NBER WORKING PAPER SERIES

A PRACTICAL GUIDE TO SHIFT-SHARE INSTRUMENTS

Kirill Borusyak
Peter Hull
Xavier Jaravel

Working Paper 33236
http://www.nber.org/papers/w33236

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
December 2024

We thank Matilde Bombardini, David Dorn, Michael Gmeiner, Paul Goldsmith-Pinkham, David McKenzie, Paul Mohnen, Jonathan Parker, Nina Pavcnik, Tim Taylor, and Heidi Williams for helpful suggestions. The views expressed herein are those of the authors and do not necessarily reflect the views of the National Bureau of Economic Research.

A Practical Guide to Shift-Share Instruments
Kirill Borusyak, Peter Hull, and Xavier Jaravel
NBER Working Paper No. 33236
December 2024
JEL No. C18, C21, C26, F16, J21, J61

ABSTRACT

A recent econometric literature shows two distinct paths for identification with shift-share instruments, leveraging either many exogenous shifts or exogenous shares. We present the core logic of both paths and practical takeaways via simple checklists. A variety of empirical settings illustrate key points.

Kirill Borusyak
Department of Agricultural
and Resource Economics
University of California, Berkeley
222 Giannini Hall
Berkeley, CA 94720
and NBER
k.borusyak@berkeley.edu

Peter Hull
Department of Economics
Box B, Brown University
Providence RI 02912
and NBER
peter_hull@brown.edu

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

Many economic studies consider units that are exposed differently to a common set of shocks. Consider, for example, Autor et al. (2013)'s influential study of how the surge in Chinese imports in the 1990s and 2000s affected US local labor markets. They measure regional exposure to this "China shock" by the extent to which workers were employed in industries that saw growing competition with China. This idea is captured by a *shift-share* explanatory variable: the average of national industry-level shifts in US imports from China, weighted by the regional shares of employment across industries. They further construct an *instrumental variable* with a similar shift-share structure: the average of industry growth in Chinese imports among non-US countries, again weighted by industry employment shares of US commuting zones. By using this instrument in a two-stage least squares regression, the authors intend to address potential endogeneity concerns: namely, that US imports from China may be affected by US-specific productivity and demand shocks.

Instruments like these, which sum a common set of shifts with heterogeneous exposure share weights, are often used in studies of labor, trade, macroeconomics, public economics, finance, and more. While such instruments date back at least to Freeman (1975, 1980), the number of papers using them has grown markedly over the last ten years (Goldsmith-Pinkham (2024)). Today, around 1/8th of all instruments featured in NBER working papers are explicitly described as shift-share, while many others implicitly have a shift-share structure.

When do such instruments successfully solve endogeneity concerns, and when might they fail? This question is challenging to answer because shift-share instruments leverage two distinct sources of variation and it is not obvious what properties of each are important. Intuitively, one might view the shifts as helpful because they represent potentially exogenous changes to the system under study. However, these shifts vary at a different level (e.g. industries) than the unit of analysis (e.g. local labor markets). Are they still useful then? In contrast, the shares vary across units but are usually predetermined (e.g., employment shares are measured in a pre-period). So how should their potential exogeneity be understood?

This article gives conceptual answers to these questions and provides practical guidance for using shift-share instruments or assessing the credibility of such instruments when used by others. We build on a recent econometric literature which suggests two distinct paths to identification. One path, developed by Borusyak et al. (2022) and Adão et al. (2019), leverages many exogenous shifts while making no assumption on the exogeneity of the shares. The second path, proposed by Goldsmith-Pinkham et al. (2020), instead focuses on share exogeneity. Each of these two approaches has distinct practical implications regarding appropriate estimators, ways to conduct valid inference, and diagnostic tests.

We begin with a discussion of broad motivations for using shift-share instruments and an overview of the core logic for both paths. We discuss how identification "from the shifts" can be understood as leveraging a shift-level natural experiment, while identification "from the shares" can be viewed as pooling together multiple difference-in-differences designs leveraging heterogeneous shock exposure. We then provide two checklists researchers can follow when implementing a shift-share design, considering the exogenous shifts and exogenous shares approaches in turn. We take an applied perspective throughout, illustrating key concepts and practical steps with examples; see Borusyak et al. (2023a) for a more technical review.

Appendix A answers further practical questions that often arise with shift-share instruments. For instance, we discuss how to interpret estimates as local averages of heterogeneous effects, how to handle multiple instruments and interaction terms, how to approach shift-share instruments where the shifts are measured in-sample (as in Bartik (1991) and Card (2009)) and whether a leave-out construction of the shifts is helpful in those cases.

Shift-Share Basics

What are Shift-Share Instruments and Where do They Come From?

Table 1 lists some prominent examples of shift-share instruments from a variety of settings. We discuss some of these examples in depth below. Here, the table is meant to illustrate some common features of a shift-share research design. Each study seeks to estimate a causal or structural relationship between two variables measured across a set of units i. The *outcome* variable is denoted y_i. Borrowing standard language from the world of causal inference, we refer to the explanatory variable x_i as the *treatment*. For example, Autor et al. (2013) seek to estimate the causal effect of growing exposure to Chinese imports x_i on the growth in local manufacturing employment y_i (among other outcomes) across US regions i. The table shows many other examples of outcomes and treatments across regions, firms, products, and individuals.

To formalize the goal in such settings, consider a model of the form:

$$y_i = \beta x_i + \gamma' w_i + \varepsilon_i, \tag{1}$$

where w_i denotes some vector of observed control variables. Here β is the parameter of interest, capturing the effect of the treatment on the outcome (which for simplicity is assumed to be the same across units). The error term ε_i captures all unobserved determinants of the outcome. We assume throughout that this outcome equation is correctly specified, focusing on consistent estimation of β rather than choosing and interpreting the specification.

Importantly, in writing equation (1), we allow for the possibility of treatment *endogeneity*: i.e., a non-zero correlation between x_i and ε_i. In the Autor et al. (2013) example this allows US regions with more exposure to Chinese imports to have different unobserved labor market conditions which would have led to lower or higher manufacturing employment growth in the absence of the China shock. Such endogeneity introduces bias in ordinary least squares estimates of equation (1). A standard solution to this challenge is to find an *instrument* z_i which is plausibly uncorrelated with the unobserved model error ε_i while nevertheless correlated with the endogenous treatment x_i. The parameter β can then be estimated by two-stage least squares.

The instruments in Table 1 are distinguished by their shift-share structure:

$$z_i = \sum_{k=1}^{K} \underbrace{s_{ik}}_{\text{Share}} \underbrace{g_k}_{\text{Shift}}, \tag{2}$$

where $(g_1, \ldots, g_K)$ is a set of shifts that is common to all units and the $(s_{i1}, \ldots, s_{iK})$ are sets of exposure shares that vary across units. In many applications, the shares sum to one for each observation such that z_i is a share-weighted average of the shifts.

In most of the Table 1 examples, the shifts are defined at a different level k than the units i. For example, Bartik (1991) and Autor et al. (2013) work with regional outcomes and industry-level shifts. Exceptions are studies of network spillovers where k indexes friends or neighbors of individuals or regions. It is also worth noting that while most examples in Table 1 use a shift-share instrument to address endogeneity in a treatment x_i, some (indicated by an asterisk in the treatment column) consider "reduced-form" regressions on z_i itself. We capture this by defining $x_i = z_i$ in such settings.

Researchers might motivate shift-share instruments in different ways. A common motivation arises when the treatment measures the growth of some economic variable over time, and can be decomposed into some start-of-period shares and over-time shifts. Suppose, for example, that $x_i = \frac{X_{i1} - X_{i0}}{X_{i0}}$ is the growth in employment X_{it} for local labor market i over two periods, $t = 0, 1$. Regional employment can be decomposed across industries: $X_{it} = \sum_k X_{ikt}$ where X_{ikt} denotes the period-t employment of industry k in local labor market i. This leads to a decomposition of regional employment growth rates in terms of period-0 industry employment shares and local industry growth rate shifts:

$$x_i = \sum_k \underbrace{\frac{X_{ik0}}{X_{i0}}}_{\text{Share}} \cdot \underbrace{x_{ik}}_{\text{Local shift}}, \qquad \text{for } x_{ik} = \frac{X_{ik1} - X_{ik0}}{X_{ik0}}. \tag{3}$$

A researcher might then construct an instrument by choosing a set of common shifts g_k to replace the local shifts. The shares from the decomposition could be kept or also replaced, e.g. with further lagged shares. Instruments constructed in this way tend to be highly correlated with the treatment.

To illustrate this motivation, consider an example inspired by the canonical shift-share instrument of Bartik (1991). The goal is to estimate the inverse elasticity of regional labor supply β relating wage growth y_i to employment growth x_i across regions i. As usual, to estimate a supply elasticity we need an instrument that shifts labor demand. Decomposition (3) captures the idea that x_i averages local employment growth across different industries, x_{ik}, using initial employment shares $s_{ik} = \frac{X_{ik0}}{X_{i0}}$ as weights. The local shifts reflect changes to both labor demand and labor supply. To isolate demand variation, we can form an instrument that keeps the local industry employment shares from the decomposition but introduces a set of common shifts. The shifts are meant to be predictive of the local industry growth rates while only capturing demand variation. Bartik (1991) defines g_k as national industry growth rates, proxying for aggregate demand shifts. One might also

Table 1: Shift-Share Instrument Examples

Study	Unit (i)	Outcome (y_i)	Treatment (x_i)	Level of shift variation (k)	Instrument (z_i): Share (s_{ik})	Instrument (z_i): Shift (g_k)
Bartik (1991)	Region	Δ Local wage	Δ Local employment	Industry	$\text{Employment}_{ik}/\text{Employment}_i$	National growth of industry employment
Miguel and Kremer (2004)	Individual	Measures of health or education	Number of neighbors selected for deworming*	Individual	Dummy(k is friend of i)	Dummy of deworming treatment
Card (2009)	Region	Relative wage of migrants vs. natives	Relative employment of migrants vs. natives	Origin country	$\text{Migrant stock}_{ik}/\text{Population}_i$	$\text{New migrants}_k/\text{Migrant stock}_k$
Autor et al. (2013)	Region	Δ Local manufacturing employment	Δ Local exposure to Chinese imports	Industry	$\text{Employment}_{ik}/\text{Employment}_i$	Δ imports from China in other countries
Hummels et al. (2014)	Worker	Wage	Imports of intermediate goods by employer	Product-by-country	$\text{Imports}_{ik}/\text{Imports}_i$	Imports from k to other countries
Nunn and Qian (2014)	Country-by-year	Conflict	Quantity of food aid (wheat) from the US	Year	Fraction of years with non-zero food aid	US wheat production in previous year
Cai et al. (2015)	Individual	Takeup of insurance	% of friends selected for an information session*	Individual	Dummy(k is friend of i)/ # of friends i has	Dummy of information session
Jaravel (2019)	Product category	Inflation and innovation	Δ Quantity demanded	Socio-demographic group	Sales of i to group k/ Total sales of i	Population change
Greenstone et al. (2020)	Region	Δ Employment	Δ Credit	Bank	Credit market share of k	Estimated credit supply shock
Aghion et al. (2022)	Firm	Δ Firm employment	Δ Firm stock of automation technologies	Technology-by-country	$\text{Imports}_{ik}/\text{Imports}_i$	Δ imports from k to other countries
Xu (2022)	Region	Δ Exports	Exposure to banking crisis*	Bank	Credit market share of k	Bankruptcy during banking crisis
Franklin et al. (2023)	Local labor market	Wage	Shift-share exposure to the intervention*	Residential neighborhood	$\text{Commuters}_{ik}/\text{Employment}_i$	Dummy of public works intervention
Mohnen (2024)	Region	Δ Young labor market outcome	Retirement rate	Age group (within 45+)	$\text{Population}_{ik}/\text{Population 45+}_i$	National retirement rate at age k

Notes: We simplify many of the settings, suppressing the time dimension (except where it is central to the design), controls and fixed effects, interaction terms, log and other transformations of the outcome and treatment, etc. Asterisks (*) indicate ordinary least squares regressions, in which the treatment itself is the shift-share with shares s_{ik} and shifts g_k.

define g_k as specific industry demand shifts, such as a change in government subsidies.

Decomposition (3) is helpful for illustrating why the shares in the definition of z_i often, but not always, sum to one for each observation. In the previous example, regional employment shares mechanically sum to one across industries. However, sometimes the instrument is constructed from shifts that could only happen in a subset of industries: say, within the manufacturing sector. Only those industries would appear in the shift-share instrument formula, and the shares would add up to a number smaller than one. We discuss the importance of this below.

Decomposing the treatment is not the only way to arrive at a shift-share instrument. Another common way is by "apportioning" some national changes to units. Appendix A.5 illustrates this approach and shows how it relates to the decomposition above. In still other cases, an instrument naturally takes a shift-share form. For instance, many reduced-form studies of how shocks propagate across a network (e.g., Cai et al. (2015)) use the fraction of unit i's friends or neighbors who have been selected for some intervention. This variable inherently has a shift-share structure: $z_i = \sum_k s_{ik} g_k$ where g_k is a dummy variable indicating that k has been selected and s_{ik} is a dummy variable indicating that k is a friend of i, scaled by the number of friends i has.

Regardless of the motivation, the core challenge in using such z_i is to argue convincingly that it is *exogenous*: i.e., uncorrelated with the model unobservable ε_i. Such arguments are typically made from contextual knowledge about the source of variation in the instrument. The unique challenge with shift-share instruments is that there are two distinct sources of variation: the shifts and the shares. Thus, to argue convincingly that these instruments are exogenous, one must explain what properties of the shifts and shares make z_i uncorrelated with ε_i (rather than simply stating the basic exogeneity restriction of $\text{Cov}\left[z_i, \varepsilon_i\right] = 0$). We next introduce the two paths for making such arguments.

