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

NBER working papers are circulated for discussion and comment purposes. They have not been peer-reviewed or been subject to the review by the NBER Board of Directors that accompanies official NBER publications.

© 2018 by Kirill Borusyak, Peter Hull, and Xavier Jaravel. All rights reserved. Short sections of text, not to exceed two paragraphs, may be quoted without explicit permission provided that full credit, including © notice, is given to the source.
ABSTRACT

Many studies use shift-share (or “Bartik”) instruments, which average a set of shocks with exposure share weights. We provide a new econometric framework for shift-share instrumental variable (SSIV) regressions in which identification follows from the quasi-random assignment of shocks, while exposure shares are allowed to be endogenous. The framework is motivated by an equivalence result: the orthogonality between a shift-share instrument and an unobserved residual can be represented as the orthogonality between the underlying shocks and a shock-level unobservable. SSIV regression coefficients can similarly be obtained from an equivalent shock-level regression, motivating shock-level conditions for their consistency. We discuss and illustrate several practical insights delivered by this framework in the setting of Autor et al. (2013).

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

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

Peter Hull
University of Chicago
5757 South University Avenue
Chicago, IL 60637
and NBER
hull@uchicago.edu

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

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

The claim for instrument validity in shift-share instrumental variable (SSIV) regressions must rely on some assumptions about the shocks, exposure shares, or both. This paper develops a novel framework for understanding such regressions as leveraging exogenous variation in the shocks, allowing the variation in exposure shares to be endogenous. Our approach is motivated by an equivalence result: the orthogonality between a shift-share instrument and an unobserved residual can be represented as the orthogonality between the underlying shocks and a shock-level unobservable. Given a first stage, it follows that the instrument identifies a parameter of interest if and only if the shocks are uncorrelated with this unobservable, which captures the average unobserved determinants of the original outcome among observations most exposed to a given shock. SSIV regression coefficients can similarly be obtained from an equivalent IV regression estimated at the level of shocks. In this regression the outcome and treatment variables are first averaged, using exposure shares as weights, to obtain shock-level aggregates. The shocks then directly instrument for the aggregated treatment. Importantly, these equivalence results only rely on the structure of the shift-share instrument and thus apply to outcomes and treatments that are not typically computed at the level of shocks.

We use these equivalence results to derive two conditions sufficient for SSIV consistency. First, we assume shocks are as-good-as-randomly assigned as if arising from a natural experiment. This is enough for the shift-share instrument to be valid: i.e. for the shocks to be uncorrelated with the relevant unobservables in expectation. Second, we assume that a shock-level law of large numbers applies—that the instrument incorporates many sufficiently independent shocks, each with sufficiently small average exposure. Instrument relevance further holds when individual units are mostly exposed to only a small number of shocks, provided those shocks affect treatment. Our two quasi-experimental conditions are similar to ones imposed in other settings where the underlying shocks are directly used as instruments, bringing SSIV to familiar econometric territory.\(^2\)

\(^1\)Observations in shift-share designs may, for example, represent regions impacted by immigration shocks from different countries (Card 2001; Peri et al. 2016), firms differentially exposed to foreign market shocks (Hummels et al. 2014; Berman et al. 2015), product groups purchased by different types of consumers (Jaravel 2019), groups of individuals facing different national income trends (Boustan et al. 2013), or countries differentially exposed to the U.S. food aid supply shocks (Nunn and Qian 2014). We present a taxonomy of existing shift-share designs, and how they relate to our framework, in Section 6.1.

\(^2\)For example, Acemoglu et al. (2016) study the impact of import competition from China on U.S. industry employment using industry (i.e. shock-level) regressions with shocks constructed similarly to those underlying the regional shift-share instrument used in Autor et al. (2013). Our framework shows that both studies can rely on similar econometric assumptions, though the economic interpretations of the estimates differ.
We extend our quasi-experimental approach to settings where shocks are as-good-as-randomly assigned only conditionally on shock-level observables, to SSIVs with exposure shares that do not add up to a constant for each observation, and to panel data. For conditional random assignment, we show that quasi-experimental shock variation can be isolated with regression controls that have a shift-share structure. Namely, it is enough to control for an exposure-weighted sum of the relevant shock-level confounders. Relatedly, in SSIVs with “incomplete shares,” where the sum of exposure shares varies across observations, it is important to control for the sum of exposure shares as the exposure-weighted sum of a constant. In panel data, we show that the SSIV estimator can be consistent both with many shocks per period and with many periods. We also show that unit fixed effects only isolate variation in shocks over time when exposure shares are time-invariant. In other extensions we show how SSIV with multiple endogenous variables can be viewed quasi-experimentally and how multiple sets of quasi-random shocks can be combined with new overidentified shock-level IV procedures.

Our framework also bears practical tools for SSIV inference and testing. Adão et al. (2019) show that conventional standard errors in SSIV regressions may be invalid because observations which similar exposure shares are likely to have correlated residuals. They are also the first to propose a solution to this inference problem in a framework based on ours, with identifying variation in shocks. We present a convenient alternative based on our equivalence result: estimating SSIV coefficients at the level of identifying variation (shocks) can yield asymptotically valid standard errors. The validity of this solution requires an additional assumption on the structure of the included controls (producing standard errors that are typically conservative otherwise). However, it offers several practical features: it can be implemented with standard statistical software, extended to various forms of shock dependence (e.g. autocorrelation), and computed in some settings where the estimator of Adão et al. (2019) fails (e.g. when there are more shocks than observations). Appropriate measures of first-stage relevance and valid falsification tests of shock exogeneity can also be obtained with conventional shock-level procedures. Monte-Carlo simulations confirm the accuracy of our asymptotic approximations in moderately-sized samples of shocks, and that the finite-sample properties of SSIV are similar to those of conventional shock-level IV regressions which use the same shocks as instruments.

We illustrate the practical insights from our framework in the setting of Autor et al. (2013), finding supporting evidence for the interpretation of their SSIV as leveraging quasi-random variation in industry-specific Chinese import shocks. This application uses a new Stata package, ssaggregate, which we have developed to help practitioners implement the appropriate shock-level analyses.3

Our quasi-experimental approach is not the only framework for SSIV identification and consistency. In related work, Goldsmith-Pinkham et al. (2020) formalize a different approach based on the exogeneity of the exposure shares, imposing no explicit assumption of shock exogeneity. This framework

---

3This Stata package creates the shock-level aggregates used in the equivalent regression. Users can install this package with the command ssc install ssaggregate. See the associated help file and this paper’s replication archive at https://github.com/borusyak/shift-share for more details.
is motivated by a different equivalence result: the SSIV coefficient also coincides with a generalized method of moments estimator, with exposure shares as multiple excluded instruments. Though exposure exogeneity is a sufficient condition for identification (and, as such, implies our shock-level orthogonality condition), we focus on plausible conditions under which it is not necessary.

Identification via exogenous shocks is attractive in many SSIV settings. Consider the Autor et al. (2013; hereafter ADH) shift-share instrument, which combines industry-specific changes in Chinese import competition (the shocks) with local exposure given by the lagged industrial composition of U.S. regions (the exposure shares). In such a setting, exogeneity of industry employment shares is difficult to justify \textit{a priori} since unobserved industry shocks (e.g., automation or innovation trends) are likely to affect regional outcomes through the same mixture of exposure shares. Our approach, in contrast, allows researchers to specify a set of shocks that are plausibly uncorrelated with such unobserved factors. Consistent with this general principle, ADH attempt to purge their industry shocks from U.S.-specific confounders by measuring Chinese import growth outside of the United States. Similarly, Hummels et al. (2014) combine country-by-product changes in transportation costs to Denmark (as shocks) with lagged firm-specific composition of intermediate inputs and their sources (as shares). They argue these shocks are “idiosyncratic,” which our approach formalizes as “independent from relevant country-by-product unobservables.” Other recent examples of where our approach may naturally apply are found, for example, in finance (Xu 2019), the immigration literature (Peri et al. 2016), and studies of innovation (Stuen et al. 2012).

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

Formally, our approach to SSIV relates to the analysis of IV estimators with many invalid instruments by Kolesar et al. (2015). Consistency in that setting follows when violations of individual instrument exclusion restrictions are uncorrelated with their first-stage effects. For quasi-experimental SSIV, the exposure shares can be thought of as a set of invalid instruments (per the Goldsmith-Pinkham et
al. (2020) interpretation), and our orthogonality condition requires their exclusion restriction violations to be uncorrelated with the shocks. Despite this formal similarity, we argue that shift-share identification is better understood through the quasi-random assignment of a single instrument (shocks), rather than through a large set of invalid instruments (exposure shares) that nevertheless produce a consistent estimate. This view is reinforced by our equivalence results, yields a natural shock-level identification condition, and suggests new validations and extensions of SSIV.

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

The remainder of this paper is organized as follows. Section 2 introduces the environment, derives our equivalence results, and motivates our approach to SSIV identification and consistency. Section 3 establishes the baseline quasi-experimental assumptions and Section 4 derives various extensions. Section 5 discusses shock-level procedures for valid SSIV inference and testing. Section 6 summarizes the types of empirical settings where our framework may be applied and illustrates its practical implications in the ADH setting. Section 7 concludes.

2 Setting and Motivation

We begin by presenting the SSIV setting and motivating our approach to identification and consistency with two equivalence results. We first show that population orthogonality of the shift-share instrument can be recast at the shock level, motivating identification by exogenous shocks when exposure shares are endogenous. We then derive a similar shock-level equivalence result for the SSIV estimator, motivating its consistency with many as-good-as-randomly assigned shocks.

2.1 The Shift-Share IV Setting

We observe an outcome $y_\ell$, treatment $x_\ell$, control vector $w_\ell$ (which includes a constant) and shift-share instrument $z_\ell$ for a set of observations $\ell = 1, \ldots, L$. We also observe a set of regression weights $e_\ell > 0$
with $\sum_\ell e_\ell = 1$ ($e_\ell = \frac{1}{L}$ covers the unweighted case). The instrument can be written as
\[ z_\ell = \sum_n s_{\ell n} g_n, \]  
for a set of observed shocks $g_n$, $n = 1, \ldots, N$, and a set of observed shares $s_{\ell n} \geq 0$ defining the exposure of each observation $\ell$ to each shock $n$. Initially we assume the sum of these exposure weights is constant across observations, i.e. that $\sum_n s_{\ell n} = 1$; we relax this assumption in Section 4.2. Although our focus is on shift-share IV, we note that the setup nests shift-share reduced-form regressions, of $y_\ell$ on $z_\ell$ and $w_\ell$, when $x_\ell = z_\ell$.

We seek to estimate the causal effect or structural parameter $\beta$ in a linear model of
\[ y_\ell = \beta x_\ell + w_\ell' \gamma + \varepsilon_\ell, \]  
where the residual $\varepsilon_\ell$ is defined to be orthogonal with the control vector $w_\ell$. For example, we might be interested in estimating the inverse labor supply elasticity $\beta$ from observations of log wage growth $y_\ell$ and log employment growth $x_\ell$ across local labor markets $\ell$. The residual $\varepsilon_\ell$ in equation (2) would then contain all local labor supply shocks, such as those arising from demographic, human capital, or migration changes, that are not systematically related to the observed controls in $w_\ell$. To estimate this $\beta$ we may use an instrument capturing local labor demand variation. For illustration, we consider a $z_\ell$ based on new government subsidies $g_n$ to the output of different industries $n$ and with $s_{\ell n}$ denoting location $\ell$’s lagged shares of industry employment. In estimating $\beta$ we may weight observations by the overall lagged regional employment, $e_\ell$. We return to this labor supply example at several points to ground the following theoretic discussion.

It is worth highlighting that in studying this setting we do not impose a typical assumption of independent and identically-distributed (iid) data $\{e_\ell, z_\ell, w_\ell, x_\ell, \varepsilon_\ell\}$, as might arise from random sampling of potential observations. Relaxing the usual iid assumption is required for us to treat the $g_n$ as random variables, which generate dependencies of the instrument (1) across observations exposed to the same random shocks. The non-iid setting further allows for unobserved common shocks, which may generate dependencies in the residual $\varepsilon_\ell$.

Given this non-iid setting, we consider IV identification of $\beta$ by the full-data moment condition
\[ E \left[ \sum_\ell e_\ell z_\ell \varepsilon_\ell \right] = 0. \]  

Footnote 3In practice shift-share instruments are sometimes presented differently, as footnote 38 below illustrates in the Autor et al. (2013) setting. We recommend that researchers follow the representation in (1) to apply our theoretical results.

Footnote 5Formally, given a linear causal or structural model of $y_\ell = \beta x_\ell + \varepsilon_\ell$ we define $\gamma = E \left[ \sum_\ell e_\ell w_\ell' w_\ell' \right]^{-1} E \left[ \sum_\ell e_\ell w_\ell' \varepsilon_\ell \right]$ and $\varepsilon_\ell = \varepsilon_\ell - w_\ell' \gamma$ as the residual from this population projection, satisfying $E \left[ \sum_\ell e_\ell w_\ell' \varepsilon_\ell \right] = 0$. Defining a unique $\gamma$ requires an implicit maintained assumption that $E \left[ \sum_\ell e_\ell w_\ell' w_\ell' \right]$ is of full rank, which holds when there is no perfect collinearity in the control vector.

Footnote 6We consider models with heterogeneous treatment effects in Appendix A.1; see footnote 15 for a summary.
This condition captures the orthogonality of the shift-share instrument with the second-stage residual, in expectation over realizations of \( \{e_\ell, z_\ell, \varepsilon_\ell\} \) for all \( \ell = 1, \ldots, L \). When such orthogonality holds the \( \beta \) parameter is identified: i.e., uniquely recoverable from observable moments, provided the instrument has a first stage.\(^7\) The full-data orthogonality condition generalizes the conventional condition of \( \mathbb{E}[z_\ell \varepsilon_\ell] = 0 \), which might be considered in an iid setting with fixed \( e_\ell \).

The moment condition (3) yields a natural estimator of \( \beta \): the coefficient on \( x_\ell \) in an IV regression of \( y_\ell \) which instruments by \( z_\ell \), controls for \( w_\ell \), and weights by \( e_\ell \). By the Frisch-Waugh-Lovell theorem, this SSIV estimator can be represented as a bivariate IV regression of outcome and treatment residuals, or as the ratio of \( e_\ell \)-weighted sample covariances between the instrument and the residualized outcome and treatment:

\[
\hat{\beta} = \frac{\sum_\ell e_\ell z_\ell y_\ell^\perp}{\sum_\ell e_\ell z_\ell x_\ell^\perp},
\]

where \( v_\ell^\perp \) denotes the residual from an \( e_\ell \)-weighted sample projection of a variable \( v_\ell \) on the control vector \( w_\ell \). Note that by the properties of residualization, it is enough to residualize \( y_\ell \) and \( x_\ell \) without also residualizing the shift-share instrument \( z_\ell \).

In our non-iid setting, we study consistency and other asymptotic properties of \( \hat{\beta} \) by considering a sequence of data-generating processes, indexed by \( L \), for the complete data \( \{e_\ell, s_n, g_n, w_\ell, x_\ell, \varepsilon_\ell\} \), for \( \ell = 1, \ldots, L, n = 1, \ldots, N \), and \( N = N(L) \). Consistency, for example, is defined as \( \hat{\beta} \xrightarrow{p} \beta \) as \( L \to \infty \) along this sequence. We do not employ conventional sampling-based asymptotic sequences (and corresponding laws of large numbers) as these are generally inappropriate in a non-iid setting where both \( z_\ell \) and \( \varepsilon_\ell \) may exhibit non-standard mutual dependencies. It is worth emphasizing that any assumptions on the data-generating sequence are useful only to approximate the finite-sample distribution of the SSIV estimator, not to define an actual process for realizations of the data. For example, we will consider below a sequence in which the number of shocks \( N \) grows with \( L \), recognizing that in reality shift-share instruments are constructed from a fixed set of shocks (e.g. subsidies across all industries) along with a fixed number of observations (e.g. all local labor markets). The assumption of growing \( N \) should here be interpreted as a way to capture the presence of a large number of shocks in a given set of observations, such that the asymptotic sequence provides a good approximation to the observed data.\(^8\)

### 2.2 A Shock-Level Orthogonality Condition

We first build intuition for our approach to satisfying the IV moment condition by showing that the structure of the shift-share instrument allows equation (3) to be rewritten as condition on the

\(^7\)Formally, when equation (3) holds the moment condition \( m(b, c) \equiv \mathbb{E}\left[ \sum_\ell e_\ell(z_\ell, w_\ell')' (y_\ell - bx_\ell - w_\ell' c) \right] = 0 \) has a unique solution of \( (\beta, \gamma) \), provided \( \mathbb{E}\left[ \sum_\ell e_\ell(z_\ell, w_\ell')' (x_\ell, w_\ell') \right] \) is of full-rank.

\(^8\)This is similar to how Bekker (1994) uses a non-standard asymptotic sequence to analyze IV estimators with many instruments: “The sequence is designed to make the asymptotic distribution fit the finite sample distribution better. It is completely irrelevant whether or not further sampling will lead to samples conforming to this sequence” (p. 638).
orthogonality of shocks $g_n$. Namely, by exchanging the order of summation across $\ell$ and $n$, we obtain

$$
\mathbb{E} \left[ \sum_{\ell} e_\ell z_{\ell} \varepsilon_{\ell} \right] = \mathbb{E} \left[ \sum_{\ell} e_\ell \sum_n s_{\ell n} g_n \varepsilon_{\ell} \right] = \mathbb{E} \left[ \sum_n s_n g_n \bar{\varepsilon}_n \right] = 0,
$$

(5)

where we define $s_n = \sum_\ell e_\ell s_{\ell n}$ and $\bar{\varepsilon}_n = \sum_{\ell_1} \bar{\varepsilon}_{\ell_1} s_{\ell_1 n}$. Just as the left-hand side of this expression captures the orthogonality of the instrument $z_{\ell}$ with the residual $\varepsilon_{\ell}$ when weighted by $e_\ell$, the right-hand side captures the orthogonality of shocks $g_n$ and $\bar{\varepsilon}_n$ when weighted by $s_n$. Since these two expressions are equivalent, equation (5) shows that such shock orthogonality is necessary and sufficient condition for the orthogonality of the shift-share instrument. As with $e_\ell$, the shock-level weights are also non-negative and sum to one, since $\sum_n s_n = \sum_\ell e_\ell (\sum_n s_{\ell n}) = 1$. The shock-level unobservables $\bar{\varepsilon}_n$ represent exposure-weighted averages of the residuals $\varepsilon_{\ell}$.

The labor supply example is useful for unpacking this first equivalence result. When $s_{\ell n}$ are lagged employment shares and $e_\ell$ are similarly lagged regional employment weights, the $s_n$ weights are proportional to the lagged industry employment.\footnote{Without regression weights (i.e. $e_\ell = \frac{1}{L}$), $s_n$ is instead the average employment share of industry $n$ across locations.} Moreover, with $\varepsilon_{\ell}$ capturing unmeasured supply shocks, $\bar{\varepsilon}_n$ is the average unobserved supply shock among regions $\ell$ that are the most specialized in industry $n$, in terms of their lagged employment $e_\ell s_{\ell n}$. Equation (5) then shows that for the shift-share instrument $z_{\ell}$ to identify the labor supply elasticity $\beta$, the industry demand shocks $g_n$ must be orthogonal with these industry-level unobservables when weighted by industry size. For example, the industries which see higher output subsidies should not face systematically different unobserved labor supply conditions (e.g., migration patterns) in their primary markets.

Shock orthogonality is a necessary condition for SSIV identification and is satisfied when, as in the preferred interpretation of Goldsmith-Pinkham et al. (2020), the exposure shares are exogenous, the data are iid, and the shocks are considered non-random.\footnote{Formally, in this framework $\mathbb{E} [e_{\ell} s_{\ell n} \varepsilon_{\ell}] = 0$ for each $(\ell, n)$, so $\mathbb{E} \left[ \sum_n s_{\ell n} g_n \bar{\varepsilon}_n \right] = \sum_n g_n \sum_\ell \mathbb{E} [e_{\ell} s_{\ell n} \varepsilon_{\ell}] = 0$.} In practice, however, this approach to SSIV identification may be untenable in many settings. In our labor supply example, the Goldsmith-Pinkham et al. (2020) approach to identification requires the (lagged) local employment share of each industry to be a valid instrument in the labor supply equation, i.e. uncorrelated with all unobserved labor supply shocks. This assumption is unlikely to hold: changes in foreign immigration, for example, are a type of a local labor supply shock which is likely related to the local industry composition (e.g., new immigrants may prefer to settle in areas with larger clusters of specific industries, such as high-tech, even conditionally on the prevailing wage). Formally, whenever the second-stage error term has a component with the shift-share structure, $\sum_n s_{\ell n} \nu_n$ for unobserved shocks $\nu_n$, then the exposure shares will be mechanically endogenous (see Appendix A.3 for a proof).\footnote{In the immigration example, $\nu_n$ is positive in high-tech industries and negative in industries that do not attract immigrants. Note that the same argument applies to migration flows within the U.S., which can similarly make local labor supply shocks related to local industry composition. Lagging local employment shares does not alleviate these threats to identification in general.}
When shares are endogenous, equation (5) suggests that identification may instead follow from the exogeneity of shocks. We formalize this approach in Section 3.1, by specifying a quasi-experimental design in which the \( g_n \) are as-good-as-randomly assigned with respect to the other terms in the expression. We show how this simple exogeneity can be relaxed with controls in Section 3.2.

2.3 Estimator Equivalence

We next build intuition for our approach to SSIV consistency by showing that the estimate \( \hat{\beta} \) is equivalently obtained as the coefficient from a non-standard shock-level IV procedure, in which \( g_n \) directly serves as the instrument. This equivalence result suggests that the large-sample properties of \( \hat{\beta} \) can be derived from a law of large numbers for the equivalent shock-level regression. An attractive feature of this approach is that it does not rely on an assumption of iid observations, which can be untenable in the presence of observed and unobserved \( n \)-level shocks. We instead place assumptions on the assignment of the equivalent IV regression’s instrument \( g_n \), similar to a more standard analysis of a randomized treatment in an experimental settings (Abadie et al. 2019).

Formally, we have the following equivalence result:

**Proposition 1** The SSIV estimator \( \hat{\beta} \) equals the second-stage coefficient from a \( s_n \)-weighted shock-level IV regression that uses the shocks \( g_n \) as the instrument in estimating

\[
\bar{y}_n = \alpha + \beta \bar{x}_n + \varepsilon_n. \tag{6}
\]

**Proof** By definition of \( z_\ell \),

\[
\hat{\beta} = \frac{\sum \epsilon_\ell \left( \sum n \ s_{\ell n} g_n \right) y_\ell}{\sum \epsilon_\ell \left( \sum n \ s_{\ell n} g_n \right) x_\ell} = \frac{\sum n \ g_n \left( \sum \epsilon_\ell s_{\ell n} y_\ell \right)}{\sum n \ g_n \left( \sum \epsilon_\ell s_{\ell n} x_\ell \right)} = \frac{\sum n \ s_n g_n \bar{y}_n}{\sum n \ s_n g_n \bar{x}_n}. \tag{7}
\]

Furthermore \( \sum_n s_n \bar{y}_n = \sum \epsilon_\ell \left( \sum_n s_{\ell n} \bar{y}_\ell \right) = \sum \epsilon_\ell \bar{y}_\ell = 0 \), since \( y_\ell \) is an \( \epsilon_\ell \)-weighted regression residual and \( \sum_n s_{\ell n} = 1 \). This and an analogous equality for \( \bar{x}_n \) imply that (7) is a ratio of \( s_n \)-weighted covariances, of \( \bar{y}_n \) and \( \bar{x}_n \) with \( g_n \). Hence it is obtained from the specified IV regression.

As with equation (5), Proposition 1 exploits the structure of the instrument to exchange orders of summation in the expression for the SSIV estimator (4). This exchange shows that SSIV estimates can also be thought to arise from variation across shocks, rather than across observations. The equivalent IV regression uses the shocks \( g_n \) directly as the instrument and shock-level aggregates of the original (residualized) outcome and treatment, \( \bar{y}_n \) and \( \bar{x}_n \). Specifically, \( \bar{y}_n \) reflects the average residualized outcome of the observations most exposed to the \( n \)th shock, while \( \bar{x}_n \) is the same weighted average of residualized treatment. The regression is weighted by \( s_n \), representing each shock’s average exposure across the observations.\(^{12}\)

\(^{12}\)In the special case of reduced-form shift-share regressions, Proposition 1 shows that the equivalent shock-level
The fact that shift-share estimates can be equivalently obtained by a shock-level IV procedure suggests a new approach to establishing their consistency. Generally, IV regressions of the form of (7) will be consistent when the instrument (here, \( g_n \)) is as-good-as randomly assigned, there is a large number of observations (here, \( N \)), the importance weights are sufficiently dispersed (here, that the \( s_n \) are not too skewed), and there is an asymptotic first stage. Consistency is then guaranteed regardless of the correlation structure of the residuals \( \overline{\epsilon}_{\ell n} \), and thus in the primitive residuals \( \epsilon_{\ell} \) and exposure shares \( s_{\ell n} \). We formalize this approach below.

Before proceeding, it is worth emphasizing that Proposition 1 is a motivating algebraic result and that the equivalent shock-level IV regression does not have independent economic content. Its outcome and treatment are unconventional shock-level objects which can, for example, be computed for outcomes and treatment that are not typically observed at the level of the shocks, such as when \( n \) indexes industries and \( y_{\ell} \) measures regional marriage rates (Autor et al. 2019). Furthermore, even if the outcome and treatment have natural measures at the shock level, \( \bar{y}_n \) and \( \bar{x}_n \) will generally not coincide with them. For example, in the labor supply setting \( \bar{y}_n \) is not industry \( n \)'s wage growth; rather, it measures the average wage growth in regions where industry \( n \) employs the most workers. Thus, while \( \hat{\beta} \) can be computed at the industry level it does not lose its interpretation as a regional SSIV estimate. In the labor supply setting, the shift-share regression estimates the elasticity of regional, rather than industry, labor supply.\(^{13}\)

3 A Quasi-Experimental SSIV Framework

We now show how SSIV identification and consistency can be satisfied by a quasi-experiment in which shocks are as-good-as-randomly assigned, mutually uncorrelated, large in number, and sufficiently dispersed in terms of their average exposure. Instrument relevance generally holds in such settings when the exposure of individual observations tends to be concentrated in a small number of shocks, and when those shocks affect treatment. We then show how this framework is naturally generalized to settings in which shocks are only conditionally quasi-randomly assigned or exhibit some forms of mutual dependence, such as clustering.

3.1 Quasi-Randomly Assigned and Mutually Uncorrelated Shocks

Our approach to SSIV consistency is based on a thought experiment in which the shocks \( g_n \) are as-good-as-randomly assigned conditional on the shock-level unobservables \( \bar{\epsilon}_n \) and exposure weights \( s_n \).

\(^{13}\)In Appendix A.2 we develop a stylized model to illustrate how the SSIV coefficient can differ from a “native” shock-level IV coefficient in the presence of local spillovers or treatment effect heterogeneity, though both parameters may be of interest. Intuitively, in the labor supply case one may estimate a low regional elasticity but a high elasticity of industry labor supply if, for example, migration is constrained but workers are mobile across industries within a region.
As motivated above, placing assumptions on this assignment process (rather than on the sampling properties of observations) has two key advantages. First, we do not rely on conventional assumptions of independent or clustered data which are generally inconsistent with the shift-share data structure when the shocks are considered random variables. Second, in conditioning on $\bar{\varepsilon} = \{\bar{\varepsilon}_n\}_n$ and $\bar{s} = \{s_n\}_n$ we place no restrictions on the dependence between the $s_{\ell n}$ and $\varepsilon_{\ell}$, allowing shock exposure to be endogenous. We first show that such endogeneity need not pose problems for SSIV identification:

**Proposition 2** The SSIV moment condition (3) is satisfied by the following condition:

**Assumption 1** (Quasi-random shock assignment): $E[g_n | \bar{\varepsilon}, \bar{s}] = \mu$, for all $n$.

**Proof** By equation (5) and the law of iterated expectations, $E[\sum_{\ell} c_{\ell} z_{\ell} | \varepsilon_{\ell}] = E[\sum_{\ell} s_n g_n \varepsilon_{\ell}] = \mu \cdot E[\sum_{\ell} s_n \bar{\varepsilon}_n]$ under Assumption 1. Furthermore, since $\sum_{\ell} s_{\ell n} = 1$ and $E[\sum_{\ell} c_{\ell} \varepsilon_{\ell}] = 0$ by construction, $E[\sum_{\ell} s_n \bar{\varepsilon}_n] = E[(\sum_{\ell} s_{\ell n}) (\sum_{\ell} c_{\ell} \varepsilon_{\ell})] = 0$.

Proposition 2 shows that the shift-share instrument is valid, in that the IV moment condition (3) holds, when the underlying shocks are as-good-as-randomly assigned: each $g_n$ has the same mean $\mu$, regardless of the realizations of the relevant unobservables $\bar{\varepsilon}$ (and average exposures $\bar{s}$). In the labor supply example this assumption would mean that industry subsidies should not have been chosen strategically, based on labor supply trends, or in a way that systematically favors certain industries (e.g. only those in the manufacturing sector).\(^{14}\)

It follows from Proposition 2 that $\beta$ is identified by Assumption 1 provided the instrument is relevant.\(^{15}\) In practice, the existence of a non-zero first stage can be inferred from the data; we discuss appropriate inferential techniques in Section 5. To illustrate how such instrument relevance might hold with quasi-experimental shocks, we consider a simple first-stage model. Consider a setting without controls ($w_{\ell} = 1$) and where treatment is a share-weighted average of shock-specific components: $x_{\ell} = \sum_n s_{\ell n} x_{\ell n}$, where $x_{\ell n} = \pi_{\ell n} g_n + \eta_{\ell n}$ with $\pi_{\ell n} \geq \bar{\pi}$ almost surely for some fixed $\bar{\pi} > 0$. In line with Assumption 1, suppose that the shocks are independent mean-zero, given the full set of exposure shares $s_{\ell n}$ and regression weights $c_{\ell}$ and the full set of $\pi_{\ell n}$ and $\eta_{\ell n}$, with variances that are bounded below by some fixed $\bar{\sigma}^2 > 0$. Then the instrument first stage is positive:

$$E \left[ \sum_{\ell} c_{\ell} z_{\ell} x_{\ell} \right] = E \left[ \sum_{\ell} c_{\ell} \left( \sum_n s_{\ell n} g_n \right) \left( \sum_n s_{\ell n} (\pi_{\ell n} g_n + \eta_{\ell n}) \right) \right] \geq \bar{\pi} \bar{\sigma}^2 E \left[ \sum_{\ell} c_{\ell} \sum_n s_{\ell n}^2 \right] > 0. \quad (8)$$

\(^{14}\)While we use output subsidies as $g_n$ in our idealized example, other shocks to industry labor demand can be similarly considered. Differences in import tariffs across industries, for example, have been used as identifying variation both directly (e.g. Fajgelbaum et al. (2020)) and in shift-share settings (e.g. by Kovak (2013) who argues that tariff changes are uncorrelated with counterfactual industry performance).

