, rather than with their first-stage effects.

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 explores the empirical performance of different shift-share instrument constructions. More broadly, our paper adds to a growing literature seeking to interpret, test, and extend the high-level assumptions of common applied research designs, including, among others, Borusyak and Jaravel (2016) for event study designs, de Chaisemartin and D'Haultfoeuille (2018) for two-way fixed effect designs, and Hull (2018) for mover designs.

The rest of this paper is organized as follows. Section 2 introduces the framework and notes the equivalence between conventional shift-share IV and a particular shock-level IV estimator. Section 3 then derives and interprets the key orthogonality condition, and Section 4 proposes quasi-experimental assumptions under which it is satisfied. Section 5 discusses inference and tests of the quasi-experimental framework, while Section 6 outlines extensions and practical implications, including the potential for inconsistency with many estimated shocks. Section 7 concludes.

[3] For example, Card (2001), Kovak (2013), Dix-Carneiro and Kovak (2017), and Acemoglu and Restrepo (2017) each construct a shift-share instrument with around twenty shocks. The Goldsmith-Pinkham et al. (2018) approach, which does not require a large shock sample, may thus be more appropriate framework for understanding these designs.

2 Shocks as Instruments in Shift-Share Designs

Suppose we observe an outcome y_ℓ, a treatment x_ℓ, and a vector of controls w_ℓ (which we assume 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 $\mathbb{E}[\varepsilon_\ell] = \mathbb{E}[w_\ell \varepsilon_\ell] = 0$.[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. In the ADH setting, for example, y_ℓ and x_ℓ denote the growth rates of employment and Chinese import penetration of local labor market ℓ, while w_ℓ contains measures of labor force demographics from a previous period.[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 realized treatment x_ℓ to be uncorrelated with the potential outcomes ε_ℓ, in which case the causal parameter β would be consistently estimable by ordinary least squares (OLS). Instead, we assume that, along with $(y_\ell, x_\ell, w_\ell')'$, we observe a set of N aggregate 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 thus be thought of as a predicted local shock for location ℓ. In the ADH design, g_n denotes industry n's average Chinese import penetration growth across eight developed non-U.S. economies, while $s_{\ell n}$ is the lagged employment share of industry n in location ℓ, as measured in a base period. The shift-share instrument z_ℓ is thus interpreted as a predicted local "China shock." Again for concreteness we refer to the space of shocks as "industries," as in Bartik (1991) and ADH. 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 the aggregate shocks g_n as known. In Section 6 we relax both of these assumptions.

The shift-share IV estimator $\hat{\beta}$ uses z_ℓ to instrument for x_ℓ in equation (1). Letting $v_\ell^\perp$ denote the

[4] It is straightforward to extend our framework to models with heterogeneous treatment effects. As shown in Appendix A, the shift-share IV estimator in general captures a convex average of location-specific linear effects β_ℓ under 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 ordinary least square shift-share regressions considered by Adao et al. (2018).

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

residual from the sample projection of a variable v_ℓ on the controls w_ℓ, we can write

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

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

To build intuition for how identification in shift-share designs may come from the aggregate shocks, we first show that $\hat{\beta}$ can also be expressed as an industry-level IV estimator that uses the shocks as an instrument. Changing the order of summation across locations and industries in both the numerator and denominator of (3), we have

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

where the weights $\hat{s}_n = \frac{1}{L}\sum_{\ell=1}^{L} s_{\ell n}$ measure the average exposure of locations to industry n and where $\bar{v}_n = \sum_{\ell=1}^{L} s_{\ell n} v_\ell / \sum_{\ell=1}^{L} s_{\ell n}$ denotes a weighted average of variable v_ℓ, with larger weights given to locations more exposed to industry n.[6] Equation (4) thus shows that the shift-share IV estimator is numerically equivalent to the coefficient from an $\hat{s}_n$-weighted regression of $\bar{y}_n^\perp$, the average residualized outcome of locations specializing in n, on the same weighted average of residualized treatment, $\bar{x}_n^\perp$, instrumented by g_n.

In the ADH example, it is expected that industries with a high non-U.S. China shock g_n would be concentrated in U.S. regions facing increasing Chinese import penetration, so that the first stage covariance $\sum_{n=1}^{N} \hat{s}_n g_n \bar{x}_n^\perp$ is positive. By forming the reduced form numerator of (4) the researcher learns whether these industries are also concentrated in areas with large declines in employment. As usual for the ratio of these covariances (and thus $\hat{\beta}$) to reveal a causal effect, a particular industry-level exclusion restriction must hold. We next derive this restriction.

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_ℓ are asymptotically correlated, controlling for w_ℓ) and validity (that z_ℓ is asymptotically uncorrelated with ε_ℓ). Since relevance can be inferred from the data, we assume throughout that it is satisfied and focus our attention on validity. Applying the logic of (4) to the population covariance

[6] The weighted 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.

between the shift-share instrument and untreated potential outcomes, we have

$$\begin{aligned}\operatorname{Cov}\left[z_\ell,\varepsilon_\ell\right] &= \mathbb{E}\left[\sum_{n=1}^{N} s_{\ell n} g_n \varepsilon_\ell\right] \\ &= \sum_{n=1}^{N} s_n g_n \phi_n,\end{aligned} \tag{5}$$

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] Given a law of large numbers, i.e. $\frac{1}{L}\sum_{\ell=1}^{L} z_\ell \varepsilon_\ell - \operatorname{Cov}\left[z_\ell,\varepsilon_\ell\right] \xrightarrow{p} 0$, the shift-share IV estimator is therefore consistent if and only if

$$\sum_{n=1}^{N} s_n g_n \phi_n \to 0. \tag{6}$$

Equation (6) is our key orthogonality condition. The left-hand side represents a covariance (measured in the set of N industries and weighted by s_n) between two variables: aggregate shocks g_n and industry-level unobservables ϕ_n. Thus in ADH, equation (6) characterizes the large-sample behavior of a covariance between industry-specific measures of Chinese import penetration (g_n) and a weighted average of unobserved factors affecting the employment growth of locations specializing in each industry (ϕ_n).