What is Identification from Many Exogenous Shifts?

One strategy to ensure that the shift-share instrument z_i is exogenous is to have exogenous g_k. For example, imagine a lottery that randomly assigns a subsidy level g_k to each industry k. In the above labor supply setting, local employment growth x_i can be instrumented for by a weighted average of the subsidies, using initial local employment shares as weights. Subsidies can be viewed as only affecting wages by shifting labor demand and do not have direct effects on labor supply. In general, exogenous shifts should be as-if randomly assigned and should only affect the outcome through the treatment (an *exclusion restriction*).

Shift-based identification stems from a simple observation: a share-weighted average of random shifts is itself as-good-as-random. This is true even if the shares are econometrically endogenous, in the sense that units with different shares may have systematically different unobservables. For instance, regions that specialize in high-skill intensive industries may experience more immigration from certain countries, such that the total employment share of high-skill intensive industries positively correlates with unobserved immigration shocks in the error term. But as long as subsidies are assigned at random across many high- and low-skill intensive industries, on average the places specializing in high-skill intensive industries will have typical values of the instrument. Thus, a shift-share research design based on experimental shifts requires no assumptions on the exogeneity of exposure shares.

While a lottery provides intuition for an idealized experiment, the necessary and sufficient condition for instrument exogeneity is a weaker condition on the shifts: g_k should be uncorrelated with an average of ε_i taken across units with weights s_{ik}. In our running example, this would mean that subsidies g_k—even if not truly randomized—are not systematically higher or lower in industries which are concentrated (in terms of employment shares s_{ik}) in regions with high vs. low labor supply shocks ε_i. Violations of this condition are the key threat to identification in the exogenous shifts approach.

Another way to understand the exogenous shifts approach is to view the shift-share instrument as a "translation device" for a set of as-good-as-random shifts to a different level of analysis. For instance, when industry subsidies are as-good-as-randomly assigned, one could imagine running an industry-level analysis which uses the subsidy g_k directly as an instrument for industry employment. Specifying the equation at the level of local labor markets may define a more interesting economic parameter, capturing spillovers when workers move across industries in response to the subsidies. However, the key identification assumption is the same, with the shift-share instrument translating the industry-level natural experiment to local labor markets.

The "weighted average of lotteries" logic highlights two other requirements of the exogenous shifts approach. First, it requires many shifts $g_1, \ldots, g_K$. Otherwise, if K is a small number, the shifts may by chance be correlated with unobservables even if they are truly random.[1] This can be viewed as an instance where the law of large numbers does not apply: there are effectively only a few exogenous comparisons, regardless of how many units are observed.

Second, the shares have to add up to one such that the shift-share instrument has an interpretation as a share-weighted *average* of shifts rather than a share-weighted *sum*. Otherwise, even if shifts are drawn fully at random, the instrument may systematically vary across units through the sum of shares. We discuss below how the exogenous shifts approach extends in this "incomplete shares" case.

What is Identification from Exogenous Shares?

A different strategy to ensure shift-share exogeneity is to have exogenous shares. What does this mean exactly? One could imagine the set of s_{ik} being as-good-as-randomly assigned to units, as if drawn in a lottery, and satisfying an exclusion restriction (that the shares affect the outcome only via the treatment of interest). Alternatively, when the outcome is measured in changes, one may interpret share exogeneity as a set of parallel trends conditions similar to ones used in difference-in-differences strategies. That is, for s_{ik} to be uncorrelated with ε_i one could assert that—if not for any change in the treatment—outcomes would have trended similarly for units that were more vs.

[1] We note that it does not matter whether the g_k take many distinct values. For instance, assigning a 10% subsidy to some of the many industries (and 0% to the rest) should be viewed as having many shifts.

less exposed to k. Shares are exogenous when such parallel trends conditions hold for each k.

To make this logic concrete, consider an example inspired by Card (2009) who estimates the (inverse) elasticity of substitution between migrant vs. native workers in labor demand, β. Here the model (1) relates changes in the relative wages of migrants vs. natives between two periods, y_i, to changes in the relative employment of these groups, x_i, across local labor markets. Suppose that between these periods we saw a sudden change in national migrant inflows from a particular origin country κ, such as the sudden inflow of Cuban immigrants following the Mariel Boatlift studied by Card (1990). One might be willing to assume that regions which were more or less exposed to this inflow, as captured by the initial share of migrants from Cuba $s_{i\kappa}$, would have seen similar trends in labor demand for migrant vs. native labor: i.e., that $\text{Cov}\left[\varepsilon_i, s_{ik}\right] = 0$. In this case, the Cuban migrant share would be a valid instrument for identifying β.

Under such share exogeneity, shift-share instruments can be viewed as combining multiple valid share instruments—each operating under the same difference-in-differences logic, but capturing different exposure variation.[2] Indeed, Goldsmith-Pinkham et al. (2020) show that shift-share estimates can generally be viewed as pooling together K "one-at-a-time" estimates each using a single s_{ik} share as the instrument. In the above example this would mean sudden changes in migrant inflows across many origin countries, to different extents. In this case, if a parallel trends condition holds with respect to each exposure share, a shift-share instrument combining them with g_k weights will also be a valid instrument.

Thus, the exogenous shares approach is appropriate when a researcher is comfortable using any of the individual shares as an exogenous instrument. The plausibility of share exogeneity depends on whether there are conceivably any unobserved shocks that affect the outcome via the same (or similar) shares as the ones used to construct the instrument. Even if shares are drawn at random from a lottery, the presence of any such shocks would always lead to parallel trend violations.

The plausibility of share exogeneity is boosted by constructing the instrument with shares which are "tailored" to the treatment of interest, in the sense of mediating only the shocks to x_i and not a broad set of shocks that might affect y_i. For example, in the literature on the effects of immigration (e.g. Card (2009)), exposure shares are tailored to the research question: they measure local migration from various origins in the past. This scenario can be contrasted with popular shift-share designs with shares reflecting local industrial composition while studying the regional impacts of specific industry shifts, such as import competition with China in Autor et al. (2013) or robotization in Acemoglu and Restrepo (2020). The industry employment shares are "generic," in that they could potentially measure an observation's exposure to other shocks (essentially, any industry shock), many of them unobserved. In studies using such shares, it would not be plausible to make a case for identification based on the exogeneity of shares. Under the exogenous shares view, Autor et al. (2013) and Acemoglu and Restrepo (2020) use essentially the same instruments (lagged employment

[2]To see this link, note that we can formalize the above example as a setting with only one non-zero immigration shift: i.e. $g_\kappa \neq 0$ and $g_k = 0$ for all other host countries $k \neq \kappa$. The resulting shift-share instrument $z_i = \sum_{k \neq \kappa} s_{ik} \cdot 0 + s_{i\kappa} g_\kappa = s_{i\kappa} g_\kappa$ is perfectly collinear with $s_{i\kappa}$, so using this single exposure share as the instrument will produce numerically the same estimate.

Table 2: Summary of Main Practical Takeaways

	Approach	
	Many exogenous shifts (1)	**Exogenous shares** (2)
Identification argument	Shifts are as-good-as-randomly assigned and only affect the outcome through the treatment	Each share satisfies parallel trends: the outcomes of units with high vs. low shares would have trended the same if not for the treatment
Estimation	Control for the sum of shares (if not one) and shift-share aggregates of any shift-level controls	Check robustness to using share instruments directly: e.g., one share at a time or pooled via two-stage least squares or limited information maximum likelihood
Statistical inference	Get exposure-robust standard errors from the equivalent shift-level instrumental variable regression	Use conventional heteroskedasticity- or cluster-robust standard errors
Balance tests	For both the shift-share instrument and the shifts	For both the shift-share instrument and the shares with high Rotemberg weights
Do not use when...	You would not want to use the shifts directly as an instrument in a shift-level regression, e.g. because they are too few or endogenous	You would not want to use shares directly as instruments, e.g. because they are "generic" (capturing the unit's exposure to many types of shocks)

shares) for different treatments.

The role of the shifts is secondary with the exogenous shares strategy: Goldsmith-Pinkham et al. (2020) show that the shifts affect the weights in their representation of shift-share as pooled one-at-a-time share-instrument estimates, but they do not affect the identification of β so long as the shares are exogenous. The choice of g_k, however, may affect the power of the shift-share instrument. Intuitively, the decomposition (3) suggests that a powerful instrument might use as the g_k the average of shifts x_{ik} across units (e.g. replacing the local growth rates of industry employment x_{ik} with the national ones g_k).

Many Exogenous Shifts in Practice

We now describe a list of practical steps for applying shift-share designs with many exogenous shifts. This checklist can also be instructive for assessing the design of existing papers using shift-share instruments. Column 1 of Table 2 summarizes some of the main practical takeaways discussed in this section. We illustrate the checklist in the labor supply setting from above, where g_k represents as-good-as-randomly assigned federal subsidies to industries k. At the end of this section, we discuss several real-world examples.

A Checklist for the Shift-Based Approach

1. Motivate the shift-share strategy with a shift-level idealized experiment Any compelling instrumental-variable design begins with thinking about what endogeneity bias is being addressed: i.e., exactly which unobserved variables (or *confounders*) are likely to bias simple ordinary least squares estimation. For example, when attempting to estimate a labor supply equation with data on local employment growth x_i and local wage growth y_i, the model error ε_i will include unobserved local labor supply shocks (e.g., immigration of foreign workers to each region). Because equilibrium employment growth arises from both labor supply and labor demand shocks it is generally correlated with ε_i, generating bias in ordinary least squares estimates. To estimate β, we need an instrument which is uncorrelated with local labor supply changes.

Once potential confounders are specified, the researcher can describe a hypothetical shift-level experiment which would generate shifts that are unrelated to these sources of bias while nevertheless generating variation in the treatment. For example, one can imagine assigning new federal subsidies at random across industries. Industries receiving larger subsidies are likely to expand their production and thus their demand for local workers, increasing local employment x_i. By virtue of random assignment, these subsidy shifts are unrelated to local labor supply conditions. The experimental ideal is thus useful to clarify exactly the type of shift-level variation one would want for identification.

2. Bridge the gap between the observed and ideal shifts The next step is to describe how the actual shift-share design used for the empirical analysis approximates the idealized experiment. This may involve (i) specifying some control variables and (ii) describing how observed shifts proxy for the ideal ones.

In our running example, imagine changes in subsidies g_k are not randomized across industries and could provide shift-level variation analogous to the randomized subsidies only conditional on some controls. There could be two types of such controls, depending on whether shift-level or unit-level confounders motivate including them. For the former, one may consider shift-level observables q_k that both correlate with the g_k and can have a direct impact on the outcome of interest. For example, one might worry that subsidies are systematically larger in skill-intensive industries and that immigration from skill-abundant countries shift labor supply in regions where those industries concentrate. In this case, one would like to control for the indicator of skill-intensive industries in the shift-share specification. But how can this be done if such q_k vary at the industry level while the specification is estimated at the regional level? Borusyak et al. (2022) show that the answer is to control for $\sum_k s_{ik} q_k$: shift-share aggregates of the industry-level confounders, with the same exposure shares as in the construction of the instrument. In the skill intensity example, this amounts to controlling for the total regional employment share of all skill-intensive industries. With this control variable, the shift-share instrument will only leverage the variation in g_k which is uncorrelated with the q_k: e.g., residual variation in subsidies after controlling for skill intensity.

Controls of the second type arise from unit-level observables which are thought to correlate both with the error term ε_i and with z_i. For example, one might expect labor markets in the US "rust belt" to experience different unobserved local labor supply shocks vs. other parts of the country, and that industries more concentrated in these states see systematically different subsidies. In this case, a straightforward solution is to control for a rust belt indicator.[3]

Even after including the controls, the shifts may only be viewed as proxies for idealized ones. In Autor et al. (2013), for example, industry-level productivity shifts in China are unobserved but proxied with the growth of imports from China in non-US countries. In Bartik (1991), labor demand shifts are proxied with national employment growth rates. In those cases, the applicability of the exogenous shifts approach depends on whether the gap between the proxy and ideal shift could be contaminated by confounders. In Appendix A.11 we show that this problem arises in Bartik (1991).

3. Include the "incomplete share" control In shift-share designs where the exposure shares s_{ik} do not add up to one—what Borusyak et al. (2022) call the "incomplete shares" case—a special control must be included: the sum of shares, $S_i = \sum_k s_{ik}$. To build intuition, recall that with "complete" shares (when $S_i = 1$), the shift-share instrument is a weighted *average* of the shifts so if shifts arise from a pure lottery then the shift-share instrument is also like a lottery outcome. This logic breaks down with incomplete shares, when z_i is a weighted *sum* of the shifts. Then, even with randomly assigned shifts—which have, say, a positive mean—an observation with a higher S_i would systematically get higher values of the instrument. The instrument is thus correlated with the sum of shares, which can in turn be correlated with the error, leading to bias. Controlling for S_i removes the problem, because units with the same S_i get different values of the shift-share only for random reasons.[4]

4. Lag shares to the beginning of the natural experiment When constructing the shift-share instrument, one needs to decide when to measure the shares. Decomposition (3) suggests measuring them at the beginning of the period of interest, but it is common in practice to lag them further. Is this practice justified?