\(^{15}\)Appendix A.1 shows how SSIV identifies a convex average of heterogeneous treatment effects (varying potentially across both $\ell$ and $n$) under a stronger notion of as-good-as-random shock assignment and a first-stage monotonicity condition. This can be seen as generalizing both the IV identification result of Angrist et al. (2000) to shift-share instruments, as well as the reduced-form shift-share identification result in Adão et al. (2019).
Given identification, SSIV consistency follows from an appropriate law of large numbers. Motivated by the estimator equivalence in Section 2.3, we consider settings in which the effective sample size of the shock-level IV regression (6) is large and the observations of the effective instrument (shocks) are mutually uncorrelated:

**Assumption 2 (Many uncorrelated shocks):** \( E \left[ \sum_{n} s_{n}^{2} \right] \to 0 \) and \( \text{Cov} \left[ g_{n}, g_{n'} \mid \bar{\varepsilon}, s \right] = 0 \) for all \((n, n')\) with \( n' \neq n \).

The first part of Assumption 2 states that the expected Herfindahl index of average shock exposure, \( E \left[ \sum_{n} s_{n}^{2} \right] \), converges to zero as \( L \to \infty \). This condition implies that the number of observed shocks grows with the sample (since \( \sum_{n} s_{n}^{2} \geq 1/N \)), and can be interpreted as requiring a large effective sample for the equivalent shock-level IV regression. An equivalent condition is that the largest importance weight in this regression, \( s_{n} \), becomes vanishingly small.\(^{16}\) The second part of Assumption 2 states that the shocks are mutually uncorrelated given the unobservables and \( s_{n} \). Both of these conditions, while novel for SSIV, would be standard assumptions to establish the consistency of a conventional shock-level IV estimator with \( g_{n} \) as the instrument and \( s_{n} \) weights.

Assumptions 1 and 2 are the baseline assumptions of our quasi-experimental framework. Given a standard relevance condition and additional regularity conditions listed in Appendix B.1, they are sufficient to establish SSIV consistency:\(^{17}\)

**Proposition 3** Suppose Assumptions 1 and 2 hold, \( \sum_{\ell} e_{\ell} z_{\ell} x_{\ell}^{\perp} \xrightarrow{p} \pi \) with \( \pi \neq 0 \), and Assumptions B1-B2 hold. Then \( \hat{\beta} \xrightarrow{p} \beta \).

**Proof** See Appendix B.1.

As before, the relevance condition merits further discussion. In our simple first-stage model, \( \sum_{\ell} e_{\ell} z_{\ell} x_{\ell}^{\perp} \) converges to \( E \left[ \sum_{\ell} e_{\ell} z_{\ell} x_{\ell} \right] \) under appropriate regularity conditions, which is bounded above zero by a term proportional to \( E \left[ \sum_{\ell} e_{\ell} \sum_{n} s_{n}^{2} \right] \). Thus, in this case, SSIV relevance holds when the \( e_{\ell} \)-weighted average of local exposure Herfindahl indices \( \sum_{n} s_{n}^{2} \) across observations does not vanish in expectation.

In our running labor supply example, where \( x_{\ell n} \) is industry-by-region employment growth, SSIV relevance generally arises from individual regions \( \ell \) tending to specialize in a small number of industries \( n \), provided subsidies have a non-vanishing effect on local industry employment.\(^{18}\) Compare this to

\(^{16}\)Goldsmith-Pinkham et al. (2020) propose a different measure of the importance of a given \( n \), termed “Rotemberg weights.” In Appendix A.4 we show the formal connection between \( s_{n} \) and these weights, and that the latter do not carry the sensitivity-to-misspecification interpretation as they do in the exogenous shares view of Goldsmith-Pinkham et al. (2020). Instead, the Rotemberg weight of shock \( n \) measures the leverage of \( n \) in the equivalent shock-level IV regression from Proposition 1. Shocks may have large leverage either because of large \( s_{n} \), as would be captured by the Herfindahl index, or because the shocks have a heavy-tailed distribution which is allowed by Assumption 2.

\(^{17}\)One high-level condition used in Proposition 3 (Assumption B2) is that the control coefficient \( \gamma \) is consistently estimated by its sample analog, \( \hat{\gamma} = \left( \sum_{\ell} e_{\ell} w_{\ell} w_{\ell}^{\prime} \right)^{-1} \sum_{\ell} e_{\ell} w_{\ell} \bar{\varepsilon}_{\ell} \) (see footnote 5). We discuss sufficient conditions for this assumption in Appendix A.5.

\(^{18}\)Note that this precludes consideration of an asymptotic sequence where \( L \) remains finite as \( N \) grows. With \( L \) (and also \( e_{1}, \ldots, e_{L} \)) fixed, Assumption 2 implies \( \sum_{n} e_{n}^{2} E \left[ \sum_{n} s_{n}^{2} \right] \to 0 \) and thus \( \text{Var} \left[ z_{\ell} \right] = \text{Var} \left[ \sum_{n} s_{n} g_{n} \right] \to 0 \) for each \( \ell \) if \( \text{Var} \left[ g_{n} \right] \) is bounded. If the instrument has asymptotically no variation it cannot have a first stage, unless the \( x_{\ell n} \) grow without bound. This result also highlights the role of picking the shares which reflect the impact of
the Herfindahl condition in Assumption 2, which instead states that the average shares of industries across locations become small. Both conditions may simultaneously hold when most regions specialize in a small number of industries, differentially across a large number of industries.

3.2 Conditional Shock Assignment and Weak Shock Dependence

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

We first relax Assumptions 1 and 2 to only hold conditionally on a vector of shock-level observables \( q_n \) (that includes a constant). For example, it may be more plausible that shocks are as-good-as-randomly assigned within a set of observed clusters \( c(n) \in \{1, \ldots, C\} \) with non-random cluster-average shocks, in which case \( q_n \) collects \( C - 1 \) cluster dummies and a constant. In the labor supply example, this may allow output subsidies to vary systematically across the agriculture, manufacturing, and service sectors, but be as-good-as-random within each of those sectors. In general, with \( q = \{q_n\}_n \), we consider the following weakened version of Assumption 1:

**Assumption 3 (Conditional quasi-random shock assignment):** \( \mathbb{E}[g_n | \bar{\epsilon}, q, s] = q_n'\mu, \) for all \( n \).

Similarly, we consider a weakened version of Assumption 2 which imposes mutual uncorrelatedness on the residual \( \tilde{g}_n = g_n - q_n'\mu \):

**Assumption 4 (Many uncorrelated shock residuals):** \( \mathbb{E}[\sum_n s_n^2] \to 0 \) and \( \text{Cov}[\tilde{g}_n, \tilde{g}_{n'} | \bar{\epsilon}, q, s] = 0 \) for all \( (n, n') \) with \( n' \neq n \).

In the shock cluster example, Assumption 4 applies with a small number of clusters, each with its own random effect, as in that case a law of large numbers may apply to the within-cluster residuals \( \tilde{g}_n \) but not the original shocks \( g_n \).

By a simple extension of the proof to Proposition 3, the SSIV estimator is consistent when these conditions replace Assumptions 1 and 2 and the residual shift-share instrument \( \tilde{z}_\ell = \sum_n s_{\ell n} \tilde{g}_n \) replaces \( z_\ell \). While this instrument is infeasible, since \( \mu \) is unknown, the following result shows that SSIV regressions that control for the exposure-weighted vector of shock-level controls, \( \tilde{w}_\ell = \sum_n s_{\ell n} q_n \), provide a feasible implementation:

\[ g_n \text{ on } x_\ell. \] Here when the shares are misspecified, i.e. when the treatment is constructed from different shares \( \tilde{s}_{\ell n} \) as \( x_\ell = \sum_n \tilde{s}_{\ell n} x_{\ell n} \), the first-stage is bounded by a term proportional to \( \mathbb{E}[\sum_\ell \epsilon_\ell \sum_n s_{\ell n}^2 \tilde{s}_{\ell n}] \), which can be arbitrarily small even if \( \mathbb{E}[\sum_\ell \epsilon_\ell \sum_n s_{\ell n}^2 \tilde{s}_{\ell n}] \to 0 \).

As an extreme example, suppose each region specializes on one industry only: \( s_{\ell n} = 1 \) \( n = n(\ell) \) for some \( n(\ell) \). Then the average local concentration index \( \sum_\ell \epsilon_\ell \sum_n s_{\ell n}^2 \tilde{s}_{\ell n} \) equals one, while Assumption 2 holds when national industry composition is asymptotically dispersed: for example, when \( \epsilon_\ell = 1/L \) and \( n(\ell) \) is drawn iid across regions and uniformly over \( 1, \ldots, N \).
Proposition 4 Suppose Assumptions 3 and 4 hold, \( \sum_{e} e_{\ell} x_{\ell}^{\perp} \xrightarrow{p} \pi \) with \( \pi \neq 0 \), and Assumptions B1-B2 hold. Then \( \hat{\beta} \xrightarrow{p} \beta \) provided \( \tilde{w}_{\ell} \) is included in \( w_{\ell} \).

**Proof** See Appendix B.1.

This result highlights a special role of controls with a shift-share structure (i.e. \( \sum_{n} s_{\ell n} q_{n} \)): besides removing confounding variation from the residual (as any \( w_{\ell} \) would do), they can also be viewed as removing such variation directly from the shocks (i.e. implicitly using \( g_{n} \) in place of \( g_{n} \)). In particular, Proposition 4 shows that controlling for each observation’s individual exposure to each cluster \( \sum_{n} s_{\ell n} 1[c(n) = c] \) isolates the within-cluster variation in shocks. This allows for a thought experiment in which shocks are drawn quasi-randomly only within observed clusters, but not across clusters with potentially different shock means. Note that Proposition 3 is obtained as a special case of Proposition 4, which sets \( q_{n} = 1 \).

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

**Assumption 5 (Many uncorrelated shock clusters):** There exists a partition of industries into clusters \( c(n) \) such that \( \mathbb{E} \left[ \sum_{n} s_{n}^{2} \right] \rightarrow 0 \) for \( s_{c} = \sum_{n: c(n) = c} s_{n} \) and \( \text{Cov} \left[ \tilde{g}_{n}, \tilde{g}_{n'} | \bar{\varepsilon}, q, s \right] = 0 \) for all \( (n, n') \) with \( c(n) \neq c(n') \);

**Assumption 6 (Many weakly correlated shocks):** For some sequence of numbers \( B_{L} \geq 0 \) and a fixed function \( f(\cdot) \) satisfying \( \sum_{r=1}^{\infty} f(r) < \infty \), \( B_{L} \mathbb{E} \left[ \sum_{n} s_{n}^{2} \right] \rightarrow 0 \) and \( |\text{Cov} \left[ \tilde{g}_{n}, \tilde{g}_{n'} | \bar{\varepsilon}, q, s \right] | \leq B_{L} \cdot f \left( |n' - n| \right) \) for all \( (n, n') \).

Assumption 5 relaxes Assumption 4 by allowing shock residuals to be grouped within mutually mean-independent clusters \( c(n) \), while placing no restriction on their within-cluster correlation. At the same time, the Herfindahl index assumption of Assumption 4 is strengthened to hold for industry clusters, with \( s_{c} \) denoting the average exposure of cluster \( c \). Assumption 6 takes a different approach, allowing all nearby shock residuals to be mutually correlated provided their covariance is bounded by a function \( B_{L} \cdot f \left( |n' - n| \right) \). This accommodates, for example, the case of first-order autoregressive time series with the covariance bound declining at a geometric rate, i.e. \( f(r) = \delta^{r} \) for \( \delta \in [0, 1) \) and constant \( B_{L} \). With \( B_{L} \) growing, stronger dependence of nearby shocks is also allowed (see Appendix B.1).

## 4 Extensions

We now present several other extensions of our quasi-experimental framework. Section 4.1 considers shift-share identification with panel data. Section 4.2 explains the need for additional controls when
the sum of exposure shares vary across observations. Section 4.3 discusses how our framework applies when the shocks are estimated within the sample, as in the canonical Bartik (1991) study. Finally, Section 4.4 extends the framework to allow for multiple treatments and shift-share instruments.

### 4.1 Panel Data

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

\[
z_{lt} = \sum_{n} s_{nt} g_{nt}
\]

is used, and the controls \( w_{lt} \) may include unit- or period-specific fixed effects.

It is straightforward to map this panel case to the previous cross-sectional setting by a simple relabeling. Let \( \hat{\ell} \in \{(\ell, t) : \ell = 1, \ldots, L; t = 1, \ldots, T\} \) and \( \hat{n} \in \{(n, t) : n = 1, \ldots, N; t = 1, \ldots, T\} \), with the time-varying outcomes now indexed as \( y_{\ell t} \) and similarly for \( x_{\ell t}, w_{\ell t}, c_{\ell t}, \) and \( g_{nt} \). Further let \( \tilde{s}_{\hat{n} \ell t} = s_{(\ell,n),(n,p)} t_{\ell} = p \) denote the exposure of observation \( \ell \) in period \( t \) to shock \( n \) in period \( p \), which is by definition zero for \( t \neq p \). The time-varying instrument (9) can then be rewritten

\[
z_{\ell t} = \sum_{n} \tilde{s}_{\hat{n} \ell t} g_{\ell t n},
\]

as in the cross-sectional case. Generalizations of the preceding quasi-experimental framework immediately follow.

Unpacking these results in the panel setting shows how standard intuitions for consistency in long and short shock-level panels translate into shift-share designs. First, consistency of the panel SSIV estimator in long panels \( (T \to \infty) \) follows even with relatively few shocks \( N \) and cross-sectional observations \( L \), provided shocks are mutually uncorrelated across units and either uncorrelated or weakly serially correlated over time (Assumptions 4 and 6, respectively). Observing many time periods yields both a growing sample size, \( LT \to \infty \), and a growing number of shocks, \( NT \to \infty \), which allows the Herfindahl condition of Proposition 4 (that \( \mathbb{E} \left[ \sum_{n} \tilde{s}_{\ell t}^2 \right] = \mathbb{E} \left[ \sum_{n} \sum_{t} \tilde{s}_{nt}^2 \right] \to 0 \) to be satisfied. This case accommodates, in particular, shift-share designs that leverage purely time-series shocks \( (N = 1, T \to \infty) \), as in Nunn and Qian (2014). Second, with many shocks observed in the cross-section \( (N \to \infty) \) and the corresponding per-period Herfindahl condition, consistency follows as \( LT \to \infty \) even with arbitrary serial correlation in the shocks (using Assumption 5, with cross-sectional units as clusters). This case accommodates short panels and repeated cross-sections \( (T \text{ fixed}, L, N \to \infty) \).

---

20 Exposure shares are typically lagged and sometimes fixed in a pre-period. Our subscript \( t \) notation indicates that these shares are used to construct the instrument for period \( t \), not that they are measured in that period.

21 Nunn and Qian (2014) estimate the impact of U.S. food aid on civil conflict, using variation in U.S. wheat production (a single “shock” per period) over a long time horizon \( (T = 36 \text{ years}) \), interacted with a country’s tendency to receive US food aid (the “exposure shares” for \( L = 125 \) countries). This case may also well-approximate settings like those of Berman et al. (2017) and Imbert et al. (2019) where \( NT \) is large despite moderate \( N \) and \( T \). Berman et al. (2017), for example, leverage price changes for \( N = 14 \) minerals over \( T = 14 \) years in a very large cross-section of spatial cells.

22 Proposition 4 also suggests an alternative way to handle serial correlation when the time series properties of shocks are known. For example given a first-order autoregressive process \( g_{nt} = \rho_0 + \rho_1 g_{nt-1} + \eta_{nt} \), controlling for the
The role of fixed effect (FE) controls in panel SSIV regressions is also worth discussing. In conventional panels, with exogenous shocks varying at the same level as the observations, unit fixed effects play a dual role: they purge time-invariant unobservables from the residual and isolate the time-varying component of the shocks. The first role directly extends to the shift-share case: Assumption 1 would thus hold if shocks are as-good-as-randomly assigned with respect to aggregated time-demeaned unobservables. However, the second role only extends when exposure shares are fixed across periods, i.e. when $s_{nt} \equiv s_{nt}\,23$ This follows from Proposition 4: the exposure-weighted averages of the shock FE which isolate the time-varying component of shocks are only subsumed by the unit FE when the exposure shares are time-invariant. A simpler argument applies to period FE, which isolate within-period shock variation when the exposure shares add up to one. In that case one can represent period FE in the original data as exposure-weighted averages of shock-level period FE in the relabeled data, i.e. $1[t = \bar{t}] = \sum_{(n,p)} \tilde{s}_{\ell(n,p)} [p = \bar{t}].\,24$

To make these insights concrete, consider our labor supply example in a panel setting with subsidies allocated to industries in each period. Imagine that certain industries get permanent subsidy shocks that are not as-good-as-randomly assigned across industries. In that case, one may prefer to only use the changes in industry subsidies over time as identifying variation. With fixed exposure shares, one way to proceed is to include region FE in the SSIV specification, which implicitly control for industry FE at the shock level while also removing any potential time-invariant regional characteristics that could yield omitted variables bias. Including period FE further removes any non-random variation in the average level of subsidies in each period.

We lastly note that while fixing exposure shares in a pre-period is useful for isolating time-varying shocks, lagging shares by many periods is likely to make the shift-share instrument less predictive of treatment, thereby reducing the first stage in panel SSIVs. Appendix A.8 formalizes this intuition with a panel extension to our simple first-stage model. In such cases it may be preferable to employ other ways of isolating time-varying shocks, which allow for updated shares. In particular, one may use the first-differenced specification

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

instrumenting $\Delta x_{t\ell}$ with $z_{t\ell,FD} = \sum_n s_{t\ell_n,t-1} \Delta g_{nt}$, where $\Delta$ is the first-differencing operator for both observations and shocks. This strategy has been employed, for example, by Autor et al. (2013) as we

---

23Shift-share IV settings with panel data and time-invariant shares include, for example, Berman et al. (2015), Berman et al. (2017), Hummels et al. (2014), Imbert et al. (2019), and Nunn and Qian (2014).

24Proposition 4 only applies under Assumption B2, which requires the control coefficient vector $\gamma$ to be consistently estimated as $LT \to \infty$, even if $w_{t\ell}$ contains an increasing number of FE (see Appendix B.1). In Appendix A.6 we show how alternative notions of shock exogeneity, which are imposed on unit-demeaned unobservables, address this incidental parameters problem. In Appendix A.7 we further discuss how these strict exogeneity assumptions can be relaxed in long panels, allowing for sequential shock generating processes.
4.2 SSIVs with Incomplete Shares

While we have so far assumed the sum of exposure shares is constant, in practice this $S_\ell = \sum_n s_{\ell n}$ may vary across observations $\ell$. For example, in the labor supply setting, the quasi-experiment in government output subsidies may only cover manufacturing industries, while the lagged manufacturing employment shares of $s_{\ell n}$ may be measured relative to total employment in region $\ell$. In this case $S_\ell$ equals the lagged total share of manufacturing employment in region $\ell$. The Autor et al. (2013) setting is an example of this scenario, as we discuss below.\(^{26}\)

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

$$z_\ell = s_{\ell 0}g_0 + \sum_{n>0} s_{\ell n}g_n,$$

where $g_0 = 0$ and $s_{\ell 0} = 1 - S_\ell$, such that $\sum_{n=0}^N s_{\ell n} = 1$ for all $\ell$. The previous quasi-experimental framework then applies to this expanded set of shocks $g_0, \ldots, g_N$. Since $g_0 = 0$, Proposition 3 requires in this case that $\mathbb{E} [g_n | s, \xi] = 0$ for $n > 0$ as well; that is, that the expected shock to each manufacturing industry is the same as the “missing” non-manufacturing shock of zero. Otherwise, even if the manufacturing shocks are random, regions with higher manufacturing shares $S_\ell$ will tend to have systematically different values of the instrument $z_\ell$, leading to bias when these regions also have different unobservables.\(^{27}\)

Cast in this way, the incomplete shares issue has a natural solution via Assumption 3. Namely, one can allow the missing and non-missing shocks to have different means by conditioning on the indicator $1[n > 0]$ in the $g_n$ vector. By Proposition 4, the SSIV estimator allows for such conditional quasi-

---

\(^{25}\)There is another argument for fixing the shares in a pre-period that applies when the current shares are affected by lagged shocks in a way that is correlated with unobservables $\xi_{\ell t}$. In the labor supply example suppose local labor markets vary in flexibility, with stronger reallocation of employment to industries with bigger subsidies in flexible markets. If subsidies are random but persistent, more subsidized industries will be increasingly concentrated in regions with flexible labor markets and Assumption 1 will be violated if such flexibility is correlated with $\xi_{\ell t}$. This concern is distinct from that in Jaeger et al. (2018) who focus on the endogeneity of shares to the lagged residuals, rather than shocks, in a setting closer to Goldsmith-Pinkham et al. (2020). Jaeger et al. also point out another issue relevant to panel SSIV, that the outcome may respond to both current and lagged shocks; we return to this issue in Section 4.4.

\(^{26}\)We note that this scenario applies to quasi-experiments in which shocks are impossible (ex ante) for some industries. In contrast, if all industries were equally likely to receive subsidies but only some did ex post, the set of $n$ should include all industries, with $S_\ell = 1$, and zero subsidies are captured by $g_n = 0$ for some $n$. In such a case, however, it is highly unlikely that all manufacturing industries receive the subsidies by chance when no non-manufacturing industries do.

\(^{27}\)Formally, if Assumptions 1 and 2 hold for all $n > 0$ we have from the proof to Proposition 3 that

$$\sum_{n=0}^N s_n g_n \xi_n = \mathbb{E} \left[ \sum_{n=0}^N s_n (g_n - \mu) \xi_n \right] + \mu \mathbb{E} [s_n \xi_n] + \sigma_p (1).$$

If $\mu \neq 0$ and the missing industry share is large ($s_0 \neq 0$) this can only converge to zero when $\mathbb{E} [s_0 \xi_n] = \mathbb{E} \left[ \sum \epsilon_{\ell s_0} \xi_0 \right]$ does, i.e. when $S_\ell$ is exogenous.
random assignment when the control vector $w_\ell$ contains the exposure-weighted average of $1[n > 0]$, which here is $\sum_{n=0}^{N} s_{\ell n} 1[n > 0] = S_\ell$. Thus, in the labor supply example, quasi-experimental variation in manufacturing shocks is isolated in regressions with incomplete shares provided one controls for a region’s lagged manufacturing share $S_\ell$.

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

### 4.3 Shift-Share Designs with Estimated Shocks

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

We leave a more general treatment of this issue to Appendix A.9 and for concreteness focus on the Bartik (1991) example. The industry growth rates $g_n$ can be written as weighted averages of the growth of each industry in each region: $g_n = \sum_\ell \omega_{\ell n} g_{\ell n}$, where the weights $\omega_{\ell n}$ are the lagged shares of industry employment located in region $\ell$, with $\sum_\ell \omega_{\ell n} = 1$ for each $n$. In a standard model of regional labor markets, $g_{\ell n}$ includes (to first-order approximation) an industry labor demand shock $g_n^*$ and a term that is proportional to the regional supply shock $\varepsilon_\ell$.\(^{28}\) We suppose that the demand shocks are as-good-as-randomly assigned across industries, such that the infeasible SSIV estimator which uses $z_\ell = \sum_n s_{\ell n} g_n^*$ as an instrument satisfies our quasi-experimental framework. The asymptotic bias of the feasible SSIV estimator which uses $z_\ell = \sum_n s_{\ell n} g_n$ then depends on the large-sample covariance

---

\(^{28}\)Appendix A.10 presents such a model, showing that $g_{\ell n}$ also depends on the regional average of $g_n^*$ (via local general equilibrium effects) and on idiosyncratic region-specific demand shocks. Both of these are uncorrelated with the error term in the model and thus do not lead to violations of Assumption 1; we abstract away from this detail here.
between the labor supply shocks $\varepsilon_\ell$ and an aggregate of the supply shock “estimation error,”

$$\psi_\ell = z_\ell - z_\ell^* \propto \sum_n s_{tn} \sum_{\ell'} \omega_{\ell' n} \varepsilon_{\ell'}. \tag{12}$$

Two insights follow from considering the bias term $\sum_\ell e_\ell \psi_\ell \varepsilon_\ell$. First, part of the covariance between $\psi_\ell$ and $\varepsilon_\ell$ is mechanical, since $\varepsilon_\ell$ enters $\psi_\ell$. In fact, if supply shocks are spatially uncorrelated this is the only source of bias from using $z_\ell$ rather than $z_\ell^*$ as an instrument. This motivates the use of a leave-one-out (LOO) shock estimator, $g_{n,-\ell} = \sum_{\ell' \neq \ell} \omega_{\ell' n} g_{\ell' n} / \sum_{\ell' \neq \ell} \omega_{\ell' n}$, and the feasible instrument $z_\ell^{LOO} = \sum_n s_{tn} g_{n,-\ell}$ to remove this mechanical covariance.\(^{29}\) Conversely, if the regional supply shocks $\varepsilon_\ell$ are spatially correlated a LOO adjustment may not be sufficient to eliminate mechanical bias in the feasible SSIV instrument, though more restrictive split-sample methods (e.g. those estimating shocks from distant regions) may suffice.

Second, in settings where there are many regions contributing to each shock estimate even the mechanical part of $\sum_\ell e_\ell \psi_\ell \varepsilon_\ell$ may be ignorable, such that the conventional non-LOO shift-share instrument $z_\ell$ (which, unlike $z_\ell^{LOO}$, has a convenient shock-level representation per Proposition 1) is asymptotically valid when $z_\ell^{LOO}$ is.\(^{30}\) In Appendix A.9, we derive a heuristic for this case, under the assumption of spatially-independent supply shocks. In a special case when each region is specialized in a single industry and there are no importance weights, the key condition is $L/N \rightarrow \infty$, or that the average number of regions specializing in the typical industry is large. With incomplete specialization or weights, the corresponding condition requires the typical industry to be located in a much larger number of regions than the number of industries that a typical region specializes in.

To illustrate the preceding points in the data, Appendix A.9 replicates the setting of Bartik (1991) with and without a LOO estimator, using data from Goldsmith-Pinkham et al. (2020). We find that in practice the LOO correction does not matter for the SSIV estimate, consistent with the findings of Goldsmith-Pinkham et al. (2020) and Adão et al. (2019), and especially so when the regression is estimated without regional employment weights. Our framework provides a explanation for this: the heuristic statistic we derive is much larger without importance weights. These findings imply that if, in the canonical Bartik (1991) setting, one is willing to assume quasi-random assignment of the underlying industry demand shocks and that the regional supply shocks are spatially-uncorrelated, one can interpret the uncorrected SSIV estimator as leveraging demand variation in large samples, as some of the literature has done (e.g. Suárez Serrato and Zidar (2016)).

\(^{29}\)This problem of mechanical bias is similar to that of two-stage least squares with many instruments (Bound et al. 1995), and the solution is similar to the jackknife instrumental variable estimate approach of Angrist et al. (1999).

\(^{30}\)Adão et al. (2019) derive the corrected standard errors for LOO SSIV and find that they are in practice very close to the non-LOO ones, in which case the SSIV standard errors we derive in the next section are approximately valid even when the LOO correction is used.
4.4 Multiple Shocks and Treatments

In some shift-share designs one may have access to multiple sets of shocks satisfying Assumptions 3 and 4 or their extensions. For example while Autor et al. (2013) construct an instrument from average Chinese import growth across eight non-U.S. countries, in principle the industry shocks from each individual country may be each thought to be as-good-as-randomly assigned. In other settings multiple endogenous variables may be required for the SSIV exclusion restriction to plausibly hold. For example Jaeger et al. (2018) show that when local labor markets respond dynamically to immigrant inflows, it may be necessary to instrument both the current and lagged immigrant growth rate with current and lagged shift-share instruments. Another example is provided by Bombardini and Li (2019), who estimate the reduced-form effects of two shocks: the regional growth of all exports and the regional growth of exports in pollution-intensive sectors. Both are shift-share variables based on the same regional employment shares across industries but different shocks: an overall industry export shock and the overall shock interacted with industry pollution intensity.31

In Appendix A.11 we show how these settings fit into our quasi-experimental framework. The key insight is that SSIV regressions with multiple instruments—with and without multiple endogenous variables—again have an equivalent representation as particular shock-level IV estimators provided the exposure shares used to construct the instruments are the same. This immediately implies extensions of the foregoing results that establish consistency for just-identified SSIV regressions with multiple instruments, such as the dynamic adjustment case of Jaeger et al. (2018). In overidentified settings, the appendix derives new shock-level IV estimators that optimally combine the quasi-experimental variation and permit omnibus tests of the identifying assumptions, via the generalized method of moments theory of Hansen (1982) and inference results discussed in the next section. For example when shocks are homoskedastic, a two-stage least squares regression of $\bar{y}_n^\perp$ on $\bar{x}_n^\perp$, weighted by $s_n$ and instrumented by multiple shocks $g_{1n}, \ldots, g_{Jn}$ yields an efficient overidentified estimate of $\beta$.