The industry-level orthogonality condition helps to formalize intuitive identification arguments in the ADH design. One may, for example, be concerned that orthogonality fails due to reverse causality: China may gain market shares in certain industries precisely because these industries, and thus locations specializing in them, are not performing well in the U.S. (as reflected in ϕ_n). Measuring import penetration outside of the U.S. addresses such a concern, to the extent that the underlying performance of U.S. and non-U.S. industries are not correlated. In addition, equation (6) helps understand potential concerns about omitted variable bias, because of either unobserved industry shocks or regional unobservables that are correlated with industrial composition. 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). These too would be captured by ϕ_n.

Before proceeding, three general points on the orthogonality condition are worth highlighting. First, note that the weight each industry receives in the covariance (5) is the expected shock exposure s_n: for example, in the ADH setting s_n is the average employment share of industry n in the pop-

[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 first drawn from a larger population. All expectations and covariances in this section should then be thought to be conditional 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 4.

ulation. In practice researchers using employment shares as the exposure measure sometimes weight shift-share IV regressions by total location employment (e.g., Card (2009) and ADH themselves). In this case s_n has a more intuitive interpretation: the lagged employment of industry n. Second, note that local variation in industry exposure plays an important role in the orthogonality condition. If all locations have the same exposure to a given industry n, i.e. $s_{\ell n} = s_n$, then $\phi_n = \mathbb{E}[\varepsilon_\ell] = 0$ and this industry will not contribute to (6). Finally, note that if the exposure measures are as-good-as-randomly assigned, i.e. if $\mathbb{E}[s_{\ell n}\varepsilon_\ell] = 0$ for all n, then $\phi_n = 0$ and equation (6) is satisfied for any shocks g_n. This is the preferred interpretation of shift-share IV in Goldsmith-Pinkham et al. (2018).

4 Quasi-Experimental Shock Assignment

When might the shift-share orthogonality condition be satisfied when exposure shares are endogenous? In this section we develop a framework in which the aggregate shocks are as-good-as-randomly assigned with respect to the relevant industry-level unobservable. We also give guidance on which individual controls a researcher might include to weaken the key quasi-experimental assumption in practice.

To formalize the shock quasi-experiment, we first take a step back. Section 3 viewed industry characteristics, including shocks, as fixed given a sample of size L. We now imagine a hierarchical data-generating process in which, given the set of industry exposure weights $s_1, \ldots, s_N$, the set of aggregate shocks g_n and industry-level unobservables ϕ_n are drawn from some distribution. Our quasi-experimental framework assumes that shocks are drawn orthogonally to the unobservables, while placing no restrictions on the joint distribution of ϕ_n itself. For example, in ADH we may imagine industry import shocks arising quasi-randomly from an unanticipated policy or productivity change in China.

Specifically, our baseline case makes two assumptions on the industry-level data-generating process:

A1. Shocks are mean-independent of ϕ_n, with the same mean: $\mathbb{E}[g_n \mid \phi_n] = \mu$;

A2. Shocks are mutually mean-independent, conditional on $\phi_1, \ldots, \phi_N$, and the Herfindahl index of industry exposure weights converges to zero: $\sum_{n=1}^N s_n^2 \to 0$.

Under these and some weak regularity conditions, the orthogonality condition holds and the shift-share instrument is valid.[8]

Our key quasi-experimental assumption A1 states that the aggregate shocks are as-good-as-randomly assigned, as formalized by mean-independence with respect to the industry-level unobservables ϕ_n.

[8] Formally, we assume that the fourth moments of each ϕ_n and $g_n - \mu$ are bounded by finite B_ϕ^4 and B_g^4, respectively. Then A1 implies $\mathbb{E}\left[\sum_n s_n\phi_n g_n\right] = \sum_n \mathbb{E}[s_n\phi_n]\mu = \mathbb{E}[\varepsilon_\ell]\mu = 0$, while by A2 and the Cauchy-Schwartz inequality $\operatorname{Var}\left[\sum_n s_n\phi_n g_n\right] = \mathbb{E}\left[\left(\sum_n s_n\phi_n(g_n-\mu)\right)^2\right] \le \sum_n \sqrt{s_n^4\mathbb{E}[\phi_n^4]\mathbb{E}\left[(g_n-\mu)^4\right]} \le B_\phi^2 B_g^2 \sum_n s_n^2 \to 0$. This guarantees L^2-convergence and thus weak convergence: $\sum_n s_n\phi_n g_n \xrightarrow{p} 0$.

The second assumption A2 then ensures that a law of large numbers applies to the weighted covariance (5), satisfying the orthogonality condition asymptotically when A1 holds. Intuitively, since $\sum_{n=1}^{N} s_n^2 \geq 1/N$, the restriction on exposure concentration in A2 implies that the number of observed industries grows with the sample, while the mutual mean-independence condition implies that additional observed shocks draw (5) closer to zero under A1.[9] Note that the concentration condition is stated in terms of the Herfindahl index but this is not the only possibility; since $\sum_{n=1}^{N} s_n^2 \leq \max_n s_n$, it is sufficient that the largest industry becomes vanishingly small as the sample grows.

As with other quasi-experimental designs, researchers may wish to assume that A1 and A2 only hold conditionally on a set of observables. For example, one could weaken the mean-independence restriction of A1 to allow the mean of shocks to depend linearly on a vector of industry-level observables q_n (which includes a constant),

$$\mathbb{E}\left[g_n \mid \phi_n, q_n\right] = q_n' \tau, \tag{7}$$

and similarly assume the mutual mean-independence condition of A2 holds for the residual $g_n^* = g_n - q_n'\tau$. In this case it is straightforward to use only the residual shock variation for identification by adding an exposure-weighted vector of industry controls $\sum_{n=1}^{N} s_{\ell n} q_n$ to the control vector w_ℓ. As shown in Appendix B, the resulting estimator is equivalent to a two-step procedure in which the residuals g_n^* are first estimated by a weighted regression of g_n on q_n and then used to construct a new shift-share instrument $\hat{z}_\ell^* = \sum_n s_{\ell n} \hat{g}_n^*$. By the Frisch-Waugh-Lovell theorem, one could also directly control for q_n in the equivalent industry-level regression (4) to weaken A1 and A2.