In the exogenous shifts approach, it is best to measure the shifts at the beginning of the natural experiment that generates them. This avoids the situation where the shifts affect the shares, potentially generating bias.[5] At the same time, shares matter for instrument power; lagging shares

[3] An alternative way would be to consider industry-level controls as described above: e.g., the share of Rust Belt regions in the industry employment. The relative merits of these two approaches remain unexplored.

[4] In Appendix A.4 we explain why this solution is typically better than renormalizing the shares to add up to one. We also note that sometimes researchers introduce the shares that add up to one across *observations* instead of shifts. In such a case, it is not appropriate to control for S_i. Instead, the shift-share instrument should be rewritten in a different way consistent with (2); see Appendix A.5.

[5] Not every response of the shares to past shifts makes the shift-share instrument endogenous: if this response is not related to the error terms, there is no problem. But it is possible to imagine situations where the bias would arise. Following footnote 32 in Borusyak et al. (2022), consider the labor supply setting and imagine that subsidies now occur in two periods. Suppose regions vary in labor market flexibility, the reallocation of employment towards industries with larger subsidies is stronger in flexible local labor markets. If subsidies are random but persistent across the two periods, industries with large subsidies will be increasingly concentrated in regions with flexible labor markets. The shift-share instrument will therefore take higher values in flexible labor markets, causing bias if flexible labor markets also have stronger employment growth for other reasons.

beyond what is necessary would typically make the instrument weaker.

What constitutes the beginning of the natural experiment? If there were no shifts correlated with g_k in the past, it is just the beginning of the period when the g_k are measured. However, if the shifts unfold over several periods in a serially correlated way, it is appropriate to lag the shares further, to the first of these periods—or alternatively to extract unpredictable shock innovations and use them to construct the shift-share instrument. Another problem that arises with serially correlated shifts is that past shifts may have direct dynamic effects on the current outcomes (see Jaeger et al. (2017)). Simply lagging the shares does not help with this problem. We discuss the problems arising in panels and possible solutions in Appendix A.1.

5. Report descriptive statistics for shifts in addition to observations Empirical papers normally present the number of observations and summary statistics for the main variables. In shift-share analyses that leverage the variation in shifts, it is important to also present such descriptive statistics for these—in the same way as one would in a non-shift-share setting at the shock level. While the mean and standard deviation of z_i is useful to know, so are the mean and standard deviation of g_k.

One detail here is that, as we show below, each shift carries an importance weight proportional to the exposure share of that shift for an average observation, $s_k = \frac{1}{N}\sum_i s_{ik}$. For example, when studying subsidy shifts across industries, the importance weights could correspond to the average industry employment share across local labor markets. Thus, it is natural to report descriptive statistics with those weights as well. For instance, the weighted version of the number of shifts is the "effective number of shifts"—the inverse of the Herfindahl index of shock importance weights, $1/\sum_k s_k^2$. When the effective number of shifts is small, a few shifts may drive the empirical analysis, potentially making the results noisy and unreliable. This is not specific to shift-shares: a similar issue can arise when running a weighted ordinary least squares regression, if some observations get disproportionately large weights.[6]

Descriptive analyses for the shifts need not be limited to their effective number and the distribution. For instance, one could also describe the distribution of the shifts after residualizing them on shift-level controls the researcher plans to include. Or one could plot the shifts on the map if they have a geographic dimension.

6. Implement balance tests for shifts in addition to the instrument In every research design, it is useful to perform balance tests: specifically, to check that the variation believed to be exogenous is indeed not correlated with proxies for confounders. In a shift-share design with exogenous shifts, this can be done in two ways: for the instrument at the level of units, and also directly for the g_k at the level of shifts.

Checking balance of the instrument at the unit level is relatively standard. For instance, a typical pre-trend test involves regressing the lagged outcome on z_i while including the controls picked in

[6] If shifts are correlated within certain clusters, the Herfindahl index can be computed at the level of such shift clusters, since having many correlated shifts may also not be enough for a reliable statistical analysis.

advance (such as the incomplete share control). The only particularity of shift-share designs in this case is that standard errors should be computed appropriately, as we discuss in the next step.

But when the identifying variation is at the shift level, it is also useful to check balance of shifts directly, with respect to shift-level variables that may proxy for unobservables. For example, in our running example with a change in industry subsidies, one could check whether the shifts correlate with variables reflecting labor supply factors, such as the composition of the workforce and the share of immigrants in the industry. This test is useful to assess whether changes in subsidies are systematically different for certain industries that would likely have been on different employment trends even absent changes in labor demand.

7. Produce the main estimates with correct standard errors and check sensitivity Valid statistical inference in shift-share designs with exogenous shifts requires a special "exposure-robust" approach. Intuitively, inference must take into account that units with similar shares mechanically have correlated z_i and may also have correlated ε_i due to their common exposure to unobserved shocks. For example, regions specializing in the same industries will be affected by the same (potentially unobserved) industry shocks. Adão et al. (2019) show with a Monte Carlo simulation that this issue can be very serious in practice.

Two solutions have been developed, both leveraging as-if random assignment of the shifts. First, Adão et al. (2019) provide a variance estimator which is asymptotically valid regardless of the correlation structure of the errors across observations, as long as the exogenous shifts are mutually uncorrelated or clustered in a known way (e.g., by group of industries). Second, Borusyak et al. (2022) show that one can simply run a particular shift-level two-stage least squares regression which always produces an identical coefficient as $\hat{\beta}$ from the shift-share regression (1) but gives valid standard errors, since it is estimated at the same level at which the shifts are assigned. In this regression, the k-level outcome and treatment are certain transformations of the original outcome and treatment, shifts g_k directly serve as a single instrument, shift-level controls q_k are directly included as controls, and estimation is weighted by average shares $s_k = \frac{1}{N}\sum_i s_{ik}$.[7] The `ssaggregate` packages in Stata and R automate the transformation of the outcome and treatment for this regression. The shift-level regression offers the flexibility to accommodate various types of dependence in the shifts: e.g., not only standard clustering but also spatial clustering and serial correlation. The equivalent regression can also be used to produce exposure-robust first-stage F-statistics to judge the instrument strength.

After producing the main shift-share estimates, it can be instructive to check their robustness to a variety of choices. For example, one may examine the stability of the estimate under alternative sets of controls which could correspond to different assumptions of conditional quasi-random shift assignment. Similarly, one may check that estimation with and without unit-level importance weights (e.g., population weights in a regional analysis) yields similar results.

[7] Specifically, the transformation of the outcome and treatment involves first residualizing them on the included i-level controls and then, for each k, averaging across observations with weights s_{ik}.

Examples of the Shift-Based Approach

We now discuss two examples, which illustrate some of the key practical insights for shift-shares with exogenous shifts. The first example focuses on how to use the shift-share design with a true experiment. The second describes a shift-share design with quasi-experimental shifts and illustrates why "incomplete shares" deserve special attention. Appendix A.1 provides an additional example leveraging time-series variation in the shifts.

Shift-share in a randomized trial Franklin et al. (2023) leverage randomized shifts in a shift-share design to estimate the indirect impacts of an intervention. They study a large public works program offering employment at high wages to low-income workers residing in specific neighborhoods in Addis Ababa, Ethiopia. The authors estimate the impact of this program on private sector wages: by increasing employment in public works, the program can reduce labor supply for other activities and increase private wages. Identification relies on the randomized rollout of the program, and the authors find large wage effects.

While the program is randomized at the level of residential neighborhoods k, it may have spillovers on wages in other neighborhoods (labor markets i) because workers can commute. Using data on the baseline-period commuting data, Franklin et al. (2023) build a measure of each labor market's exposure to the randomized roll-out: for each labor market, the shift-share treatment takes an average of intervention dummies across places of residence (the shifts) weighted by the share of workers who commute from those places of residence (exposure shares which sum to one).

In this setting, if the shifts are simply randomly assigned, there is no need to introduce controls. Imagine, however, that some residential neighborhoods k were ineligible for randomization. Then, the total share of commuters from eligible areas is less than one, and controlling for this total is necessary. With this control, and assuming that commuting shares correctly capture the structure of spillovers, the shift-share design identifies the causal impact of the program.

Shift-share without an experiment Autor et al. (2013) study the impact of import competition with China on US employment. While this relationship could be analyzed across industries, they adopt the "local labor markets approach" (following, e.g., Topalova (2010)). Simplifying details, they define the outcome as the employment change in a US local labor market (commuting zone) and the treatment as the change in local exposure to import competition. Local exposure is measured as the average of national industry changes of imports from China (in dollars per US worker), weighted by local employment weights of different industries.[8] Ordinary least squares estimates may be biased if, for example, high productivity growth in China happens in industries with systematically different productivity or demand trends in the US or if US consumers substitute to Chinese goods in industries where US productivity is lagging.

[8] This approach is meant to account for important spillovers across industries: if workers can move from an industry affected by import competition to another one, declines in industry employment are not informative of the aggregate effects of import competitions. Spillovers across commuting zones are likely more limited.

In this setting, the idealized experiment would be to assign observed productivity shifts at random across manufacturing industries in China. These shifts would have different incidence across US commuting zones given the pre-determined industrial composition of each area. In practice, productivity changes are unobserved and, as a proxy, Autor et al. (2013) use the observed growth of imports from China in industry k in eight high-income countries excluding the US. Measuring imports in those countries ensures that demand and supply shocks that are idiosyncratic to the US cannot bias the results.

An important feature of this setting is that the exposure shares do not sum to one, since only manufacturing industries are exposed to trade with China. Locations with a larger total share of employment in manufacturing are likely on different potential outcome trends, e.g. because of the secular decline in manufacturing (which can have many causes other than trade). To address this issue, it is necessary to control for the sum of exposure shares in each location. Note that the appropriate control equals the total regional share of manufacturing employment in the period in which the shares are measured. Since Autor et al. (2013) lag the shares by a decade relative to the period of the outcome and treatment, the incomplete share control should be lagged as well.

A further adjustment is called for because Autor et al. (2013) conduct the analysis in a repeated cross-section over two ten-year periods, and the average shifts are different in the two periods. Here, leveraging shock variation across industries only within periods requires controlling for the interaction of the sum of exposure shares with period indicators. This control prevents the bias that would arise if the manufacturing sector as a whole (and regions specializing in manufacturing industries) declined at different rates in the two periods for reasons unrelated to trade.

To assess the plausibility of the design, it is instructive to conduct industry- and commuting zone-level balance tests. At the industry level, it could be that China specializes in certain industries (e.g., low-skill industries) that could have been on different employment trends in the US absent trade shocks. To speak to this concern, one can correlate the shift g_k with industry-level variables reflecting the structure of employment and technologies—such as the skill and labor intensity, average wages, and investment in new technologies (e.g., computers) in a pre-period. Using the data from Autor et al. (2013), Borusyak et al. (2022) find that the shifts are balanced across these dimensions.

Correlating the regional shift-share instrument with potential commuting zone-level confounders is also instructive. Such a correlation would arise if China specializes in industries that are located in commuting zones with unusual observed characteristics, which can raise concerns they are on different potential employment trends, too. One can regress commuting zone-level predetermined variables—such as the lagged fraction of population who is college-educated, foreign-born, female, or working in routine occupations—on the shift-share instrument, controlling for the sum of shares interacted with period fixed effects. One can also implement a standard "pre-trend" test, with lagged commuting zone outcome on the left-hand side. Autor et al. (2013) and Borusyak et al. (2022) find that most of these tests pass.

Using this shift-share strategy, Autor et al. (2013) document a substantial decrease in both manufacturing employment and total employment in local labor markets that were more exposed to import competition from China. Introducing incomplete share controls interacted with period indicators, Borusyak et al. (2022) find smaller effects, especially for total employment.

Exogenous Shares in Practice

We now provide a list of practical steps to determine whether and how to use the exogenous-share approach to shift-share designs. As before, these steps can also serve as a blueprint for readers of papers using these designs. A summary of the key takeaways is given in Column 2 of Table 2. We develop this checklist with the immigration setting from above, where s_{ik} represents the lagged share of immigrants from country k in region i. We discuss several applied examples at the end of this section.

A Checklist for the Share-Based Approach

1. Determine whether the exposure shares are potentially suitable instruments Like before, the researcher can start by motivating the outcome equation and describing the main sources of treatment endogeneity. With the exogenous shares approach in mind, the researcher would then provide reasons why the shares may be useful instruments to address the corresponding threats. While we illustrate how this can done with detailed empirical examples below, here we highlight two general guiding principles.

First, the instrument exogeneity argument requires shares to be "tailored" to the treatment. Recall that the shares cannot be exogenous instruments if they capture the exposure of the outcome to some unobserved shocks. This rules out cases where the shares are "generic," in the sense of capturing exposure to many shocks, while the treatment only captures one such mechanism (e.g., import competition in Autor et al. (2013)) and it is not feasible to control for the effects of all other shocks. Conversely, in our running migration example, it is conceivable that the share of migrants from a certain origin only captures the region's exposure to migration shocks, making such shares potentially exogenous.

Second, the identification strategy can be strengthened by exploiting a source of variation in the initial shares that is more likely to satisfy exogeneity. Terry et al. (2023), for instance, study the effects of migration on innovation and worry that the initial composition of migrants may be correlated with labor demand factors. For instance, migrants from certain origins may have settled in regions with persistently strong labor demand, which could directly impact innovation. They address this issue by replacing the shares with their component arising from specific historical quasi-experiments, leveraging how the timing of historic waves of immigration coincided with the timing of growth across US regions.