5 Shock-Level Inference and Testing

A shock-level view also brings new insights to SSIV inference and testing. In this section we first show how a problem with conventional SSIV inference, first studied by Adão et al. (2019), has a convenient solution based on our estimator equivalence result. In particular, we show that conventional standard error calculations may be asymptotically valid when SSIV coefficients are estimated at the level of identifying variation (shocks). We then discuss how other novel shock-level procedures can be used to assess first-stage relevance and to implement valid falsification tests of shock exogeneity. Lastly, we

31 The instruments here are $\sum_n \bar{r}_n g_n$ and $\sum_n \bar{r}_n g_n q_n$ where $q_n$ denotes industry $n$’s pollution intensity. Our framework applies in this case even if $q_n$ is not randomly assigned: as long as the export shock $g_n$ satisfies an appropriate version of Assumption 1, $E[g_n | \bar{r}, q, s] = \mu$, the interacted shock satisfies $E[g_n q_n | \bar{r}, q, s] = \mu q_n$, i.e. Assumption 3. The natural extension of Proposition 4 to multiple instruments applies as long as $\sum_n \bar{r}_n q_n$, a measure of pollution intensity of the region, is controlled for, as Bombardini and Li (2019) do in some specifications.
summarize a variety of Monte-Carlo simulations illustrating the finite-sample properties of SSIV and relating them to conventional shock-level analyses.

5.1 Exposure-Robust Standard Errors

As with consistency, SSIV inference is complicated by the fact that the observed shocks $g_n$ and any unobserved shocks $\nu_n$ induce dependencies in the instrument $z_\ell$ and residual $\varepsilon_\ell$ across observations with similar exposure shares. This problem can be understood as an extension of the standard clustering concern (Moulton 1986), in which the instrument and residual are correlated across observations within predetermined clusters, with the additional complication that in SSIV every pair of observations with overlapping shares may have correlated $(z_\ell, \varepsilon_\ell)$. Adão et al. (2019) show the magnitude of the problem: in Monte-Carlo simulations, they find that tests based on conventional standard errors with nominal 5% significance can reject a true null in 55% of placebo shock realizations. They then develop a novel approach to conducting valid inference in presence of exposure-based clustering, building on our quasi-experimental framework for identification.

Our equivalence result in Section 2.3 motivates a convenient alternative to valid SSIV inference. By estimating SSIV coefficients with an equivalent shock-level IV regression, one directly obtains valid (“exposure-robust”) standard errors under the assumptions in Adão et al. (2019) and an additional condition on the controls that we discuss below. This solution generalizes a well-known approach to addressing conventional group clustering (Angrist and Pischke 2008, p. 313): by estimating a regression at the level of as-good-as-random variation (here, shocks) one avoids inferential biases due to clustering (here, by shock exposure). Like in Adão et al. (2019), the validity of this approach derives from the properties of the shocks and does not restrict the correlation structure of $\varepsilon_\ell$.

We first present our formal result and then discuss its advantages and limitations relative to the original solution of Adão et al. (2019).

**Proposition 5** Consider $s_n$-weighted IV estimation of the second stage equation

$$\tilde{y}_{n1} = \alpha + \beta \tilde{x}_{n1} + q_n^T \tau + \tilde{\varepsilon}_{n1}$$

where $\tilde{w}_\ell = \sum_n s_{\ell n} q_n$ is included in the control vector $w_\ell$ used to compute $\tilde{y}_{n1}$ and $\tilde{x}_{n1}$, and $\tilde{x}_{n1}$ is instrumented by $g_n$. The IV estimate of $\beta$ is numerically equivalent to the SSIV estimate $\hat{\beta}$. Furthermore, when Assumptions B3–B6 in Appendix B.2 hold and $\sum_\ell e_\ell x_\ell^T z_\ell \overset{P}{\to} \pi$ for $\pi \neq 0$, the conventional heteroskedasticity-robust standard error for $\hat{\beta}$ yields asymptotically-valid confidence intervals for $\beta$.

---

32It is easy to verify that the regression from our Proposition 5 is indeed the group-level regression of Angrist and Pischke (2008, p. 313) when the exposure shares are binary, i.e. for each $\ell$, $s_{\ell n} = 1 [n = n(\ell)]$ for some $n(\ell) \in \{1, \ldots, N\}$, and there are no controls. In this case $z_\ell = g_{n(\ell)}$, such that the shift-share instrument is constant among observations belonging to the same group $n(\ell)$. 

20
Equation (13) extends the previous shock-level estimating equation (6) by including a vector of controls $q_n$ which, as in Proposition 4, are included in the SSIV control vector $w_\ell$ as exposure-weighted averages. The first result in Proposition 5 is that the addition of these controls does not alter the coefficient equivalence established in Proposition 1. The second result states conditions, which strengthen those of Proposition 4, under which conventional shock-level standard errors from estimation of (13) yield valid asymptotic inference on $\beta$. Appendix B.2 further establishes two related results. First, absent any controls, i.e. with $w_\ell = q_n = 1$, the shock-level robust standard error for $\beta$ is numerically equivalent to the standard error formula that Adão et al. (2019) propose. Second, an alternative inference procedure that Adão et al. (2019) propose to improve finite-sample performance, by imposing the null hypothesis $\beta = \beta_0$ when estimating the residual $\varepsilon_\ell$, is also readily obtained from the shock-level equation (13).

A limitation of Proposition 5, relative to the original result in Adão et al. (2019), concerns the control vector $w_\ell$. While our previous results do not restrict the structure of this vector, Assumption B4 allows for only two types of controls: those with a shift-share structure (i.e. $\sum_n s_{\ell,n} q_n$, as in Proposition 4) that capture all sources of shock-level confounding, and other controls that are not asymptotically correlated with the instrument but may increase the asymptotic efficiency of the estimator. Adão et al. (2019) further allow controls that are necessary for identification and have an “approximate” shift-share structure: i.e. $\sum_n s_{\ell,n} \tilde{q}_n + u_\ell$, where $\tilde{q}_n$ is unobserved but the average variance of the noise $u_\ell$ is asymptotically small. Appendix B.2 shows that in the presence of such approximate shift-share controls, the standard errors obtained from estimating equation (13) will tend to be asymptotically conservative. Valid shift-share inference with general control vectors, of the kind we allow in our discussion of SSIV identification, remains an open problem.

At the same time, our shock-level approach offers three practical features. First, it can be performed with standard statistical software packages given an simple initial transformation of the data (i.e. to obtain $\tilde{y}_n$, $\tilde{x}_n$, and $s_n$), for which we have released a Stata package ssaggregate (see footnote 3). Second, it is readily extended to settings where shocks are clustered (a case also considered by Adão et al. (2019)) or autoregressive, as in Assumptions 5 and 6 respectively. Conventional cluster-robust or heteroskedastic-and-autocorrelation-consistent (HAC) standard error calculations applied to equation (13) are then valid. Third, the shock-level inference approach works in some cases where the original standard error calculation procedure of Adão et al. (2019) fails. Their procedure involves projecting $z^\perp_\ell$ on the vector of shares in order to account for the shock-level confounders underlying the approximate shift-share controls. When $N > L$, or when some exposure shares are collinear, such projection is infeasible. Our alternative approach avoids this issue, in settings without approximate

---

33 In contrast, Appendix B.2 also shows that when the two inference procedures are asymptotically equivalent, standard errors from Proposition 5 are likely to be smaller than those of Adão et al. (2019) in finite samples. We find that case to be relevant in the application in Section 6.2.4.

34 This is an empirically relevant issue: for instance, employment shares of some industries are collinear in the Autor et al. (2013) setting. To give one example, SIC code 2068 “Salted and roasted nuts and seeds” was part of code 2065
shift-share controls.

5.2 Falsification and Relevance Tests

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

As usual, the validity of \( z_{t} \) cannot be tested directly. However, indirect falsification tests may be conducted given an observed variable \( r_{t} \) thought to proxy for the unobserved residual \( \varepsilon_{t} \). Namely one may test whether \( r_{t} \) is uncorrelated with the shift-share instrument \( z_{t} \), while controlling for \( w_{t} \). Examples of such an \( r_{t} \) may include a baseline characteristic realized prior to the shocks or a lagged observation of the outcome \( y_{t} \) (resulting in what is often called a “pre-trend” test). To interpret the magnitude of the reduced form falsification regression coefficient, researchers may also scale it by the first stage regression of \( x_{t} \), yielding a placebo SSIV coefficient.

The insights of Adão et al. (2019) and the previous section apply directly to such tests. For exposure-robust inference on a placebo \( e_{t}-\)weighted SSIV regression of \( r_{t} \) on \( x_{t} \), instrumented by \( z_{t} \), one may use the conventional standard errors from an \( s_{n}-\)weighted regression of \( \tilde{r}_{n}^{\perp} \) on \( \tilde{x}_{n}^{\perp} \), instrumenting by \( g_{n} \) and controlling for any shock-level covariates \( q_{n} \). Similarly, for valid inference from a reduced form regression of \( r_{t} \) on \( z_{t} \), one may use the conventional standard errors from an IV regression of \( \tilde{r}_{n}^{\perp} \) on \( \tilde{z}_{n}^{\perp} \), with the same instrument, weights, and controls (see footnote 12). If a researcher observes a shock-level confounder \( r_{n} \), they can construct its observation-level average \( r_{t} = \sum_{n} s_{tn} r_{n} \) and perform a similar test. A simpler test regresses \( r_{n} \) on \( g_{n} \) directly (weighting by \( s_{n} \) and controlling for \( q_{n} \)).

Unlike exclusion, the SSIV relevance condition can be evaluated directly, via OLS regressions of \( x_{t} \) on \( z_{t} \) that control for \( w_{t} \). For exposure-robust inference on this OLS coefficient one may again use an equivalent shock-level IV regression: of \( \tilde{x}_{n}^{\perp} \) on \( \tilde{z}_{n}^{\perp} \), instrumenting by \( g_{n} \), weighting by \( s_{n} \), and controlling for \( q_{n} \). The first stage \( F \)-statistic, which is a common heuristic for instrument relevance, is then obtained as a squared \( t \)-statistic. We generalize this result to the case of multiple shift-share instruments in Appendix A.11 by detailing the appropriate construction of the “effective” first-stage
5.3 Monte-Carlo Simulations

Though the exposure-robust standard errors obtained from estimating equation (13) are asymptotically valid, it is useful to verify that they offer appropriate coverage with a finite number of observations and shocks. Of interest especially is whether the finite-sample performance of the equivalent regression (13) is comparable to that of more conventional shock-level IV regressions, in which the outcome and instrument are not aggregated from a common set of \( y_\ell \) and \( x_\ell \).

In Appendix A.12 we provide Monte-Carlo evidence suggesting that the finite-sample properties of SSIV and conventional shock-level IV regressions are similar, and that the variance of SSIV estimators is well-approximated by our exposure-robust standard errors. Specifically, we compare the finite-sample behavior of two estimators which both measure the effects of import competition with China on U.S. employment and use the same industry variation to construct the instruments. The first estimator is an SSIV and takes regional growth of manufacturing employment as the outcome (following Autor et al. (2013)) while the second is a conventional IV where the outcome is industry employment growth (similar to Acemoglu et al. (2016)). While finite sample properties of SSIV are not always perfect, e.g. there is moderate overrejection with heavy-tailed shocks, traditional industry-level IV suffers from the same problems. Our results further indicate that the conventional rule of thumb to detect weak instruments—the Montiel Olea and Pflueger (2013) first-state \( F \)-statistic—applies equally well to both the SSIV and the traditional IV estimators when computed for SSIV as described in Appendix A.11. Finally, our results also show that with the Herfindahl concentration index \( \sum_n s_n^2 \) as high as \( 1/20 \) (i.e. with an effective sample size of 20 in the equivalent shock-level regression) the asymptotic approximation for the SSIV estimator is still reasonable, with rejection rates in the vicinity of 7\% for tests with a nominal size of 5\% despite heavy tails of the shock distribution. Together, these results indicate that a researcher who is comfortable with the finite-sample performance of a shock-level analysis with some set of \( g_n \) should also be comfortable using such shocks in SSIV, provided there is sufficient variation in exposure shares to yield a strong SSIV first stage.\(^{37}\)

6 Shift-Share IV in Practice

We now summarize and illustrate the practical implications of our econometric framework. We first characterize the kinds of empirical settings to which the foregoing framework may be applied. We then apply the framework to the influential setting of Autor et al. (2013).

\(^{37}\)Naturally, these simulation results may be specific to the data-generating process we consider here, modeled after the “China shock” setting of Autor et al. (2013). In practice, we recommend that researchers perform similar simulations based on their data if they are concerned with the quality of asymptotic approximation—a suggestion that of course applies to conventional shock-level IV analyses as well.
6.1 A Taxonomy of SSIV Settings

Our framework can be applied to various empirical settings. To characterize these settings, we distinguish between three cases of SSIVs employed in the literature.

In the first case, the shift-share instrument is based on a set of shocks which can itself be thought of as an instrument. For example, the $g_n$ which enter $z_\ell$ might correspond to a set of observed growth rates that could be plausibly thought of as being randomly assigned to a large number of industries. Our framework shows how the shift-share instrument maps these shocks to the level of observed outcomes and treatments (e.g., geographic regions). A researcher who is comfortable with the identification conditions and finite-sample performance of an industry-level analysis based on $g_n$ should generally also be comfortable applying our framework, provided there is sufficient variation in the exposure shares and treatment to yield a strong first stage. Autor et al. (2013) and the corresponding industry-level analysis conducted by Acemoglu et al. (2016) give a prime example of this case, as we show below.

Empirical settings covered by this first case belong to various fields in economics, with outcomes and shocks defined at levels ($\ell$ and $n$, respectively) different than regions and industries. In international trade, Hummels et al. (2014) estimate the wage effects of offshoring across Danish importing firms $\ell$. They leverage a shift-share instrument for offshoring based on shocks to export supply by type of intermediate inputs and origin country; titanium hinges from Japan is an example of an $n$. While they translate these shocks to the firm level by using the lagged composition of firm imports as the shares, one could imagine an analysis of Danish imports at the input-by-country level directly that would leverage the same supply shocks. In finance, Xu (2019) examines the long-term effects of financial shocks on exports across countries $\ell$. Her shift-share instrument is based on a disruption that affected some but not all London-based banks $n$ in 1866, with country-specific exposure shares measuring pre-1866 market shares of those banks in each country. In line with considering bank shocks as-good-as-randomly assigned, she reports that affected and unaffected banks were balanced on various observable characteristics. In the immigration literature, Peri et al. (2016) estimate the effect of immigrant STEM workers on the labor market outcomes of natives across U.S. cities $\ell$. They exploit variation in the supply of STEM workers across migration origin countries $n$ and over time that arises from plausibly exogenous shifts to national H1-B policy. Similarly, in a literature on innovation, Stuen et al. (2012) leverage education policy shocks in foreign countries as a supply shock to U.S. doctoral programs.

In the second case, a researcher does not directly observe a large set of quasi-experimental shocks, but can still conceive of an underlying set of $g_n$ which if observed would be a useful instrument. Constructing the instrument then requires an initial step where these shocks are estimated in-sample, potentially introducing mechanical biases. In the canonical setting of Bartik (1991) and Blanchard and Katz (1992), for example, a local labor demand instrument is sought, with the ideal $g_n$ measuring
an aggregate change in industry labor demand that may be assumed orthogonal to local labor supply shocks. Aggregate demand changes are however not directly observed and must be estimated from national industry employment growth (often using leave-out corrections, as in Autor and Duggan (2003) and Diamond (2016)). We have discussed how our framework generalizes to this more involved setting in Section 4.3, showing the additional assumptions required for the estimation error to be asymptotically ignorable. While this case differs from the first in terms of the instrument construction, the underlying logic of our framework still applies. This case also covers instruments in the immigration literature, as in Card (2001) and Card (2009), where latent shocks to out-migration from foreign countries can be thought to be as-good-as-randomly assigned but are estimated from aggregate immigration flows in the U.S.

The third case is conceptually distinct, in that the $g_n$ underlying the (perhaps idealized) instrument cannot be naturally viewed as an instrument itself. This could either be because it is not plausible that these shocks are as-good-as-randomly assigned, even conditionally on shock-level observables, or because there are too few shocks. Identification in this case may instead follow from exogeneity of the exposure shares, as suggested by Goldsmith-Pinkham et al. (2020).

Share exogeneity may be a more plausible approach in the third case when the exposure shares are “tailored” to the specific economic question, and to the particular endogenous variable included in the model. In this case, the scenario considered in Section 2.2—that there are unobserved shocks $\nu_n$ which enter $\varepsilon_{\ell}$ through the shares—may be less of a concern. Mohnen (2019), for example, uses the age profile of older workers in local labor markets as the exposure shares of a shift-share instrument for the change in the local elderly employment rate in the following decade. He argues, based on economic intuition, that these tailored shares are uncorrelated with unobserved trends in youth employment rates. This argument notably does not require one to specify the age-specific shocks $g_n$, which only affect power of the instrument (in fact, the shocks are dispensed with altogether in robustness checks that directly instrument with the shares). Similarly, Algan et al. (2017) use the lagged share of the construction sector in the regional economy as an instrument for unemployment growth during the Great Recession, arguing that it does not predict changes in voting outcomes in other ways. With a single industry the identification assumption reduces to that of conventional difference-in-differences with continuous treatment intensity and our framework cannot be applied.

In contrast, our framework may be more appropriate in settings where shocks are tailored to a specific question while the shares are “generic,” in that they could conceivably measure an observation’s exposure to multiple shocks (both observed and unobserved). Both Autor et al. (2013) and Acemoglu and Restrepo (Forthcoming), for example, build shift-share instruments with similar lagged employment shares but different shocks—rising trade with China and the adoption of industrial robots, respectively. According to the Goldsmith-Pinkham et al. (2020) view, these papers use essentially the same instruments (lagged employment shares) for different endogenous variables (growth of import
competition and growth of robot adoption), and are therefore mutually inconsistent. Our framework helps reconcile these identification strategies, provided the variation in each set of shocks can be described as arising from a natural experiment. In principle, shares and shocks may simultaneously provide valid identifying variation, but in practice it would seem unlikely for both sources of variation to be \textit{a priori} plausible in the same setting.

This discussion highlights that plausibility of our framework, as with the alternative framework of Goldsmith-Pinkham et al. (2020), depends on the details of the SSIV application. We encourage practitioners to use our framework only after establishing an \textit{a priori} argument for the plausibility of exogenous shocks. Various diagnostics on the extent of shock variation and falsification of this assumption may then be conducted to assess \textit{ex post} the plausibility of exogenous shocks. We next illustrate this approach in the Autor et al. (2013) setting.

### 6.2 Application to Autor, Dorn, and Hanson (2013)

Our application to Autor et al. (2013, henceforth ADH) aims to illustrate our theoretical framework only, and not to reassess their substantive findings. In line with this goal, we first describe how the ADH instrument could be thought to leverage quasi-experimental shocks and discuss potential threats to this identification strategy. We then illustrate the tools and lessons that follow from our framework, demonstrating steps that researchers can emulate in their own SSIV applications. Specifically, we analyze the distribution of shocks to assess the plausibility of Assumption 4 (many conditionally uncorrelated shocks), use balance tests to corroborate the plausibility of Assumption 3 (conditional quasi-random shock assignment), use equivalent shock-level IV regressions to obtain exposure-robust inference, and analyze the sensitivity of the results to the inclusion of different shock-level controls. This analysis shows how our quasi-experimental framework can help understand the identifying variation in the ADH SSIV design.

#### 6.2.1 Setting and Intuition for Identification

ADH use a shift-share IV to estimate the causal effect of rising import penetration from China on U.S. local labor markets. They do so with a repeated cross section of 722 commuting zones $\ell$ and 397 four-digit SIC manufacturing industries $n$ over two periods $t$, 1990-2000 and 2000-2007. In these years U.S. commuting zones were exposed to a dramatic rise in import penetration from China, a historic change in trade patterns commonly referred to as the “China shock.” Variation in exposure to this change across commuting zones results from the fact that different areas were initially specialized in different industries which saw different changes in the aggregate U.S. growth of Chinese imports. ADH combine import changes across industries in eight comparable developed economies (as shocks) with lagged industry employment (as exposure shares) to construct their shift-share instrument.

To illustrate our framework in this setting we focus on ADH’s primary outcome of the change
in total manufacturing employment as a percentage of working-age population during period $t$ in location $\ell$, which we write as $y_{\ell t}$. The treatment variable $x_{\ell t}$ measures local exposure to the growth of imports from China in $1,000$ per worker. The vector of controls $w_{\ell t}$, which comes from the preferred specification of ADH (Column 6 of their Table 3), contains start-of-period measures of labor force demographics, period fixed effects, Census region fixed effects, and the start-of-period total manufacturing share to which we return below. The shift-share instrument is $z_{\ell t} = \sum_n s_{\ell nt} g_{nt}$, where $s_{\ell nt}$ is the share of manufacturing industry $n$ in total employment in location $\ell$ (measured a decade before each period $t$ begins) and $g_{nt}$ is industry $n$’s growth of imports from China in the eight comparable economies over period $t$ (also expressed in $1,000$ per U.S. worker). Importantly, the sum of lagged manufacturing shares across industries ($S_{\ell t} = \sum_n s_{\ell nt}$) is not constant across locations and periods, placing the ADH instrument in the “incomplete shares” class discussed in Section 4.2. All regressions are weighted by $e_{\ell t}$, which measures the start-of-period population of the commuting zone, and all variables are measured in ten-year equivalents.

To see how the ADH instrument can be viewed as leveraging quasi-experimental shocks, it is useful to imagine an idealized experiment generating random variation in the growth of imports from China across industries. One could imagine, for example, random variation in industry-specific productivities in China affecting import growth in both the U.S. and in comparable economies. This would yield a set of observed productivity changes $g_{nt}$ which would plausibly satisfy our Assumption 1. Assumption 2 would further hold when the productivity shocks are idiosyncratic across many industries, with small average exposure to each shock across commuting zones. Weaker versions of this experimental ideal, in which productivity shocks can be partly predicted by industry observables and are only weakly dependent across industries, are accommodated by the extensions in Section 3.2. For example, in ADH’s repeated cross section one might invoke Assumption 3 in allowing the average shock to vary across periods, in recognition that the 1990s and 2000s were very different trade environments, with China joining the World Trade Organization in 2001. Here $q_{nt}$ would indicate periods. Other extensions would allow the industry shocks to be serially correlated and clustered by larger groups.

ADH’s approach can be seen as approximating this idealized experiment by using observed changes in trade patterns between China and a group of developed countries outside the United States. The use of non-U.S. data can be seen as avoiding a potential concern of omitted variables bias: trade between the U.S. and China depends on changes in U.S. supply and demand conditions, which may...
have direct effects on employment dynamics in U.S. regions. In contrast, variation in the ADH $g_{nt}$ reflects only Chinese productivity shocks and the various supply and demand shocks in the non-U.S. developed countries. In this way, the ADH strategy can be understood as eliminating bias from shocks that are specific to the United States.

This discussion gives an a priori justification for thinking of the ADH instrument as leveraging quasi-experimental shocks within the two periods. Nonetheless, since the ADH shocks are not truly randomized, one may still worry that they are confounded by other unobserved characteristics. For example, China’s factor endowment may imply that it specializes in low-skill industries, which could have been on different employment trends in the U.S. even absent increased trade with China. Similarly, one can imagine a common component of import growth in the U.S. and the group of comparable developed economies due to correlated technological shocks in those countries, which may have a direct effect on U.S. labor markets. Given these potential concerns, it will be important assess the plausibility of Assumption 3 in this setting by conducting within-period falsification tests of the kind we describe in Section 5.2. It will also be important to access whether there is sufficient variation in the ADH shocks for Assumption 4 to hold, allowing for serial correlation and clustering.

Before applying these tests, it is worth highlighting that the assumption of exogenous exposure shares, as discussed by Goldsmith-Pinkham et al. (2020), is likely to be a priori implausible in the ADH setting. As indicated in Section 2.2, any unobserved shocks $\nu_n$ invalidate the share exogeneity assumption if they enter the error $\varepsilon_t$ in a manner which is correlated with the shares. Because ADH use generic manufacturing employment shares to instrument for a specific treatment variable, the possibility of other industry shocks entering $\varepsilon_t$ looms large. These unobserved shocks could take many forms, for example heterogeneous speeds of automation, secular changes in consumer demand, or changes in factor prices which differentially affect industries based on their skill intensity.

With a plausible justification of our framework in hand, we next illustrate its application.

### 6.2.2 Properties of Industry Shocks and Exposure Shares

Our quasi-experimental view of the ADH research design places particular emphasis on the variation in Chinese import growth rates $g_{nt}$ and their average exposure $s_{nt}$ across industries and periods. With few or insufficiently-variable shocks, or highly concentrated shocks exposure, the large-$N$ asymptotic approximation developed in Section 3.1 is unlikely to be a useful tool for characterizing the finite-sample behavior of the SSIV estimator. We thus first summarize the distribution of $g_{nt}$, as well as the industry-level weights from our equivalence result, $s_{nt} \propto \sum_t e_{lt} s_{nt}$ (normalized to add up to one in the entire sample).\(^{39}\)

In summarizing the industry-level variation it is instructive to recall that the ADH instrument is constructed with “incomplete” manufacturing shares. Per the discussion in Section 4.2, this means that

\(^{39}\)Note that $s_{nt}$ would be proportional to lagged industry employment if the ADH regression weights $e_{lt}$ were lagged regional employment. ADH however use a slightly different $e_{lt}$: the start-of-period commuting zone population.
absent any regression controls the SSIV estimator uses variation not only in manufacturing industry shocks but also implicitly the variation in the 10-year lagged total manufacturing share $S_{lt}$ across commuting zones and periods. In practice, ADH control for the start-of-period manufacturing share, which is highly—though not perfectly—correlated with $S_{lt}$. We thus summarize the ADH shocks both with and without the “missing” shock $g_{0t} = 0$, which here represents the lack of a “China shock” in service (i.e. non-manufacturing) industries. Moreover, in recognition of the fact that the Chinese trade environment was very different in the 1990s and 2000s, so shock exogeneity may only plausibly hold across industries, within periods, we focus on the within-period variation in manufacturing shocks.

Table 1 reports summary statistics for the ADH shocks $g_{nt}$ computed with importance weights $s_{nt}$, and characterizes these weights. Column 1 includes the “missing” service industry shock of zero in each period. It is evident that with this shock the distribution of $g_{nt}$ is unusual: for example, its interquartile range is zero. This is because the service industry accounts for a large fraction of total employment ($s_{ot}$ is 71.9% of the period total in the 1990s and 79.5% in the 2000s). As a result we see a high concentration of industry exposure as measured by the inverse of its Herfindahl index (HHI), $1/\sum_{n,t} s_{nt}^2$, which corresponds to the effective sample size of our equivalent regression and plays a key role in Assumption 2. With the “missing” shock included, the effective sample size is only 3.5. For an HHI computed at the level of three-digit industry codes $\sum_c s_c^2$, where $s_c$ aggregates exposure across the two periods and industries within the same 3-digit group $c$, it is even lower, at 1.7. This suggests even less industry-level variation is available when shocks are allowed to be serially correlated or clustered by groups. Furthermore, the mean of manufacturing shocks is significantly different from the zero shock of the missing service industry. Together, these analyses suggest that the service industry should be excluded from the identifying variation, because it is likely to violate both Assumption 1 ($\mathbb{E}[g_{nt} | \bar{\varepsilon}, s] \neq g_{0t} = 0$) and Assumption 2 ($\sum_{n,t} s_{nt}^2$ is not close to zero).

Column 2 of Table 1 therefore summarizes the sample with the service industry excluded. The distribution of shocks is now much more regular, with an average of 7.4, a standard deviation of 20.9 and an interquartile range of 6.6. The inverse HHI of the $s_{nt}$ is also relatively high: 191.6 across industry-by-period cells and 58.4 when exposure is aggregated by SIC3 group. The largest shock weights in this column are only 3.4% across industry-by-periods and 6.5% across SIC3 groups. This suggests a sizable degree of variation at the industry level, consistent with Assumption 2. In general, we recommend that researchers report the inverse of the HHI of shock-level average exposure as a simple way of describing their effective sample size. A first-stage $F$-statistic, which we discuss appropriate computation of in Section 5.2, will provide a formal test of the power of the shock variation.

Finally, column 3 of Table 1 summarizes the distribution of within-period manufacturing shocks, which would be leveraged by an assumption of conditional quasi-experimental assignment (Assumption 3). The column confirms that even conditional on period there is sizable residual shock variation.

---

40The weighted mean of manufacturing shocks is 7.4, with a standard error clustered at the 3-digit SIC level (as in our analysis below) of 1.3.
The standard deviation and interquartile range of shock residuals (obtained from regressing shocks on period fixed effects with $s_{nt}$ weights) are only mildly smaller than in Column 2, despite the higher mean shock in the later period, at 12.6 versus 3.6.

Besides the condition on the effective sample size, Assumption 2 (and its clustered version in Assumption 5) requires the shocks to be sufficiently mutually uncorrelated. To assess the plausibility of this assumption and choose the appropriate level of clustering for exposure-robust standard errors, we next analyze the correlation patterns of shocks across manufacturing industries using available industry classifications and the time dimension of the pooled cross section. In particular, we compute intra-class correlation coefficients (ICCs) of shocks within different industry groups, as one might do to correct for conventional clustering parametrically (e.g. Angrist and Pischke (2008, p. 312)).

These ICCs come from a random effects model, providing a hierarchical decomposition of residual within-period shock variation:

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

(14)

where $\mu_t$ are period fixed effects; $a_{ten(n),t}$, $b_{sic2(n),t}$, and $c_{sic3(n),t}$ denote time-varying (and possibly auto-correlated) random effects generated by the ten industry groups in Acemoglu et al. (2016), 20 groups identified by SIC2 codes, and 136 groups corresponding to SIC3 codes, respectively; and $d_n$ is a time-invariant industry random effect (across our 397 four-digit SIC industries). Following convention, we estimate equation (14) as a hierarchical linear model by maximum likelihood, assuming Gaussian residual components.