An intuitive application of this result is found in the case where industries are grouped into bigger clusters $c(n) \in \{1, \ldots, C\}$; for example, detailed industries in ADH can be grouped into larger industrial sectors. Here researchers may be more willing to assume that the aggregate shocks are exogenously assigned within clusters, but that the cluster-average shock is endogenous. With q_n including a set of cluster indicators, the shift-share IV would then be valid with the researcher controlling for the individual exposure to each cluster (e.g. the local employment shares of each sector in ADH), $s_{\ell c} = \sum_{n=1}^{N} s_{\ell n} \mathbf{1}\left[c(n) = c\right]$.[10] Interestingly, the same approach also relaxes A2. If shocks have a cluster-specific component, $g_n = g_{c(n)} + g_n^*$, and the number of clusters is small, then the law of large numbers can apply to g_n^* but not to g_n, even if the g_c are also as-good-as-randomly assigned. There are therefore two distinct reasons to control for cluster exposure in shift-share IV: either to remove a non-random component of the shocks or to remove a random component that causes shocks to be too correlated with one another asymptotically.

When shocks are clustered but the set of cluster exposures $s_{\ell c}$ are not controlled for, consistency follows from a modified version of A2:

[9] Note that while A1 and A2 only restrict industry-level variables, we still also require a law of large numbers to hold in the sample of locations; namely, we need both $L \to \infty$ and $N \to \infty$ for consistency.

[10] Recently, Jaravel (2017) and Garin and Silverio (2017) follow this strategy.

A2′. Cluster shocks g_c are mutually mean-independent conditional on $\phi_1, \ldots, \phi_N$, and the Herfindahl index of expected cluster exposure converges to zero: $\sum_{c=1}^{C} s_c^2 \to 0$, where $s_c = \mathbb{E}\left[s_{\ell c}\right]$.

This assumption implies instead a cluster-level law of large numbers, where the number of observed clusters C increases with the sample and no one cluster drives variation in the shift-share instrument asymptotically.[11] As we discuss in the next section, this "random cluster effect" approach has implications for shift-share IV inference.

To conclude this section, we note that the quasi-experimental assumptions generalize to a panel setting in which outcomes, treatment, and shocks are observed for the same location over multiple periods. This setup, which we formally develop in Appendix C, delivers two insights. First, when the analogs of assumptions A1 and A2 hold (implying that shocks are mean-independent across both locations and time periods), identification can come either from having many industries N or from observing a finite number of industries over many periods. Second, researchers can relax the two key assumptions by including location and time fixed effects in the regression: aggregate shocks are then allowed to be correlated with time-invariant aggregated local unobservables (even if exposure varies over time) and may have time-varying means. Controls can also be constructed to extract the idiosyncratic component from serially correlated shocks when their stochastic process belongs to a known parametric class, such as an autoregressive model.

5 Inference and Testing

When the aggregate shocks are random and growing in number, as in A1 and A2, conventional 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 $N \ll L$, the covariance of g_n and ϕ_n may thus be more important than the location sampling variation targeted by conventional standard errors.

In simulations Adao et al. (2018) show that failing to account for this covariance can lead to large distortions in the coverage probabilities of standard shift-share IV confidence intervals. They then derive corrected standard error formulas under the 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. While restrictive, this framework captures two main reasons for including controls in shift-share designs: either because shocks g_n satisfy A1 only conditionally on the industry-level vector q_n (as in equation (7)) or because the $w_{\ell j}$ absorb some of the residual variance in outcomes, thereby

[11] Indeed, denoting $N(c) = \{n \mid c(n) = c\}$, $\text{Var}\left[\left(\sum_c \sum_{n \in N(c)} s_n \phi_n g_n\right)^2\right] = \sum_c s_c^2 \mathbb{E}\left[\left(\sum_{n \in N_c} \frac{s_n}{s_c} \phi_n g_n\right)^2\right] = \sum_c s_c^2 \sum_{n,m \in N(c)} \frac{s_n}{s_c} \frac{s_m}{s_c} \mathbb{E}\left[\phi_n \phi_m g_n g_m\right] \leq \sum_c s_c^2 \sum_{n,m \in N(c)} \frac{s_n}{s_c} \frac{s_m}{s_c} \sqrt[4]{\mathbb{E}\left[\phi_n^4\right] \mathbb{E}\left[\phi_m^4\right] \mathbb{E}\left[g_n^4\right] \mathbb{E}\left[g_m^4\right]} \leq B_\phi^2 B_g^2 \sum_c s_c^2 \to 0$, when A2′ holds and, as before, we assume the fourth moments of each ϕ_n and $g_n - \mu$ are bounded.

increasing the efficiency of the estimator. Note that with $\sum_{n=1}^{N} s_{\ell n} = 1$, the constant in the location-level control vector w_ℓ corresponds to a constant in q_n.

Our equivalence result (4) suggests a convenient implementation of the Adao et al. (2018) standard errors, and thus a straightforward path to correct quasi-experimental inference via standard statistical software. In Appendix D we first show that when w_ℓ contains only a constant, the conventional heteroskedastic-robust standard error formula from the industry-level regression of $\bar{y}_n^\perp$ on $\bar{x}_n^\perp$ and a constant, instrumenting with g_n and weighting by $\hat{s}_n$, matches the corresponding formula of Adao et al. (2018). More generally, the appendix shows that when controls satisfy the Adao et al. (2018) condition one can obtain valid standard errors from the industry-level regression by controlling for the set of q_{nj}. Location-level controls included only for efficiency do not require special adjustment. These results also extend to applications of cluster-robust standard errors in the industry-level regression: shocks are then allowed to have a random cluster component, as in A2$'$. Notably, the Adao et al. (2018) standard errors are valid even when the structural errors are themselves clustered by location, such as with the state-level clustering scheme of ADH.

In practice researchers may wish to exploit these results by estimating shift-share IV coefficients with the equivalent industry-level regression. An added benefit of such an approach is that it also sidesteps inferential issues with standard IV diagnostics, such as tests of instrument relevance. For example, the usual first stage F-statistic for weak instruments may be misleading when estimated at the location level, while conventional standard errors applied to the $\hat{s}_n$-weighted first stage regression of $\bar{x}_n^\perp$ on g_n and q_n are correct, as is the corresponding F-statistic.