A simpler strategy of lagging the shares can also sometimes help, but it does not by itself guarantee exogeneity. Lagging the shares will typically weaken the instrument, so it's important to explain why it is plausible that lagging the shares reduces their covariance with the error term by more than it reduces the covariance with the treatment. For example, in studies of the effect of immigration on local labor markets, lagging the shares to an earlier decade is helpful if (i) labor demand shocks that attract migrants are transitory, and (ii) new migrants persistently go to places where migrants from the same origin arrived earlier.

2. Choose the necessary unit-level controls Even if the shares are tailored, their exogeneity is a nontrivial assumption, similar to any parallel trends assumption. As usual, exogeneity can be relaxed by including control variables. For example, the researcher can control for certain sums of shares to only leverage share variation conditional on these sums. In the migration setting, controlling for the initial total immigrant share would mean that the shift-share would leverage variation in the composition of migrants across locations, avoiding comparisons between regions with high and low migration intensity overall.

3. Characterize which shares matter the most for the estimates When viewing the shift-share estimate as a pooled version of K one-at-a-time share-instrument estimates, it can be important to understand whether a small subset of these instruments drive the results. If this is the case, the researcher can use those shares to explain how the identification strategy works and can focus on them for the balance tests described below.

Goldsmith-Pinkham et al. (2020) show how to measure the importance weight of each share instrument, which they refer to as "Rotemberg weights" (referencing Rotemberg (1983)). They are based on a decomposition of the shift-share estimator into a weighted sum of individual-share-instrument estimators with weights that add up to one, although some can be negative. These weights are larger for shares that are exposed to a bigger g_k and that are more predictive of treatment. The Rotemberg weights can be interpreted as measuring the sensitivity of the shift-share estimate to violations of exogeneity by each share instrument. The `bartik_weight` command in Stata and R provided by Goldsmith-Pinkham et al. (2020) computes these weights.[9]

4. Implement balance tests for individual shares in addition to the instrument Like in any design, it is worth checking balance of the instrument on observable variables that may be expected to correlate with the error term. Different variables can serve for useful balance tests: pre-period changes in the outcome variable (corresponding to a pre-trends test), unit characteristics measured at the beginning of the period, or contemporaneous changes in placebo outcomes that are not expected to be causally affected by the treatment.

The special feature of the exogenous-share approach is that balance tests can also be performed on individual shares, since each of them is assumed to be exogenous. To avoid issues with testing many hypotheses, it is natural to focus on the subset of shares that are most important for the resulting estimate as measured by the Rotemberg weights. We note that outcome pre-trends are more likely uncorrelated with individual shares when either the shares have changed drastically

[9] Unfortunately, the Rotemberg weights are not unique when the shares add up to one. This is because the shares—and thus individual-share estimators—are perfectly multicollinear. In this case, Goldsmith-Pinkham et al. (2020) recommend choosing the Rotemberg weights that correspond to the demeaned shifts.

since the pre-period or there were no shocks of the same nature as g_k in the past (see Jaeger et al. (2017)).

5. Check sensitivity to how share instruments are combined When shares are exogenous, the parameter β is *overidentified*: any individual share or linear combination of shares is itself a valid instrument. The shift-share instrument is one such combination but, since many others are available, it is instructive to check that the shift-share estimate would not be too dependent on the researcher's choice. Here we review several such tests and discuss what their failure may indicate.

A standard statistical test for whether using each of the individual shares as an instrument yields statistically indistinguishable estimates of β is the Sargan-Hansen overidentification test. (Wooldridge, 2002, Ch. 6.2.2). Graphical procedures aid the interpretation of this test. A conventional "visual instrument variable" procedure (Angrist and Pischke, 2008, p.103) plots K reduced-form coefficients (from regressions of the outcome on a given share, including all controls) against corresponding first-stage coefficients (from similar regressions of the treatment). Since individual-share estimates of β are given by the ratio of reduced-form and first-stage coefficients, the points in this plot should all lie on a single ray from the origin when all share instruments estimate the same parameter (i.e. when the overidentification test "passes").[10] An alternative graph is proposed by Goldsmith-Pinkham et al. (2020): a scatter plot of the K estimates of β, each using one of the shares as an instrument, against their respective first-stage F-statistics. Here one hopes to see that all estimates are similar, especially those with high F-statistics and large Rotemberg weights.

Since individual share instruments may not be very strong, it is also useful to check the sensitivity of the β estimates to alternative combinations of multiple share instruments. One may examine whether the estimate changes when using only a few shares—e.g., those with the largest Rotemberg weights. Another approach is to keep all shares for higher precision but combine them in a different way. When K is small relative to the sample size, two-stage least squares is the natural estimator to report; an efficient generalized method of moments estimator is another option. With many shares, two-stage least squares suffers from bias but several estimators robust to "many weak instruments" are available instead: jackknife instrumental variables (Angrist et al. (1999)), limited information maximum likelihood, the heteroskedasticity-robust Fuller estimator (Hausman et al. (2012)), and modified bias-corrected two-stage least squares (Kolesar et al. (2015)).[11]

It is comforting if all of the above checks indicate robustness of the shift-share estimate; but what should the researcher conclude if not? The answer depends on whether the causal effect of x_i on y_i can vary across units i (making the constant-β model (1) misspecified). When the effects are homogeneous, the failure of the above tests indicates that the share exogeneity assumption is violated. This need not be the case with heterogeneous effects, as different combinations of share instruments may estimate different combinations of causal effects even when all share instruments

[10] Formally, Appendix B.2 shows that the shift-share estimate of β equals the coefficient from a regression of reduced-form coefficients on first-stage coefficients, with no intercept and with particular weights related to the Rotemberg weights.

[11] The shift-share estimator also requires a similar bias correction when the shifts are estimated from the sample, as in Bartik (1991). In Appendix A.12 we discuss the leave-out shift-share estimator that helps in this scenario.

are exogenous. Still, sensitivity of the estimates to the choice of the instruments is a cause for concern: the interpretation of shift-share estimates (and those from the alternative estimators) can be challenging under such effect heterogeneity. They may not represent the average effect for some subpopulation of "compliers," especially because share instruments are correlated with each other (Mogstad et al. (2021)).

Examples of the Share-Based Approach

We now use two examples to illustrate the exogenous-share approach in the contexts of the labor market responses to migration and retirement rates. Both examples show the tight conceptual link between the exogenous shares approach and difference-in-differences research designs.

Labor market effects of immigration We first consider the design of Card (2009, Table 6) and its re-analysis by Goldsmith-Pinkham et al. (2020). The goal is to estimate the (inverse) elasticity of substitution between immigrant workers and native workers in labor demand, i.e. the relationship between the log wage gap between immigrant and native workers and the ratio of immigrant to native hours worked. Simplifying details, the analysis considers a cross section of outcomes (in levels) in 2000 across 124 cities, separately for high-school and college-educated workers.

As with any demand equation, ordinary least squares estimates may be biased: a positive labor demand shock for migrants would draw more immigrants into a location and at the same time increase their wages relative to natives. An instrument is needed that shifts the relative supply of migrant and native workers. Card (2009) proposes a shift-share instrument, leveraging immigration patterns from 38 countries indexed by k. Here s_{ik} is the share of immigration group k in the population of city i in 1980; note that these shares add up to the initial migration share rather than one (and the initial migration rate is not controlled for). The g_k is the number of migrants in group k moving to the US from 1990 to 2000, normalized by the national stock of migrants from k already in the US in 1990.

This shift-share strategy alleviates some endogeneity concerns, as the shares are uncorrelated with some relative labor demand factors. Specifically, transitory regional labor demand shocks (which attract migrants to a particular location in the current period only) would be a problem for ordinary least squares but not for Card's instrument, since the migrant shares are measured before these shocks are realized. In contrast, persistent regional labor demand factors (e.g., characteristics that always make immigrants more productive relative to native workers, such as the prevalence of certain languages like Spanish) would remain a problem for both ordinary least squares and the shift-share approach, since these factors impact the beginning-of-period migrant share while also entering contemporaneous labor demand in the error term.

Some of the potential limitations of Card's instrument can be addressed by simple adjustments to the empirical strategy. In particular, estimating the outcome equation in differences would alleviate concerns about time-invariant regional confounders. Moreover, controlling for the total initial share of migrants would make the shift-share leverage the *composition* of migration origins. This would address labor demand shocks that affect all migrants equally.

Working with the original Card (2009) setting, Goldsmith-Pinkham et al. (2020) compute the Rotemberg weights to show which shocks matter most for the estimates. They show that Mexico receives half of the weight in the sample of high school equivalent workers. Thus, for these workers one can largely think of the research design as using the initial Mexican immigrant share as the instrument. Indeed, Card (2009) notes that the shift-share is highly correlated with the initial fraction of Mexican migrants. For college equivalent workers, Goldsmith-Pinkham et al. (2020) document that the top country is the Philippines, receiving 15% of the total weight.

Goldsmith-Pinkham et al. (2020) also perform balance tests for share instruments with high Rotemberg weights. They report that the 1980 Mexican immigrant share does not predict relative wages in 1980 or 1990, but does in 2000 (the year of analysis). While the patterns for Mexico are comforting, the 1980 share of immigrants from the Philippines correlates with the native-immigrant wage gap in all three periods. Other countries also feature statistically significant violations of pre-trend balance, raising concerns about share exogeneity. In part, the correlation between pre-period outcomes and certain origin shares could arise because pre-period outcomes are affected by pre-period immigration rates. Including lagged immigration rates in the model could help pre-trend tests pass while also making causal estimates more credible.

With these caveats, shift-share estimates suggest that when the ratio of immigrant to native hours worked increases by 10% (because of supply shocks), the wage of migrants relative to native workers falls by 4% for high school graduates, and by 7% for college graduates. This implies that migrant and native workers are more substitutable for low-skill groups. Goldsmith-Pinkham et al. (2020) also report the results obtained with alternative estimators, such as two-stage least squares. They find that the point estimates remain very similar. Similarly, plotting the estimates using individual shares as instruments against their respective F-statistics, they find little variation in the estimates, especially for the strong instruments. In Appendix Figure 1 we report the visual instrumental variables graph. Since individual estimates are similar for all origin countries, the estimates lie near the ray through the origin with the slope equal to the shift-share estimate. These tests demonstrate the robustness of the baseline estimate to alternative ways of combining the share instruments and indicate that treatment effect heterogeneity is a limited concern in this setting.

Labor market effects of retirement Mohnen (2024) studies the impact of the retirement rate of older generations on labor market outcomes for younger generations in the US, conducting the analysis at the commuting zone level. Specifically, the author relates 10-year differences in labor market outcomes for the young (unemployment rate, share working in high-skilled jobs, etc.) to retirement rates over 10 years in the commuting zone. The specification further includes start-of-period regional controls: employment share of manufacturing and routine occupations, unemployment rate, etc. Still, ordinary least squares estimates may be biased because strong labor demand in some commuting zones may explain both low retirement rates among older workers and low unemployment rates for younger workers.

The author addresses this identification challenge with a shift-share strategy that leverages cross-area variation in age composition among the older population. Specifically, the instrument for the 10-year retirement rate in commuting zone i uses the local share of age k among the population aged 45 to 80 as s_{ik} (such that shares sum to one in each commuting zone), and the national 10-year retirement rate by age as g_k. The age composition predicts retirement rates because older workers are more likely to retire, giving the shift-share instrument power. The identification assumption is that the age shares (among people above 45) are all valid instruments conditional on the controls.

To better understand the source of variation, the author describes which age shares matter the most in driving the estimates. He reports Rotemberg weights, documenting that they are close to proportional to the 10-year national retirement rates by age.

Shift-share estimates suggest that the retirement slowdown in the US in recent decades was detrimental to career outcomes for the youth. In places where fewer workers retire, young workers have lower wages and are more likely to have low-skill jobs, and their job mobility falls, although their unemployment does not increase.

To assess whether the results might depend on how the share instruments are combined, the author performs an overidentification test which passes. He also reports alternative estimators: using a particular combination of shares (the initial share of the population age 52-59 as a fraction of the population above 45), or all detailed age shares as separate instruments via generalized method of moments. All estimates suggest similar results, lending support to the validity and robustness of the design.

Conclusion

We have reviewed two frameworks for shift-share research designs, which include sufficient conditions for instrument validity, narratives for interpreting these conditions intuitively, balance tests for the assumptions, and various practical recommendations. Table 2 summarizes the key practical insights for the two approaches, leveraging either exogenous shifts or shares.

How can one pick between the two approaches? In some settings one approach is a "non-starter": e.g., the exogenous shifts approach with too few shifts or the exogenous shares approach when the treatment is specific while the shares are generic. In other settings, it may be productive to think through the potential bias and efficiency properties of the instruments each approach would suggest. For instance, when estimating the local demand elasticity for migrant labor, can a plausibly exogenous supply shift ("push factor") with a strong effect on migration be found? Or is it plausible that there are no national demand shifts for migrants of any origins—in which case a (likely stronger) share-based instrument may be convincing enough? We hope our review will help researchers assess such tradeoffs.

References

Acemoglu, Daron and Pascual Restrepo, "Robots and Jobs: Evidence from US Labor Markets," *Journal of Political Economy*, 2020, *128* (6), 2188–2244.

Adão, Rodrigo, Costas Arkolakis, and Federico Esposito, "General Equilibrium Effects in Space: Theory and Measurement," *Working Paper*, 2023.

_ , Michal Kolesar, and Eduardo Morales, "Shift-Share Designs: Theory and Inference," *Quarterly Journal of Economics*, 2019, *134* (4), 1949–2010.