Table 2 reports estimated ICCs from equation (14), summarizing the share of the overall shock residual variance due to each random effect. These reveal moderate clustering of shock residuals at the industry and SIC3 level (with ICCs of 0.169 and 0.073, respectively). At the same time, there is less evidence for clustering of shocks at a higher SIC2 level and particularly by ten cluster groups (ICCs of 0.047 and 0.016, respectively, with standard errors of comparable magnitude). This supports the assumption that shocks are mean-independent across SIC3 clusters, so it will be sufficient to cluster standard errors at the level of SIC3 groups, as Acemoglu et al. (2016) do in their conventional industry-level IV regressions. The inverse HHI estimates in Table 1 indicate that at this level of shock clustering there is still an adequate effective sample size.

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

In particular we estimate an unweighted mixed-effects regression using Stata’s `mixed` command, imposing an exchangeable variance matrix for $\left(a_{ten(n),1}, a_{ten(n),2}, b_{sic2(n),1}, b_{sic2(n),2}, c_{sic3(n),1}, c_{sic3(n),2}\right)$.  

30
6.2.3 Falsification Tests

We next implement falsification tests of ADH shock orthogonality, which provide a way of assessing the plausibility of Assumption 3. Following Section 5.2, we do this in two ways, both different from conventional falsification tests sometimes run in SSIV settings. First, we regress potential proxies for the unobserved residual (i.e., any unobserved industry labor demand or labor supply shock) on the instrument $z_{t}$ but use exposure-robust inference that takes into account the inherent dependencies of the data. Second, we regress potential industry-level confounders directly on the shocks (while again clustering by SIC3). While this second type of falsification tests would be standard in industry-level analyses, such as Acemoglu et al. (2016), it has rarely been used to assess the plausibility of SSIV designs (with Xu (2019), mentioned above, being a rare exception).

Choosing the set of potential confounders for these exercises is a context-specific issue, which should be justified separately in every application. To discipline our illustrative exercise, we use the industry-level production controls in Acemoglu et al. (2016) and the regional controls in ADH. Consistent with our a priori view of the quasi-experiment, we maintain only the period fixed effects as controls when evaluating balance on these other observables. For the industry-level balance test this amounts to regressing each potential confounder on the manufacturing shocks (normalized to have a unit variance) and period fixed effects, weighting by average industry employment shares. Regional balance coefficients are obtained by regressing each potential confounder on the shift-share instrument (normalized to have a unit variance) and the share-weighted average of period effects (i.e., the period-interacted sum-of-shares), since ADH is a setting with incomplete shares. To obtain exposure-robust standard errors, we implement these regressions at the shock level, as discussed above.

Panel A of Table 3 reports the results of our industry-level balance tests. The five Acemoglu et al. (2016) production controls are an industry’s share of production workers in employment in 1991, the ratio of its capital to value-added in 1991, its log real wages in 1991, the share of its investment devoted to computers in 1990, and the share of its high-tech equipment in total investment in 1990.43 Broadly, these variables reflect the structure of employment and technology across industries. If the ADH shocks are as-good-as-randomly assigned to industries within periods, we expect them to not predict these predetermined variables. Panel A shows that there is indeed no statistically significant correlation within periods, consistent with Assumption 3.

Panel B of Table 3 reports the results of our regional balance tests. The five ADH controls are the fraction of a commuting zone’s population who is college-educated, the fraction of its population who is foreign-born, the fraction of its workers who are female, its fraction of employment in routine occupations, and the average offshorability index of its occupations. Broadly, these variables reflect the composition of a region’s workforce. We again find no statistically significant relationships between

---

43The last two controls are missing for five out of 397 industries. We impute the missing values by the medians in the SIC3 industry group or, when not available, in the SIC2 group.
these variables and the shift-share instrument within periods, except for the foreign-born population fraction. Locations exposed to a large ADH trade shock tend to have a higher fraction of immigrants, suggesting that they may be subject to different labor supply dynamics. We explore the importance of this imbalance for the SSIV coefficient estimate in sensitivity tests below.

Finally, the last two rows of the same panel conduct a regional “pre-trends” analysis. We regress the pre-trend variables from ADH—manufacturing employment growth in the 1970s and 1980s—on the shift-share instrument, using the same specification as in the previous rows. We find no relationship between the shift-share instrument and manufacturing employment growth in the 1980s, but there is a positive statistically significant relationship with manufacturing employment growth in the 1970s. Both findings are similar to those from ADH’s pre-trend analysis.

Overall, we fail to reject imbalance in ten out of the twelve potential confounders at conventional levels of statistical significance. How to proceed when some balance tests fail is a general issue in quasi-experimental analyses and has to be decided in the context of an application. One might view the balance failures as sufficient evidence against Assumption 3 to seek alternative shocks or more appropriate shock-level controls. Alternatively, one may argue that the observed imbalances are unlikely to invalidate the research design. ADH, for example, note that the positive relationship they find between the shift-share instrument and manufacturing employment in the 1970s occurs in the distant past, while the insignificant relationship in the 1980s demonstrates that the relationship between rising China trade exposure and declining manufacturing employment was absent in the decade immediately prior to China’s rise. Similarly the imbalance of the foreign-born share that we observe need not generate a bias in the estimate if it is not strongly correlated with the second-stage residual. To gauge this potential for OVB one can include such variables as controls in the SSIV specification and check sensitivity of the coefficient; we report results of this exercise next.

### 6.2.4 Main Estimates and Sensitivity Analyses

We next estimate the effects of import competition on local labor market outcomes, leveraging within-period exogeneity of the industry shocks $g_{nt}$. We then check sensitivity of results to inclusion of the Autor et al. (2013) regional controls and Acemoglu et al. (2016) industry-level controls.

Table 4 reports SSIV coefficients from regressing regional manufacturing employment growth in the U.S. on the growth of import competition from China, instrumented by predicted Chinese import growth.$^{44}$ Per the results in Section 5.1, we estimate these coefficients with equivalent industry-level regressions in order to obtain valid exposure-robust standard errors. Consistent with the above analysis of shock ICCs, we cluster standard errors at the SIC3 level. We also report first-stage $F$-statistics with corresponding exposure-robust inference. As discussed in Section 5.2, these come from industry-level IV regressions of the aggregated treatment and instrument (i.e. $\bar{x}_{nt}$ on $\bar{z}_{nt}$), instrumented with shocks

---

$^{44}$ Appendix Table C1 reports estimates for other outcomes in ADH: growth rates of unemployment, labor force non-participation, and average wages, corresponding to columns 3 and 4 of Table 5 and column 1 of Table 6 in ADH.
and weighting by $s_{nt}$. The $F$-statistics are well above the conventional threshold of ten in all columns of the table.

Column 1 first replicates column 6 of Table 3 in Autor et al. (2013) by including in $w_{lt}$ period fixed effects, Census division fixed effects, start-of-period conditions (% college educated, % foreign-born, % employment among women, % employment in routine occupations, and the average offshorability index), and the start-of-period manufacturing share. The point estimate is -0.596, with a corrected standard error of 0.114.\footnote{Appendix Table C2 implements three alternative methods for conducting inference in Table 4, reporting conventional state-clustered standard errors as in ADH (which are not exposure-robust), the Adão et al. (2019) standard errors (which are asymptotically equivalent to ours but differ in finite samples), and null-imposed confidence intervals obtained from shock-level Lagrange multiplier tests (which may have better finite-sample properties). Consistent with the theoretical discussion in Appendix B.2, the conventional standard errors are generally too low, while the Adão et al. (2019) standard errors are slightly larger than those from Table 4 in most columns. Imposing the null widens the confidence interval more substantially, by 30–50%, although more so on the left end, suggesting that much larger effects are not rejected by the data. This last finding is consistent with Adão et al. (2019), except that we use the equivalent industry-level regression to compute the null-imposed confidence interval.}

As noted, the ADH specification in column 1 does not include the lagged manufacturing share control $S_{lt}$, which is necessary to solve the incomplete shares issue in Section 4.2, though it does include a highly correlated control (start-of-period manufacturing share). In column 2 of Table 4 we isolate within-manufacturing variation in shocks by replacing the latter sum-of-share control with the former. The SSIV point estimate remains almost unchanged, at -0.489 (with a standard error of 0.100). Here exposure-robust standard errors are obtained from an industry-level regression that drops the implicit service sector shock of $g_{0t} = 0$.

Isolating the within-period variation in manufacturing shocks requires further controls in the incomplete shares case, as discussed in Section 4.2. Specifically, column 3 controls for lagged manufacturing shares interacted with period indicators, which are the share-weighted sums of period effects in $q_{nt}$. This is equivalent to the use of period fixed effects in the industry-level analysis of Acemoglu et al. (2016). With these controls the SSIV point estimate is -0.267 with an exposure-robust standard error of 0.099.\footnote{Appendix Figure C1 reports binned scatter plots that illustrate the first-stage and reduced-form industry-level relationships corresponding to the column 3 specification. This estimate can be interpreted as a weighted average of two period-specific shift-share IV coefficients. Column 1 of Appendix Table C3 shows the underlying estimates, from a just-identified IV regression where both treatment and the instrument are interacted with period indicators (as well as the manufacturing share control, as in column 3), with exposure-robust standard errors obtained by the equivalent industry-level regression discussed in Section 5. 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.} While the coefficient remains statistically and economically significant, it is smaller in magnitude than the estimates in columns 1 and 2. The difference stems from the fact that 2000–07 saw both a faster growth in imports from China (e.g., due to its entry to the WTO) and a faster decline in U.S. manufacturing. The earlier columns attribute the faster manufacturing decline to increased trade with China, while the specification in Column 3 controls for any unobserved shocks specific to the manufacturing sector overall in the 2000s (e.g., any demand or supply shock affecting the manufacturing sector, which could include automation, innovation, falling consumer demand due to income effects, etc.). Conventional industry-level IV regressions control for such unobserved shocks

45Appendix Table C2 implements three alternative methods for conducting inference in Table 4, reporting conventional state-clustered standard errors as in ADH (which are not exposure-robust), the Adão et al. (2019) standard errors (which are asymptotically equivalent to ours but differ in finite samples), and null-imposed confidence intervals obtained from shock-level Lagrange multiplier tests (which may have better finite-sample properties). Consistent with the theoretical discussion in Appendix B.2, the conventional standard errors are generally too low, while the Adão et al. (2019) standard errors are slightly larger than those from Table 4 in most columns. Imposing the null widens the confidence interval more substantially, by 30–50%, although more so on the left end, suggesting that much larger effects are not rejected by the data. This last finding is consistent with Adão et al. (2019), except that we use the equivalent industry-level regression to compute the null-imposed confidence interval.

46Appendix Figure C1 reports binned scatter plots that illustrate the first-stage and reduced-form industry-level relationships corresponding to the column 3 specification. This estimate can be interpreted as a weighted average of two period-specific shift-share IV coefficients. Column 1 of Appendix Table C3 shows the underlying estimates, from a just-identified IV regression where both treatment and the instrument are interacted with period indicators (as well as the manufacturing share control, as in column 3), with exposure-robust standard errors obtained by the equivalent industry-level regression discussed in Section 5. 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.
with period fixed effects, as in Table 3, column 1 of Acemoglu et al. (2016). The translation of their industry specification into the regional setup of ADH requires interacting the lagged manufacturing shares with period indicators, a simple but important insight of our framework.

Column 4 implements a simple sensitivity test to assess the stability of the results when the controls from ADH are omitted. This test is motivated by the result of the balance test in panel B of Table 3, which indicated that the shift-share instrument was correlated with the share of foreign born population. It is therefore instructive to see whether the headline regression coefficient is sensitive to the inclusion of this and other controls. In fact, we find that the results remain very similar without controls, with a point estimate of -0.314 and an exposure-robust standard error of 0.134. We proceed by keeping the ADH controls for the remainder of the analysis.

Further columns of Table 4 parallel the specifications of Acemoglu et al. (2016, Table 3) that include further industry-level controls. This illustrates how our framework makes it straightforward to introduce more detailed industry-level controls in SSIV, which are commonly used in industry-level studies of the China shock. The validity of these estimates relies on weaker versions of conditional random assignment (Assumption 3), and robustness of the coefficients is therefore reassuring. Specifically, Acemoglu et al. (2016) control for fixed effects of ten broad industry groups (one-digit manufacturing sectors) in column 2 of their Table 3. By Proposition 4, we can exploit shock variation within these industry groups in the SSIV design by controlling for the lagged shares of exposure to these industry groups (and including fixed effects of these groups in the equivalent industry regressions for correct inference, per Section 5.1). The resulting point estimate in column 5 of Table 4 remains very similar to that of column 3, at -0.310 with a standard error of 0.134.

Column 6 instead parallels the specification of Acemoglu et al. (2016) that includes production controls, which we used for the balance tests in Panel A of Table 3. This is done by controlling for the regional share-weighted sums of those controls. The results remain virtually unchanged, with a regression coefficient of -0.293 and an exposure-robust standard error of 0.125.

Finally, column 7 instead introduces industry fixed effects, again following Acemoglu et al. (2016). This specification is more ambitious because it isolates changes in trade with China within each four-digit SIC industry, across the two periods. To translate the industry fixed effects to the location-level setup, we control for the lagged location-specific share of exposure to each industry. The magnitude of the regression coefficient increases, to -0.432, with an exposure-robust standard error of 0.205.

47In principle, China could have affected the path of the U.S. manufacturing sector as a whole, and thus the variation in the average China shock across periods may be informative about the effects of interest. However, because of the multiplicity of shocks that may affect the manufacturing sector as a whole in a given period, this variation cannot be viewed as a quasi-experimental source of variation for the impact of trade with China on employment and other outcomes. This is why industry-level studies of the China shock use period fixed effects, possibly reducing power but substantially improving robustness of the estimates. In the ADH application the estimation power is not actually reduced, as the Table 4 column 3 standard error is even slightly smaller than that in the previous columns.

48If the shares used to construct the instrument were time-invariant, a more conventional and intuitive way to exploit over-time variation in the shocks would be by including the regional fixed effects in the regression, as Section 4.1 explained. In the ADH setting where the shares vary over time, they need to be controlled for directly.
Broadly, these results demonstrate the stability of the SSIV regression coefficient under alternative sets of controls, corresponding to different assumptions of conditional quasi-random shock assignment.

The appendix reports estimates from additional specifications. Appendix Table C4 includes additional controls corresponding to other specifications of Acemoglu et al. (2016), Table 3: for example, controlling for observed changes in employment in the pre-periods or combining multiple sets of controls. The regression coefficients remain stable across all specifications. Appendix Table C5 instead shows robustness of the coefficients to using overidentified SSIV procedures (leveraging variation in eight country-specific Chinese import growth, instead of the ADH total), illustrating the theoretical results of Section 4.4. The table also reports a p-value for the shock-level overidentification test of 0.142, providing further support to the identification assumptions.

6.2.5 Discussion

Taken together, the sensitivity, falsification, and overidentification exercises suggest that the ADH approach can be reasonably viewed as leveraging exogenous shock variation via our framework. This is notably in contrast to the analysis of Goldsmith-Pinkham et al. (2020), who find the ADH exposure shares to be implausible instruments via different balance and overidentification tests. This contrast should perhaps come as no surprise. As mentioned, the exogeneity of industry employment shares is an ex ante implausible research design, because it is invalidated by any unobserved labor demand or supply shocks across industries (which we view as an inherent feature of the economy).

In contrast, our approach relies on the exogeneity of the specific ADH trade shocks, allowing for endogenous exposure shares. With this view, the potential confounders are a more specific set of unobserved industry shocks (namely, unobserved shocks that would correlate with the ADH shocks), rather than any unobserved shocks. In principle, the conditions for shock orthogonality could still fail because of these specific unobserved shocks. In practice, our balance tests indicate that there is little evidence to suggest that the ADH shocks are confounded.

Our ADH application therefore illustrates two points. First, the assumptions of our framework are plausible, both ex ante and ex post, in an influential empirical setting, where an alternative SSIV framework is inapplicable. Second, our framework helps researchers translate shock-level identifying assumptions to appropriate SSIV regression controls, falsify those assumptions with appropriate balance tests, and perform correct inference.

7 Conclusion

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

Through various extensions and illustrations, we show how our quasi-experimental SSIV framework can guide empirical work in practice. By controlling for exposure-weighted averages of shock-level confounders, researchers can isolate more plausibly exogenous variation in shocks, such as over time or within narrow industry groups. By estimating SSIV coefficients, placebo regressions, and first stage $F$-statistics at the level of shocks, researchers can conveniently perform exposure-robust inference that accounts for the inherent non-standard clustering of observations with common shock exposure. Our shock-level analysis also raises new concerns: SSIV designs with few or insufficiently dispersed shocks may have effectively small samples, despite there being many underlying observations, and instruments constructed from exposure shares that do not add up to a constant require appropriate controls in order to isolate quasi-random shock variation. We illustrate these practical implications in an application to the influential study of Autor et al. (2013).

In sum, our analysis formalizes the claim that SSIV identification and consistency may arise from the exogeneity of shocks, while providing new guidance for SSIV estimation and inference that may be applied across a number of economic fields, including international trade, labor economics, urban economics, macroeconomics, and public finance. Our shock-level assumptions connect SSIV in these settings to conventional shock-level IV estimation, bringing shift-share instruments to more familiar econometric territory and facilitating the assessment of SSIV credibility in practice.
Table 1: Shock Summary Statistics in the Autor et al. (2013) Setting

|                           | (1)  | (2)  | (3)  |
|---------------------------|------|------|------|
| Mean                      | 1.79 | 7.37 | 0    |
| Standard deviation        | 10.79| 20.92| 20.44|
| Interquartile range       | 0    | 6.61 | 6.11 |

**Specification**

- Excluding service industries: ✓ ✓
- Residualizing on period FE: ✓

**Effective sample size ($1/HHI$ of $s_{nt}$ weights)**

- Across industries and periods: 3.5 191.6 191.6
- Across SIC3 groups: 1.7 58.4 58.4

**Largest $s_{nt}$ weight**

- Across industries and periods: 0.398 0.035 0.035
- Across SIC3 groups: 0.757 0.066 0.066

**Observation counts**

- # of industry-period shocks: 796 794 794
- # of industries: 398 397 397
- # of SIC3 groups: 137 136 136

Notes: This table summarizes the distribution of China import shocks $g_{nt}$ across industries $n$ and periods $t$ in the Autor et al. (2013) application. Shocks are measured as the total flow of imports from China in eight developed economies outside of the United States. All statistics are weighted by the average industry exposure shares $s_{nt}$; shares are measured from lagged manufacturing employment, as described in Section 6.2.1. Column 1 includes the non-manufacturing industry aggregate in each period with a shock of 0, while columns 2 and 3 restrict the sample to manufacturing industries. Column 3 residualizes manufacturing shocks on period indicators. We report the effective sample size (the inverse renormalized Herfindahl index of the $s_{nt}$ weights, as described in Section 6.2.2) with and without the non-manufacturing industry, at the industry-by-period level and at the level of SIC3 groups (aggregated across periods), along with the largest $s_{nt}$.
Table 2: Shock Intra-Class Correlations in the Autor et al. (2013) Setting

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

Notes: This table reports intra-class correlation coefficients for the Autor et al. (2013) manufacturing shocks, estimated from the hierarchical model described in Section 6.2.2. Estimates come from a maximum likelihood procedure with an exchangeable covariance structure for each industry and sector random effect and with period fixed effects. Robust standard errors are reported in parentheses.
Table 3: Shock Balance Tests in the Autor et al. (2013) Setting

Panel A: Industry-Level Balance

| Balance variable                                                | Coef. | SE    |
|-----------------------------------------------------------------|-------|-------|
| Production workers’ share of employment, 1991                   | -0.011| (0.012)|
| Ratio of capital to value-added, 1991                          | -0.007| (0.019)|
| Log real wage (2007 USD), 1991                                 | -0.005| (0.022)|
| Computer investment as share of total, 1990                     | 0.750 | (0.465)|
| High-tech equipment as share of total investment, 1990          | 0.532 | (0.296)|

# of industry-periods  794

Panel B: Regional Balance

| Balance variable                                                      | Coef. | SE    |
|----------------------------------------------------------------------|-------|-------|
| Start-of-period % of college-educated population                      | 0.915 | (1.196)|
| Start-of-period % of foreign-born population                          | 2.920 | (0.952)|
| Start-of-period % of employment among women                          | -0.159| (0.521)|
| Start-of-period % of employment in routine occupations               | -0.302| (0.272)|
| Start-of-period average offshorability index of occupations          | 0.087 | (0.075)|
| Manufacturing employment growth, 1970s                               | 0.543 | (0.227)|
| Manufacturing employment growth, 1980s                               | 0.055 | (0.187)|

# of region-periods  1,444

Notes: Panel A of this table reports coefficients from regressions of the industry-level covariates in Acemoglu et al. (2016) on the Autor et al. (2013) shocks, controlling for period indicators and weighting by average industry exposure shares. Standard errors are reported in parentheses and allow for clustering at the level of three-digit SIC codes. Panel B reports coefficients from regressions of commuting zone-level covariates and pre-trends from Autor et al. (2013) on the shift-share instrument, controlling for period indicators interacted with the lagged manufacturing share. Balance variables (the first five rows of this panel) vary across the two periods, while pre-trends (the last two rows) do not. SIC3-clustered exposure-robust standard errors are reported in parentheses and obtained from equivalent industry-level IV regressions as described in Section 6.2.3. Independent variables in both panels are normalized to have a variance of one in the sample.
Table 4: Shift-Share IV Estimates of the Effect of Chinese Imports on Manufacturing Employment

|                  | (1)  | (2)  | (3)  | (4)  | (5)  | (6)  | (7)  |
|------------------|------|------|------|------|------|------|------|
| Coefficient      | -0.596 | -0.489 | -0.267 | -0.314 | -0.310 | -0.290 | -0.432 |
|                  | (0.114) | (0.100) | (0.099) | (0.107) | (0.134) | (0.129) | (0.205) |
| Regional controls|      |      |      |      |      |      |      |
| Autor et al. (2013) controls | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Start-of-period mfg. share | ✓ |      |      |      |      |      |      |
| Lagged mfg. share | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Period-specific lagged mfg. share | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ | ✓ |
| Lagged 10-sector shares | ✓ |      |      |      |      |      |      |
| Local Acemoglu et al. (2016) controls |      | ✓ |      |      |      |      |      |
| Lagged industry shares | ✓ |      |      |      |      |      |      |
| SSIV first stage F-stat. | 185.6 | 166.7 | 123.6 | 272.4 | 64.6 | 63.3 | 27.6 |
| # of region-periods | 1,444 | 1,444 | 1,444 | 1,444 | 1,444 | 1,444 | 1,444 |
| # of industry-periods | 796 | 794 | 794 | 794 | 794 | 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, instrumented with predicted China import growth as described in Section 6.2.1. Column 1 replicates column 6 of Table 3 in Autor et al. (2013) by controlling for period fixed effects, Census division fixed effects, start-of-period conditions (% college educated, % foreign-born, % employment among women, % employment in routine occupations, and the average offshorability index), and the start-of-period manufacturing share. Column 2 replaces the start-of-period manufacturing shares control with the lagged manufacturing shares underlying the instrument, while column 3 interacts this control with period indicators. Column 4 removes the Census division fixed effects and start-of-period covariates. Columns 5–7 instead add exposure-weighted sums of industry controls from Acemoglu et al. (2016): indicators of 10 industry sectors (column 5), production controls (column 6), and indicators of 397 industries (column 7). Production controls are: employment share of production workers, ratio of capital to value-added, log real wage (all measured in 1991); and computer investment as share of total and high-tech equipment as share of total employment (both measured in 1990). Exposure-robust standard errors (reported in parentheses) and first-stage $F$-statistics are obtained from equivalent industry-level IV regressions, as described in the text, allowing for clustering of shocks at the level of three-digit SIC codes. For commuting zone controls which have a shift-share structure (all controls starting with the lagged manufacturing share), we include the corresponding $q_{it}$ controls in the industry-level IV regression. The sample in columns 2–7 includes 722 locations (commuting zones) and 397 industries, each observed in two periods; the estimate in column 1 implicitly includes an additional two observations for the non-manufacturing industry with a shock of zero in each period.
References

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

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

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

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

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

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

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

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

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

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

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

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

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

Berman, Nicolas, Antoine Berthou, and Jérôme Héricourt. 2015. “Export Dynamics and Sales at Home.” Journal of International Economics 96:298–310.

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

Blanchard, Olivier Jean, and Lawrence F. Katz. 1992. “Regional Evolutions.” Brookings Papers on Economic Activity: 1–75.
Bombardini, Matilde, and Bingjing Li. 2019. “Trade, Pollution and Mortality in China.” Working Paper.

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

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

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

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

Card, David. 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 99:1–21.

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

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

Fajgelbaum, Pablo D., Pinelopi K. Goldberg, Patrick J. Kennedy, and Amit K. Khandelwal. 2020. “The Return to Protectionism.” Quarterly Journal of Economics 135:1–55.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. “Bartik Instruments: What, When, Why, and How.” American Economic Review 110:2586–2624.

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

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

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

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

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

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

Jaeger, David A, Joakim Ruist, and Jan Stuhler. 2018. “Shift-Share Instruments and Dynamic Adjustments: The Case of Immigration.” Working Paper.
Jaravel, Xavier. 2019. “The Unequal Gains from Product Innovations: Evidence from the U.S. Retail Sector.” *The Quarterly Journal of Economics* 134:715–783.

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

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

Mohsen, Paul. 2019. “The Impact of the Retirement Slowdown on the U.S. Youth Labor Market.” Working Paper.

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

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

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

Peri, Giovanni, Kevin Shih, and Chad Sparber. 2016. “STEM workers, H-1B Visas, and productivity in US cities.” *Journal of Labor Economics* 49:277–307.

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

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

Xu, Chenzi. 2019. “Reshaping Global Trade: The Immediate and Long-Run Effects of Bank Failures.” Working Paper.
Appendix to “Quasi-Experimental Shift-Share Research Designs”

Kirill Borusyak, UCL
Peter Hull, University of Chicago
Xavier Jaravel, London School of Economics
August 2020

Contents

A Appendix Results 45
A.1 Heterogeneous Treatment Effects 45
A.2 Comparing SSIV and Native Shock-Level Regression Estimands 46
A.3 Unobserved \( n \)-level Shocks Violate Share Exogeneity 48
A.4 Connection to Rotemberg Weights 49
A.5 Consistency of Control Coefficients 50
A.6 SSIV Consistency in Short Panels 51
A.7 Sequential Exogeneity in Long Panels 52
A.8 SSIV Relevance with Panel Data 53
A.9 Estimated Shocks 54
A.10 Equilibrium Industry Growth in a Model of Local Labor Markets 57
A.11 SSIV with Multiple Endogenous Variables or Instruments 59
A.12 Finite-Sample Performance of SSIV: Monte-Carlo Evidence 63

B Appendix Proofs 69
B.1 Proposition 4 and Extensions 69
B.2 Proposition 5 and Related Results 72
B.3 Proposition A1 77
B.4 Proposition A2 78
B.5 Proposition A3 80
B.6 Proposition A4 81
B.7 Proposition A5 81
B.8 Proposition A6 83

C Appendix Figures and Tables 84
### A Appendix Results

#### A.1 Heterogeneous Treatment Effects

In this appendix we consider what a linear SSIV identifies when the structural relationship between $y_\ell$ and $x_\ell$ is nonlinear. We show that under a first-stage monotonicity condition the large-sample SSIV coefficient estimates a convexly weighted average of heterogeneous treatment effects. This holds even when the instrument has different effects on the outcome depending on the underlying realization of shocks, for example when

$$y_\ell = \sum_n s_{tn} \tilde{\beta}_{tn} x_{tn} + \varepsilon_\ell$$

with $\tilde{\beta}_{tn}$ capturing the effects of (possibly unobserved) observation- and shock-specific treatments $x_{tn}$ making up the observed $x_\ell = \sum_n s_{tn} x_{tn}$.

Consider a general structural outcome model of

$$y_\ell = y(x_{\ell 1}, \ldots, x_{\ell R}, \varepsilon_\ell),$$

where the $R$ treatments are given by $x_{\ell r} = x_r(g, \eta_{\ell r})$ with $g$ collecting the vector of shocks $g_n$ and with $\eta = (\eta_{\ell 1}, \ldots, \eta_{\ell R})$ capturing first-stage heterogeneity. We consider an IV regression of $y_\ell$ on some aggregated treatment $x_\ell = \sum_r \alpha_{\ell r} x_{\ell r}$ with $\alpha_{\ell 1} \geq 0$. Note that this nests the case of a single aggregate treatment ($R = 1$ and $\alpha_{\ell 1} = 1$) with arbitrary effect heterogeneity, as well as the special case above ($R = N$ and $\alpha_{\ell r} = s_{tn}$). We abstract away from controls $w_\ell$ and assume each shock is as-good-as-randomly assigned (mean-zero and mutually independent) conditional on the vector of second-stage unobservables $\varepsilon_\ell$ and the matrices of first-stage unobservables $\eta_{\ell r}$, exposure shares $s_{tn}$, importance weights $e_\ell$, and aggregation weights $\alpha_{\ell r}$, collected in $I = \{\varepsilon_\ell, e_\ell, \{\eta_{\ell r}, \alpha_{\ell r}\}_r, \{s_{tn}\}_n\}_\ell$. This assumption is stronger than Assumption 3 but generally necessary in a non-linear setting while still allowing for the endogeneity of exposure shares. For further notational simplicity we assume that $y(\cdot, \varepsilon_\ell)$ and each $x_r(\cdot, \eta_{\ell r})$ are almost surely continuously differentiable, such that $\beta_{\ell r}(\cdot) = \frac{\partial}{\partial x_r} y(\cdot, \varepsilon_\ell)$ captures the effect, for observation $\ell$, of marginally increasing treatment $r$ on the outcome and $\pi_{\ell nr}(\cdot) = \frac{\partial}{\partial g_n} x_r(\cdot, \eta_{\ell r})$ captures the effect of marginally increasing the $n$th shock on the $r$th treatment at $\ell$.