The industry-level approach is also useful for validating the identification assumptions A1 and A2 (or A2$'$). Suppose a researcher observes a variable r_ℓ that is plausibly correlated with untreated potential outcomes ε_ℓ and wishes to indirectly verify the orthogonality condition by testing $\text{Cov}\left[z_\ell, r_\ell^\perp\right] = 0$. For example, Appendix C shows that lagged outcomes are a natural candidate for r_ℓ in the panel setting, leading to familiar pre-trend tests. A valid industry-level balance test would regress the exposure-weighted average, $\bar{r}_n^\perp = \sum_{\ell=1}^{L} s_{\ell n} r_\ell^\perp / \sum_{\ell=1}^{L} s_{\ell n}$, on shocks and industry-level controls, weighting by $\hat{s}_n$ and using conventional industry-level standard errors. Of course, instrument balance on industry-level observables thought to be correlated with ϕ_n may also be tested this way.

When there are naturally-occurring clusters of industries, researchers may also validate A2 at the industry level by testing that the intra-class correlation of shocks is zero. When intra-class correlation is present, they may either control for cluster exposure or invoke A2$'$, per the previous section. As noted, clustered industry-level standard errors must be used in the latter case.

6 Extensions

In this section we consider three extensions to the basic shift-share IV approach. First, we relax the assumption that the exposure weights sum to one in each location. Second, we consider the case where multiple sets of shocks satisfy the industry-level orthogonality condition. Finally, we note the possibility of inconsistency in shift-share designs when the aggregate shocks are estimated in the IV sample, as well as a possible solution via split-sample estimation. In each case we highlight implications for practitioners and links to existing shift-share applications.

6.1 Incomplete Shares

In some settings the sum of exposure measures $S_\ell = \sum_{n=1}^{N} s_{\ell n}$ varies across locations. For example, in ADH the set of $s_{\ell n}$ correspond to lagged employment shares of manufacturing industries only, so that S_ℓ denotes the lagged share of manufacturing in local labor market ℓ. We show that to invoke the quasi-experimental framework in such cases, a researcher must either include S_ℓ into the set of location controls w_ℓ or first de-mean the aggregate shocks.

To develop intuition for this result, suppose that the expected value of shocks, μ, is positive. Then locations with higher S_ℓ will tend to have higher values of the instrument $z_\ell = \sum_{n=1}^{N} s_{\ell n} g_n$, even when shocks are randomly assigned. If places with high S_ℓ (i.e. those with more manufacturing in ADH) are differentially affected by other factors, the orthogonality condition may fail. Showing the issue another way, note that the shift-share instrument can always be written $z_\ell = \sum_{n=1}^{N} s_{\ell n} g_n + s_{\ell 0} g_0$, where $s_{\ell 0} = 1 - S_\ell$ is location ℓ's exposure to the missing industry (non-manufacturing in ADH), $g_0 = 0$ is the shock to that industry, and the total exposure sums to one: $\sum_{n=1}^{N} s_{\ell n} + s_{\ell 0} = 1$. When the other shocks g_n are not mean-zero and the missing industry's weight $s_0 = E[s_{\ell 0}]$ remains large, its contribution to the industry-level covariance (5) violates the orthogonality condition. Controlling for the share of the missing industry, or equivalently for S_ℓ, solves this problem, as does recentering shocks to be mean-zero.

Formally, the orthogonality condition (6) requires $\sum_{n=1}^{N} s_n g_n \phi_n$ to converge to zero. Assumption A2 implies that as the sample grows, this sum converges to its expectation, which under A1 is

$$\begin{aligned} \mathbb{E}\left[\sum_{n=1}^{N} s_n g_n \phi_n\right] &= \mu \cdot \sum_{n=1}^{N} s_n \frac{\mathbb{E}\left[s_{\ell n}\varepsilon_\ell\right]}{\mathbb{E}\left[s_{\ell n}\right]} \\ &= \mu \cdot \mathbb{E}\left[S_\ell \varepsilon_\ell\right]. \end{aligned} \tag{8}$$

The expectation in (8) is zero only when the structural residual ε_ℓ is uncorrelated with the total exposure S_ℓ or when the shock mean μ is zero. The former is mechanically guaranteed when $S_\ell = 1$ or when S_ℓ is included in the list of controls, while the latter holds when shocks are first recentered.

Although in principle it is not important how shocks are de-meaned in the latter case, the numerical equivalence result (4) again suggests a straightforward approach via an industry-level regression. Namely, including a constant in the $\hat{s}_n$-weighted and g_n-instrumented regression of $\bar{y}_n^\perp$ on $\bar{x}_n^\perp$ automatically re-centers shocks to also have a weighted mean of zero, thus solving the incomplete shares issue. In general we recommend researchers include a constant whenever estimating industry-level regressions. Without incomplete shares ($S_\ell = 1$) or when S_ℓ varies but is controlled for, it is innocuous: the coefficient estimate will not change since the weighted means of both $\bar{y}_n^\perp$ and $\bar{x}_n^\perp$ are then zero. Note that including a constant also ensures standard software packages produce correct standard errors (see Appendix D).

6.2 Multiple Shocks

In some shift-share designs researchers may have access to multiple aggregate shocks plausibly satisfying A1 and A2. For example, while ADH measure Chinese import penetration by its average across eight non-U.S. countries, one may think that each set of country-specific growth rates g_{nk} for $k = 1, \ldots, 8$ 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_n. As usual, such overidentification of β raises the possibility of an efficient GMM estimator which optimally combines the quasi-experimental variation, as well as a Hansen (1982) omnibus test of the identifying assumptions. Here we show that in the quasi-experimental framework these estimators and tests differ from those usually run at the location level, but are again easily produced with certain industry-level regressions.

Suppose a researcher uses multiple shocks to form K shift-share instruments $z_{\ell k} = \sum_{n=1}^{N} s_{\ell n} g_{nk}$. A typical overidentified IV procedure for estimating β 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 the instrument-specific estimates and is thus consistent when the orthogonality condition (6) holds for each of the K sets of shocks. However, it will only be the efficient estimate when the vector of structural errors $\varepsilon = (\varepsilon_1, \ldots, \varepsilon_L)'$ 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).[12]

To characterize the optimal GMM estimator of β, we again note and leverage a numerical equivalence. Namely, the location-level moment function based on the validity of the shift-share instrument

[12] 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_ℓ on x_ℓ controlling for w_ℓ and instrumenting by $s_{\ell 1}, \ldots, s_{\ell N}$ 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.