Aghion, Philippe, Celine Antonin, Simon Bunel, and Xavier Jaravel, "Modern Manufacturing Capital, Labor Demand, and Product Market Dynamics: Evidence from France," *Working Paper*, 2022.

Angrist, J D, Guido W. Imbens, and AB Krueger, "Jackknife instrumental variables estimation," *Journal of Applied Econometrics*, 1999, *67* (95), 57–67.

Angrist, Joshua D. and JS Pischke, *Mostly harmless econometrics: An empiricist's companion*, Princeton University Press, 2008.

Autor, David H. and Mark G. Duggan, "The rise in the disability rolls and the decline in unemployment," *Quarterly Journal of Economics*, 2003, *118* (1), 157–205.

_ , David Dorn, and Gordon H. Hanson, "The China Syndrome: Local Labor Market Impacts of Import Competition in the United States," *American Economic Review*, 2013, *103* (6), 2121–2168.

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

Berman, Nicolas, Antoine Berthou, and Jerome Hericourt, "Export Dynamics and Sales at Home," *Journal of International Economics*, 2015, *96* (2), 298–310. Publisher: Elsevier B.V.

Blanchard, Olivier Jean and Lawrence F. Katz, "Regional Evolutions," *Brookings Papers on Economic Activity*, 1992, (1), 1–75.

Bombardini, Matilde and Bingjing Li, "Trade , Pollution and Mortality in China," *Journal of International Economics*, 2020.

Borusyak, Kirill and Matan Kolerman-Shemer, "Regression Discontinuity Aggregation, with an Application to the Union Effects on Inequality," *Mimeo*, 2024.

_ and Peter Hull, "Non-Random Exposure to Exogenous Shocks: Theory and Applications," *Mimeo*, 2021.

_ and _ , "Non-Random Exposure to Exogenous Shocks," *Econometrica*, 2023, *91* (6), 2155–2185.

_ , _ , and Xavier Jaravel, "Quasi-Experimental Shift-Share Research Designs," *Review of Economic Studies*, 2022, *89* (1), 181–213.

__ , __ , and __ , “Design-Based Identification with Formula Instruments: A Review,” *Working Paper*, 2023.

__ , Xavier Jaravel, and Jann Spiess, “Revisiting Event Study Designs: Robust and Efficienct Estimation,” *Working Paper*, 2023.

Cai, Jing, Alain De Janvry, and Elisabeth Sadoulet, “Social Networks and the Decision to Insure,” *American Economic Journal: Applied Economics*, April 2015, *7* (2), 81–108.

Card, David, “The Impact of the Mariel Boatlift on the Miami Labor Market,” *ILR Review*, 1990, *43* (2), 245–257.

__ , “Immigration and Inequality,” *American Economic Review: Papers & Proceedings*, 2009, *99* (2), 1–21.

Christian, Paul and Christopher B Barrett, “Spurious Regressions and Panel IV Estimation: Revisiting the Causes of Conflict,” *The Economic Journal*, 2024, *134* (659), 1069–1099.

Clots-Figueras, Irma, “Women in politics,” *Journal of Public Economics*, August 2011, *95* (7-8), 664–690.

Dauth, Wolfgang, Sebastian Findeisen, and Jens Suedekum, “The rise of the east and the far east: German labor markets and trade integration,” *Journal of the European Economic Association*, 2014, *12* (6), 1643–1675. ISBN: 9783863041267.

de Chaisemartin, Clement and Xavier D’Haultfoeuille, “Two-way fixed effects estimators with heterogeneous treatment effects,” *American Economic Review*, 2020, *110* (9), 2964–2996. arXiv: 1803.08807v1.

__ and Ziteng Lei, “More Robust Estimators for Instrumental-Variable Panel Designs, With An Application to the Effect of Imports from China on US Employment,” *Working Paper*, 2023.

Franklin, Simon, Clement Imbert, Girum Abebe, and Carolina Mejia-Mantilla, “Urban Public Works in Spatial Equilibrium: Experimental Evidence from Ethiopia,” *Working Paper*, 2023.

Freeman, Richard B., “Overinvestment in College Training?,” *Journal of Human Resources*, 1975, *10* (3), 287–311.

__ , “An Empirical Analysis of the Fixed Coefficient ”Manpower Requirements” Model, 1960-1970,” *Journal of Human Resources*, 1980, *15* (2), 176–199.

Goldsmith-Pinkham, Paul, “Tracking the Credibility Revolution across Fields,” *arXiv preprint arXiv:2405.20604*, 2024.

__ , Isaac Sorkin, and Henry Swift, “Bartik Instruments: What, When, Why, and How,” *American Economic Review*, 2020, *110* (8), 2586–2624.

Greenstone, Michael, Alexandre Mas, and Hoai-Luu Nguyen, "Do Credit Market Shocks Affect the Real Economy? Quasi-Experimental Evidence from the Great Recession and "Normal" Economic Times," *American Economic Journal: Economic Policy*, 2020, *12* (1), 200–225.

Hausman, Jerry A., Whitney K. Newey, Tiemen Woutersen, John C. Chao, and Norman R. Swanson, "Instrumental variable estimation with heteroskedasticity and many instruments: Instrumental variable estimation," *Quantitative Economics*, July 2012, *3* (2), 211–255.

Hummels, David, Rasmus Jorgensen, Jakob Munch, and Chong Xiang, "The Wage Effects of Offshoring: Evidence From Danish Matched Worker-Firm Data," *American Economic Review*, 2014, *104* (6), 1597–1629. arXiv: 1011.1669v3 ISBN: 9788578110796.

Imbens, Guido W. and Joshua D. Angrist, "Identification and Estimation of Local Average Treatment Effects," *Econometrica*, 1994, *62* (2), 467–475. ISBN: 00129682.

Jaeger, David A, Joakim Ruist, and Jan Stuhler, "Shift-Share Instruments and the Impact of Immigration," *Working Paper*, 2017.

Jaravel, Xavier, "The Unequal Gains from Product Innovations: Evidences from the US Retail Sector," *Quarterly Journal of Economics*, 2019, *134* (2), 715–783.

Kolesar, Michal, Raj Chetty, John Friedman, Edward Glaeser, and Guido W. Imbens, "Identification and Inference With Many Invalid Instruments," *Journal of Business and Economic Statistics*, 2015, *33* (4), 474–484. ISBN: 0735-0015.

Miguel, Edward and Michael Kremer, "Worms: Identifying impacts on education and health in the presence of treatment externalities," *Econometrica*, 2004, *72* (1), 159–217.

Mogstad, Magne, Alexander Torgovitsky, and Christopher R. Walters, "The Causal Interpretation of Two-Stage Least Squares with Multiple Instrumental Variables," *American Economic Review*, November 2021, *111* (11), 3663–3698.

Mohnen, Paul, "The Impact of the Retirement Slowdown on the U.S. Youth Labor Market," *Journal of Labor Economics*, 2024.

Nunn, Nathan and Nancy Qian, "US Food Aid and Civil Conflict," *American Economic Review*, 2014, *104* (6), 1630–1666.

Rotemberg, Julio, "Instrument variable estimation of misspecified models," *MIT Sloan Working Paper 1508-83*, 1983.

Terry, Stephen J, Thomas Chaney, and Konrad B Burchardi, "Immigration, Innovation, and Growth," *Working Paper*, 2023.

Topalova, Petia, "Factor immobility and regional impacts of trade liberalization: Evidence on poverty from India," *American Economic Journal: Applied Economics*, 2010, *2* (4), 1–41.

Wooldridge, Jeffrey M., *Econometric analysis of cross section and panel data MIT press*, MIT Press, 2002. Publication Title: Cambridge, MA.

Xu, Chenzi, "Reshaping Global Trade: The Immediate and Long-Run Effects of Bank Failures," *Quarterly Journal of Economics*, 2022, *137* (4), 2107–2161.

A Other Practical Concerns with Shift-Share IVs

In this appendix we discuss some questions which come up frequently in shift-share instrumental variable (IV) designs. These twelve questions are: What if shares are observed in a panel? What kind of local average treatment effect does the shift-share IV estimate? Do the "shares" have to really be shares, between zero and one? Should the shares be normalized to add up to one? Can shift-share instruments be constructed by apportioning national changes to units? Can the shifts be unit-specific? Can one take a shift-share average of shift-share IVs? What if a log, or another transformation, of a shift-share variable is used? Can one use multiple shift-share instruments? What about interaction terms in shift-share regressions? Should the instruments in Card (2009) and Bartik (1991), which measure the shifts as the national growth of some equilibrium outcome (industry employment or total migration by origin), be viewed through the lens of exogenous shifts or exogenous shares? And what is the role of leave-one-out construction of shifts?

A.1 What if shocks are observed over multiple periods in a panel?

While our main discussion considered a single cross-section (typically, with first-differenced outcomes), in many applications researchers have access to shifts g_{kt} happening in multiple periods t. In a panel of units i over periods $t = 1, \ldots, T$, one may consider an IV specification

$$y_{it} = \beta x_{it} + \gamma' w_{it} + \varepsilon_{it}, \tag{4}$$

with a shift-share IV $z_{it} = \sum_k s_{ikt} g_{kt}$ and some controls w_{it}, potentially including unit and period fixed effects. Here we indexed the shares s_{ikt} by the period when they are used, not when they are measured: for instance, s_{ikt} can be time-invariant, fixed in an early "base" period.

The panel setting offers new possibilities. In the exogenous shifts approach, if a natural experiment generates exogenous shifts in several periods, "stacking" them provides more estimation power. Moreover, panels with *many* periods make it possible to apply the many exogenous shifts approach even when K is small, thanks to the time-series variation in the shifts. In the simplest case, there may be just one shift in each period and heterogeneous unit exposure to this shift, such that $z_{it} = s_{it} g_t$ (where s_{it} is often time-invariant). When many periods are observed, exogeneity of the time series variation in g_t is sufficient for consistent estimation. Exposure-robust inference also follows from the time-series properties of the shifts. The shift-level equivalent IV regression in this case is just a time-series regression regardless of the number of units in the panel, and standard errors should correspondingly be clustered in a time-series way (e.g., by period or allowing for serial correlation in the shifts). Below we provide a detailed illustration of these points in the setting of Nunn and Qian (2014).

In the exogenous shares approach, the number of available instruments is also larger in a panel, as primitive instruments are individual shares interacted with period dummies, all assumed to be valid. Correspondingly, Rotemberg weights are computed for each pair of k and t, although Goldsmith-Pinkham et al. (2020) recommend reporting their sums over time for clarity.

Panel data also pose new challenges. We focus on the exogenous shifts approach in this discussion. First, the shifts may have different means in different periods. In conventional panel models, time-varying means are addressed by including period fixed effects (FEs), γ_t. Correspondingly, in shift-share designs, time-varying shift means are addressed by a share-weighted aggregate of period FEs: $\sum_k s_{ikt}\gamma_t$. With complete shares, i.e. when $\sum_k s_{ikt} = 1$, this control coincides with the period FEs. But in the incomplete shares case, the sum of shares control needs to be interacted with period FEs. The setting of Autor et al. (2013) discussed in the main text illustrates this point.

Second, shifts can be serially correlated, in which case each period cannot be viewed as a separate natural experiment. Then, as mentioned in Step 4 of the exogenous shifts checklist, the static specification in (4) suffers from an omitted variables bias problem when there are dynamic causal effects, i.e. if lagged shifts affect current outcomes (Jaeger et al., 2017). Intuitively, the estimated coefficient for the treatment in specification (4) is biased because it also includes the dynamic causal effect of past treatments. Moreover, if the shares can respond to past shifts which are correlated with contemporaneous shifts, the shares cannot be viewed as measured before the natural experiment in shifts began (see footnote 5).

There are two solutions to the problems of serial correlation in shifts. One involves estimating richer specifications which include the relevant lagged treatments, as well as lagging the shares underlying the shift-share IV further. The shares need to be measured at a date before the sequence of serially correlated shifts began if such a date exists. Jaeger et al. (2017), for instance, show that migration rates by country of origin are very serially correlated since 1970s, but not correlated with those from earlier decades. Thus, year 1970 can be viewed as the beginning of the natural experiment in their setting.

An alternative solution is based on isolating the unpredictable component of the contemporaneous shifts before constructing the shift-share IV. For instance, if the shifts follow a first-order autoregressive process, one can control for the lagged shifts (by controlling for a share-aggregated version of them at the unit level). If the shift-share IV leverages the idiosyncratic component of shifts, the issues stemming from serial correlation disappear.[12] This approach only yields the contemporaneous effect but does not require a correct specification of the dynamic effects.

[12] There are different ways of extracting the idiosyncratic component of shifts. Instead of controlling for lagged shifts, another natural approach could be to control for the time-invariant component of shifts. Implementing this strategy is easy when time-invariant shares are used: then including *unit* fixed effects in the control vector w_{it} implicitly removes any shift-level confounders α_k, since the corresponding share-aggregated control $\sum_k s_{ikt}\alpha_k$ is time-invariant.

Example of shift-share IV with time-series variation Nunn and Qian (2014) study the impact of US food aid on conflicts in a long panel of recipient countries. Simple ordinary least square (OLS) estimates, or even those with country fixed effects, are subject to several potential biases: the presence of conflict may increase the demand for food aid; there might be many omitted variables—such as political and economic crises—affecting both conflict and food aid; or donors may decide to reduce food aid to countries engaged in conflict.