Under an appropriate law of large numbers, the shift-share IV estimator approximates the IV estimand:

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

Given this, we have the following result:

**Proposition A1** When $\pi_{\ell nr}(\tilde{g}; g_{-n}) \geq 0$ almost surely for all $\tilde{g} \in \mathbb{R}$, equation (16) can be written

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

(17)

45
where $\omega_{\ell n r}(\bar{y}) \geq 0$ almost surely and

$$\beta_{\ell n r}(\bar{y}) = \frac{\beta_{\ell r}(x_1(\bar{y}; g_n, \eta_1), \ldots, x_R(\bar{y}; g_n, \eta_R))}{\alpha_{\ell r}}$$

(18)

is a rescaled treatment effect, evaluated at $(x_1(\bar{y}; g_n), \ldots, x_R(\bar{y}; g_n), \eta_R)$ for $[\bar{y}; g_n] = (\bar{y}^1, \ldots, \bar{y}^n, g, g+1, \ldots, gN)'$.

**Proof** See Appendix B.3.

This shows that in large samples $\hat{\beta}$ estimates a convex average of rescaled treatment effects, $\beta_{\ell n r}(\bar{y})$, when the first stage is monotone in each shock. Appendix B.3 shows that the weights $\omega_{\ell n r}(\bar{y})$ are proportional to the first-stage effects $\pi_{\ell n r}(\bar{y}; g_n)$, exposure shares $s_{\ell n}$, regression weights $\epsilon_\ell$, treatment aggregation weights $\alpha_{\ell r}$, and a function of the shock distribution. In the case without aggregation, i.e. $R = \alpha_{\ell r} = 1$, there is no rescaling in the $\beta_{\ell n r}(\bar{y})$. Equation (17) then can be seen as generalizing the result of Angrist et al. (2000), on the identification of heterogeneous effects of continuous treatments, to the continuous shift-share instrument case. Intuition for the $\omega_{\ell n r}(\bar{y})$ weights follows similarly from this connection. With aggregation—that is, when the realization of shocks may have heterogeneous effects on $y_\ell$ holding the aggregated $x_\ell$ fixed—equation (17) shows that SSIV captures a convex average of treatment effects per aggregated unit. Thus in the leading example of $y_\ell = \sum_n s_{\ell n} \beta_{\ell n} x_{\ell n} + \epsilon_\ell$ and $x_\ell = \sum_n s_{\ell n} x_{\ell n}$, this result establishes identification of a convex average of the $\beta_{\ell n}$. In this way the result generalizes Adão et al. (2019), who establish the identification of convex averages of rescaled treatment effects in reduced form shift-share regressions.

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

In this appendix we illustrate economic differences between the estimands of two regressions that researchers may consider: SSIV using outcome and treatment observations $y_\ell$ and $x_\ell$ (which we show in Proposition 1 are equivalent to certain shock-level IV regressions), and more conventional shock-level IV regressions using “native” $y_n$ and $x_n$. These outcomes and treatments capture the same economic concepts as the original $y_\ell$ and $x_\ell$, in contrast to the constructed $\bar{y}_n$ and $\bar{x}_n$ discussed in Section 2.3.

In line with the labor supply and other key SSIV examples, we will for concreteness refer to the $\ell$ and $n$ as indexing regions and industries, respectively. We consider the case where both the outcome and treatment can be naturally defined at the level of region-by-industry cells (henceforth, cells)—$y_{\ell n}$ and $x_{\ell n}$, respectively—and thus suitable for aggregation across either dimension with some weights $E_{\ell n}$ (e.g., cell employment growth rates aggregated with lagged cell employment weights): $y_\ell = \sum_n s_{\ell n} y_{\ell n}$ for $s_{\ell n} = \frac{E_{\ell n}}{\sum_r E_{\ell r}}$ and $y_n = \sum_\ell s_{\ell n} y_{\ell n}$ for $\omega_{\ell n} = \frac{E_{\ell n}}{\sum_r E_{\ell r}}$, with analogous expressions for $x_\ell$ and $x_n$.

We further define $E_\ell = \sum_n E_{\ell n}$ and $E_n = \sum_\ell E_{\ell n}$.

This formulation nests reduced-form shift-share regressions when $x_{\ell n} = g_n$ for each $\ell$. The labor supply example of Section 2.1 fits only partially in this formal setup because the industry or regional wage growth $y_n$ is not equal to a
We consider the estimands of two regression specifications: $\beta$ from the regional level model (2), instrumented by $z_\ell$ and weighted by $e_\ell = E_\ell / E$ for $E = \sum_\ell E_\ell$, and $\beta_{\text{ind}}$ from a simpler industry-level IV regression of

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

(19)

instrumented by the industry shock $g_n$ and weighted by $s_n = E_n / E$. For simplicity we do not include any controls in either specification and implicitly condition on $\{E_{\ell n}\}_{\ell, n}$ (and some other variables as described below), viewing them as non-stochastic.\(^{50}\)

We show that $\beta$ and $\beta_{\text{ind}}$ generally differ when there are within-region spillover effects or when treatment effects are heterogeneous. We study these cases in turn, maintaining several assumptions:

(i) a first stage relationship analogous to the one considered in Section 3.1:

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

(20)

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

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

$$y_{tn} = \beta_0 x_{tn} - \beta_1 \sum_{n'} s_{\ell n'} x_{\ell n'} + \varepsilon_{tn},$$

(21)

Here $\beta_0$ captures the direct effect of the shock on the cell outcome, and $\beta_1$ captures a within-region spillover effect. The local employment effects of industry demand shocks from the model in Appendix A.10 fit in this framework, see equation (49).\(^{51}\) The following proposition shows that the SSIV estimand $\beta$ captures the effect of treatment net of spillovers (i.e. $\beta_0 - \beta_1$), whereas $\beta_{\text{ind}}$ subtracts the spillover only partially; this is intuitive since the spillover effect is fully contained within regions but not within industries.

**Proposition A2** Suppose equation (21) holds and the average local concentration index $H_L = \sum_{\ell, n} e_\ell s_{\ell n}^2$ is bounded from below by a constant $\bar{H}_L > 0$. Further assume $\pi_{\ell n} = \bar{\pi}$ and $\text{Var}[g_n] = \sigma_g^2$ for all $\ell$ and $n$. Then the SSIV estimator satisfies

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

(22)

\(^{50}\)Note that we thereby condition on the shares $s_{\ell n}$ and importance weights $e_\ell$. Yet we still allow for share endogeneity by not restricting $\mathbb{E}[\varepsilon_{tn}]$ to be zero.

\(^{51}\)In the labor supply example from the main text $y_{tn}$ is the cell wage, which is equalized within the region, and $x_{\ell n}$ is cell employment. Equation (21) therefore holds for $\beta_0 = 0$ and $-\beta_1$ being the inverse labor supply elasticity.
while the native industry-level IV estimator satisfies

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

(23)

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

**Proof** See Appendix B.4.

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

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

(24)

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

**Proposition A3** In the casual model (24),

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

(25)

and

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

(26)

**Proof** See Appendix B.5.

**A.3 Unobserved $n$-level Shocks Violate Share Exogeneity**

In this appendix, we show that the assumption of SSIV share exogeneity from Goldsmith-Pinkham et al. (2020) is violated when there are unobserved shocks $\nu_n$ that affect outcomes via the exposure shares $s_{\ell n}$, i.e. when the residual has the structure

$$
\varepsilon_{\ell} = \sum_n s_{\ell n} \nu_n + \hat{\varepsilon}_{\ell}. 
$$

(27)
We consider large-sample violations share exogeneity in terms of the asymptotic non-ignorability of the \( \bar{\varepsilon}_n \) terms in the equivalent moment condition (5). It is intuitive that the cross-sectional dependence between \( s_{\ell n} \) and \( \varepsilon_\ell \) will not asymptotically vanish when \( N \) is fixed (as in Goldsmith-Pinkham et al. (2020)) and each \( \nu_n \) shock contributes significantly to the residual, causing \( \varepsilon_n \not\to 0 \) for some or all \( n \). We next prove this result and show that it generalizes to the case of increasing \( N \), where the contribution of each \( \nu_n \) to the variation in \( \varepsilon_\ell \) becomes small. The intuition here is that the SSIV relevance condition generally requires individual observations to be sufficiently concentrated in a small number of shocks (see Section 3.1), and under this condition the share exogeneity violations remain asymptotically non-ignorable even as \( N \to \infty \).

We define share endogeneity as non-vanishing \( \text{Var} [\bar{\varepsilon}_n] \) at least for some \( n \). This will tend to make the SSIV estimator inconsistent, unless shocks are as-good-as-randomly assigned (Assumption 1), even if the importance weights of individual shocks, \( s_n \), converge to zero (Assumption 2). Here we treat \( e_\ell \) and \( s_{\ell n} \) as non-stochastic to show this result with simple notation.

**Proposition A4** Suppose condition (27) holds with the \( \nu_n \) mean-zero and uncorrelated with the \( \bar{\varepsilon}_\ell \) and with each other, and with \( \text{Var} [\nu_n] = \sigma_n^2 \geq \sigma_\nu^2 \) for a fixed \( \sigma_\nu^2 > 0 \). Also assume \( H_L = \sum_\ell e_\ell \sum_n s_{\ell n}^2 \to \bar{H} > 0 \) such that first-stage relevance can be satisfied. Then there exists a constant \( \delta > 0 \) such that \( \max_n \text{Var} [\bar{\varepsilon}_n] > \delta \) for sufficiently large \( L \).

**Proof** See Appendix B.6.

---

**A.4 Connection to Rotemberg Weights**

In this appendix we rewrite the decomposition of the SSIV coefficient \( \hat{\beta} \) from Goldsmith-Pinkham et al. (2020) that gives rise to their “Rotemberg weight” interpretation, and show that these weights measure the leverage of shocks in our equivalent shock-level IV regression. We then show that, in our framework, skewed Rotemberg weights do not measure sensitivity to misspecification (of share exogeneity) and do not pose a problem for SSIV consistency. We finally discuss the implications of high-leverage observations for SSIV inference.

Proposition 1 implies the following decomposition:

\[
\hat{\beta} = \sum_n s_n g_n \frac{y_n}{\bar{x}_n} = \sum_n \alpha_n \hat{\beta}_n, \tag{28}
\]

where

\[
\hat{\beta}_n = \frac{\bar{y}_n}{\bar{x}_n} - \frac{\sum_\ell e_\ell s_{\ell n} \bar{y}_\ell}{\sum_\ell e_\ell s_{\ell n} \bar{x}_\ell} \tag{29}
\]

and

\[
\alpha_n = \frac{s_n g_n \bar{x}_n}{\sum_{n'} s_{n''} g_{n''} \bar{x}_{n''}}. \tag{30}
\]
This is a shock-level version of the decomposition discussed in Goldsmith-Pinkham et al. (2020): \( \hat{\beta}_n \) is the IV estimate of \( \beta \) that uses share \( s_{\ell n} \) as the instrument, and \( \alpha_n \) is the so-called Rotemberg weight.

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

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

(31)

In this way, \( \alpha_n \) measures the sensitivity of \( \hat{\beta} \) to \( \hat{\beta}_n \).

In the preferred interpretation of Goldsmith-Pinkham et al. (2020), exposure to each shock is a valid instrument such that \( \hat{\beta}_n \overset{D}{\to} \beta \) for each \( n \). However, in our framework deviations of \( \hat{\beta}_n \) from \( \beta \) reflect nonzero \( \bar{\varepsilon}_n \) in large samples, and such share endogeneity is not ruled out; thus \( \alpha_n \) does not have the same sensitivity-to-misspecification interpretation. Moreover, a high leverage of certain shocks (“skewed Rotemberg weights,” in the language of Goldsmith-Pinkham et al. (2020)) is not a problem for consistency in our framework, provided it results from a heavy-tailed and high-variance distribution of shocks (that still satisfies our regularity conditions, such as finite shock variance), and each \( s_n \) is small as required by Assumption 2.

Nevertheless, skewed \( \alpha_n \) may cause issues with SSIV inference, as would high leverage observations in any regression. In general, the estimated residuals \( \hat{\varepsilon}_n^\perp \) of high-leverage observations will tend to be biased toward zero, which may lead to underestimation of the residual variance and too small standard errors (e.g., Cameron and Miller 2015). This issue can be addressed, for instance, by computing confidence intervals with the null imposed, as Adão et al. (2019) recommend and as we discuss in Section 5.1. In practice our Monte-Carlo simulations in Appendix A.12 find that the coverage of conventional exposure-robust confidence intervals to be satisfactory even with Rotemberg weights as skewed as those reported in the applications of Goldsmith-Pinkham et al. (2020) analysis.

### A.5 Consistency of Control Coefficients

This appendix shows how the control coefficient \( \gamma \), defined in footnote 5, can be consistently estimated as required in Proposition 3 (Assumption B2). We discuss conditions for \( \sum_{\ell} e_{\ell w_{\ell} \varepsilon_{\ell}} \overset{D}{\to} \mathbb{E} \left[ \sum_{\ell} e_{\ell w_{\ell}} \varepsilon_{\ell} \right] \), where by definition \( \mathbb{E} \left[ \sum_{\ell} e_{\ell w_{\ell} \varepsilon_{\ell}} \right] = 0 \). Consistency of the estimator \( \hat{\gamma} = \gamma + \left( \sum_{\ell} e_{\ell w_{\ell}} w_{\ell}' \right)^{-1} \sum_{\ell} e_{\ell w_{\ell} \varepsilon_{\ell}} \) follows, provided the elements of \( \left( \sum_{\ell} e_{\ell w_{\ell}} w_{\ell}' \right)^{-1} \) are stochastically bounded (i.e., \( O_p(1) \)). For simplicity we consider control vectors \( w_{\ell} \) of fixed length.

The argument for convergence of \( \sum_{\ell} e_{\ell w_{\ell} \varepsilon_{\ell}} \) depends on the source of randomness in \( w_{\ell} \) and \( \varepsilon_{\ell} \). We consider two characteristic cases. In the first case, \( \left( e_{\ell}, w_{\ell}', \varepsilon_{\ell} \right)' \) can be viewed as \( iid \) or clustered in a conventional way. For example, \( w_{\ell} \) and \( \varepsilon_{\ell} \) may contain observed and unobserved local labor supply shocks which are uncorrelated across markets, clusters of markets (e.g. states), or beyond a given distance threshold. In this case conventional laws of large numbers can be used to establish
\[ \sum_{t} \epsilon_{t} w_{t} \varepsilon_{t} \xrightarrow{p} 0. \] For instance if \((\epsilon_{t}, w_{t}', \varepsilon_{t})'\) is iid then \(\sum_{t} \epsilon_{t} w_{t} \varepsilon_{t}\) gives a vector of sample averages of mutually uncorrelated mean-zero random variables, which weakly converge to zero when the \(\epsilon_{t}\) weights are asymptotically dispersed \((\mathbb{E} \left[ \sum_{t} \epsilon_{t}^{2} \right] \rightarrow 0)\) and when \(\mathbb{E} \left[ w_{t}^{2} \varepsilon_{t}^{2} \mid \epsilon \right]\) is uniformly bounded.

In the second case, either \(w_{t}\) or \(\varepsilon_{t}\) has a shift-share structure like \(z_{t}^{*}\): i.e. \(w_{t} = \sum_{n} s_{tn} q_{n}\) for an observed \(q_{n}\) (in line with our Proposition 3) or \(\varepsilon_{t} = \sum_{n} s_{tn} \nu_{n}\) for an unobserved \(\nu_{n}\) (capturing, for example, a set of unobserved industry-level factors averaged with the employment weights \(s_{tn}\)). In this case convergence of \(\sum_{t} \epsilon_{t} w_{t} \varepsilon_{t}\) can be shown to follow similarly to the convergence of the sample analog of the instrument moment condition (3). If, for instance, \(\varepsilon_{t} = \sum_{n} s_{tn} \nu_{n}\) with \(\mathbb{E} [\nu_{n} \mid s, w] = 0\) and \(\text{Cov} [\nu_{n}, \nu_{m} \mid s, w] = 0\) for \(w = \{w_{n}\}_{n}\), then for each control \(\sum_{t} \epsilon_{t} w_{tm} \varepsilon_{t} = \sum_{n} s_{tn} \nu_{n} \bar{w}_{nm}\) weakly converges when the \(s_{n}\) weights are dispersed \((\mathbb{E} \left[ \sum_{n} s_{n}^{2} \right] \rightarrow 0)\) and both \(\text{Var} [\nu_{n} \mid s, w]\) and \(\mathbb{E} \left[ \bar{w}_{nm}^{2} \mid s \right]\) are uniformly bounded. This argument can be extended to the case where either \(w_{t}\) or \(\varepsilon_{t}\) is formed from different exposure shares \(\hat{s}_{tk}\), perhaps defined over a different range of \(K\) observed \(q_{k}\) or unobserved \(\nu_{k}\), and when the \(q_{k}\) or \(\nu_{k}\) are clustered or otherwise weakly mutually correlated.

More generally, the two cases can be combined to settings where \(w_{t} = \sum_{k} s_{tk} q_{k} + \hat{w}_{t}\) and \(\varepsilon_{t} = \sum_{k} s_{tk} \nu_{k} + \hat{\varepsilon}_{t}\) where \((\epsilon_{t}, \hat{w}_{t}', \hat{\varepsilon}_{t})'\) is iid or conventionally clustered and where \(q_{k}\) and \(\nu_{k}\) are many weakly correlated random shocks or, even more generally, allowing for multiple shift-share terms with different exposure shares.

### A.6 SSIV Consistency in Short Panels

This appendix shows how alternative shock exogeneity assumptions imply the consistency of panel SSIV regressions with many fixed effect coefficients. We consider the incidental parameters problem in “short” panels, with fixed \(T\) and \(L \rightarrow \infty\) and with unit fixed effects, in which case the control coefficient \(\gamma\) cannot be consistently estimated with the fixed effects included in \(w_{t}\). We show how an analog of Assumption 3 can be instead applied to a demeaned shock-level unobservable thatpartials out the fixed effect nuisance coefficients. A similar argument applies to period fixed effects in the fixed \(L\) and \(T \rightarrow \infty\) asymptotic.

Suppose for the linear causal model \(y_{it} = \beta x_{it} + \epsilon_{it}\) and control vector \(w_{it}\) (which includes unit FE), we define \(\gamma = \mathbb{E} \left[ \sum_{t} e_{it} w_{it}^{\Delta} \varepsilon_{it}^{\Delta}' \right]^{-1} \mathbb{E} \left[ \sum_{t} e_{it} w_{it}^{\Delta} e_{it}^{\Delta} \right]\) where \(\varepsilon_{it}^{\Delta}\) is a subvector of the (weighted) unit-demeaned observation of variable \(v_{it}\), \(v_{it} - \sum_{t'} e_{it} v_{it'} / \sum_{t'} e_{it'}\), that drops any elements that are identically zero (e.g. those corresponding to the unit FE in \(w_{it}\)). Note we have assumed no perfect multicollinearity in the remaining elements such that \(\mathbb{E} \left[ \sum_{t} e_{it} w_{it}^{\Delta} \varepsilon_{it}^{\Delta}' \right]\) is invertible. We can then write \(y_{it}^{\Delta} = \beta x_{it}^{\Delta} + w_{it}^{\Delta'} \gamma + e_{it}^{\Delta}\). Suppose also that \(\sum_{t} \epsilon_{it} x_{it}^{\Delta} \xrightarrow{p} \gamma \) for some \(\pi \neq 0\) and the analog of Assumption B2 for unit-demeaned controls holds. Then, following the proof to Proposition 3, \(\hat{\beta}\) is
consistent if and only if

$$\sum_{n,t} s_{nt} g_{nt} \varepsilon_{nt}^\Delta \rightarrow 0,$$

(32)

where

$$s_{nt} = \sum_l e_{lt} s_{lnt}$$

and

$$\bar{\varepsilon}_{nt} = \sum_l e_{lt} s_{lnt} \varepsilon_{lt}.$$

This condition is satisfied when analogs of Assumptions 1, 2, and B1 hold, or under the various extensions discussed in Section 3. In particular when \(w_{lt}\) contains \(t\)-specific FE the key assumption of quasi-experimental shock assignment is \(E[g_{nt} | \bar{e}_t, s] = \mu_t\), for all \(n\) and \(t\), allowing endogenous period-specific shock means \(\mu_t\) via Proposition 4. This assumption avoids the incidental parameters problem by considering shocks as-good-as-randomly assigned given the set of unobserved \(\bar{\varepsilon}_{nt}\), each of which is a function of the time-varying \(\varepsilon_{lt}\) across all periods \(p\).

An intuitive special case is when the exposure shares and importance weights are time-invariant:

$$s_{nt} = s_{n0}$$

and

$$e_{lt} = e_{l0}.$$

Then the weights in (32) are also time-invariant, \(s_{nt} = s_{n0}\), and

$$\bar{\varepsilon}_{nt} = \frac{\sum_l e_{lt} s_{lnt} \varepsilon_{lt}}{\sum_l e_{lt} s_{lnt}} = \frac{\sum_l e_{lt} s_{lnt} \varepsilon_{lt} - \frac{1}{T} \sum_s \varepsilon_{lt}}{\sum_l e_{lt} s_{lnt}} = \bar{\varepsilon}_{nt} - \frac{1}{T} \sum_{l'} \bar{\varepsilon}_{nl'},$$

(33)

where \(\bar{\varepsilon}_{nt} = \frac{\sum_l e_{lt} s_{lnt} \varepsilon_{lt}}{\sum_l e_{lt} s_{lnt}}\) is an aggregate of period-specific unobservables \(\varepsilon_{lt}\). It is then straightforward to extend Propositions 3 and 4 under a shock-level assumption of strong exogeneity, i.e. that \(E[g_{nt} | \bar{e}, s] = \mu_n + \tau_t\) for all \(n\) and \(t\). Here endogenous \(n\)-specific shock means are permitted by the observation in Section 4.1, that share-weighted \(n\)-specific FEAs at the shock level are subsumed by \(t\)-specific FEAs in the SSIV regression when shares and weights are time-invariant.

A.7 Sequential Exogeneity in Long Panels

This appendix shows how the strict exogeneity condition of Assumption 3 can be relaxed to a sequential exogeneity assumption in long panels (with \(T \rightarrow \infty\) and \(L\) fixed). In contrast to Appendix A.6 we assume the unit FEAs can be consistently estimated and included in \(g\).\(^{52}\)

To see this extension simply, we consider a scenario where \(q_n = 1\) (such that Assumption 3 coincides with Assumption 1) and \(\mu = 0\). Rather than imposing \(E[g_{nt} | \bar{e}, s] = 0\), we make a weaker assumption \(E[g_{nt} | I_{nt}] = 0\) for \(I_{nt} = \{\{\bar{\varepsilon}_{nt'}, s_{nt'}\}_{t' \geq t}, \{\bar{\varepsilon}_{nt'}, s_{nt'}\}_{n' \neq n, t'}\}\). This assumption captures the possibility of a sequential shock generating process in which draws of \(g_{nt}\) do not on average depend on present or future realizations of the unit’s unobservables, or the realizations of unobservables at other

\(^{52}\)Arguments similar to those in Appendix A.5 imply consistency of the fixed effects. Some sequences with both \(L\) and \(T\) growing can also be accommodated as long as the fixed effects are estimated sufficiently precisely, such that Assumption B2 holds.
units, but may depend on past realizations $\varepsilon_{nt'}$ for $t' < t$. Similarly, we suppose $\text{Cov} \left[ g_{nt}, g_{nt'} \mid I_{nt} \right] = 0$ for $t' \geq t$ and $\text{Cov} \left[ g_{nt}, g_{nt'} \mid I_{nt} \right] = 0$ for $n' \neq n$ and all $t'$. Following the proof to Proposition 3, we continue to have $\mathbb{E} \left[ \sum_{n,t} s_{nt} g_{nt} \varepsilon_{nt} \right] = 0$. Furthermore,

$$\text{Var} \left[ \sum_{n,t} s_{nt} g_{nt} \varepsilon_{nt} \right] = \mathbb{E} \left[ \left( \sum_{n,t} s_{nt} g_{nt} \varepsilon_{nt} \right)^2 \right] = \sum_{n,t} \sum_{n',t'} \mathbb{E} \left[ s_{nt} s_{n't'} g_{nt} g_{n't'} \varepsilon_{nt} \varepsilon_{n't'} \right]
\begin{split}
&= \sum_{n,t} \mathbb{E} \left[ s_{nt}^2 g_{nt}^2 \varepsilon_{nt}^2 \right] + \sum_{n,t} \sum_{t'>t} \mathbb{E} \left[ s_{nt} s_{nt'} g_{nt} g_{nt'} \varepsilon_{nt} \varepsilon_{nt'} \right] \\
&\quad + \sum_{n,t} \sum_{t' \geq t} \mathbb{E} \left[ s_{nt} s_{nt'} \varepsilon_{nt} \mathbb{E} \left[ g_{nt} g_{nt'} \mid I_{nt} \right] \varepsilon_{nt} \varepsilon_{nt'} \right] \\
&\quad + \sum_{n,t} \sum_{n' \neq n, t'} \mathbb{E} \left[ s_{nt} s_{n't'} \varepsilon_{nt} \mathbb{E} \left[ g_{nt} g_{n't'} \mid I_{nt} \right] \varepsilon_{nt} \varepsilon_{n't'} \right] \\
&= \sum_{n,t} \mathbb{E} \left[ s_{nt}^2 g_{nt}^2 \varepsilon_{nt}^2 \right] + \sum_{n,t} \sum_{t' < t} \mathbb{E} \left[ s_{nt} s_{nt'} g_{nt} g_{nt'} \varepsilon_{nt} \varepsilon_{nt'} \right]. \quad (34)
\end{split}
$$

The first term of this expression is $o(1)$ when $\mathbb{E} \left[ \sum_{n,t} s_{nt}^2 \right] \to 0$, as in Assumption 2, and when $\mathbb{E} \left[ g_{nt}^2 \mid \varepsilon_{nt}, s_{nt} \right]$ and $\mathbb{E} \left[ \varepsilon_{nt}^2 \mid s_{nt} \right]$ are uniformly bounded. The weak convergence of $\sum_{n,t} s_{nt} g_{nt} \varepsilon_{nt}$ thus depends on whether the second term of equation (34) is also $o(1)$. One way this convergence would hold is when $N$ is fixed (with $T \to \infty$) and $|\mathbb{E} \left[ g_{nt} g_{nt'} \mid \varepsilon, s \right]| \leq f \left( t - t' \right)$ for all $n$ and $t' < t$ given a fixed function $f(\cdot)$ satisfying $\sum_{r=1}^{\infty} f(r) < \infty$. This approach can be seen to extend Assumption 6 to relax strict exogeneity in the sense of $\mathbb{E} \left[ g_{nt} \mid \varepsilon, s \right] \neq 0$. Another way this convergence would hold is when $N \to \infty$, units are sufficiently dispersed ($\mathbb{E} \left[ \sum_{n} \left( \sum_{t} s_{nt} \right)^2 \right] \to 0$) and $\mathbb{E} \left[ g_{nt} g_{nt'} \mid \varepsilon, s \right]$ is uniformly bounded. This approach can be seen to similarly relax Assumption 1 by imposing Assumption 5, with clusters given by the units observed over time.

### A.8 SSIV Relevance with Panel Data

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

$$\mathbb{E} \left[ \sum_{\ell} \sum_{t} e_{lt} z_{lt} x_{lt} \right] \geq \bar{\sigma}_g^2 \bar{\pi} \sum_{\ell} \sum_{t} e_{lt} \sum_{n} s_{nt \ell} s_{nt \ell 0}. \quad (35)$$

53
For panel SSIV relevance we require the $e_{lt}$-weighted average of $\sum_n s_{tn}s_{tn}$ to not vanish asymptotically. With updated shares this is satisfied when the Herfindahl index of an average observation-period (across shocks) is non-vanishing, while in the fixed shares case the overlap of shares in periods 0 and $t$, $\sum_n s_{tn_0}s_{tn}$, may become weak or even vanish as $T \to \infty$, on average across observations.

A.9 Estimated Shocks

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

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

$$g_n = \sum_{l} \omega_{ln}g_{ln}. \quad (36)$$

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

$$g_{n,-l} = \frac{\sum_{l' \neq l} \omega_{ln}'g_{ln}'}{\sum_{l' \neq l} \omega_{ln}}. \quad (37)$$

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

$$g_{ln} = g^*_n + \psi_{ln}, \quad (38)$$

where $g^*_n$ satisfies Assumptions 1 and 2 and $\psi_{ln}$ is estimation error (in Section 4.3 we considered the special case of $\psi_{ln} \propto \epsilon_l$). This implies a feasible shift-share instrument of $z_l = z^*_l + \psi_l$ and its LOO version $z^*_l = z^*_l + \psi^*_l$, where $z^*_l = \sum_n s_{tn}g^*_n$, $\psi_l = \sum_n s_{tn} \sum_{l'} \omega_{ln}'z_{ln}'$, and $\psi^* = \sum_n s_{tn} \sum_{l' \neq l} \omega_{ln}'z_{ln}'$. Consistency with these instruments, given a first stage, requires that $\sum_{l} e_{lt}\epsilon_{lt}\psi_{lt} \sim 0$ and $\sum_{l} e_{lt}\epsilon_{lt}\psi^*_{lt} \sim 0$ respectively.

We now present three sets of results. First, we establish a simple sufficient condition under which the LOO instrument satisfies $\sum_{l} e_{lt}\epsilon_{lt}\psi^*_{lt} \sim 0$. We also propose stronger conditions that guarantee consistency of LOO-SSIV. Second, we explore the conditions under which the covariance between $\epsilon_{lt}$ and $\psi_{lt}$ is ignorable, i.e. asymptotically does not lead to a “mechanical” bias of the conventional non-leave-one-out estimator. We propose a heuristic measure that is large when the bias is likely to be small. Lastly, we apply these ideas to the setting of Bartik (1991) using the data from Goldsmith-Pinkham et al. (2020). In line with previous appendices, we condition on $s_{tn}$, $\omega_{ln}$, and $e_{lt}$ and treat
them as non-stochastic for notational convenience. We also assume the SSIV regressions are estimated without controls $w_t$.