vector $z_\ell = (z_{\ell 1}, \dots z_{\ell K})'$ can be rewritten

$$\begin{aligned} m(b) &= \sum_{\ell=1}^{L} z_\ell (y_\ell^\perp - b x_\ell^\perp) \\ &= \sum_{n=1}^{N} \hat{s}_n g_n (\bar{y}_n^\perp - b \bar{x}_n^\perp), \end{aligned} \tag{9}$$

where 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 g_n and industry-level unobservables ϕ_n. The optimal GMM estimator using $m(b)$ is then given by

$$\hat{\beta}^* = \arg\min_b m(b)' W^* m(b), \tag{10}$$

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

As in Section 5, the asymptotic theory of Adao et al. (2018) can be used to characterize W^*. In the simple case of no controls and homoskedastic shocks, i.e. $\mathrm{Var}\left[g_n \mid \phi_n\right] = \mathrm{Var}\left[g_n\right]$, that theory shows that the optimal moment-weighting matrix is proportional to the inverse variance of shocks. In this case rearranging (10) shows that the optimal shift-share estimator is equivalent to an unweighted industry-level 2SLS regression of $\hat{s}_n \bar{y}_n^\perp$ on $\hat{s}_n \bar{x}_n^\perp$, using shocks as instruments.[13] Naturally, when $K = 1$ and the model is just-identified this reduces to the earlier IV estimator (4).

The industry-level regression interpretation of $\hat{\beta}^*$ extends when there are controls or when shocks are heteroskedastic or clustered, as in the modified A2$'$. Without homoskedasticity, the GMM-IV estimator of White (1982) takes the place of 2SLS. In all cases, the minimized criteria function in (10) yields an omnibus chi-squared overidentification test, with $K - 1$ degrees of freedom.[14] As before, these estimates and test statistics are straightforward to compute with standard statistical software at the industry level.

6.3 Estimated Shocks

So far we have assumed that the researcher directly observes the set of aggregate shocks. In practice, however, shocks are typically estimated, often within the IV estimation sample. 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, within the same location sample. Here we

[13] Letting P_g denote the industry-level matrix projecting onto the vector of shocks and a constant, we have $\hat{\beta}^* = \arg\min_b \left(\sum_n \left(\hat{s}_n \bar{y}_n^\perp - b \hat{s}_n \bar{x}_n^\perp\right) g_n'\right) W^* \left(\sum_n g_n \left(\hat{s}_n \bar{y}_n^\perp - b \hat{s}_n \bar{x}_n^\perp\right)\right) = \left(\bar{x}^{\perp\prime} \hat{s} P_g \hat{s} \bar{x}^\perp\right)^{-1} \bar{x}^{\perp\prime} \hat{s} P_g \hat{s} \bar{y}^\perp$ when W^* estimates the inverse sample variance of shocks, where here $\bar{x}^\perp$ and $\bar{y}^\perp$ collect observations of $\bar{x}_n^\perp$ and $\bar{y}_n^\perp$ and where $\hat{s}$ is a $N \times N$ diagonal matrix of the $\hat{s}_n$. This corresponds to the formula for the above industry-level 2SLS regression.

[14] As usual, overidentification test rejections may come either from invalid instruments or from different weightings of heterogeneous treatment effects across the instruments. See Appendix A for a derivation of the shift-share IV weights.

show that with the many shocks required by the quasi-experimental approach, such two-step estimation may lead to inconsistency of the shift-share IV estimator even if the orthogonality condition holds. 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.

To illustrate the issue simply, we return to the case of a single set of shocks g_n, which we suppose the researcher estimates via a weighted average of observed variables $g_{\ell n}$.[15] That is, given weights $\omega_{\ell n}$, she computes

$$\hat{g}_n = \sum_{\ell=1}^{L} \frac{\omega_{\ell n} g_{\ell n}}{\sum_{m=1}^{L} \omega_{mn}}. \tag{11}$$

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. The researcher then uses $\hat{g}_n$ to form a feasible shift-share instrument $\hat{z}_\ell = \sum_{n=1}^{N} s_{\ell n}\hat{g}_n = z_\ell + \psi_\ell$, where we define

$$\psi_\ell = \sum_{n=1}^{N} s_{\ell n}(\hat{g}_n - g_n) \tag{12}$$

as a weighted average of industry-level estimation error $\hat{g}_n - g_n$. When the orthogonality condition for the infeasible shift-share instrument z_ℓ holds, i.e. $\text{Cov}\left[z_\ell, \epsilon_\ell\right] \to 0$, validity of the feasible instrument $\hat{z}_\ell$ requires an additional condition:

$$\text{Cov}\left[\psi_\ell, \varepsilon_\ell\right] \to 0, \tag{13}$$

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

When might this condition fail? Note that we can rewrite (12) as the sum of two terms,

$$\psi_\ell = \sum_{n=1}^{N} s_{\ell n} \frac{\omega_{\ell n}(g_{\ell n} - g_n)}{\sum_{m=1}^{L} \omega_{mn}} + \sum_{n=1}^{N} s_{\ell n} \frac{\sum_{k \neq \ell} \omega_{kn}(g_{kn} - g_n)}{\sum_{m=1}^{L} \omega_{mn}}, \tag{14}$$

where the first term captures location ℓ's own contribution to estimation error $\hat{g}_n - g_n$ and the second is the contribution of all other locations. For $\text{Cov}\left[\psi_\ell, \varepsilon_\ell\right] \not\to 0$, 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 aggregate shocks grows large, the number of observations determining the $\hat{g}_n$ 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 (13) 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 the first place.

[15] For notational simplicity, we also return to the convention of treating the shocks g_n as fixed.