To resolve these issues, the authors leverage exogenous time variation in US wheat production over time. Due to price stabilization policies requiring the US government to buy wheat from US farmers at a set price, the US government accumulates excess reserves in high production years, which is shipped to developing countries as food aid. The shift-share design leverages these time series shifts, using as exposure weights a country's likelihood of being a US food aid recipient. Specifically, the quantity of wheat aid shipped from the US to recipient i in year t is instrumented by $z_{it} = s_i g_t$, where g_t is the amount of US wheat production in the previous year and s_i is the fraction of years that recipient country i receives a positive amount of US food aid during the sample period, 1971–2006.

How can one follow our exogenous shifts checklist in this context? For steps 1–2, the researcher would clarify whether all time series variation in wheat production is considered as-good-as-random. This requires an exclusion restriction, that US wheat production affects conflict in other countries only through US aid. Moreover, if US wheat production is correlated with key economic indicators such as oil prices, which can have a direct effect on conflict, the researcher would need to control for these variables interacted with s_i. Indeed, the interaction of the oil price with s_i is one of the controls Nunn and Qian (2014) include. They also include other controls, such as dummies for six geographic regions of the world interacted with year dummies. For step 3, the incomplete share control here is simply s_i, the time-invariant exposure to US aid, since each observation is exposed to only one shift; in Nunn and Qian's regression it is absorbed by country fixed effects. For step 4, one would measure s_i before, rather than during, the sample period. For step 5, it would be useful to plot the time series of wheat prices, which serves as identifying variation. Christian and Barrett (2024, Fig. 3) finds strong serial correlation and an inverse U-shaped trend in wheat prices. In this case, it may be appropriate to analyze dynamic causal effects or extract an unpredictable component of the time series of US wheat production. For step 6, one can check whether the time series of wheat production is correlated with potential confounders, such as the aforementioned oil prices. At the country-by-year level, an IV regression with lagged conflict as the outcome would constitute a standard pre-trend test. Finally, for step 7, one can cluster standard errors at the level of identifying variation, i.e. by year (rather than by country, which is more conventional in panel regressions). Given the shifts (and, likely, errors) are serially correlated, heteroskedasticity and autocorrelation-consistent standard errors may be more appropriate. Indeed, Christian and Barrett (2024) show that conventional standard errors can lead to spuriously significantly relationships in this setting. These standard errors are easier to obtain from the time-series regression equivalent to the original panel regression.

A.2 Does shift-share IV estimate a LATE when the effects are heterogeneous?

With many exogenous shifts, yes, and different units and shifts receive a different weight in shift-share IV regressions. Otherwise, a local average treatment effect (LATE) interpretation is more challenging.

Provided the shifts are as-good-as-randomly assigned and mutually uncorrelated, as if arising from a lottery, Adão et al. (2019) and Borusyak et al. (2022) prove that shift-share regressions (both IV and OLS) identify convex averages of heterogeneous treatment effects under a monotonicity condition similar to the one imposed by Imbens and Angrist (1994) to establish identification of LATEs in standard IV regressions. Note that as-good-as-random assignment here is formally stronger than the necessary condition on shift exogeneity described in the main text. For example, it requires the shifts to be independent of treatment effect heterogeneity as well.

What makes the shift-share setting unique is that effect heterogeneity can arise in two dimensions: across units i and shifts k. Thus, the shift-share IV estimate can be interpreted as averaging across both dimensions, with certain weights. We derive and interpret these weights in a heterogeneous-effects causal model inspired by the decomposition formula (3).[13] For concreteness, we consider our labor supply example. The model is as follows:

$$\begin{aligned} x_{ik} &= \pi_{ik} g_k + u_{ik}, \\ x_i &= \sum_k s_{ik} x_{ik}, \\ y_i &= \sum_k \beta_{ik} s_{ik} x_{ik} + \varepsilon_i. \end{aligned}$$

Here x_{ik} are changes in employment by region and industry—the local shifts. They are affected by the industry subsidies g_k with a coefficient of π_{ik}. Regional employment growth x_i is an aggregate of industry-by-region growth rates x_{ik} weighted by regional employment shares s_{ik}, as in equation (3). But the effects of x_{ik} on the wage change y_i are not necessarily proportional to s_{ik}, as captured by the heterogeneous effects β_{ik}. Variation in β_{ik} across i captures the idea that the local labor supply elasticity may depend on the region. In turn, variation in β_{ik} across k reflects the scenario in which employment changes coming different industries (say, tradable and nontradable ones) would have different wage impacts through labor supply.

Following the logic of Adão et al. (2019) and Borusyak et al. (2022), it is easy to show that, when the shifts g_k have mean zero, variance σ_k^2, and no mutual correlation conditional on all other

[13] This formulation generalizes Proposition 2 of Adão et al. (2019) to IV, rather than reduced-form regressions. Footnote 16 in Adão et al. (2019) considers IV regression but does not allow the effects β_{ik} to vary by k. The analysis in Appendix A.1 of Borusyak et al. (2022) is very general, allowing further for nonlinear effects, but they do not discuss the intuition for the resulting LATE. One limitation of our formulation is that employment growth in industry k is not allowed to be affected by subsidies to other industries; the model in Appendix A.7 of Borusyak et al. (2022) relaxes that assumption but does not study heterogeneous effects.

sources of unobserved heterogeneity $(u_{ik}, \varepsilon_i, \pi_{ik}, \beta_{ik})$, the shift-share IV estimand equals

$$\beta = \frac{\sum_i \sum_k \pi_{ik} s_{ik}^2 \sigma_k^2 \cdot \beta_{ik}}{\sum_i \sum_k \pi_{ik} s_{ik}^2 \sigma_k^2}.$$

That is, heterogeneous effects β_{ik} are averaged with weights proportional to: (1) the strength of the first-stage effect π_{ik}, (2) the local share s_{ik} squared, and (3) the shift variance σ_k^2. Evidently, the weights are non-negative under a monotonicity condition $\pi_{ik} \geq 0$ (which holds trivially in shift-share OLS regressions, which correspond to $\pi_{ik} = 1$, $u_{ik} = 0$).

To gain more intuition for the weights, suppose $\pi_{ik} \equiv \pi$ and $\sigma_k^2 \equiv \sigma^2$. Then, if all heterogeneity is by region ($\beta_{ik} = \beta_i$), the weight on β_i is equal to the Herfindahl–Hirschman index of local industry concentration, $\sum_k s_{ik}^2$. A region exposed to many different industry shifts will not be useful for the regression because the law of large numbers eliminates most variation in the shift-share instrument. Conversely, when all heterogeneity is by industry ($\beta_{ik} = \beta_k$), the weight on β_k is equal to $\sum_i s_{ik}^2$.[14] Naturally, this weight is higher for larger industries. More interestingly, it is also higher when the local shares of this industry are very unequal across regions. For instance, tradable industries will play a larger role than nontradable industries of a similar national size, as typical tradable industries concentrate in a small number of regions while nontradable ones are present in every region with relatively homogeneous shares.

The heterogeneous-effect interpretation of shift-share IVs without as-if random shifts is less established. de Chaisemartin and Lei (2023) raise concerns of non-convex weighting of unit-specific causal effects when shift-share IVs are justified by parallel trend assumptions, with respect to either shares (as in the exogenous shares approach) or shifts (i.e., with a weaker restriction on the shift exogeneity), adding to a large literature noting similar issues for popular two-way fixed effect specifications (e.g. de Chaisemartin and D'Haultfoeuille (2020) and Borusyak et al. (2023b)). Part of this issue is apparent in the Goldsmith-Pinkham et al. (2020) Rotemberg weight decomposition since, as Goldsmith-Pinkham et al. (2020) note, some weights may be negative. To the best of our knowledge, the case of heterogeneous causal effects across shifts has not—to our knowledge—been studied without as-if random shifts.

A.3 Do the "shares" have to really be shares?

No, they can be any exposure weights.

In most applications s_{ik} is non-negative and typically they are some initial shares; notably this is the case when the shift-share IV follows from the decomposition (3). But econometric results go through when s_{ik} are any weights that measure the exposure of observation i's treatment to the shift g_k.

As an example, consider the Miguel and Kremer (2004) study of spillover effects of deworming. In their OLS specification, the key explanatory variable z_i is the number of student i's neighbors who have received a randomized deworming treatment. Upon inspection, one may notice that this

[14] Note that this is not a Herfindahl–Hirschman index because the shares add up to one across industries, not regions.

is a shift-share variable: $z_i = \sum_{k=1}^{N} s_{ik} g_k$ where k indexes all students, s_{ik} is a dummy that equals one if students i and k are neighbors, and g_k is a dummy that student k has been selected for deworming. Here the exposure weights are not shares of anything: they take values of zero and one and their total is the number of neighbors student i has. There is no problem with this, as long as the sum of shares (i.e., the number of student's neighbors) is controlled for.

A.4 Should I normalize the shares to add up to one?

You could, but controlling for the sum of shares is probably a better solution.

From our earlier discussion of how the incomplete shares case requires extra care (specifically, picking appropriate controls in the many exogenous shifts approach), one might conclude that this case is something to be avoided. This can be done by constructing the instrument using shares normalized to add up to one. For instance, while Autor et al. (2013) define s_{ik} as employment shares of manufacturing industry k relative to total employment in labor market i, one could consider redefining the shares to have local manufacturing employment in the denominator.

Such a conclusion would be misguided, however. First consider IV regressions, where the treatment x_i is given by the economic question. Then the researcher needs to choose the best shift-share IV z_i, and in particular the shares, to maximize instrument strength. Whether identification leverages exogenous shifts or exogenous shares, power is maximized when the shares reflect the relationship between the treatment and the shifts, e.g. following the treatment decomposition (3). For example, in the Autor et al. (2013) setting, using the local manufacturing employment in the denominator would reduce power because the shift-share instrument would exhibit large variation even in areas where manufacturing is a low share of total employment and the treatment (import competition) is close to zero. Including appropriate controls is a better way to avoid OVB while retaining statistical power, compared to modifying the shares.

Second, consider OLS shift-share analyses, such as spillover regressions, where the researcher is deciding on the right-hand side variable $x_i = z_i$. This choice is about specifying the most plausible functional form for how the shifts affect the outcome, such that the coefficient is economically meaningful. Again, this is achieved by setting the shares to reflect the exposure of observations to the exogenous shifts. For example, the fraction of treated friends, as in Cai et al. (2015), is a shift-share variable with the shares adding up to one, while the number of treated friends, as in Miguel and Kremer (2004), is an incomplete shares example. Still, if the researcher believes that the outcome is determined by the *number* of treated friends, they should use that specification, and include appropriate controls to avoid bias.

A.5 Can shift-share instruments be constructed by apportioning national changes to units?

Yes, and in fact Bartik (1991), Card (2009), and Autor et al. (2013) all derived their instruments this way. However, to apply the tools from this paper correctly, the resulting instruments must be rewritten with different shares and shifts, as in equation (2).

We illustrate the apportioning logic with the labor supply example. Recall that the percent change in regional employment is an aggregate across industries: $x_i = \left(\sum_k \Delta X_{ik}\right)/X_{i0}$. The researcher can then replace the local industry employment change (in levels), ΔX_{ik}, with a prediction that allocates the national change in the industry employment, ΔX_k, to regions proportionally to the initial regional composition of the industry, $\frac{X_{ik0}}{X_{k0}}$. Region i therefore "gets" $\frac{X_{ik0}}{X_{k0}} \cdot \Delta X_k$ workers in industry k. Adding up such predictions and rescaling them by the initial regional employment yields the instrument:

$$z_i = \frac{\sum_k \frac{X_{ik0}}{X_{k0}} \cdot \Delta X_k}{X_{i0}}. \tag{5}$$

While this expression *looks* different from the shift-share instrument $\sum_k \frac{X_{ik0}}{X_{i0}} \cdot \frac{\Delta X_k}{X_{k0}}$ that follows from the decomposition of x_i in equation (3), a simple rearrangement of terms shows that they are actually the same:

$$\frac{\sum_k \frac{X_{ik0}}{X_{k0}} \Delta X_k}{X_{i0}} = \frac{\sum_k X_{ik0} \cdot \frac{\Delta X_k}{X_{k0}}}{X_{i0}} = \sum_k \frac{X_{ik0}}{X_{i0}} \cdot \frac{\Delta X_k}{X_{k0}}. \tag{6}$$

This rearranging step is crucial for applying the theoretical results and taking the practical steps in both exogenous shifts and exogenous shares approaches. The left-hand side of (6) is based on employment shares relative to the industry total, whereas the shares on the right-hand side are relative to the regional total. The left-hand side suggests that the national shifts are industry employment changes in levels, ΔX_k, although this leaves the denominator unaccounted for by either shares or shifts; on the right-hand side of (6), the shifts are relative changes in national industry employment.

Both conceptual and practical issues arise if the apportioning formula (5) is used without rewriting it as in (6). In the exogenous shifts approach, assuming ΔX_k is as-good-as-randomly assigned is untenable, as larger industries of course get larger employment changes on average (provided national employment is growing).[15] This assumption is also not sufficient because the denominator X_{i0} in (5) is ignored, while it affects the identification conditions. Measuring shifts in relative terms instead makes their as-if random assignment a more plausible assumption. In the exogenous shares approach, using the shares relative to the industry total, X_{ik}/X_k, as instruments is the same as using initial employment levels X_{ik}, since the share denominators in (5) do not vary across observations. Thus, variation in the local industry size is used instead of the local composition of industries that is usually intended in shift-share IV designs. Moreover, since the remaining terms in the summation, $\Delta X_k/X_{i0}$, vary across i, z_i cannot be viewed as pooling variation in the shares (relative to the industry total).