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

**Proposition A5**

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

2. If $E\left[\left(\varepsilon_\ell, \psi_{\ell,n}\right) \mid \left(\varepsilon_{\ell'}, \psi_{\ell',n'}\right)_{\ell' \neq \ell, n'}\right] = 0$ for all $\ell$ and $n$, then the LOO estimator is consistent, provided it has a first stage and two regularity conditions hold: $E\left[\sum_\ell e_\ell \varepsilon_\ell \psi_{\ell,n_1} \psi_{\ell,n_2}\right] \leq B$ for a constant $B$ and all $(\ell_1, \ell_2, \ell'_1, \ell'_2, n_1, n_2)$ and

$$
\sum_{(\ell_1, \ell_2, \ell'_1, \ell'_2) \in J} e_{\ell_1} e_{\ell_2} s_{\ell_1,n_1} s_{\ell_2,n_2} \frac{\omega_{\ell_1,n_1}}{\sum_{\ell \neq \ell_1} \omega_{\ell,n_1}} \frac{\omega_{\ell_2,n_2}}{\sum_{\ell \neq \ell_2} \omega_{\ell,n_2}} \to 0, \tag{39}
$$

with $J$ denoting the set of tuples $(\ell_1, \ell_2, \ell'_1, \ell'_2)$ for which one of the two conditions hold: (i) $\ell_1 = \ell_2$ and $\ell'_1 = \ell'_2 \neq \ell_1$, (ii) $\ell_1 = \ell'_2$ and $\ell_2 = \ell'_1 \neq \ell_1$.

3. Condition (39) is satisfied if $\frac{N}{n} \to 0$ in the special case where each region is specialized in one industry, i.e. $s_{tn} = 1[n = n(\ell)]$ for some $n(\cdot)$, there are no importance weights ($e_\ell = \frac{1}{L}$), and shocks estimated by simple LOO averaging among observations exposed to a given shock $(\omega_{tn} = \frac{1}{L_n}$ for $L_n = \sum_\ell 1[n(\ell) = n]$), assuming further that $L_n \geq 2$ for each $n$ so that the LOO estimator is well-defined.

**Proof** See Appendix B.7.

The condition in the first part of Proposition A5 would be quite innocuous in random samples of $\ell$ – the environment in which leave-one-out adjustments are often considered (e.g. Angrist et al. (1999)) – but is strong without random sampling. It requires $\varepsilon_\ell$ and $\psi_{\ell,n}$ to be uncorrelated for $\ell' \neq \ell$, which may easily be violated when both $\ell$ and $\ell'$ are exposed to the same shocks—a situation in which excluding own observation is not sufficient. Moreover, since we have conditioned on the exposure shares throughout, $E[\varepsilon_\ell \psi_{\ell,n}] = 0$ generally requires either $\varepsilon_\ell$ or $\psi_{\ell,n}$ to have a zero conditional mean—the share exogeneity assumption applied to either the residuals or the estimation error. At the same time, this condition does not require $E[\varepsilon_\ell \psi_{\ell,n}] = 0$ for $\ell = \ell'$, which reflects the benefit of LOO: eliminating the mechanical bias from the residual directly entering shock estimates.
Heuristic for Importance of LOO Correction  We now return to the non-LOO SSIV estimator. As in Proposition A5, we assume that \( \mathbb{E} [\varepsilon_{t} \psi_{tn}] = 0 \) for \( t' \neq t \) and all \( n \), so the LOO estimator is consistent under the additional regularity conditions. We also assume, without loss of generality, that \( z_t \) is mean-zero. Then the “mechanical bias” mentioned in Section 4.3 is the only potential problem: under appropriate regularity conditions (similar to those in part 2 of Proposition A5),

\[
\hat{\beta} - \beta = \frac{\mathbb{E} \left[ \sum_{t} e_{t} \varepsilon_{t} \psi_{tn} \right]}{\mathbb{E} \left[ \sum_{t} e_{t} \varepsilon_{t} \bar{x}_{t} \right]} + o_{p}(1) \\
= \frac{\sum_{t,n} e_{t} s_{tn} \omega_{tn} \mathbb{E} [\varepsilon_{t} \psi_{tn}]}{\mathbb{E} \left[ \sum_{t} e_{t} \varepsilon_{t} \bar{x}_{t} \right]} + o_{p}(1). \tag{40}
\]

With \( |\mathbb{E} [\varepsilon_{t} \psi_{tn}]| \) bounded by some \( B_1 > 0 \) for all \( t \) and \( n \), the numerator of (40) is bounded by \( H_N B_1 \), for an observable composite of the relevant shares \( H_N = \sum_{t,n} e_{t} s_{tn} \omega_{tn} \). The structure of the shares also influences the strength of the first stage in the denominator. Imposing our standard model of the first stage from Section 3.1 (but specified based on the latent shock \( g_{n}^{*} \)), i.e. \( x_{tn} = \pi_{tn} g_{n}^{*} + \eta_{tn} \), \( \eta_{tn} \) mean-zero and uncorrelated with \( g_{n}^{*} \) for all \( t, n, n' \), \( \text{Var}[g_{n}^{*}] = \sigma_{g}^{2} > 0 \) and \( \pi_{tn} \geq \pi > 0 \), yields:

\[
\mathbb{E} \left[ \sum_{t} e_{t} z_{t} \bar{x}_{t} \right] = \sum_{t} e_{t} \mathbb{E} \left[ \left( \sum_{n} s_{tn} \left( \pi_{tn} g_{n}^{*} + \psi_{tn} \right) \right) \left( \sum_{n'} s_{t'n'} \left( \pi_{tn'} g_{n'}^{*} + \eta_{tn'} \right) \right) \right] \\
= \sum_{t,n} e_{t} s_{tn}^{2} \pi_{tn} \text{Var}[g_{n}^{*}] + \sum_{t} e_{t} \sum_{n,n'} s_{tn} s_{tn'} \mathbb{E} [\psi_{tn} (\pi_{tn} g_{n'}^{*} + \eta_{tn'})]. \tag{41}
\]

Excepting knife-edge cases where the two terms in (41) cancel out, \( \mathbb{E} \left[ \sum_{t} e_{t} z_{t}^{2} \bar{x}_{t} \right] \neq 0 \) provided \( H_L = \sum_{t,n} e_{t} s_{tn}^{2} \geq \bar{H} \) for some fixed \( \bar{H} > 0 \).

We thus define the following heuristic:

\[
H = \frac{H_L}{H_N} = \frac{\sum_{t,n} e_{t} s_{tn}^{2}}{\sum_{t,n} e_{t} s_{tn} \omega_{tn}}. \tag{42}
\]

When \( H \) is large, we expect the non-LOO SSIV estimator to be relatively insensitive to the mechanical bias generated by the average covariance between \( \psi_{tn} \) and \( \varepsilon_{t} \), and thus similar to the LOO estimator.

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

\[
H_N = \sum_{t,n} \frac{E_{t}}{E} \frac{E_{tn}}{E_{tn}} \frac{E_{tn}}{E_{tn}} = \sum_{t,n} \frac{E_{n}}{E} \left( \frac{E_{tn}}{E_{tn}} \right)^{2} = \sum_{n} s_{n} \sum_{t} \omega_{tn}^{2}, \tag{43}
\]

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

The formula simplifies further with \( E_{\ell n} = 1 \) if \( n = n(\ell) \) for all \( \ell, n \), corresponding to the case of complete specialization of observations in shocks with no regression or shock estimation weights, as in part 3 of Proposition A5. In that case,

\[
H = \frac{1}{\sum_{\ell} \frac{1}{L_n(\ell)}} = \frac{1}{L} \sum_{n} \sum_{\ell: n(\ell)=n} \frac{1}{L_n} = \frac{L}{N}. \tag{44}
\]

Our heuristic is therefore large when there are many observations per estimated shock.53

**Application to Bartik (1991)** We finally apply our insights to the Bartik (1991) setting, using the Goldsmith-Pinkham et al. (2020) replication code and data. Table C6 reports the results. Column 1 shows the estimates of the inverse local labor supply elasticity using SSIV estimators with and without the LOO correction and using population weights, replicating Table 3, column 2, of Goldsmith-Pinkham et al. (2020) except with employment on the left-hand side and wages on the right-hand side. Column 2 repeats the analysis without the population weights.54 We find all estimates to range between 1.2 and 1.3, showing that in practice for Bartik (1991) the LOO correction does not play a substantial role.

This is however especially true without weights, where the LOO and conventional SSIV estimators are 1.30 and 1.29, respectively. Our heuristic provides an explanation: \( H \) is almost 8 times bigger when computed without weights. The intuition is that large commuting zones, such as Los Angeles and New York, may constitute a substantial fraction of employment in industries of their comparative advantage. This generates a potential for the mechanical bias: labor supply shocks in those regions affect shock estimates; this bias is avoided by LOO estimators. However, the role of the largest commuting zones is only significant in weighted regressions (by employment or, as in Goldsmith-Pinkham et al. (2020), population).

**A.10 Equilibrium Industry Growth in a Model of Local Labor Markets**

This appendix develops a simple model of regional labor supply and demand, similar to the model in Adão et al. (2020). Our goal is to show how the national growth rate of industry employment can be viewed as a noisy version of the national industry-specific labor demand shocks, and how regional labor supply shocks (along with some other terms) generate the “estimation error.”

53Here \( 1/H = N/L \) is proportional to the “bias” of the non-LOO estimator, which is similar to how the finite-sample bias of conventional 2SLS is proportional to the number of instruments over the sample size (Nagar 1959).

54Goldsmith-Pinkham et al. (2020) estimate the inverse labor supply elasticity. By properties of IV estimation, our coefficient is the inverse of theirs.

55Industry growth shocks in this column are the same as in Column 1, again estimated with employment weights.
Consider an economy that consists of a set of $L$ regions. In each region $\ell$ there is a prevailing wage $W_\ell$, and labor supply has constant elasticity $\phi$:

$$E_\ell = M_\ell W_\ell^\phi,$$

(45)

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

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

(46)

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

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

(47)

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

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

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

(48)

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

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

(49)

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

**Proof** See Appendix B.8.

The first term in (49) justifies our interpretation of the observed industry employment growth as a noisy estimate of the latent labor demand shock $g^*_n$. The other terms constitute the “estimation error.” The first of them is proportional to the residual of the labor supply equation, $\varepsilon_\ell$; we have previously established the conditions under which it may or may not confound SSIV estimation. The other terms, that we abstracted away from in Section 4.3, include the idiosyncratic demand shock $\hat{\xi}_{\ell n}$.
and shift-share averages of both national and idiosyncratic demand shocks. If the model is correct, all of these are uncorrelated with \( \varepsilon_{\ell} \), thus not affecting Assumption 1.

A.11 SSIV with Multiple Endogenous Variables or Instruments

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

Generalized Equivalence and SSIV Consistency

We consider a class of SSIV estimators of an outcome model with multiple treatment channels,

\[
y_{\ell} = \beta'x_{\ell} + \gamma'w_{\ell} + \varepsilon_{\ell},
\]

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

\[
\hat{\beta} = (cz'e_{\ell}x_{\ell})^{-1}cz'e_{\ell}y_{\ell},
\]

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

As in Proposition 1, \( \hat{\beta} \) can be equivalently obtained by a particular shock-level IV regression. Intuitively, when the shares are the same, \( cz_{\ell} \) also has a shift-share structure based on a linear combination of shocks \( cg_{jn} \), and thus Proposition 1 extends. Formally, write \( z = sg \) where \( s \) is an \( L \times N \) matrix of exposure shares and \( g \) stacks observations of the shock vector \( g_{jn}' \); then,

\[
\hat{\beta} = (cg's'x_{\ell})^{-1}cg's'ey_{\ell}
= (cg'Sx_{\ell})^{-1}(cg'Sy_{\ell}),
\]

59
where $S$ is an $N \times N$ diagonal matrix with elements $s_n$, $\bar{x}$ is an $N \times K$ matrix with elements $\bar{x}_{kn}$, and $\bar{y}$ is an $N \times 1$ vector of $\bar{y}_n$. This is the formula for an $s_n$-weighted IV regression of $\bar{y}_n$ on $\bar{x}_{1n}, \ldots, \bar{x}_{Kn}$ with shocks as instruments, no constant, and the same $c$ matrix. Furthermore, as in Proposition 1,

$$\iota' S \bar{y} = \sum_n s_n \bar{y}_n = \sum_n \varepsilon \left( \sum_n s_{\ell n} \right) y_{\ell} = \sum_n \varepsilon \bar{y}_{\ell} = 0,$$

(53)

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

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

$$\sum_n s_{\ell n} g_{kn} \bar{\varepsilon}_n \overset{p}{\to} 0$$

for each $k$. Then under the relevance condition from Proposition 2, again appropriately generalized, the SSIV estimates are consistent: $\hat{\beta} \overset{p}{\to} \beta$.

**Effective First-Stage $F$-statistics** With one endogenous variable and multiple instruments, the Montiel Olea and Pflueger (2013) effective first-stage $F$-statistic provides a state-of-art heuristic for detecting a weak first-stage. Here we describe a correction to it for SSIV that generalizes the $F$-statistic in the single instrument case discussed in Section 5.2. The Stata command `weakssivtest`, provided with our replication archive, implements this correction.\(^{56}\)

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

$$x_{\ell} = \pi' z_{\ell} + \rho w_{\ell} + \eta_{\ell}. \quad (54)$$

Suppose each of the shocks satisfies Assumption 3, i.e. $E [g_{jn} \mid \bar{\varepsilon}, q, s] = \mu_j' q_n$, where $\sum_n s_{\ell n} q_n$ is

\(^{56}\)Our package extends the `weakivtest` command developed by Pflueger and Wang (2015).
included in \( w_\ell \), and the residual shocks \( g^*_n = g_n - \mu'_q g_q \) are independent from \( \{ \eta_\ell \} _\ell \). The Montiel Ollea and Pflueger (2013) effective \( F \)-statistic for the 2SLS regression of \( y_\ell \) on \( x_\ell \), instrumenting with \( z_{1\ell}, \ldots, z_{J\ell} \), controlling for \( w_\ell \), and weighting by \( e_\ell \), is given by

\[
F_{\text{eff}} = \left( \sum_\ell e_\ell x_\ell \perp z_\ell \perp \right)' \left( \sum_\ell e_\ell x_\ell \perp z_\ell \perp \right) \frac{\text{tr}(\hat{V})}{\text{tr}(\hat{V}^2)} \tag{55}
\]

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

\[
\sum_\ell e_\ell x_\ell \perp z_\ell \perp = \sum_\ell e_\ell x_\ell \perp z_\ell \perp = \sum_n s_n g_n \bar{x}_n \perp = \sum_n s_n g_n \perp \bar{x}_n \perp \tag{56}
\]

where \( g_n \perp \) is the residuals from an \( s_n \)-weighted projection of \( g_n \) on \( q_n \), which consistently estimates \( g^*_n \). A natural extension of Proposition 5 to many mutually-uncorrelated shocks further implies that \( V \) is well-approximated by

\[
\hat{V} = \sum_n s_n^2 g_n \perp g_n \perp \bar{\eta}_n^2 \tag{57}
\]

where, per the discussion in Section 5.2, \( \bar{\eta}_n \) denotes the residuals from an IV regression of \( \bar{x}_n \perp \) on \( \bar{z}^1_{1n}, \ldots, \bar{z}^J_{Jn} \), instrumented with \( g_{1n}, \ldots, g_{Jn} \), weighted by \( s_n \) and controlling for \( q_n \). Plugging this \( \hat{V} \) into (55) yields the corrected effective first-stage \( F \)-statistic.

### Efficient Shift-Share GMM

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

We first note that the equivalence result (52) applies to SSIV-GMM estimators as well:

\[
\hat{\beta} = \arg \min_b (y \perp - x \perp b)' e z W z' e (y \perp - x \perp b) = \arg \min_b (\bar{y} \perp - \bar{x} \perp b)' S g W g' S (\bar{y} \perp - \bar{x} \perp b) \tag{58}
\]

where \( W \) is an \( J \times J \) moment-weighting matrix. This leads to an IV estimator with \( c = \bar{x} \perp S g W \). For 2SLS estimation, for example, \( W = (z \perp e z \perp)^{-1} \). Under appropriate regularity conditions, the efficient choice of \( W^* \) consistently estimates the inverse asymptotic variance of \( z' e (y \perp - x \perp \hat{\beta}) = g' S \bar{\varepsilon} + o_p(1) \). Generalizations of results in Adão et al. (2019) can then be used to characterize this \( W^* \) when shocks are as-good-as-randomly assigned with respect to \( \bar{\varepsilon} \). Given an estimate \( \hat{W}^* \), an
efficient coefficient estimate $\hat{\beta}^*$ is given by shock-level IV regressions (52) that set $c^* = \bar{x}^\perp S \hat{W}^*$. A $\chi^2_{J-K}$ test statistic based on the minimized objective in (58) can be used for specification testing.

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

$$\text{Var} \left[ g' S (\bar{y}^\perp - \bar{x}^\perp \beta) \right] = \mathbb{E} [\bar{\epsilon}' S gg' S \bar{\epsilon}] = \text{tr} (\mathbb{E} [\bar{\epsilon}' SG S \bar{\epsilon}]) = kG$$

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

$$\hat{\beta}^* = (\bar{x}^\perp S \hat{g} \hat{G}^{-1} g' S \bar{x}^\perp)^{-1} (\bar{x}^\perp S \hat{g} \hat{G}^{-1} g' S \bar{y}^\perp) = \left( (P_g \bar{x}^\perp)' S \bar{x}^\perp \right)^{-1} (P_g \bar{x}^\perp)' S \bar{y}^\perp,$$

where $P_g = g(g' S g)^{-1} g' S$ is an $s_n$-weighted shock projection matrix. This is the formula for an $s_n$-weighted IV regression of $\bar{g}^*_n$ on the fitted values from projecting the $\bar{x}^\perp_{kn}$ on the shocks, corresponding to the 2SLS regression above. Straightforward extensions of this equivalence between optimally-weighted estimates of $\beta$ and shock-level overidentified IV procedures follow in the case of heteroskedastic or clustered shocks, in which case the 2SLS estimator (60) is replaced by the estimator of White (1982). We emphasize that these shock-level estimators are generally different than 2SLS or White (1982) estimators at the level of original observations, which are optimal under conditional homoskedasticity and independence assumptions placed on the residual $\epsilon_\ell$ (assumptions which are generally violated in our quasi-experimental framework).

**Many Shocks in Autor et al. (2013)** Appendix Table C5 illustrates different shock-level overidentified IV estimators in the setting of Autor et al. (2013), introduced in Section 6.2.1. ADH construct their shift-share instrument based on the growth of Chinese imports in eight economies comparable to the U.S., together. We separate them to produce eight sets of industry shocks $g_{jn}$, $j = 1, \ldots, 8$, each reflecting the growth of Chinese imports in one of those countries. As in Section 6.2, the outcome of interest is a commuting zone’s growth in total manufacturing employment with the single treatment variable measuring a commuting zone’s local exposure to the growth of imports from China (see footnote 38 for precise variable definitions). The vector of controls coincides with that of column 3 of
Table 4, isolating within-period variation in manufacturing shocks. Per Section 5.1, exposure-robust standard errors are obtained by controlling for period main effects in the shock-level IV procedures, and we report corrected first stage $F$-statistics constructed as detailed above.

Column 1 reports estimates of the ADH coefficient $\beta$ using the industry-level two-stage least squares procedure (60). At -0.238, this estimate is very similar to the just-identified estimate in column 3 of Table 4. Column 2 shows that we also obtain a very similar coefficient of -0.247 with an industry-level limited information maximum likelihood (LIML) estimator. Finally, in column 3 we report a two-step optimal IV estimate of $\beta$ using an industry-level implementation of the White (1982) estimator. Both the coefficient and standard error fall somewhat, with the latter consistent with the theoretical improvement in efficiency relative to columns 1 and 2. From this efficient estimate we obtain an omnibus overidentification test statistic of 10.92, distributed as chi-squared with seven degrees of freedom under the null of correct specification. This yields a $p$-value for the test of joint orthogonality of all eight ADH shocks of 0.142. Table C5 also reports the corrected effective first-stage $F$-statistic which measures the strength of the relationship between the endogenous variable and the eight shift-share instruments across regions. At 15.10 it is substantially lower than with one instrument in column 3 of Table 4 but still above the conventional heuristic threshold of 10.

A.12 Finite-Sample Performance of SSIV: Monte-Carlo Evidence

In this appendix we study the finite-sample performance of the SSIV estimator via Monte-Carlo simulation. We base this simulation on the data of Autor et al. (2013), as described in Section 6.2. For comparison, we also simulate more conventional shock-level IV estimators, similar to those used in Acemoglu et al. (2016), which also estimate the effects of import competition with China on U.S. employment. We begin by describing the design of these simulations and the benchmark Monte-Carlo results. We then explore how the simulation results change with various deviations from the benchmark: with different levels of industry concentration, different numbers of industries and regions, and with many shock instruments. Besides showing the general robustness of our framework, these extensions allow us to see how informative some conventional rules of thumb are on the finite-sample performance of shift-share estimators.

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

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

\[ y_{lt} = \hat{\beta}x_{lt} + \hat{\gamma}'w_{lt} + \hat{\varepsilon}_{lt}, \]  
\[ x_{lt} = \hat{\pi}z_{lt} + \hat{\rho}'w_{lt} + \hat{\upsilon}_{lt}. \]  

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

We estimate SSIV specifications that parallel (61)-(62) from the simulated data

\[ y_{lt}^* = \beta^*x_{lt}^* + \gamma'^*w_{lt} + \varepsilon_{lt}^*, \]  
\[ x_{lt}^* = \pi^*z_{lt}^* + \rho'^*w_{lt} + \upsilon_{lt}^*. \]  

using the original weights \( e_{lt} \) and controls \( w_{lt} \). We then test the (true) hypothesis \( \beta^* = 0 \) using either the heteroskedasticity-robust standard errors from the equivalent industry-level regression or their version with the null imposed, as in Section 5.1.\(^{57}\) As in column 3 of Table 4, we control for period indicators as \( q_{nt} \) in the industry-level regression.

Our comparison estimator is a conventional industry-level IV inspired by Acemoglu et al. (2016). However, we try to keep the IV regression as similar to the SSIV as possible, thus diverging from Acemoglu et al. (2016) in some details. Specifically, the outcome \( y_{nt} \) is the industry employment growth as measured by these authors. It is defined for 392 out of the 397 industries in Autor et al. (2013), so we drop the remaining five industries in each period. The endogenous regressor \( x_{nt} \equiv g_{nt}^{US} \) (growth of U.S. imports from China per worker) and the instrument \( g_{nt} \) (growth of China imports into eight developed economies) are those from which we built the shift-share endogenous regressor and treatment, respectively (see footnote 38). Construction of those variables differ from Acemoglu et al. (2016) who measure imports relative to domestic absorption rather than employment. We also follow our SSIV analysis in using period indicators as the only industry-level control variables \( q_{nt} \) and taking identical regression importance weights \( s_{nt} \).

\(^{57}\)Note that there is no need for clustering since we generate the shocks independently across industries in all simulations. We have verified, however, that allowing for correlation in shocks within industry groups and using clustered standard errors yields similar results.
The Monte-Carlo strategy for the conventional shock-level IV parallels the one for SSIV; we obtain an estimated industry-level second and first stage of

\[ y_{nt} = \hat{\beta}_{\text{ind}} x_{nt} + \gamma' q_{nt} + \varepsilon_{nt}, \]  
\[ x_{nt} = \hat{\pi}_{\text{ind}} g_{nt} + \rho' q_{nt} + \hat{u}_{nt}. \]  

(65)  

(66)

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

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

**Main results** Table C7 reports the rejection rates for shift-share IV (columns 1 and 2) and conventional industry-level IV (columns 3 and 4) in various simulations. Specifically, column 1 corresponds to using exposure-robust standard errors from the equivalent industry-level IV, and column 2 implements the version with the null hypothesis imposed. Columns 3 and 4 parallel columns 1 and 2 when applied to conventional IV: the former uses heteroskedasticity-robust standard errors and the latter tests \( \beta_{\text{ind}} = 0 \) with the null imposed, which amounts to using the Lagrange multiplier test.

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

This simulation may not approximate the data-generating process well because of heteroskedasticity: smaller industries have more volatile shocks.\(^{58}\) To match unrestricted heteroskedasticity, in row (b) we use wild bootstrap, generating \( g_{nt}^* = g_{nt} \nu_{nt}^* \) by multiplying the year-demeaned observed shocks \( g_{nt} \) by \( \nu_{nt}^* \sim \mathcal{N}(0, 1) \) (Liu 1988). This approach also provides a better approximation for the marginal distribution of shocks than the normality assumption. Here the relative performance of SSIV is even better: the rejection rate is 8.0% vs. 14.2% for conventional IV.

We now depart from the row (b) simulation in several directions, as a case study for the sensitivity

---

\(^{58}\)This is established by unreported regressions of \( |g_{nt}| \) on \( s_{nt} \), for year-demeaned \( g_{nt} \) from ADH, with or without weights. The negative relationship is significant at conventional levels.
of the asymptotic approximation to different features of the SSIV setup. Specifically, we study the role of the Herfindahl concentration index across industries, the number of regions and industries, and the many weak instrument bias. We uniformly find that the performance of the SSIV estimator is similar to that of industry-level IV. Our results also suggest that the Herfindahl index is a useful statistic for measuring the effective number of industries in SSIV, and the first-stage $F$-statistic is informative about the weak instrument bias, as usual.

**The Role of Industry Concentration** Since Assumption 2 requires small concentration of industry importance weights, measured using the Herfindahl index $\sum_{n,t} s_{nt}^2 / \left( \sum_{n,t} s_{nt} \right)^2$, Panel B of Table C7 studies how increasing the skewness of $s_{nt}$ towards the bigger industries affects coverage of the tests. For conventional IV this simply amounts to reweighting the regression. Specifically, for a parameter $\alpha > 1$, we use weights

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

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

$$\frac{\sum_{n,t} (\tilde{s}_{nt})^2}{\left( \sum_{n,t} \tilde{s}_{nt} \right)^2} = \frac{\tilde{HHI}}{HHI}. \quad (67)$$

Matching the Herfindahl index in SSIV is more complicated since we need to choose how exactly to amend shares $\tilde{s}_{nt}$ and regional weights $\tilde{e}_{lt}$ that would yield $\tilde{s}_{nt}$ from (67). We proceed as follows: we consider the lagged level of manufacturing employment by industry $E_{lnt} = e_{lt} s_{lnt}$ and the total regional non-manufacturing employment $E_{l0t} = e_{lt} (1 - \sum_n s_{lnt})$. We then define $\tilde{E}_{lnt} = E_{lnt} \cdot \tilde{s}_{nt}$ for manufacturing industries (and leave non-manufacturing employment unchanged, $\tilde{E}_{l0t} = E_{l0t}$). This increases employment in large manufacturing industries proportionately in all regions, while reducing it in smaller ones. We then recompute shares $\tilde{s}_{lnt}$ and weights $\tilde{e}_{lt}$ accordingly:

$$\tilde{e}_{lt} = \sum_{n=0}^N \sum_t \tilde{E}_{lnt},$$

$$\tilde{s}_{lnt} = \frac{\tilde{E}_{lnt}}{\tilde{e}_{lt}}.$$

Rows (c)-(e) of Table C7 Panel B implement this procedure for target Herfindahl levels of 1/50, 1/20, and 1/10, respectively. For comparison, the Herfindahl in the actual ADH data is 1/191.6 (Table 1, column 2). The table finds that even with the Herfindahl index of 1/20 (corresponding to the “effective” number of shocks of 20 in both periods total) the rejection rate is still around 7%, a level

---

59 Note that in ADH $\sum_n s_{lnt}$ equals the lagged share of regional manufacturing employment, which is below one. We thus renormalize the shares when computing the Herfindahl.

60 The interpretation of $E_{lnt}$ as the lagged level is approximate since $e_{lt}$ is measured at the beginning of period in ADH, while $s_{lnt}$ is lagged.
that may be considered satisfactory. It also shows that the rejection rate grows when the Herfindahl is even higher, at $1/10$, suggesting that the Herfindahl can be used as an indicative rule of thumb. More importantly, the rejection rates are similar for SSIV and conventional industry-level IV, as before.

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

First, to consider the case of small $N$, we aggregate industries in a natural way: from 397 four-digit manufacturing SIC industries into 136 three-digit ones and further into 20 two-digit ones and reconstruct the endogenous right-hand side variable and the instrument using aggregated data. $^{61}$ Rows (f) and (g) of Table C7 Panel C report simulation results based on the aggregated data. They show that rejection rates are similar to the case of detailed industries, and between SSIV and conventional IV. This does not mean that disaggregated data are not useful: the dispersion of the simulated distribution (not reported) increases with industry aggregation, reducing test power. However, standard errors correctly reflect this variability, resulting in largely unchanged test coverage rates.

Second, to study the implications of having fewer regions $L$, we select a random subset of them in each simulation. The results are presented in Rows (h) and (i) of Panel C for $L = 100$ and $25$, compared to the original $L = 722$, respectively. $^{62}$ They show once again that rejection rates are not significantly affected (even though unreported standard errors expectedly increase).

**Many Weak Instruments** In this final simulation we return to the question of SSIV bias. Since our previous simulations confirm that just-identified SSIV is median-unbiased, we turn to the case of multiple instruments. We show that the problem of many weak instruments is similar between SSIV and conventional IV, and that first-stage $F$-statistics, when properly constructed, can serve as useful heuristics.

For clarity, we begin by describing the procedure for the conventional shock-level IV that is a small departure from Column 3 of Table C7. For a given number of instruments $J \geq 1$, in each simulation we generate $g_{jnt}^*, j = 1, \ldots, J$, independently across $j$ using wild bootstrap (as in Table C7 Row (b)). $^{63}$ We make only the first instrument relevant by setting $x_{nt}^* = \tilde{x}_{ind}g_{1nt}^* + \sum_{j=2}^{J} 0 \cdot g_{jnt}^* + \tilde{u}_{nt}$. We then estimate the IV regression of $y_{nt}^* \equiv \hat{\varepsilon}_{nt}$ on $x_{nt}^*$, instrumenting with $g_{1nt}^*, \ldots, g_{Jnt}^*$, controlling for $q_{nt}$, and weighting by $s_{nt}$. We use robust standard errors and compute the effective first-stage

---

$^{61}$Specifically, we aggregate imports from China to the U.S. and either developed economies as well as the number of U.S. workers by manufacturing industry to construct the new $g_{nt}$ and $y_{nt}^{US}$. We then aggregate the shares $s_{nt}$ and $s_{nt}^{current}$ to construct $x_{nt}$ and $z_{it}$ (see footnote 38 for formulas). We do not change the regional outcome, controls, or importance weights. For conventional IV, we additionally reconstruct the outcome (industry employment growth) by aggregating employment levels by year in the Acemoglu et al. (2016) data and measuring growth according to their formulas.