Inconsistency from many estimated shocks is analogous to the bias of conventional 2SLS estimation with many instruments. Appendix E makes this link explicit by considering 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, so that 2SLS can be thought of as a shift-share IV estimator in which shocks are given by group-specific expectations of treatment.[16] Here $|\text{Cov}[\psi_\ell, \varepsilon_\ell]|$ 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.[17]

Fortunately, as with the conventional bias of 2SLS (Angrist and Krueger (1995); Angrist et al. (1999)), this issue may have a simple solution in the form of sample splitting. Rather than using all observations to both estimate shocks and the shift-share IV coefficient, suppose the researcher randomly partitions the sample for these two distinct purposes. At the extreme, we could imagine using leave-one-out estimates of the shocks

$$\tilde{g}_{\ell n} = \frac{\sum_{k \neq \ell} \omega_{kn} g_{kn}}{\sum_{k \neq \ell} \omega_{kn}} \tag{15}$$

to form a leave-one-out shift-share instrument $\tilde{z}_\ell = \sum_{n=1}^N s_{\ell n} \tilde{g}_{\ell n}$. Then, under independent sampling,

$$\begin{aligned} \text{Cov}[\tilde{z}_\ell, \varepsilon_\ell] &= \sum_{n=1}^N \mathbb{E}\left[s_{\ell n} \varepsilon_\ell \mathbb{E}\left[\tilde{g}_{\ell n} \mid s_{\ell n}, \varepsilon_\ell\right]\right] \\ &= \sum_{n=1}^N \hat{s}_n \mathbb{E}\left[\tilde{g}_{\ell n}\right] \phi_n, \end{aligned} \tag{16}$$

so that the validity condition for the feasible shift-share IV estimator is the same as the quasi-experimental orthogonality condition (6), with $\mathbb{E}[\tilde{g}_{\ell n}]$ replacing g_n. Of course when the leave-one-out shock estimator is unbiased for g_n, as in Bartik (1991), these conditions are the same.

This discussion of many-shock bias provides a formal justification for leave-one-out shock estimation in shift-share designs, a practice that has become common – tracing back at least as far as Autor and Duggan (2003) – though often with little theoretical underpinning. The split-sample solution also highlights a virtue of shift-share designs in which the aggregate shocks are measured in a separate

[16] Recent examples of examiner designs include Chetty et al. (2011), Maestas et al. (2013), Doyle et al. (2015), 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.

[17] Note that for the asymptotic variance of the shift-share instrument to be non-degenerate, exposure shares must be sufficiently concentrated in a small number of industries.

sample for other substantive reasons, as with the non-U.S. shocks in ADH. It is worth emphasizing that this issue does not affect the consistency of the feasible shift-share IV estimator when, as in Goldsmith-Pinkham et al. (2018), the number of industries is fixed. For the shocks-as-instruments interpretation, however, the number of industries must grow large, and split-sample shock estimation may guard against inconsistency in otherwise valid shift-share designs.

7 Conclusion

Shift-share instruments combine variation in the local exposure to aggregate shocks with variation in the shocks *per se*. We provide a general framework for understanding the validity of these instruments, while focusing on the shock variation. Our framework is motivated by a simple equivalence result: shift-share IV estimates can be reframed as coefficients from weighted industry-level regressions, which use shocks to 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, we argue that the latter may be more plausible and better aligned with researchers' motivations in many settings, such as Autor et al. (2013). The quasi-experimental approach assumes shocks are drawn as-good-as-randomly and independently across industries, perhaps conditional on observables, with the average exposure to any one industry becoming small as the sample grows.

We then outline various tests and extensions of the quasi-experimental shift-share framework. Several of these – such as the checks of instrument relevance and balance or the handling of controls, exposure weights that do not sum to one, and multiple shock instruments – are easily implemented with industry-level regressions and standard statistical software. In practice researchers may therefore wish to conduct shift-share inference and validation at the industry level. In terms of other practical recommendations, we argue that researchers should adopt the quasi-experimental mindset and take a stand on which variation in growth rates is plausibly random: for instance, within industry clusters, across clusters, or both. This choice matters for what industry- and location-level controls to include and, in the case of clusters, how to compute valid standard errors. Finally, we recommend researchers use leave-one-out or other types of split sample methods for estimating shocks, as the failure to do so may cause quasi-experimental shift-share IV estimates to be inconsistent. Each of these recommendations draw on intuitions that applied researchers are likely to have from other quasi-experimental settings, bringing shift-share IV estimators to familiar econometric territory.

A 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 aggregate 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, assume 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.

B Controlling for Industry Observables

This section shows that adding an exposure-weighted vector of industry-level controls $\sum_n s_{\ell n} q_n$ as a location-level control is equivalent to a two-step procedure in which industry-level residuals g_n^* are first estimated by a matrix-weighted regression of g_n on q_n and then used to construct a new shift-share instrument $\hat{z}_\ell^* = \sum_n s_{\ell n} \hat{g}_n^*$. For simplicity, we abstract from other location-level controls. By the Frisch-Waugh-Lovell theorem, the first shift-share IV estimator is

$$\hat{\beta} = \frac{\sum_{\ell=1}^{L} \hat{z}_\ell^* y_\ell}{\sum_{\ell=1}^{L} \hat{z}_\ell^* x_\ell}, \tag{23}$$

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

$$z_\ell = \sum_{n=1}^{N} s_{\ell n} q_n' \tau + z_\ell^*. \tag{24}$$

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

$$\sum_{n=1}^{N} s_{\ell n} g_n = \sum_{n=1}^{N} s_{\ell n} q_n' \tau + \sum_{n=1}^{N} s_{\ell n} z_\ell^*. \tag{25}$$

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 (25) as

$$\begin{aligned} \hat{\tau} &= \left((sq)'sq\right)^{-1} (sq)' sg \\ &= \left(q' \left(s's\right) q\right)^{-1} q' \left(s's\right) g, \end{aligned} \tag{26}$$

which is a matrix-weighted projection of g_n on q_n, with weight matrix $s's$. Thus

$$\begin{aligned} \hat{z}_\ell^* &= z_\ell - \sum_{n=1}^{N} s_{\ell n} q_n' \hat{\tau} \\ &= \sum_{n=1}^{N} s_{\ell n} \left(g_n - q_n' \hat{\tau}\right). \end{aligned} \tag{27}$$

This shows that $\hat{\beta}$ is equivalent to a shift-share IV regression coefficient that uses a modified aggregate shock $g_n - q_n'\hat{\tau}$. This modified shock reflects the residual from the industry-level projection (26).

C Shift-Share Instruments in Panels

In this section we consider a panel extension of the cross-sectional shift-share IV setting. We derive the orthogonality condition and quasi-experimental assumptions, paralleling Sections 3 and 4. We also show how these conditions are relaxed by including location and time fixed effects or with assumptions on the stochastic process for aggregate shocks, and propose a simple pre-trend test.