More practically, applying the checklists above to the wrong shares and shifts would lead to incorrect controls (e.g., incomplete share controls) and diagnostic tests (e.g., based on wrong Rotemberg weights). In (5), it looks like there is an incomplete share problem, while (6) makes it clear there is not (since $\sum_k \frac{X_{ik0}}{X_{k0}} \neq 1$ while $\sum_k \frac{X_{ik0}}{X_{i0}} = 1$).

[15] In Appendix A.11 we argue that the exogenous shifts lens may not be appealing for the Bartik (1991) instrument. However, the issues we discuss here are not specific to that application, and they arise similarly with the Autor et al. (2013) instrument.

A.6 Can the shifts be unit-specific?

Yes.

While we introduced shift-share variables as combining heterogeneous shares with a common set of shifts, the econometric framework also nests settings where each unit is exposed to a distinct set of shifts. One can define k to index the shifts to all observations and redefine the shares such that the exposure of a unit to another unit's shift is zero.

A set of examples is considered by Borusyak and Kolerman-Shemer (2024) who study "regression discontinuity aggregation" designs in which a shift-share treatment aggregates policy discontinuities defined at smaller geographic units. For instance, Clots-Figueras (2011) estimates the effect of the fraction of women in state legislatures in India, using the fraction of women who won against a man in a close election as the IV. Although each state has a distinct set of constituencies, this instrument is a shift-share where each state has non-zero exposure only to its own constituencies' shifts.

A.7 Can I take a shift-share average of shift-share IVs?

Yes, and the result is also a shift-share IV, with the same shifts but more complicated shares.

This situation commonly arises when studying spillovers from treatments (or instruments) that already have a shift-share structure. Adão et al. (2023), for instance, study spatial spillovers from regional import competition with China. Let $z_i = \sum_k s_{ik} g_k$ be the Autor et al. (2013) instrument, capturing direct exposure of commuting zone i to Chinese imports based on industry shifts g_k and local employment shares s_{ik}. Slightly simplifying, Adão et al. (2023) define the indirect exposure of commuting zone j as the inverse-distance weighted average of direct exposures of all other commuting zones: $z_j^* = \sum_i s_{ji}^* z_i$, where the shares s_{ji}^* decay with the distance between j and i (and $s_{jj}^* = 0$). One can see that this variable can be rewritten as $z_j^* = \sum_k s_{jk}^{**} g_k$ with compound shares $s_{jk}^{**} = \sum_i s_{ji}^* s_{ik}$ and original shifts g_k.

Representing the shift-share instrument with the resulting shares and shifts, in one step, makes it clear that exogeneity of g_k is still sufficient for identification. It also yields appropriate incomplete share and other share-aggregated controls, and correct standard errors.

A.8 What if I take logs of a shift-share?

A log — or any other nonlinear transformation — of a shift-share variable is not a shift-share variable. This may or may not complicate IV exogeneity.

In the exogenous shares approach, which views the shift-share IV as a particular function of the shares (where the shifts serve as weights), a nonlinear function of a shift-share IV is just another function of the same shares. If all individual shares are exogenous instruments, i.e. $\mathbb{E}[\varepsilon_i \mid s_{i1}, \ldots, s_{iK}] = 0$, then any function of them is exogenous, too.

On the contrary, shift exogeneity does not imply exogeneity of nonlinear transformations of the shift-share IV, such as taking the log; such transformations can lead to a new type of bias. To see this, imagine the shares add up to one and the exogenous shifts are assigned in a lottery with positive values. Then, regardless of how the shares are correlated with the error term, the share-weighted average of the lottery shifts $z_i = \sum_k s_{ik} g_k$ is not correlated with the error. That logic fails for $\log z_i$: because of Jensen's inequality, units with dispersed shares will on average have a higher $\log z_i$ than units with concentrated shares, potentially leading to bias. Similar issues arise with other transformations of shift-share IVs, e.g. using a dummy that a shift-share variable is in the lowest quartile of its distribution, as in Greenstone et al. (2020).

There are two ways to avoid this bias. First, Borusyak and Hull (2023) propose a "recentering" adjustment to the nonlinear instrument, such as $\log z_i$, based on rerandomizing the shifts, e.g. by permuting them. Second, putting the log inside the sum, i.e. replacing $\log \sum_k s_{ik} g_k$ with $\sum_k s_{ik} \log g_k$, yields an actual shift-share IV with shares s_{ik} and shifts $\log g_k$.

For a concrete example, Berman et al. (2015) estimate the effects of log firm exports on the log of its domestic sales to measure returns to scale. While our discussion so far has focused on outcomes and treatments measured as changes, consistent with the decomposition (3), Berman et al. (2015) perform the analysis in logs of levels, using a panel of firms and controlling for firm fixed effects. They instrument log exports with $z_{it} = \log \sum_k s_{ik} G_{kt}$, where k denotes product-by-country pairs, s_{ik} is the share of this pair in firm's exports (on average across periods), and G_{kt} is the total world exports of this product to this country. Leveraging exogeneity of G_{kt}, or the log-changes in G_{kt} over time, would require the corrections discussed above.

Borusyak and Hull (2021, footnote 82) show an additional problem with this IV: it implicitly uses shares that are not the s_{ik} and may not capture the intended economic intuition. For instance, one may think that for firm i that has 50% of initial exports in a certain product-country cell k ($s_{ik} = 0.5$), a 10% increase of world exports in that cell raises z_{it} by approximately 0.05. This is not the case. To see the issue, suppose changes in G_{kt} over time are sufficiently small and consider how $z_{it} = \log \sum_k s_{ik} G_{kt}$ changes in response, relative to some base period 0 (recalling that, with firm fixed effects, changes over time play the key role). It is easy to show that

$$z_{it} - z_{i0} \approx \sum_k \frac{s_{ik} G_{k0}}{\sum_{k'} s_{ik'} G_{k'0}} \left(\log G_{kt} - \log G_{k0}\right) \neq \sum_k s_{ik} \left(\log G_{kt} - \log G_{k0}\right). \tag{7}$$

Thus, the response of z_{it} to a 10% shift to G_{kt} is determined not only by the share of k in firm i's initial exports but also by the world supply of k in the initial period — which was presumably not intended when constructing the instrument. To avoid this issue, one can replace $\log \sum_k s_{ik} G_{it}$ with $\sum_k s_{ik} \log G_{kt}$.

A.9 What if I have multiple shift-share instruments?

This is fine, both when multiple shift-share variables instrument for a single treatment and when multiple IVs are necessitated by multiple treatments. One should just perform the relevant steps for each of the shift-share IVs: e.g., include incomplete share controls in the exogenous shifts approach and check sensitivity to how shares are combined in the exogenous shares approach.

Getting exposure-robust standard errors may be more challenging in this case. When the shares are the same but there are several sets of exogenous shifts, Borusyak et al. (2022) show how the shift-level equivalent IV regression extends in this case, yielding correct standard errors. Appendix B.1 below extends this result by allowing for several shift-share IVs that use different shares and different shifts, as long as all shifts are defined at the same "level" k. We derive an equivalent shift-level representation of the estimator in terms of a set of moment conditions (but no longer as a simple IV). This equivalence result yields exposure-robust standard errors. A Stata example is available in our GitHub repository, https://github.com/borusyak/shift_share_jep.

We give two examples. First, Dauth et al. (2014) consider the impacts of two import competition shifts in Germany, originating from the growth of China and from the accession of Eastern European countries into the European Union. Both are shift-share variables that combine the local employment shares of different industries with two national industry import competition shifts.

Second, including both direct and spillover effects of a certain treatment in the same specification can be viewed as using two shift-share variables with the same shifts but different shares. For instance, the right-hand side variables in Miguel and Kremer (2004) are the student i's own deworming dummy and the number of her dewormed friends. We explained above how their spillover treatment is a shift-share IV that uses deworming dummies as the shifts g_k and the patterns of friendship as exposure weights. Mechanically, one's own deworming status is also a shift-share with the exposure weight being one for $i = k$ and zero otherwise.

A.10 What if I have interaction terms in a shift-share regression?

This is similar to having multiple shift-share variables.

There can be two types of interaction terms in shift-share regressions. A more conventional one interacts z_i with some unit-level variable a_i. For instance, in the Autor et al. (2013) context, one may be interested in understanding whether labor market responses to import competition vary by the share of college graduates in the region. This interaction can be written as a shift-share IV with the same shifts and different exposure weights: $a_i z_i = \sum_k (a_i s_{ik}) g_k$.

The second type — albeit not exactly an interaction — aims to identify the heterogeneous responses to different groups of *shifts*. For instance, Bombardini and Li (2020) consider the health effects of two treatments: regional exposure to the national industry growth of exports for all industries and for pollution-intensive industries in particular. The former is a standard shift-share variable $z_i = \sum_k s_{ik} g_k$ while the latter can be written as $z_i' = \sum_k s_{ik}(b_k g_k)$ where b_k is industry's pollution intensity.[16] This z_i' is a shift-share IV with shares s_{ik} and shifts $b_k g_k$.[17] We refer the reader to Appendices A.9 and B.1 for a discussion of incomplete share and other appropriate controls, as well as exposure-robust standard errors with multiple shift-share instruments.

[16] Note that z_i' is not the same as the interaction of z_i with the regional share of pollution-intensive industries, which would be an interaction term of the first type.

[17] It can also be viewed as a shift-share with shares $s_{ik} b_k$ and shifts g_k. Both interpretations lead to the same practical conclusions, in different ways. For instance, with as-good-as-random g_k, one needs to control for $\sum_k s_{ik} b_k$. In the former interpretation this follows because the shifts $g_k b_k$ can be considered as-good-as-random only controlling for b_k (while the shares add up to one). In the latter interpretation this follows because the shares $s_{ik} b_k$ add up to $\sum_k s_{ik} b_k$ (while the shifts are already as-good-as-random).

A.11 Can the instruments in Bartik (1991) and Card (2009) be valid without exogenous shares?

This depends on the underlying model, but probably not. We consider shift-share instruments with shifts constructed as national averages of endogenous local shifts correlated with the error terms, as in Bartik (1991) and Card (2009). The researcher might argue that these shifts proxy for some latent exogenous shifts. Here we show that the proxy error in the shifts is innocuous when the local *shares* are exogenous (and if there are many more observations than shifts), while otherwise the proxy error typically makes the instrument invalid. Thus, there is little value in focusing on the shifts for justifying the validity of Bartik (1991) and Card (2009) type instruments, except in a special case discussed below.

For concreteness, we illustrate the general insight in the setting of Bartik (1991); we discuss Card (2009) at the end. Bartik (1991) estimated the (inverse) elasticity of local labor supply by using a shift-share instrument that combined local employment shares of different industries with the national growth rate of employment in each industry (see Table 1). For the Bartik (1991) instrument to be valid, it has to capture labor demand conditions. Interestingly, Blanchard and Katz (1992) focus on the exogeneity of the shifts, rather than local employment shares, when introducing the Bartik (1991) instrument: *"This series will be valid for our purposes [of isolating a labor demand shift] as long as the national growth rates are not correlated with labor supply shifts in the state"* (p. 25). Is the exogenous shifts approach appropriate in this setting? In particular, is it a problem that the shifts are equilibrium outcomes which may also be affected by labor supply factors?

We give intuition before the formal analysis. Suppose high net migration—internal or foreign—into a region makes employment in all local industries grow. Then, industries that are concentrated in regions with growing net migration will systematically have higher employment growth in most regions, and therefore nationally. That, however, is precisely the situation when the industry growth rate shifts are econometrically endogenous. The main (although not the only) case when this does not happen is if no industry is concentrated in regions with growing or falling net migration. But that corresponds to the case where the local shares of all industries are exogenous with respect to the local net migration rate. It is further required that industries are not too concentrated in a small number of regions, such that random local migration shocks do not have a big impact on national industry growth rates.

We now formally characterize how labor supply shocks affect national industry growth rates. We model employment growth by region and industry as

$$x_{ik} = g_k^* + u_{ik},$$

where g_k^* is the latent national industry labor demand shock and u_{ik} captures labor supply factors.[18] Since labor demand conditions are unobserved, Bartik (1991) proxy for them by the national industry employment growth rate g_k when constructing the shift-share instrument. Denoting by E_{ik},

[18] The results extend directly to the case where labor demand shifts vary across regions.

E_i, and E_k the initial employment levels by region-industry, region, and industry respectively, we have:

$$g_k = \sum_i \frac{E_{ik}}{E_k} x_{ik} = g_k^* + \tilde{g}_k,$$

where the proxy error is given by

$$\tilde{g}_k = \frac{\sum_i E_{ik} u_{ik}}{E_k}. \tag{8}$$

We consider a favorable case where the labor demand conditions g_k^* are exogenous, and thus the only concern is whether $\tilde{g}_k$ affects the exogeneity of the instrument.

As discussed in the main text, the necessary and sufficient condition for the shift-share instrument to be valid is that the measured shifts g_k have no covariance with a particular industry confounder. Specifically, provided the regional analysis is performed with initial employment E_i as importance weights, as is commonly done, this confounder is the average of regional error terms ε_i weighted by initial employment in industry k:[19]

$$\bar{\varepsilon}_k = \frac{\sum_i E_{ik} \varepsilon_i}{E_k}. \tag{9}$$

The expressions for the proxy noise (8) and the confounder (9) exhibit a striking similarity: if some labor supply conditions in ε_i affect employment local in all industries (u_{ik}), we may expect employment-weighted averages of those shocks to be correlated, too. Indeed, in the simple model of local labor markets in Appendix A.7 of Borusyak et al. (2022), regional labor supply shocks affect industry employment growth rates equally, such that $u_{ik} = \gamma \varepsilon_i$ for some $\gamma > 0$. In that model, $\tilde{g}_k$ and $\bar{\varepsilon}_k$ would be perfectly correlated.