$^{62}$When we select regions, we always keep observations from both periods for each selected region. We keep the second- and first-stage coefficients from the full sample to focus on the noise that arises from shock randomness.

$^{63}$For computational reasons we perform only $15,000/J$ simulations when $J > 1$ (but $10,000$ for $J = 1$ as before).
The procedure for SSIV is more complex but as usual parallels the one for the conventional shock-level IV as much as possible. Given simulated shocks $g_{jnt}^*$, we construct shift-share instruments $z_{jnt}^* = \sum_\ell s_{nt} g_{\ell nt}^*$ and make only the first of them relevant, $x_{\ell t}^* = \pi z_{1\ell t}^* + \sum_{j=2}^J 0 \cdot z_{jnt}^* + \hat{u}_{\ell t}$. Since the equivalence result from Section 2.3 need not hold for overidentified SSIV, we rely on the results in Appendix A.11: we estimate $\beta^*$ from the industry-level regression of $\bar{y}_{nt}^\perp$ (based on $y_{\ell t}^* = \hat{\varepsilon}_{\ell t}$ as before) on $\bar{x}_{nt}^\perp$ by 2SLS, instrumenting by $g_{1nt}^*, \ldots, g_{Jnt}^*$, controlling for $q_{nt}$ and weighting by $s_{nt}$. We compute robust standard errors from this regression to test $\beta^* = 0$. For effective first-stage $F$-statistics, we follow the procedure described in Appendix A.11 and implemented via our weakssivtest command in Stata.

Table C8 reports the result for $J = 1, 5, 10, 25,$ and $50$, presenting the rejection rate corresponding to the 5% nominal, the median bias as a percentage of the simulated standard deviation, and the median first-stage $F$-statistic. Panel A corresponds to SSIV and Panel B to the conventional shock-level IV. For higher comparability, we adjust the first-stage coefficient $\pi_{\text{ind}}$ in the latter in order to make the $F$-statistics approximately match between the two panels. We find that the median bias is now non-trivial and grows with $J$, at the same time as the $F$-statistic declines. However, the level of bias is similar for the two estimators. The rejection rates tend to be higher for conventional IV than SSIV, although they converge as $J$ grows.
B Appendix Proofs

B.1 Proposition 4 and Extensions

This section proves Proposition 4 and extensions that allow for certain forms of mutual shock dependence (Assumptions 5 and 6). Proposition 3 is obtained as a special case, where \( q_n = 1 \). In addition to Assumptions 3 and 4 and the relevance condition of \( \sum_\ell e_\ell z_\ell x_\ell^+ \xrightarrow{p} \pi \) with \( \pi \neq 0 \), the proof of Proposition 4 uses two regularity conditions:

**Assumption B1**: \( E \left[ \tilde{g}_n^2 | \bar{\varepsilon}, q, s \right] \) and \( E \left[ \tilde{\varepsilon}^2_n | s \right] \) are uniformly bounded by some fixed \( B_{\tilde{g}} \) and \( B_{\tilde{\varepsilon}} \).

**Assumption B2**: \( \| \sum_\ell e_\ell w_\ell \varepsilon_\ell \|_1 = o_p(1) \), \( \max | (\sum_\ell e_\ell w_\ell w'_\ell) - \| | = O_p(1) \), and \( \max | \sum_\ell e_\ell w_\ell z_\ell | = O_p(1) \).

The first of these is a weak condition on the second moments of shocks and shock-level unobservables which we show below permits a shock-level law of large numbers. The second condition ensures the consistency of the IV estimate of the control coefficient, \( \hat{\gamma} = \left( \sum_\ell e_\ell w_\ell w'_\ell \right)^{-1} \sum_\ell e_\ell w_\ell \varepsilon_\ell = \gamma + (\sum_\ell e_\ell w_\ell w'_\ell)^{-1} \sum_\ell e_\ell w_\ell \varepsilon_\ell \) (see footnote 5), and stochastic boundedness of the weighted average \( \sum_\ell e_\ell w_{\ell m} z_\ell \), while generally allowing the length of the control vector to increase with \( L \). We discuss low-level conditions for the consistency of \( \hat{\gamma} \) in Appendix A.5.

To prove Proposition 4, we first note that under Assumption B2,

\[
\sum_n s_n g_n \bar{\varepsilon}_n - \sum_n s_n g_n \bar{\varepsilon}_n^+ = \sum_\ell e_\ell z_\ell \left( \bar{\varepsilon}_\ell - \bar{\varepsilon}_\ell^+ \right) \\
= \left( \sum_\ell e_\ell z_\ell w'_\ell \right) \left( \hat{\gamma} - \gamma \right) \\
= \left( \sum_\ell e_\ell z_\ell w'_\ell \right) \left( \sum_\ell e_\ell w_\ell w'_\ell \right)^{-1} \sum_\ell e_\ell w_\ell \varepsilon_\ell \xrightarrow{p} 0, \tag{68}
\]

so that, when the relevance condition holds,

\[
\hat{\beta} - \beta = \frac{\sum_n s_n g_n \bar{\varepsilon}_n^+}{\sum_n s_n g_n \bar{\varepsilon}_n} \\
= \pi^{-1} \sum_n s_n g_n \bar{\varepsilon}_n \left( 1 + o_p(1) \right). \tag{69}
\]

Furthermore, since \( \sum_n s_\ell n = 1 \), we also have under Assumption B2 that

\[
\sum_n s_n q'_n \mu \bar{\varepsilon}_n = \left( \sum_\ell e_\ell w_\ell \varepsilon_\ell \right)' \mu \xrightarrow{p} 0. \tag{70}
\]
Thus

\[
\sum_n s_n g_n \bar{\varepsilon}_n = \sum_n s_n \tilde{g}_n \bar{\varepsilon}_n + o_p(1),
\]

(71)

with

\[
E \left[ \sum_n s_n \tilde{g}_n \bar{\varepsilon}_n \right] = 0
\]

(72)

under Assumption 3.

To prove consistency of \( \hat{\beta} \), it remains to show that \( \text{Var} \left[ \sum_n s_n \tilde{g}_n \bar{\varepsilon}_n \right] \rightarrow 0 \) such that \( \sum_n s_n g_n \bar{\varepsilon}_n \xrightarrow{P} 0 \).

Since

\[
E \left[ \tilde{g}_n \tilde{g}_{n'} | \bar{\varepsilon}, q, s \right] = \text{Cov} \left[ \tilde{g}_n, \tilde{g}_{n'} | \bar{\varepsilon}, q, s \right] = 0
\]

(73)

under Assumptions 3 and 4,

\[
\text{Var} \left[ \sum_n s_n \tilde{g}_n \bar{\varepsilon}_n \right] = E \left[ \left( \sum_n s_n \tilde{g}_n \bar{\varepsilon}_n \right)^2 \right]
\]

\[
= \sum_n \sum_{n'} E \left[ s_n s_{n'} \tilde{g}_n \tilde{g}_{n'} \bar{\varepsilon}_n \bar{\varepsilon}_{n'} \right]
\]

\[
= \sum_n \left[ s_n^2 E \left[ E \left[ \tilde{g}_n^2 | \bar{\varepsilon}, q, s \right] \bar{\varepsilon}^2_n | s \right] \right].
\]

(74)

Then, by Assumption B1 and the Cauchy-Schwartz inequality:

\[
\text{Var} \left[ \sum_n s_n \tilde{g}_n \bar{\varepsilon}_n \right] \leq B_g B_x E \left[ \sum_n s_n^2 \right] \rightarrow 0.
\]

(75)
**Extensions**  Similar steps establish equation (75) when Assumption 4 is replaced by either Assumption 5 or 6. Under Assumption 5 we have, for \( N(c) = \{ n : c(n) = c \} \),

\[
\text{Var} \left[ \sum_n s_n \tilde{g}_n \tilde{\varepsilon}_n \right] = \mathbb{E} \left[ \left( \sum_n \sum_{n \in N(c)} s_n \tilde{g}_n \tilde{\varepsilon}_n \right)^2 \right] \\
= \mathbb{E} \left[ \sum_n s_n^2 \mathbb{E} \left[ \left( \sum_{n \in N(c)} \frac{s_n \tilde{g}_n \tilde{\varepsilon}_n}{s_n} \right)^2 \bigg| s \right] \right] \\
= \mathbb{E} \left[ \sum_n s_n^2 \sum_{n, n' \in N(c)} \frac{s_n s_{n'}}{s_n s_{n'}} \mathbb{E} \left[ \tilde{g}_n \tilde{g}_{n'} \tilde{\varepsilon}_n \tilde{\varepsilon}_{n'} \bigg| s \right] \right] \\
\leq B_g B_{\tilde{g}} \mathbb{E} \left[ \sum_n s_n^2 \right] \to 0. \tag{76}
\]

Here the last line used Assumption B1 and the Cauchy-Schwartz inequality twice: to establish, for \( n, n' \in N(c) \),

\[
\mathbb{E} \left[ \tilde{g}_n \tilde{g}_{n'} \mid \tilde{\varepsilon}, q, s \right] \leq \sqrt{\mathbb{E} \left[ \tilde{g}_n^2 \mid \tilde{\varepsilon}, q, s \right] \mathbb{E} \left[ \tilde{g}_{n'}^2 \mid \tilde{\varepsilon}, q, s \right]} \\
\leq B_{\tilde{g}} \tag{77}
\]

and

\[
\mathbb{E} \left[ |\tilde{\varepsilon}_n| \mid |\tilde{\varepsilon}_{n'}|, s \right] \leq \sqrt{\mathbb{E} \left[ \tilde{\varepsilon}_n^2 \mid s \right] \mathbb{E} \left[ \tilde{\varepsilon}_{n'}^2 \mid s \right]} \\
\leq B_{\tilde{\varepsilon}}. \tag{78}
\]

If we instead replace Assumption 4 with Assumption 6, we have

\[
\text{Var} \left[ \sum_n s_n \tilde{g}_n \tilde{\varepsilon}_n \right] = \mathbb{E} \left[ \left( \sum_n s_n \tilde{g}_n \tilde{\varepsilon}_n \right)^2 \right] \\
= \sum_n \sum_{n'} \mathbb{E} \left[ s_n s_{n'} \mathbb{E} \left[ \tilde{g}_n \tilde{g}_{n'} \mid \tilde{\varepsilon}, q, s \right] \tilde{\varepsilon}_n \tilde{\varepsilon}_{n'} \right] \\
\leq B_L \sum_n \sum_{n'} f \left( |n' - n| \right) \mathbb{E} \left[ |s_n \tilde{\varepsilon}_n s_{n'} \tilde{\varepsilon}_{n'}| \right] \\
= B_L \left( \sum_n \mathbb{E} \left[ (s_n \tilde{\varepsilon}_n)^2 \right] f(0) + 2 \sum_{r=1}^{N-1} \sum_{n=1}^{N-r} \mathbb{E} \left[ |s_{n+r} \tilde{\varepsilon}_{n+r}| \cdot |s_n \tilde{\varepsilon}_n| \right] f(r) \right) \\
\leq \left( B_L \sum_n \mathbb{E} \left[ s_n^2 \mathbb{E} \left[ \tilde{\varepsilon}_n^2 \mid s \right] \right] \right) \left( f(0) + 2 \sum_{r=1}^{N-1} f(r) \right) \\
\leq B_{\tilde{\varepsilon}} \left( f(0) + 2 \sum_{r=1}^{N-1} f(r) \right) \left( B_L \mathbb{E} \left[ \sum_n s_n^2 \right] \right) \to 0. \tag{79}
\]
using \( \mathbb{E} \left[ \varepsilon_n^2 \mid s_n \right] < B_\varepsilon \) in the last line. Here the second-to-last line follows because for any sequence of numbers \( a_1, \ldots, a_N \) and any \( r > 0 \),
\[
\sum_n a_n^2 \geq \frac{1}{2} \left( \sum_{n=1}^{N-r} a_n^2 + \sum_{n=1}^{N-r} a_{n+r}^2 \right) \\
= \frac{1}{2} \sum_{n=1}^{N-r} (a_n - a_{n+r})^2 + \sum_{n=1}^{N-r} a_n a_{n+r} \\
\geq \sum_{n=1}^{N-r} a_n a_{n+r}, \tag{80}
\]
and the same is true in expectation if \( a_n = |s_n \varepsilon_n| \) are random variables. We note that allowing \( B_L \) to grow in the asymptotic sequence imposes much weaker conditions on the correlation structure of shocks. For example, with shock importance weights \( s_n \) approximately equal, i.e. \( \sum_n s_n^2 = O_p(1/N) \), it is enough to have \( \text{Cov} \left[ \tilde{g}_n, \tilde{g}_n' \mid \varepsilon, q, s \right] \leq B_1/N^\alpha \) for any \( \alpha > 0 \): in this case one can satisfy Assumption 6 by setting \( B_L = B_1 N^{1-\alpha/2} \) and \( f(r) = r^{-1-\alpha/2} \).

### B.2 Proposition 5 and Related Results

This section proves Proposition 5 and establishes several additional results mentioned in Section 5.1. First, we show the heteroskedasticity-robust standard error from estimating equation (13) is numerically equivalent to the baseline IV standard error of Adão et al. (2019) when \( w_\ell \) contains only a constant. Second, we discuss the likely difference between our standard error estimates and those of Adão et al. (2019), both without and with approximate shift-share controls. Finally, we show how the alternative null-imposed inference procedure of Adão et al. (2019) is also conveniently obtained from our equivalent shock-level regression.

We prove Proposition 5 in a restricted version of our baseline framework that largely follows Adão et al. (2019), imposing the following assumptions:

**Assumption B3:** The first stage satisfies \( x_\ell = \sum_n s_{\ell n} \pi_{\ell n} g_n + \eta_\ell \), for all \( \ell \).

**Assumption B4:** The control vector can be partitioned as \( w_\ell = [\tilde{w}_\ell', w_\ell']', \) for \( \tilde{w}_\ell = \sum_n s_{\ell n} q_n \).

The vector \( q_n \) captures all sources of shock confounding: \( \mathbb{E} \left[ g_n \mid I_L \right] = q_n \mu \), for all \( n \) and \( I_L = \{ q_n \}, \{ u_\ell, \varepsilon_\ell, \eta_\ell, \{ s_{\ell n}, \pi_{\ell n} \}_n, e_\ell \} \).

**Assumption B5:** The \( g_n \) are mutually independent given \( I_L \), \( \max_n s_n \to 0 \), and \( \max_n \frac{s_n^2}{\sum_n s_n^2} \to 0 \).

**Assumption B6:** \( \mathbb{E} \left[ |g_n|^{4+v} \mid I_L \right] \) is uniformly bounded for some \( v > 0 \) and \( \sum_\ell e_\ell \sum_n s_{\ell n}^2 \text{Var} \left[ g_n \mid I_L \right] \pi_{\ell n} \neq 0 \) almost surely. The support of \( \pi_{\ell n} \) is bounded, the fourth moments of \( \varepsilon_\ell, \eta_\ell, \tilde{u}_\ell, \) and \( \tilde{q}_n \) exist and are uniformly bounded, and \( \sum_\ell e_\ell w_\ell w_\ell' \overset{p}{\to} \Omega_{ww} \) for positive definite \( \Omega_{ww} \). The control vector \( \gamma \) is consistently estimated by \( \hat{\gamma} = \left( \sum_\ell e_\ell w_\ell w_\ell' \right)^{-1} \sum_\ell e_\ell w_\ell  \varepsilon_\ell \).
We note that Assumption B5 both strengthens our baseline Herfindahl index condition in Assumption 4 and implicitly treats the set of $s_n$ as non-stochastic, following Assumption 2 of Adão et al. (2019). The regularity condition B6 includes the relevant conditions from Assumptions 4 and A.3 of Adão et al. (2019). These assumptions strengthen those of Proposition 4: Assumptions B3–B6 imply our Assumptions 3, 4, B1, and B2. Relative to Adão et al. (2019), we do not impose that $L > N$ or that the shares are non-collinear.

To establish the equivalence of IV coefficients in Proposition 5, note that when $\sum_n s_n q_n$ is included in $w_\ell$

$$\sum_n s_n q_n \bar{y}_n = \sum_\ell e_\ell y_\ell (\sum_n s_\ell q_n) = 0$$

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

To establish validity of the standard errors, note that the conventional heteroskedasticity-robust standard error from for the $s_n$-weighted shock-level IV regression of $\bar{y}_n$ on $\bar{x}_n$ and $q_n$, instrumented by $g_n$, is given by

$$\hat{s}_\text{equiv} = \sqrt{\sum_n s_n^2 \hat{\varepsilon}_n^2 \hat{g}_n^2} \left/ \sum_n s_n \bar{x}_n \bar{g}_n \right. ,$$

where $\hat{\varepsilon}_n = \bar{y}_n - \hat{\beta} \bar{x}_n$ is the estimated shock-level regression residual (where we used the fact that the estimated coefficients on $q_n$ in that regression are numerically zero) and $\hat{g}_n = g_n - \hat{\mu} q_n$, where $\hat{\mu} = (\sum_n s_n q_n q_n')^{-1} \sum_n s_n q_n g_n$, is the residual from a projection of the instrument in equation (13) on the control vector $q_n$. By Proposition 1, $\hat{\varepsilon}_n$ coincides with the share-weighted aggregate of the SSIV estimated residuals $\hat{\varepsilon}_\ell = y_\ell - \hat{\beta} x_\ell$:

$$\hat{\varepsilon}_n = \sum_\ell e_\ell s_\ell t_n y_\ell - \hat{\beta} \cdot \sum_\ell e_\ell s_\ell t_n x_\ell = \sum_\ell e_\ell s_\ell t_n \hat{\varepsilon}_\ell.$$  (83)

The squared numerator of (82) can thus be rewritten

$$\sum_n \sum_\ell e_\ell s_\ell t_n \hat{\varepsilon}_n = \sum_n \left( \sum_\ell e_\ell s_\ell t_n \hat{\varepsilon}_\ell \right)^2 \hat{g}_n^2.$$  (84)

The expression in the denominator of (82) estimates the magnitude of the shock-level first-stage covariance, which matches the $e_\ell$-weighted sample covariance of $x_\ell$ and $z_\ell$:

$$\sum_n s_n \bar{x}_n g_n = \sum_n \left( \sum_\ell e_\ell s_\ell t_n x_\ell \right) g_n = \sum_\ell e_\ell x_\ell z_\ell.$$  (85)
Thus
\[ \hat{s}_{\text{equiv}} = \sqrt{\frac{\sum_n (\sum_\ell e_\ell s_\ell \hat{\epsilon}_\ell)^2 \hat{g}_n^2}{\sum_\ell e_\ell x_\ell^2 z_\ell}}. \] (86)

We now compare this expression to the standard error formula from Adão et al. (2019), incorporating the \( e_\ell \) importance weights. Equation (39) in their paper yields
\[ \hat{s}_{\text{AKM}} = \sqrt{\frac{\sum_n (\sum_\ell e_\ell s_\ell \hat{\epsilon}_\ell)^2 \hat{g}_n^2}{\sum_\ell e_\ell x_\ell^2 z_\ell}}, \] (87)

where \( \hat{g}_n \) denotes the coefficients from regressing the residualized instrument \( z_\ell^+ \) on all shares \( s_\ell n \), without a constant; note that to compute this requires \( L > N \) and that the matrix of exposure shares \( s_\ell n \) is full rank. The formulas for \( \hat{s}_{\text{equiv}} \) and \( \hat{s}_{\text{AKM}} \) therefore differ only in the construction of shock residuals, \( \hat{g}_n \) versus \( \hat{g}_n \).

We establish the general asymptotic equivalence of \( \hat{s}_{\text{equiv}}^2 \) and \( \hat{s}_{\text{AKM}}^2 \), and thus the asymptotic validity of \( \hat{s}_{\text{equiv}} \), by showing that both capture the conditional asymptotic variance of \( \hat{\beta} \) given \( I_L \) under Assumptions B3-B6. Both of the resulting confidence intervals are then asymptotically valid unconditionally, since if \( \text{Pr}(\beta \in \hat{C} I \mid I_L) = \alpha \) then \( \text{Pr}(\beta \in C I) = E \left[ \text{Pr}[\beta \in \hat{C} I \mid I_L] \right] = \alpha \) by the law of iterated expectations. Under Assumptions B3-B6, Proposition A.1 of Adão et al. (2019) applies and shows that
\[ \sqrt{r_L}(\hat{\beta} - \beta) = N \left( 0, r_L \frac{\mathcal{V}_L}{\pi^2} \right) + o_p(1) \] (88)

where \( r_L = 1/ (\sum_n s_n^2) \) and \( \mathcal{V}_L = \sum_n (\sum_\ell e_\ell s_\ell \hat{\epsilon}_\ell)^2 \text{Var}[g_n \mid I_L] \), provided \( r_L \mathcal{V}_L \) converges to a non-random limit. To establish the asymptotic validity of \( \hat{s}_{\text{AKM}} \), i.e. that \( r_L \left( \sum_n (\sum_\ell e_\ell s_\ell \hat{\epsilon}_\ell)^2 \hat{g}_n^2 - \mathcal{V}_L \right) \overset{p}{\to} 0 \), Adão et al. (2019) further assume that \( L \geq N \), the matrix of \( s_\ell n \) is always full rank, and additional regularity conditions (see their Proposition 5). We establish \( r_L \left( \sum_n (\sum_\ell e_\ell s_\ell \hat{\epsilon}_\ell)^2 \hat{g}_n^2 - \mathcal{V}_L \right) \overset{p}{\to} 0 \) without imposing those assumptions.

To start, we write \( g_n^* = g_n - q_n \mu \) and decompose
\[ r_L \left( \sum_n \left( \sum_\ell e_\ell s_\ell \hat{\epsilon}_\ell \right)^2 \hat{g}_n^2 - \mathcal{V}_L \right) = r_L \left( \sum_n \left( \sum_\ell e_\ell s_\ell \hat{\epsilon}_\ell \right)^2 \hat{g}_n^2 - \mathcal{V}_L \right) + r_L \sum_n \left( \left( \sum_\ell e_\ell s_\ell \hat{\epsilon}_\ell \right)^2 - \left( \sum_\ell e_\ell s_\ell \hat{\epsilon}_\ell \right)^2 \right) \hat{g}_n^2 + r_L \sum_n \left( \sum_\ell e_\ell s_\ell \hat{\epsilon}_\ell \right)^2 (\hat{g}_n^2 - \hat{g}_n^2). \] (89)

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

\begin{equation}
 r_L \sum_{\ell} \sum_{\ell'} \sum_{n} e_{\ell} e_{\ell'} s_{\ell n} s_{\ell'n} A_{L\ell} B_{L\ell'} C_{L_n} = O_p(1). \tag{90}
\end{equation}

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

\begin{equation}
\left( \sum_{\ell} e_{\ell} s_{\ell n} \hat{e}_{\ell} \right)^2 = \left( \sum_{\ell} e_{\ell} s_{\ell n} \varepsilon_{\ell} \right)^2 + 2 \sum_{\ell} \sum_{\ell'} e_{\ell} e_{\ell'} s_{\ell n} s_{\ell'n} D_\ell' (\theta - \hat{\theta}) \varepsilon_{\ell'},
\end{equation}

and both \( D_\ell \) and \( \varepsilon_{\ell} \) have bounded fourth moments by the assumption of bounded fourth moments of \( e_{\ell}, \eta_{\ell}, u_{\ell}, \) and \( p_n \) in Assumption B6. Thus by the lemma

\begin{equation}
 r_L \sum_n \left( \left( \sum_{\ell} e_{\ell} s_{\ell n} \hat{e}_{\ell} \right)^2 - \left( \sum_{\ell} e_{\ell} s_{\ell n} \varepsilon_{\ell} \right)^2 \right) \tilde{g}_n^2 = 2 \left( \theta - \hat{\theta} \right)' \left( r_L \sum_{\ell} \sum_{\ell'} \sum_{n} e_{\ell} e_{\ell'} s_{\ell n} s_{\ell'n} \tilde{g}_n^2 D_\ell D_{\ell'} \right) \left( \theta - \hat{\theta} \right) + \left( \theta - \hat{\theta} \right)' O_p(1) + \left( \theta - \hat{\theta} \right)' O_p(1) \left( \theta - \hat{\theta} \right), \tag{92}
\end{equation}

which is \( o_p(1) \) by the consistency of \( \hat{\theta} \) (implied by Assumptions B3-B6). Adão et al. (2019) further show the first term of equation (89) is \( o_p(1) \), without using the additional regularity conditions of their Proposition 5.

It thus remains for us to show the third term of (89) is also \( o_p(1) \). Note that

\begin{equation}
\tilde{g}_n^2 = (g_n - q'_n \hat{\mu})^2 = \tilde{g}_n^2 + (q'_n (\hat{\mu} - \mu))^2 - 2 \tilde{g}_n q'_n (\hat{\mu} - \mu), \tag{93}
\end{equation}

so that

\begin{equation}
 r_L \sum_n \left( \sum_{\ell} e_{\ell} s_{\ell n} \hat{e}_{\ell} \right)^2 \left( \tilde{g}_n^2 - \tilde{g}_n'^2 \right)
= r_L \sum_n \left( \sum_{\ell} e_{\ell} s_{\ell n} \hat{e}_{\ell} \right)^2 \left( q'_n (\hat{\mu} - \mu) - 2 \tilde{g}_n \right) q'_n (\hat{\mu} - \mu)
= r_L \sum_n \left( \sum_{\ell} e_{\ell} s_{\ell n} \hat{e}_{\ell} \right)^2 \left( q'_n (\hat{\mu} - \mu) - 2 \tilde{g}_n \right) q'_n (\hat{\mu} - \mu)
+ r_L \sum_n \left( \left( \sum_{\ell} e_{\ell} s_{\ell n} \hat{e}_{\ell} \right)^2 - \left( \sum_{\ell} e_{\ell} s_{\ell n} \varepsilon_{\ell} \right)^2 \right) \left( q'_n (\hat{\mu} - \mu) - 2 \tilde{g}_n \right) q'_n (\hat{\mu} - \mu). \tag{94}
\end{equation}
Using the previous lemma, the first term of this expression is $O_p(1)(\hat{\mu} - \mu)$ since $\varepsilon_{\ell}, p_n$, and $\tilde{g}_n$ have bounded fourth moments under Assumption B6. The second term is similarly $O_p(1)(\hat{\mu} - \mu)$ by the lemma and the decomposition used in equation (92). Noting that $\hat{\mu} - \mu = (\sum_n s_n q_n q_n')^{-1} \sum_n s_n q_n \tilde{g}_n \overset{p}{\to} 0$ under the assumptions completes the proof.

**The Case of No Controls** We show that when there are no controls besides a constant, i.e. $w_{\ell} = g_n = 1$, the standard errors are numerically the same. To prove this, it suffices to show that $\tilde{g}_n = \hat{g}_n$. Absent controls, $\hat{g}_n = g_n - \sum_n s_n g_n$ is the $s_n$-weighted demeaned shock. The $\tilde{g}_n$ are obtained as the projection coefficient of $z_{\ell}^\perp = z_{\ell} - \sum_{\ell} e_{\ell} z_{\ell}$ on the $N$ shares. Note that

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

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

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

(96)

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

**Comparison of Standard Errors** The characterization of the standard errors in equations (86)–(87) offers insights into how these standard errors may differ in presence of controls. We argue that under the conditions of Proposition 5, our standard errors are likely smaller in finite samples. However, when Assumption B3 is violated because some controls have the approximate shift-share structure allowed by Adão et al. (2019), our standard errors are likely asymptotically conservative.

To see this, consider versions of the two standard error formulas obtained under shock homoskedasticity (i.e. Var $[g_n | I_L] = \sigma^2_g$):

$$\hat{s}_e^{\text{homo}}_{\text{equiv}} = \sqrt{\frac{\sum_n s_n^2 g_n^2}{\left| \sum_n s_n \bar{y}_n g_n \right|}}$$

(97)

$$\hat{s}_e^{\text{homo}}_{\text{AKM}} = \sqrt{\frac{\sum_n s_n^2 g_n^2}{\left| \sum_{\ell} e_{\ell} \bar{x}_{\ell} z_{\ell} \right|}},$$

(98)

which differ by a factor of $\sqrt{\sum_n s_n^2 \tilde{g}_n^2 / \sum_n s_n^2 g_n^2}$.

When the SSIV controls have an exact shift-share structure, $w_{\ell} = \sum_n s_{\ell n} q_n$, the share projection producing $\tilde{g}_n$ has exact fit such that one can represent $\tilde{g}_n = g_n - q_n \hat{\mu}_{\text{AKM}}$ for some $\hat{\mu}_{\text{AKM}}$. In this case the $s_n$-weighted sum of squares of shock residuals is lower in our equivalent regression by construction of $\hat{\mu}$: $\sum_n s_n \tilde{g}_n^2 \leq \sum_n s_n g_n^2$ (with strict inequality when $\hat{\mu}_{\text{AKM}} \neq \hat{\mu}$). Similarly, when $w_{\ell}$ instead contains controls that are included for efficiency only and are independent of the shocks, projection of $z_{\ell}$ on the shares produces a noisy estimate of $g_n - \sum_n s_n g_n$, which again has a higher
weighted sum of squares.

In contrast, if some controls have an approximate shift-share structure, \( w_\ell = \sum_n s_\ell n \tilde{q}_n + u_\ell \), where \( \tilde{q}_n \) is unobserved and correlated with \( g_n \) and the variance of the noise \( u_\ell \) converges to zero, then \( \tilde{g}_n \) asymptotically residualizes \( g_n \) on \( \tilde{q}_n \) while \( \hat{g}_n \) does not. Therefore, the sum of squares of \( \tilde{g}_n \) is asymptotically smaller than that of \( \hat{g}_n \). Since the Adão et al. (2019) standard errors are asymptotically correct, our formula is asymptotically conservative in this scenario.