Suppose we observe T repeated observations t of outcomes, treatment, exposure, and shocks over time. We continue with a constant effects model for outcomes and treatment, but now decompose the structural error term into a fixed location component and its residual: $\varepsilon_{\ell t} = \alpha_\ell + \nu_{it}$ and

$$y_{\ell t} = \beta x_{\ell t} + w_{\ell t}'\gamma + \alpha_\ell + \nu_{\ell t}, \tag{28}$$

where α_ℓ denotes the location-specific mean of $\varepsilon_{\ell t}$. We also construct a time-varying shift-share instrument

$$z_{\ell t} = \sum_{n=1}^{N} s_{\ell t n} g_{nt}, \tag{29}$$

where g_{nt} is now the shock to industry n in time t. Note that this can be rewritten

$$z_{\ell t} = \sum_{n=1}^{N} \sum_{p=1}^{T} s_{\ell t n p} g_{np}, \tag{30}$$

where here $s_{\ell t n p} = s_{\ell t n} \mathbf{1}[n = p]$ denotes the exposure of location ℓ in time t to industry n in time p, which is zero for $n \neq p$.

With this expanded share notation, the key orthogonality condition for a the validity of a fixed effects shift-share IV regression of $y_{\ell t}$ on $x_{\ell t}$, controlling for $w_{\ell t}$ and location fixed effects, is

$$\text{Cov}\left[z_{\ell t}, \nu_{\ell t}\right] = \sum_{n=1}^{N} \sum_{t=1}^{T} s_{nt} g_{nt} \phi_{nt} \to 0, \tag{31}$$

where $s_{np} = \mathbb{E}\left[s_{\ell t n p}\right]$ and $\phi_{np} = \mathbb{E}\left[s_{\ell t n p} \nu_{\ell t}\right] / \mathbb{E}\left[s_{\ell t n p}\right]$. This is a weighted covariance of the aggregate shocks g_{nt}, now time-varying, and a time-varying measure of relevant unobservables ϕ_{nt}. Due to the location fixed effects, the industry-level unobservables here reflect a weighted average of only the time-varying component of structural residuals, $\nu_{\ell t}$.

The quasi-experimental assumptions A1 and A2 now map easily to the panel setting. The mean-independence condition is $\mathbb{E}\left[g_{nt} \mid \phi_{nt}\right] = \mu$, so that A1 is satisfied if the time varying shocks are mean-independent of the time-varying component unobservables. Importantly here shocks need not be as-good-as-randomly assigned with respect to the exposure-weighted averages of time-invariant heterogeneity, $\mathbb{E}\left[s_{\ell ntp}\alpha_\ell\right]/\mathbb{E}\left[s_{\ell tnp}\right]$. As before, A1 can be further weakened with the inclusion of industry-by-time observables. In the panel setting, a natural choice is a set of time fixed effects; by the equivalence result in Appendix B, including time fixed effects in the shift-share IV regression allows the mean of the aggregate shocks to vary over periods.

In the panel setting, assumption A2 requires the time-varying shocks to be mutually mean-independent, conditional on the set of ϕ_{nt}., with $\sum_{n=1}^N \sum_{t=1}^T s_{nt}^2 \to 0$. Note that in a balanced panel, $s_{np} = \mathbb{E}\left[s_{\ell tnp}\mathbf{1}[t=p]\right] = \mathbb{E}\left[s_{\ell tnp}\right]/T$. Thus the latter condition would be satisfied either when the number of periods T is fixed and $\max_{n,p} s_{np} \to 0$, or when N is fixed but $T \to \infty$; in long panels, shift-share IV may be consistent even with only a small number of industries.

The assumption of mutually mean-independent shocks in the panel setting rules out autoregressive shock processes: for example g_{nt} can not be conditionally correlated with $g_{n,t-1}$. One may imagine replacing assumption A2 to allow for strongly mixing or ergodic quasi-experimental shocks, along the lines of the modified A2$'$. We leave formalizing this approach for future work, noting here that given a particular time series model for g_{nt} one could use the previous result on industry-level controls to satisfy A2. For example, suppose the researcher assumes a first-order autoregressive process:

$$g_{nt} = \rho_0 + \rho_1 g_{n,t-1} + g_{nt}^*, \tag{32}$$

where the residuals g_{nt}^* are idiosyncratic. Then a researcher may choose to control for $\sum_{n=1}^N s_{\ell n,t} g_{n,t-1}$ in the panel IV specification, to only use variation in g_{nt}^* asymptotically.

With assumptions A1 and A2 holding either on the original shocks or on their idiosyncratic residual, researchers can validate the panel identifying assumptions by testing for pre-trends. This entails correlating the residualized outcome $y_{\ell t}^\perp$ with leads of the shift-share instrument, for example with $z_{\ell,t+1} = \sum_{n=1}^N s_{\ell n,t+1} g_{n,t+1}$. Under A1 and A2, we should expect

$$\mathrm{Cov}\left[y_{\ell t}^\perp, z_{\ell,t+1}\right] = \sum_{n=1}^N g_{n,t+1} \mathrm{Cov}\left[s_{\ell n,t+1}, y_{\ell t}\right] \to 0. \tag{33}$$

As discussed in section 5, validating this pre-trend condition is a special case of our test for assumption A1, where a lagged outcome is used as an observable proxy for the current-period error term.

D Industry-Level Regression 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

$$\widehat{se}_{\text{reg}} = \frac{\sqrt{\sum_{n=1}^N \hat{s}_n^2 \hat{\varepsilon}_n^2 \left(g_n - \bar{g}\right)^2}}{\left|\sum_{n=1}^N \hat{s}_n \bar{x}_n^\perp g_n\right|}, \tag{34}$$

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

$$\begin{aligned} \hat{\varepsilon}_n &= \frac{\sum_{\ell=1}^L s_{\ell n}\left(y_\ell^\perp - \hat{\beta} x_\ell^\perp\right)}{\hat{s}_n} \\ &= \frac{\sum_{\ell=1}^L s_{\ell n}\hat{\varepsilon}_\ell}{\hat{s}_n}, \end{aligned} \tag{35}$$

where $\hat{\varepsilon}_\ell$ is the estimated residual from the location-level shift-share IV regression by the equivalence (4). Therefore, the numerator of (34) can be rewritten

$$\sqrt{\sum_{n=1}^N \hat{s}_n^2 \hat{\varepsilon}_n^2 \left(g_n - \bar{g}\right)^2} = \sqrt{\sum_{n=1}^N \left(\sum_{\ell=1}^L s_{\ell n}\hat{\varepsilon}_\ell\right)^2 \left(g_n - \bar{g}\right)^2}. \tag{36}$$

The expression in the denominator of (34) estimates the magnitude of the industry-level first-stage covariance, which matches the covariance at the location level:

$$\begin{aligned} \sum_{n=1}^N \hat{s}_n \bar{x}_n^\perp g_n &= \sum_{n=1}^N \left(\sum_{\ell=1}^L s_{\ell n} x_\ell^\perp\right) g_n \\ &= \sum_{\ell=1}^L x_\ell^\perp z_\ell. \end{aligned} \tag{37}$$

Thus

$$\widehat{se}_{reg} = \frac{\sqrt{\sum_{n=1}^N \left(g_n - \bar{g}\right)^2 \left(\sum_{\ell=1}^L s_{\ell n}\hat{\varepsilon}_\ell\right)^2}}{\left|\sum_{\ell=1}^L x_\ell^\perp z_\ell\right|}. \tag{38}$$

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

$$\widehat{se}_{AKM} = \frac{\sqrt{\sum_{n=1}^N \ddot{g}_n^2 \left(\sum_{\ell=1}^L s_{\ell n}\hat{\varepsilon}_\ell\right)^2}}{\left|\sum_{\ell=1}^L x_\ell^\perp z_\ell\right|}, \tag{39}$$

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

$$
\begin{aligned}
\frac{1}{L}\sum_{\ell=1}^{L} z_\ell &= \frac{1}{L}\sum_{\ell=1}^{L}\sum_{n=1}^{N} s_{\ell n} g_n \\
&= \sum_{n=1}^{N} \hat{s}_n g_n \\
&= \bar{g},
\end{aligned} \tag{40}
$$

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

$$
\begin{aligned}
z_\ell - \frac{1}{L}\sum_{\ell=1}^{L} z_\ell &= \sum_{\ell=1}^{L} s_{\ell n} g_n - \bar{g} \\
&= \sum_{\ell=1}^{L} s_{\ell n}\left(g_n - \bar{g}\right).
\end{aligned} \tag{41}
$$

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 (38) instead of $(g_n - \bar{g})^2$. 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 of Section 8 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 (7).

In this case, the residuals $\ddot{g}_n$ in equation (39) 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.[18] 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

[18] 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 $x_\ell^\perp$, as, for example $\sum_n q_n s_n \bar{y}_n^\perp = \sum_n q_n \sum_\ell s_{\ell n} y_\ell^\perp = \sum_\ell y_\ell^\perp w_\ell^1 = 0$. 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 (38) 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.

E Many-Shock Bias in the Examiner Case

This section shows how the asymptotic bias of shift-share IV with many estimated shocks reduces to that of Nagar (1959) in an examiner design. Exposure shares in this setting are binary, $s_{\ell n} \in \{0,1\}$, with shocks given by the group-level expectation of treatment: $g_n = \mathbb{E}\left[x_\ell \mid s_{\ell n} = 1\right]$. The researcher estimates shocks by the corresponding sample average,

$$\hat{g}_n = \frac{\sum_{\ell=1}^L s_{\ell n} x_\ell}{\sum_{\ell=1}^L s_{\ell n}}, \tag{42}$$

so that in terms of the general equation (11), $g_{\ell n} = x_\ell$ and $\omega_{\ell n} = s_{\ell n}$.

For simplicity here we assume the control vector w_ℓ contains only a constant. Projecting treatment and the structural error onto the shares, we have

$$x_\ell = \sum_{n=1}^N g_n s_{\ell n} + \eta_\ell \tag{43}$$

$$\varepsilon_\ell = \sum_{n=1}^N \phi_n s_{\ell n} + \nu_\ell, \tag{44}$$

where $\mathbb{E}\left[\eta_\ell \mid s_{\ell 1}, \ldots, s_{\ell N}\right] = \mathbb{E}\left[\nu_\ell \mid s_{\ell 1}, \ldots, s_{\ell N}\right] = 0$ by construction. Let s_ℓ be a $N \times 1$ vector collecting $s_{\ell n}$, s be a $L \times N$ matrix collecting s'_ℓ, g be an $N \times 1$ vector collecting g_n, x be an $L \times 1$ vector collecting x_ℓ, and η be an $L \times 1$ vector collecting η_ℓ. Then in matrix form the measurement error term (12) is

$$\begin{aligned} \psi_\ell &= s'_\ell \left(\left(s's\right)^{-1} s'x - g \right) \\ &= s'_\ell \left(s's\right)^{-1} s'\eta. \end{aligned} \tag{45}$$

Under independent sampling and conditional homoskedasticity of (η_ℓ, ν_ℓ), we then have

$$\begin{aligned} \operatorname{Cov}\left[\psi_\ell, \varepsilon_\ell\right] &= \mathbb{E}\left[s'_\ell \left(s's\right)^{-1} s'\eta\varepsilon_\ell\right] \\ &= \mathbb{E}\left[s'_\ell \left(s's\right)^{-1} s_\ell \operatorname{Cov}\left[\varepsilon_\ell, \eta_\ell \mid s\right]\right] \\ &= \mathbb{E}\left[s'_\ell \left(s's\right)^{-1} s_\ell \sigma_{\eta\nu}\right] \\ &= \frac{N}{L}\sigma_{\eta\nu}, \end{aligned} \tag{46}$$

where $\sigma_{\eta\nu} = \operatorname{Cov}\left[\eta_\ell, \nu_\ell\right] = \operatorname{Cov}\left[\varepsilon_\ell, \eta_\ell\right]$.

References

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): "Inference in Shift-Share Designs: Theory and Inference," Working Paper.

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. 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 (2016): "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.

——— (2009): "Immigration and Inequality," *American Economic Review: Papers & Proceedings*, 99, 1–21.

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.

Dix-Carneiro, R. and B. K. Kovak (2017): "Trade Liberalization and Regional Dynamics," *American Economic Review*, 107, 2908–2946.

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.

GARIN, A. AND F. SILVERIO (2017): "How Does Firm Performance Affect Wages? Evidence from Idiosyncratic Export Shocks," Working Paper.

GOLDSMITH-PINKHAM, P., I. SORKIN, AND H. SWIFT (2018): "Bartik Instruments: What, When, Why, and How," Working Paper.

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.

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.

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," .

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.

WHITE, H. (1982): "Instrumental Variables Regression with Independent Observations," *Econometrica*, 50, 483–499.

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