**Null-Imposed Inference Procedure** Finally, our shock-level equivalence provides a convenient implementation for the alternative inference procedure that may have superior finite-sample performance. Adão et al. (2019) show how standard errors that impose a given null hypothesis \( \beta = \beta_0 \) in estimating the residual \( \varepsilon_\ell \) can generate confidence intervals with better coverage in situations with few shocks (and a similar argument can be made in the case of shocks with a heavy-tailed distribution). Building on Proposition 5, such confidence intervals can be constructed in the same way as in any regular shock-level IV regression. To test \( \beta = \beta_0 \), one regresses \( \bar{y}_n - \beta_0 \bar{x}_n \) on the shocks \( g_n \) (weighting by \( s_n \) and including any relevant shock-level controls \( q_n \)) and uses a null-imposed residual variance estimate. This procedure corresponds to the standard shock-level Lagrange multiplier test for \( \beta = \beta_0 \) that can be implemented by standard statistical software.

**B.3 Proposition A1**

We consider each expectation in equation (16) in turn. For each \( n \), write

\[
\kappa_n(g_n, \varepsilon_\ell, \eta_\ell) = \lim_{g_n \to -\infty} y(x_1([g_n; g_n], \eta_1), \ldots, x_{1R}([g_n; g_n], \eta_{1R}), \varepsilon_\ell) \tag{99}
\]

such that

\[
s_{\ell n} \ell / g_n y_\ell = s_{\ell n} \ell / g_n \kappa_n(g_n, \varepsilon_\ell, \eta_\ell) \tag{100}
\]

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

\[
E[s_{\ell n} \ell / g_n \kappa_n(g_n, \varepsilon_\ell, \eta_\ell)] = E[s_{\ell n} \ell / g_n | s, \varepsilon, \varepsilon_n, \eta_\ell] \kappa_n(g_n, \varepsilon_\ell, \eta_\ell) = 0, \tag{101}
\]

---

\[64\] As explained by Adão et al. (2019), the problem that this “AKM0” confidence interval addresses generalizes the standard finite-sample bias of cluster-robust standard errors with few clusters (Cameron and Miller 2015). With few or heavy-tailed shocks, estimates of the residual variance will tend to be biased downwards, leading to undercoverage of confidence intervals based on standard errors that do not impose the null.

\[65\] For example in Stata one can use the ivreg2 overidentification test statistic from regressing \( \bar{y}_n - \beta_0 \bar{x}_n \) on \( q_n \) with no endogenous variables and with \( g_n \) specified as the instrument (again with \( s_n \) weights).
while the expectation of the second is

\[
E \left[ s_{tn}e_{\ell}g_n \int_{-\infty}^{g_n} \frac{\partial}{\partial g_n} y(x_1(\gamma; g_n), \eta_{\ell 1}), \ldots, x_R(\gamma; g_n), \eta_{\ell R}, \epsilon_\ell) d\gamma \right] \\
= E \left[ s_{tn}e_{\ell} \int_{-\infty}^{g_n} g_n \frac{\partial}{\partial g_n} y(x_1(\gamma; g_n), \eta_{\ell 1}), \ldots, x_R(\gamma; g_n), \eta_{\ell R}, \epsilon_\ell) d\gamma dF_n(g_n | I) \right] \\
= E \left[ s_{tn}e_{\ell} \int_{-\infty}^{\infty} g_n dF_n(g_n | I) d\gamma \right] \\
= E \left[ s_{tn}e_{\ell} \int_{-\infty}^{\infty} \frac{\partial}{\partial g_n} y(x_1(\gamma; g_n), \eta_{\ell 1}), \ldots, x_R(\gamma; g_n), \eta_{\ell R}, \epsilon_\ell) \int_{\gamma}^{\infty} g_n dF_n(g_n | I) d\gamma \right] \\
= \sum_r E \left[ \int_{-\infty}^{\infty} s_{tn}e_{\ell} \alpha_{\ell r} \pi_{\ell r n}(\gamma; g_n) \mu_n(\gamma | I) \beta_{\ell r n}(\gamma) d\gamma \right] \\
(102)
\]

where \( F_n(\cdot | I) \) denotes the conditional distribution of \( g_n \). Thus

\[
E \left[ s_{tn}e_{\ell}g_n y_\ell \right] = E \left[ s_{tn}e_{\ell} \int_{-\infty}^{\infty} \frac{\partial}{\partial g_n} y(x_1(\gamma; g_n), \eta_{\ell}), \ldots, x_R(\gamma; g_n), \eta_{\ell R}, \epsilon_\ell) \mu_n(\gamma | I) d\gamma \right] \\
= \sum_r \int_{-\infty}^{\infty} s_{tn}e_{\ell} \alpha_{\ell r} \pi_{\ell r n}(\gamma; g_n) \mu_n(\gamma | I) \beta_{\ell r n}(\gamma) d\gamma \\
(103)
\]

where

\[
\mu_n(\gamma | I) \equiv \int_{\gamma}^{\infty} g_n dF_n(g_n | I) = (E[g_n | g_n \geq \gamma, I] - E[g_n | g_n < \gamma, I]) Pr(g_n \geq \gamma | I) (1 - Pr(g_n \geq \gamma | I)) \geq 0 \ a.s. \\
(104)
\]

Similarly

\[
E \left[ s_{tn}e_{\ell}g_n x_\ell \right] = \sum_r E \left[ \int_{-\infty}^{\infty} s_{tn}e_{\ell} \alpha_{\ell r} \pi_{\ell r n}(\gamma; g_n) \mu_n(\gamma | I) d\gamma \right] \\
(105)
\]

Combining equations (103) and (105) completes the proof, with

\[
\omega_{\ell r n}(\gamma) = s_{tn}e_{\ell} \alpha_{\ell r} \mu_n(\gamma | I) \pi_{\ell r n}(\gamma; g_n) \geq 0 \ a.s. \\
(106)
\]

**B.4 Proposition A2**

To prove (22), we aggregate (21) across industries within a region using \( E_{\ell n} \) weights:

\[
y_\ell = (\beta_0 - \beta_1) x_\ell + \epsilon_\ell, \\
(107)
\]
where \( \varepsilon_\ell = \sum_n s_{\ell n} \varepsilon_{\ell n} \). The shift-share instrument \( z_\ell \) is relevant because

\[
E \left[ \sum_\ell e_\ell x_\ell z_\ell \right] = \sum_\ell e_\ell E \left[ \sum_n s_{\ell n} (\bar{\pi} g_n + \eta_{\ell n}) \cdot \sum_{n'} s_{\ell n'} g_{n'} \right] \\
= \sum_\ell e_\ell s_{\ell n}^2 \bar{\pi} \sigma_y^2 \\
\geq H_L \bar{\pi} \sigma_y^2, \tag{108}
\]

while exclusion holds because

\[
E \left[ \sum_\ell e_\ell z_\ell \varepsilon_\ell \right] = \sum_\ell e_\ell E \left[ \sum_n s_{\ell n} \varepsilon_{\ell n} \cdot \sum_{n'} s_{\ell n'} g_{n'} \right] \\
= 0. \tag{109}
\]

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

To study \( \hat{\beta}_{\text{ind}} \), we aggregate (21) across regions (again with \( E_{\ell n} \) weights):

\[
y_n = \beta_0 x_n - \beta_1 \sum_\ell \omega_{\ell n} \sum_{n'} s_{\ell n'} x_{\ell n'} + \varepsilon_n, \tag{110}
\]

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

\[
\hat{\beta}_{\text{ind}} - \beta_0 = \frac{\sum_n s_{n} y_n g_n}{\sum_n s_{n} x_n g_n} - \beta_0 \\
= \frac{\sum_n s_{n} (-\beta_1 \sum_\ell \omega_{\ell n} \sum_{n'} s_{\ell n'} x_{\ell n'} + \varepsilon_n) g_n}{\sum_n s_{n} x_n g_n}. \tag{111}
\]

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

\[
E \left[ \sum_n s_{n} x_n g_n \right] = \sum_n s_{n} E \left[ \sum_\ell \omega_{\ell n} (\bar{\pi} g_n + \eta_{\ell n}) g_n \right] \\
= \sum_n s_{n} \omega_{\ell n} \bar{\pi} \sigma^2 \\
= \sum_n \frac{E_{\ell n}}{E} \cdot \frac{E_{\ell n}}{E} \bar{\pi} \sigma^2 \\
= \bar{\pi} \sigma^2, \tag{112}
\]

while the expected numerator is

\[
E \left[ \sum_n s_{n} \left( -\beta_1 \sum_\ell \omega_{\ell n} \sum_{n'} s_{\ell n'} x_{\ell n'} + \varepsilon_n \right) g_n \right] = -\beta_1 \sum_{n,\ell} s_{n} \omega_{\ell n} s_{\ell n} \bar{\pi} \sigma^2 \\
= -\beta_1 H_L \bar{\pi} \sigma^2, \tag{113}
\]
where the last equality follows because

\[ \sum_{n, \ell} s_n \omega_{\ell n} s_{\ell n} = \sum_{n, \ell} \frac{E_n E_{\ell n} E_{\ell n}}{E_n E_{\ell}} \]

\[ = \sum_{n, \ell} \frac{E_{\ell n}}{E_\ell} \left( \frac{E_{\ell n}}{E_\ell} \right)^2 \]

\[ = \sum_{n, \ell} e_{\ell n} s_{\ell n}^2 \]

\[ = H_L. \quad (114) \]

Thus by an appropriate law of large numbers,

\[ \hat{\beta}_{\text{ind}} = \beta_0 - \beta_1 H_L + o_p(1). \quad (115) \]

### B.5 Proposition A3

By appropriate laws of large numbers,

\[ \hat{\beta} = \frac{E \left[ \sum_\ell E_\ell \left( \sum_n s_{\ell n} y_{\ell n} \right) \left( \sum_{n'} s_{\ell n'} g_{n'} \right) \right]}{E \left[ \sum_\ell E_\ell \left( \sum_n s_{\ell n} x_{\ell n} \right) \left( \sum_{n'} s_{\ell n'} g_{n'} \right) \right]} + o_p(1) \]

\[ = \frac{\sum_\ell n E_\ell s_{\ell n}^2 \pi_{\ell n} \sigma_{\ell n}^2 \beta_{\ell n}}{\sum_\ell s_{\ell n}^2 \pi_{\ell n} \sigma_{\ell n}^2} + o_p(1) \]

\[ = \frac{\sum_\ell n E_\ell s_{\ell n} \pi_{\ell n} \sigma_{\ell n}^2 \beta_{\ell n}}{\sum_\ell s_{\ell n} \pi_{\ell n} \sigma_{\ell n}^2} + o_p(1) \quad (116) \]

while

\[ \hat{\beta}_{\text{ind}} = \frac{\sum_n E_n y_n g_n}{\sum_n E_n x_n g_n} \]

\[ = \frac{E \left[ \sum_n E_n \left( \sum_\ell \omega_{\ell n} y_{\ell n} \right) g_n \right]}{E \left[ \sum_n E_n \left( \sum_\ell \omega_{\ell n} x_{\ell n} \right) g_n \right]} + o_p(1) \]

\[ = \frac{\sum_\ell n E_\ell \omega_{\ell n} \pi_{\ell n} \sigma_{\ell n}^2 \beta_{\ell n}}{\sum_\ell \omega_{\ell n} \pi_{\ell n} \sigma_{\ell n}^2} + o_p(1) \]

\[ = \frac{\sum_\ell n E_\ell \pi_{\ell n} \sigma_{\ell n}^2 \beta_{\ell n}}{\sum_\ell \pi_{\ell n} \sigma_{\ell n}^2} + o_p(1). \quad (117) \]
B.6 Proposition A4

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

$$
\bar{\varepsilon}_n = \frac{\sum \ell e_{\ell s} (\sum_{n'} s_{\ell n'} \nu_{n'} + \bar{\varepsilon}_\ell)}{\sum \ell e_{\ell s}}
\equiv \sum \alpha_{nn'} \nu_{n'} + \bar{\varepsilon}_n,
$$

(118)

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

$$
\text{Var}[\bar{\varepsilon}_n] = \sum_{n'} \sigma_{n'}^2 \alpha_{nn'}^2 + \text{Var}[\bar{\varepsilon}_n]
\geq \sigma_p^2 \alpha_{nn}^2,
$$

(119)

and

$$
\max_n \text{Var}[\bar{\varepsilon}_n] \geq \sigma_p^2 \max_n \alpha_{nn}^2.
$$

(120)

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

$$
\sum_n s_n \alpha_{nn} = \sum_n s_n \frac{\sum \ell e_{\ell s}^2}{s_n} = H_L.
$$

(121)

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

B.7 Proposition A5

We prove each part of this proposition in turn.

1. Expanding the moment condition yields:

$$
\mathbb{E} \left[ \sum_{\ell} e_{\ell} \bar{\varepsilon}_\ell \psi_{\ell,\text{LOO}} \right] = \sum_{\ell} \mathbb{E} \left[ e_{\ell} \bar{\varepsilon}_\ell \sum_n s_{\ell n} \frac{\sum_{\ell' \neq \ell} \omega_{\ell'n} \psi_{\ell'n}}{\sum_{\ell' \neq \ell} \omega_{\ell'n}} \right]
= \sum_{\ell} \mathbb{E} \left[ e_{\ell} \sum_n s_{\ell n} \frac{\sum_{\ell' \neq \ell} \omega_{\ell'n} \mathbb{E} \left[ \bar{\varepsilon}_\ell \psi_{\ell'n} \right]}{\sum_{\ell' \neq \ell} \omega_{\ell'n}} \right]
= 0.
$$

(122)

2. The assumption of part (1) is satisfied here, so $\mathbb{E} \left[ \sum_{\ell} e_{\ell} \bar{\varepsilon}_\ell \psi_{\ell,\text{LOO}} \right] = 0$. We now establish that $\mathbb{E} \left[ (\sum_{\ell} e_{\ell} \bar{\varepsilon}_\ell \psi_{\ell,\text{LOO}})^2 \right] \to 0$, which implies $\sum_{\ell} e_{\ell} \bar{\varepsilon}_\ell \psi_{\ell,\text{LOO}} \not\to 0$ and thus consistency of the LOO
SSIV estimator provided it has a first stage:

$$\mathbb{E} \left[ \left( \sum_{\ell} e_{\ell \ell_1} s_{\ell_1 n_1} s_{\ell_2 n_2} \right)^2 \right] = \sum_{(\ell_1, \ell_2, \ell_1', \ell_2') \in \mathcal{J}, n_1, n_2} \frac{\omega_{\ell_1 n_1} \omega_{\ell_2 n_2}}{\sum_{\ell_1} \omega_{\ell_1 n_1} \sum_{\ell_2} \omega_{\ell_2 n_2}} \mathbb{E} \left[ e_{\ell_1} e_{\ell_2} \psi_{\ell_1 n_1} \psi_{\ell_2 n_2} \right]$$

$$\leq \sum_{(\ell_1, \ell_2, \ell_1', \ell_2') \in \mathcal{J}, n_1, n_2} \frac{\omega_{\ell_1 n_1} \omega_{\ell_2 n_2}}{\sum_{\ell_1} \omega_{\ell_1 n_1} \sum_{\ell_2} \omega_{\ell_2 n_2}} \cdot B \to 0.$$  

(123)

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

3. We show that under the given assumptions on $s_{\ell n}$, $e_{\ell}$, and $\omega_{\ell n}$, the expression in (39) is bounded by $4N/L$:

$$\sum_{(\ell_1, \ell_2, \ell_1', \ell_2') \in \mathcal{J}, n_1, n_2} e_{\ell_1} e_{\ell_2} s_{\ell_1 n_1} s_{\ell_2 n_2} \frac{\omega_{\ell_1 n_1} \omega_{\ell_2 n_2}}{\sum_{\ell_1} \omega_{\ell_1 n_1} \sum_{\ell_2} \omega_{\ell_2 n_2}}$$

$$= \sum_{(\ell_1, \ell_2, \ell_1', \ell_2') \in \mathcal{J}} \frac{1}{L^2} \frac{\omega_{\ell_1 n(\ell_1)} \omega_{\ell_2 n(\ell_2)}}{\sum_{\ell_1} \omega_{\ell_1 n(\ell_1)} \sum_{\ell_2} \omega_{\ell_2 n(\ell_2)}}$$

$$= \frac{1}{L^2} \sum_{n} \frac{1}{L_{n(\ell_1)} - 1} \frac{1}{L_{n(\ell_2)} - 1}$$

$$= \frac{1}{L^2} \sum_{n} \frac{2L_n (L_n - 1)}{(L_n - 1)^2} \leq \frac{4N}{L}.$$  

(124)

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

82
B.8 Proposition A6

National industry employment satisfies $E_n = \sum \ell E_{\ell n}$; log-linearizing this immediately implies (48). To solve for $g_{\ell n}$, log-linearize (45), (46), and (47):

\[
\hat{E}_\ell = \phi \hat{W}_\ell + \varepsilon_\ell, \\
g_{\ell n} = g^*_{n} + \hat{\xi}_{\ell n} - \sigma \hat{W}_\ell, \\
\hat{E}_\ell = \sum_n s_{\ell n} g_{\ell n}.
\]

Solving this system of equations yields

\[
\hat{W}_\ell = \frac{1}{\sigma + \phi} \left( \sum_n s_{\ell n} \left( g^*_{n} + \hat{\xi}_{\ell n} \right) - \varepsilon_\ell \right)
\]

and expression (49).
C Appendix Figures and Tables

Figure C1: Industry-Level Variation in the Autor et al. (2013) Setting

Notes: This figure shows binned scatterplots of shock-level outcome and treatment residuals, $\bar{y}_{nt}^\bot$ and $\bar{x}_{nt}^\bot$, corresponding to the SSIV specification in column 3 of Table 4. The manufacturing industry shocks, $g_{nt}$, are residualized on period indicators (with the full-sample mean added back) and grouped into fifty weighted bins, with each bin representing around 2% of total share weight $s_{nt}$. Lines of best fit, indicated in red, are weighted by the same $s_{nt}$. The slope coefficients equal $5.71 \times 10^{-3}$ and $-1.52 \times 10^{-3}$, respectively, with the ratio (-0.267) equalling the SSIV coefficient in column 3 of Table 4.
|                          | (1)    | (2)    | (3)    | (4)    | (5)    | (6)    | (7)    |
|--------------------------|--------|--------|--------|--------|--------|--------|--------|
| Unemployment growth      | 0.221  | 0.217  | 0.063  | -0.014 | 0.104  | 0.107  | 0.235  |
|                          | (0.049)| (0.046)| (0.060)| (0.079)| (0.079)| (0.083)| (0.178)|
| NILF growth               | 0.553  | 0.534  | 0.098  | 0.149  | 0.142  | 0.117  | 0.187  |
|                          | (0.185)| (0.183)| (0.133)| (0.083)| (0.155)| (0.161)| (0.297)|
| Log weekly wage growth   | -0.759 | -0.607 | 0.227  | 0.320  | 0.145  | 0.063  | -0.211 |
|                          | (0.258)| (0.226)| (0.242)| (0.209)| (0.264)| (0.260)| (0.651)|
| # of industry-periods    | 796    | 794    | 794    | 794    | 794    | 794    | 794    |
| # of region-periods      | 1,444  | 1,444  | 1,444  | 1,444  | 1,444  | 1,444  | 1,444  |

Notes: This table extends the analysis of Table 4 to different regional outcomes in Autor et al. (2013): unemployment growth, labor force non-participation (NILF) growth, and log average weekly wage growth. The specifications are otherwise the same as in the corresponding columns of Table 4. SIC3-clustered exposure-robust standard errors are computed using equivalent industry-level IV regressions and reported in parentheses.
Table C2: Alternative Standard Errors in the Autor et al. (2013) Setting

|                  | (1)   | (2)   | (3)   | (4)   | (5)   | (6)   | (7)   |
|------------------|-------|-------|-------|-------|-------|-------|-------|
| Coefficient      | -0.596| -0.489| -0.267| -0.314| -0.310| -0.290| -0.432|
| Table 4 SE       | (0.114)| (0.100)| (0.099)| (0.107)| (0.134)| (0.129)| (0.205)|
| State-clustered SE | (0.099)| (0.086)| (0.086)| (0.097)| (0.104)| (0.101)| (0.193)|
| Adão et al. (2019) SE | (0.126)| (0.116)| (0.113)| (0.107)| (0.143)| (0.140)| (0.192)|
| Confidence interval with the null imposed | [-1.059, -0.396] | [-0.832, -0.309] | [-0.568, -0.028] | [-0.637, -0.018] | [-0.705, -0.002] | [-0.699, 0.002] | [-1.207, 0.122] |

Notes: This table extends the analysis of Table 4 by reporting conventional state-clustered standard errors, the Adão et al. (2019) SIC3-clustered standard errors, and confidence intervals based on the equivalent industry-level IV regression with the null imposed, as discussed in Section 5.1. The specifications are the same as those in the corresponding columns of Table 4; for comparison we repeat the coefficient estimates and exposure-robust standard errors from that table.
|                  | (1)  | (2)  | (3)  | (4)  |
|------------------|------|------|------|------|
| Mfg. emp.        | -0.491| 0.329| 1.209| -0.649|
| Unemp.           | (0.266)| (0.155)| (0.347)| (0.571)|
| NILF             | -0.225| 0.014| -0.109| 0.391|
| Wages            | (0.103)| (0.083)| (0.123)| (0.288)|

Notes: This table reports coefficient estimates for versions of the shift-share IV specification in column 3 of Tables 4 and C1, 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. The controls are the same as in column 3 of Table 4. SIC3-clustered exposure-robust standard errors are obtained by the equivalent shock-level regressions and reported in parentheses.
### Table C4: Robustness to Acemoglu et al. (2016) Controls in the Autor et al. (2013) Setting

|                  | (1)  | (2)  | (3)  | (4)  |
|------------------|------|------|------|------|
| Coefficient      | -0.200 | -0.293 | -0.241 | -0.232 |
| (0.093)          | (0.125) | (0.115) | (0.122) |
| Regional controls ($w_{rt}$) |      |      |      |      |
| Autor et al. (2013) controls | ✓  | ✓  | ✓  | ✓  |
| Period-specific lagged mfg. share | ✓  | ✓  | ✓  | ✓  |
| Lagged 10-sector shares | ✓  | ✓  | ✓  | ✓  |
| Local Acemoglu et al. (2016) controls | ✓  | ✓  |      |      |
| Local Acemoglu et al. (2016) pre-trends | ✓  | ✓  | ✓  |      |
| SSIV first stage $F$-stat. | 118.9 | 53.3 | 65.9 | 56.6 |
| # of region-periods | 1,444 |      |      |      |
| # of industry-periods |       | 794  |      |      |

Notes: This table extends Table 4 by adding exposure-weighted sums of the other industry-level controls in Table 3 of Acemoglu et al. (2016). Pre-trends controls refer to the changes in industry log average wages and in the industry share of total U.S. employment over 1976–91; see the notes to Table 4 notes for details on the other controls and calculation of the SIC3-clustered exposure-robust standard errors (in parentheses) and first-stage $F$-statistics.
Table C5: Overidentified Shift-Share IV Estimates of the Effect of Chinese Imports on Manufacturing Employment

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

Notes: Column 1 of this table reports an overidentified estimate of the coefficient corresponding to column 3 of Table 4, obtained from a two-stage least squares regression of shock-level average manufacturing employment growth residuals $\bar{y}_{nt}$ on shock-level average Chinese import competition growth residuals $\bar{x}_{nt}$, instrumenting by the growth of imports (per U.S. worker) in each of the eight non-U.S. countries from ADH, $g_{nk}$ for $k = 1, \ldots, 8$, controlling for period fixed effects $q_{nt}$, and weighting by average industry exposure $s_{nt}$. Column 2 reports the corresponding limited information maximum likelihood estimate, while column 3 reports a two-step optimal generalized method of moments estimate. Standard errors, the optimal weight matrix, and the Hansen (1982) $\chi^2$ test of overidentifying restrictions all allow for clustering of shocks at the SIC3 industry group level. The first-stage $F$-statistic is computed by a shift-share version of the Montiel Olea and Pflueger (2013) method described in Appendix A.11.
Table C6: Bartik (1991) Application

|                              | (1)   | (2)   |
|------------------------------|-------|-------|
| Leave-one-out estimator      | 1.277 | 1.300 |
|                              | (0.150)| (0.124)|
| Conventional estimator       | 1.215 | 1.286 |
|                              | (0.139)| (0.121)|
| $H$ heuristic                | 1.32  | 10.50 |
| Population weights           | ✓     |       |
| # of region-periods          | 2,166 |       |

Notes: Column 1 replicates column 2 of Table 3 from Goldsmith-Pinkham et al. (2020), reporting two SSIV estimators of the inverse labor supply elasticity, with and without the leave-one-out adjustment. Regions are U.S. commuting zones; periods are 1980s, 1990s, and 2000s; all specifications include controls for 1980 regional characteristics interacted with period indicators (see Goldsmith-Pinkham et al. (2020) for more details). Standard errors allow for clustering by commuting zones. Column 1 uses 1980 population weights, while column 2 repeats the same analysis without population weights. The table also reports the $H$ heuristic for the importance of the leave-one-out adjustment proposed in Appendix A.9 (equation (42)).
Table C7: Simulated 5% Rejection Rates for Shift-Share and Conventional Shock-Level IV

|               | SSIV Exposure-Robust SE | Shock-level IV Robust SE |
|---------------|-------------------------|--------------------------|
|               | Null not Imposed        | Null not Imposed         |
| (1)           |                         |                          |
| (2)           |                         |                          |
| (3)           |                         |                          |
| (4)           |                         |                          |

Panel A: Benchmark Monte-Carlo Simulation

(a) Normal shocks 7.6% 5.2% 6.8% 5.0%
(b) Wild bootstrap (benchmark) 8.0% 4.9% 14.2% 4.0%

Panel B: Higher Industry Concentration

(c) $1/HHI = 50$ 5.6% 4.9% 8.4% 6.1%
(d) $1/HHI = 20$ 7.3% 5.5% 7.0% 10.7%
(e) $1/HHI = 10$ 9.0% 8.2% 14.8% 23.8%

Panel C: Smaller Numbers of Industries or Regions

(f) $N = 136$ (SIC3 industries) 5.4% 4.5% 7.7% 4.3%
(g) $N = 20$ (SIC2 industries) 7.7% 3.7% 7.9% 3.2%
(h) $L = 100$ (random regions) 9.7% 4.5% N/A
(i) $L = 25$ (random regions) 10.4% 4.3% N/A

Notes: This table summarizes the results of the Monte-Carlo analysis described in Appendix A.12, reporting the rejection rates for a nominal 5% level test of the true null that $\beta^* = 0$. In all panels, columns 1 and 2 are simulated from the SSIV design based on Autor et al. (2013), as in column 3 of Table 4, while columns 3 and 4 are based on the conventional industry-level IV in Acemoglu et al. (2016). Column 1 uses exposure-robust standard errors from the equivalent industry-level IV and column 2 implements the version with the null hypothesis imposed. Columns 3 and 4 parallel columns 1 and 2 when applied to conventional IV. In Panel A, the simulations approximate the data-generating process using a normal distribution in row (a), with the variance matched to the sample variance of the shocks in the data after de-meaning by year, while wild bootstrap is used in row (b), following Liu (1988). Panel B documents the role of the Herfindahl concentration index across industries, varying $1/HHI$ from 50 to 10 in rows (c) to (e), compared with 191.6 for shift-share IV and 189.7 for conventional IV. Panel C documents the role of the number of regions and industries. We aggregate industries from 397 four-digit manufacturing SIC industries into 136 three-digit industries in row (f) and further into 20 two-digit industries in row (g). In rows (h) and (i), we select a random subset of region in each simulation. See Appendix A.12 for a complete discussion.
| Number of Instruments | 1 | 5 | 10 | 25 | 50 |
|-----------------------|---|---|----|----|----|
| (1)                   |   |   |    |    |    |
| (2)                   |   |   |    |    |    |
| (3)                   |   |   |    |    |    |
| (4)                   |   |   |    |    |    |
| (5)                   |   |   |    |    |    |

**Panel A: SSIV**

|                         | 8.0% | 8.9% | 11.5% | 15.0% | 23.0% |
|-------------------------|------|------|-------|-------|-------|
| 5% rejection rate       | 0.3% | 14.6%| 28.3% | 43.2% | 72.1% |
| Median bias, % of std. dev. | 54.3 | 14.8 | 9.1  | 6.4  | 7.7  |

**Panel B: Conventional Shock-Level IV**

|                         | 13.6% | 13.9% | 14.9% | 17.7% | 22.0% |
|-------------------------|-------|-------|-------|-------|-------|
| 5% rejection rate       | -0.3% | 10.1% | 27.1% | 57.0% | 80.2% |
| Median bias, % of std. dev. | 59.4 | 19.4 | 13.2 | 10.0 | 11.2 |

| Number of simulations   | 10,000 | 3,000 | 1,500 | 500  | 300  |

Notes: This table reports the results of the Monte-Carlo analysis with many weak instruments, described in Appendix A.12. Panel A is simulated from the SSIV design based on Autor et al. (2013), as in column 3 of Table 4, while Panel B is based on the conventional industry-level IV in Acemoglu et al. (2016). The five columns increase the number of shocks \( J = 1, 5, 10, 25, \) and 50, with only one shock relevant to treatment. The table reports the rejection rates corresponding to a nominal 5% level test of the true null that \( \beta^* = 0 \), the median bias of the estimator as a percentage of the simulated standard deviation, and the median first-stage \( F \)-statistic obtained via the Montiel Olea and Pflueger (2013) method (extended to shift-share IV in Panel A, following Appendix A.11). See Appendix A.12 for a complete discussion.
References

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

Adão, Rodrigo, Costas Arkolakis, and Federico Esposito. 2020. “General Equilibrium Indirect Effects in Space: Theory and Measurement.” Working paper.

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

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

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

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

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

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

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. “Bartik Instruments: What, When, Why, and How.” American Economic Review 110:2586–2624.

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

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

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

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

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

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

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

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