There are, however, some special cases in which the problem does not arise, both linked to the properties of the local employment shares. First, if the shares of all industries are exogenous with respect to the regional error term ε_i and the number of industries is small, $\bar{\varepsilon}_k \xrightarrow{p} 0$ for each k. In this case, the exogeneity of the shifts is not required so any proxy noise is fine (Goldsmith-Pinkham et al., 2020).[20]

Second, if for each industry the share s_{ik} is exogenous with respect to the local employment change in that industry due to labor supply, u_{ik}, the proxy noise will average out: $\tilde{g}_k \xrightarrow{p} 0$. This is the case, in particular, for labor supply shocks that induce reallocation of workers across domestic regions in the sample without changing industry. Naturally, such reallocation does not affect the national industry employment growth. However, it seems unlikely that this scenario constitutes the only source of local shift endogeneity, particularly since industry switching (or, similarly, international migration or mobility out of unemployment or non-employment) is necessary to generate nontrivial national industry growth rates to begin with. In that case, the Bartik (1991) instrument

[19]The weights in this averaging combine the shares underlying the instrument, $s_{ik} = E_{ik}/E_i$, and importance weights E_i.

[20]One can see is that $\bar{\varepsilon}_k$ is the weighted covariance between s_{ik} and ε_i weighted by E_i and rescaled by E_k; see Appendix A.2 in Borusyak et al. (2022) for further details on how $\bar{\varepsilon}_k \xrightarrow{p} 0$ constitutes the relevant notion of share exogeneity.

cannot be valid without shares being exogenous (with respect to ε_i) too.

Analogous issues are likely present in Card (2009), who constructs the shifts as the national growth rate of migration from origin country k which are aggregates of local migration rates from that country. Here problems arise if the local migration rates by origin are endogenous, i.e. correlated with the error term—which for Card (2009) reflects the relative demand for migrant vs. native labor. As long as the local migration rates from all origins respond to the same local relative demand conditions, the resulting national shifts cannot be viewed as exogenous without the local shares of migration from different origins being exogenous, too.

We close by noting that our discussion here concerned problems with shifts constructed as national averages of endogenous local shifts, but in practice researchers often use leave-one-out averages.We discuss the role of leave-one-out adjustments in the next section.

A.12 What is the role of leave-one-out construction of shifts?

This is a useful way of mitigating bias when pooling variation from many exogenous share instruments. It is an open question whether this practice can help to extract exogenous latent shifts when the shares are endogenous.

In settings like Bartik (1991) where, as explained above, the shifts can be mechanically confounded by the errors, it is common since Autor and Duggan (2003) to use "leave-out" constructions of shift-share instruments: $z_i = \sum_k s_{ik} g_{k,-i}$, where $g_{k,-i}$ is, say, the industry growth rate in all regions except i (or perhaps except nearby regions, too).[21] With many exogenous *shares*, Borusyak et al. (2022, Appendix A.6) show that using leave-one-out means to construct the national growth rates is useful to address the finite sample bias that can mechanically arise when using own-observation information. This approach is similar to how jackknife instrument variable estimators avoid bias of 2SLS in presence of many instruments (Angrist et al., 1999).

In practice, Autor and Duggan (2003) observed that including own region in shift construction made the IV substantially stronger, raising concerns about the mechanical relationship. Other authors (e.g., Goldsmith-Pinkham et al. (2020)) found that the leave-out correction is empirically minor when the measured shifts average over sufficiently many observations.

It is an open question whether leave-one-out constructions can help address the problem of proxy bias in the shifts discussed in Appendix A.11 when the shares are endogenous. On the one hand, the leave-one-out construction can be viewed as similar to jackknife instrumental variable estimation which Kolesar et al. (2015) show can be consistent under a particular orthogonality condition even when there are many invalid instruments. On the other hand, the error terms of observations with shares similar to i can be correlated with ε_i, in which case leaving out i may not suffice. In a Monte Carlo simulation available by request, we confirm that leave-one-out need not fully eliminate the bias when the shares are endogenous.

We finally note that the leave-out constructions of shift-shares are distinct from a practice of

[21] Strictly speaking, such z_i is not a shift-share as defined by equation (2), since $g_{k,-i}$ has some variation across units.

measuring the shifts from entirely different data, e.g. in different countries. Autor et al. (2013), for instance, instrument regional exposure to Chinese imports in the US using industry shifts measured in other developed countries; Hummels et al. (2014) and Aghion et al. (2022) use similar approaches when instrumenting firm-level imports. Unlike leave-out constructions, here the shifts g_k are the same for all units in the sample. Moreover, the mechanical correlation between the error term and the shifts does not arise, such that the exogenous shifts approach can be applied under the appropriate assumptions (e.g., that import demand shifts in the US and other developed economies are uncorrelated in the Autor et al. (2013) context).[22]

B Theoretical Results

In this appendix, we present three new theoretical results.

B.1 Exposure-Robust Standard Errors for Shift-Share IV Regressions with Multiple Treatments

We derive exposure-robust standard errors for a IV (or OLS) regression with multiple shift-share instruments, by recasting the estimator as a method of moments estimator at the level of shifts. This result builds on Borusyak et al. (2022), Proposition 5.

Consider a just-identified shift-share IV regression:

$$y_i = \beta_1 x_{1i} + \cdots + \beta_R x_{Ri} + \gamma' w_i + \varepsilon_i \tag{10}$$

where $x_{1i}, \ldots, x_{Ri}$ are instrumented with a set of shift-shares $z_{1i}, \ldots, z_{Ri}$ for $z_{ri} = \sum_{k=1}^{K} s_{rik} g_{rk}$. Both the shares and the shifts can differ across r but we require the shifts to vary at the same level for all r (and thus with the same number of shifts K). Assuming the shifts g_{rk} are as-good-as-randomly assigned after controlling for some vector of shift-level controls q_{rk} (which can vary across r), we require the vector of controls w_i to include $\sum_k s_{rik} q_{rk}$ for each r. The vector w_i further includes the intercept and possibly other controls.

The IV estimator $\hat{\beta}$ for $\beta = (\beta_1, \ldots, \beta_R)'$ in (10) satisfies a system of R equations:

$$\frac{1}{N} \sum_i \left(y_i^{\perp} - \sum_{j=1}^{R} \hat{\beta}_j x_{ji}^{\perp} \right) z_{ri} = 0, \qquad r = 1, \ldots, R,$$

where for any variable v_i we let $v_i^{\perp}$ denote the in-sample projection of v_i on w_i. Expanding the expression for z_{ri}, exchanging the order of summation, denoting $\tilde{v}_k^{(r)} = \frac{1}{N} \sum_i s_{rik} v_i^{\perp}$, and combining

[22] In settings like Hummels et al. (2014), the researcher may therefore entertain two options: to measure the shifts in a different country and follow the exogenous shifts approach, or to measure the shifts in the country of interest in a leave-out way and follow the exogenous shares approach.

terms yields a set of R shift-level moment conditions satisfied by $\hat{\beta}$:

$$\sum_k \left(\tilde{y}_k^{(r)} - \sum_{j=1}^R \hat{\beta}_j \tilde{x}_{jk}^{(r)} \right) g_{rk} = 0, \qquad r = 1, \dots, R.$$

Letting $\tilde{g}_{rk}$ be the projection of g_{rk} on q_{rk} weighted by $s_{rk} = \frac{1}{N} \sum_i s_{rik}$ and noting that $\sum_k \tilde{v}_k^{(r)} q_{rk} = 0$ since w_i includes $\sum_k s_{rik} q_{rk}$, we further have a set of R equations on the residualized shifts:

$$\sum_k \tilde{\psi}_k^{(r)} = 0, \quad \text{for } \tilde{\psi}_k^{(r)} = \left(\tilde{y}_k^{(r)} - \sum_{j=1}^R \hat{\beta}_j \tilde{x}_{jk}^{(r)} \right) \tilde{g}_{rk}.$$

In matrix form, this can be rearranged as

$$\Omega \hat{\beta} = M,$$

where $\Omega_{rj} = \sum_k \tilde{x}_{jk}^{(r)} \tilde{g}_{rk}$ and $M_r = \sum_k y_k^{(r)} \tilde{g}_{rk}$. Thus, $\hat{\beta} = \Omega^{-1} M$. Moreover, since (10) implies $\tilde{y}_k^{(r)} - \sum_{j=1}^R \beta_j \tilde{x}_{jk}^{(r)} = \tilde{\varepsilon}_k^{(r)}$ for true β, we also have

$$\hat{\beta} - \beta = \Omega^{-1} E \qquad \text{for } E_r = \sum\nolimits_k \varepsilon_k^{(r)} \tilde{g}_{rk}.$$

We assume that the appropriate relevance condition holds and suppose that vectors of shift residuals $\tilde{g}_k = (\tilde{g}_{1k}, \dots, \tilde{g}_{Rk})'$ are asymptotically independent across some shift clusters c. Letting $\tilde{\psi}_k = \left(\tilde{\psi}_k^{(1)}, \dots \tilde{\psi}_k^{(R)} \right)'$, we then have an asymptotic approximation of the exposure-robust variance-covariance matrix of $\hat{\beta}$:

$$\begin{aligned} \mathrm{Var}\left[\hat{\beta} - \beta\right] &\approx \Omega^{-1} \mathrm{Var}\left[E\right] (\Omega^{-1})' \\ &\approx \Omega^{-1} \left(\sum_c \left(\sum_{k \in c} \tilde{\psi}_k \right) \left(\sum_{k \in c} \tilde{\psi}_k \right)' \right) (\Omega^{-1})'. \end{aligned} \tag{11}$$

We note that the derivation here simplifies if the shares are the same for all shift-share instruments (and q_{rk} are also the same) and only the shifts vary across r. In this case, the coefficients and exposure-robust standard errors can be obtained by an IV estimator at the shift-level, as shown by Borusyak et al. (2022). This includes instruments constructed as $\sum_k s_{ik} b_k g_k$ for different "interaction" variables b_k, as discussed in Appendix A.10.

B.2 Visual IV weights with Share Exogeneity

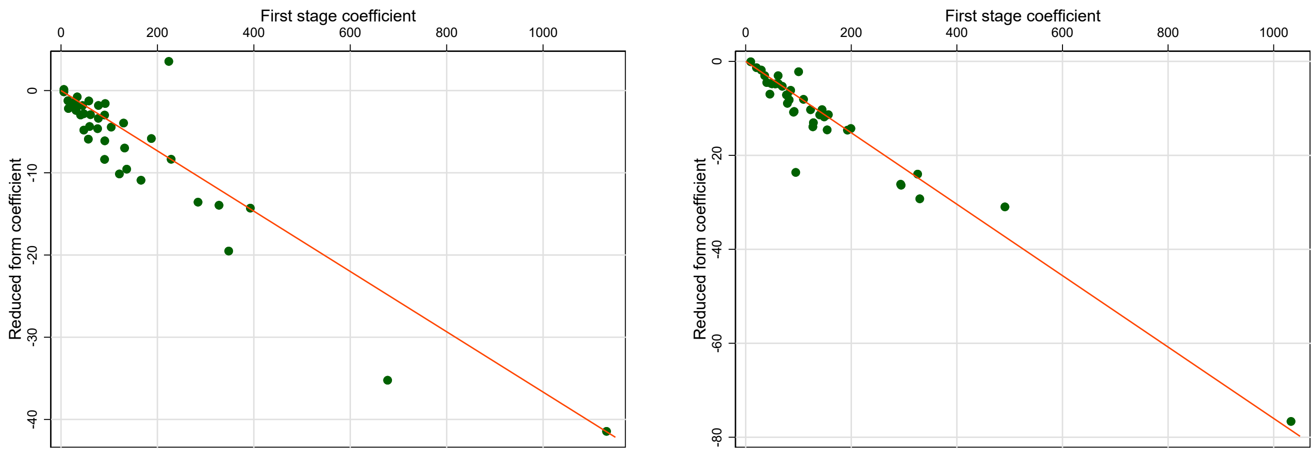
In this appendix, we show the shift-share IV estimate equals the slope of the regression line through the points on the visual IV graph for the exogenous shares approach, without intercept and with appropriate weights.

Let $\hat{\beta}_k$ be the share-IV estimate for industry k and let ω_k denote the Rotemberg weights, which

Figure 1: Visual IV for Exogenous Shares, Applied to Card (2009)

A: High-school graduates

B: College graduates



Notes: The visual IV graph in the setting of Card (2009) using the replication data from Goldsmith-Pinkham et al. (2020), estimating the relationship between the log wage gap between immigrant and native workers (as the outcome) and the ratio of immigrant to native hours worked (as the treatment). Card (2009) instruments the local ratio of immigrant to native hours with a shift-share instrument, leveraging immigration patterns from 38 countries. We plot the reduced-form coefficient against the first-stage coefficient for each share IV, using the immigration shares from each of the 38 countries one at a time as instruments. Panel A focuses on high-school graduates while Panel B considers college graduates. The shift-share IV estimate is visualized as the slope of the ray through the origin.

sum to one and are such that the shift-share IV coefficient is $\hat{\beta} = \sum_k \omega_k \hat{\beta}_k$.[23] Write $\hat{\beta}_k = \hat{\rho}_k / \hat{\pi}_k$, where $\hat{\rho}_k$ and $\hat{\pi}_k$ are reduced-form and first-stage estimates for the kth share-IV. Then we have:

$$\hat{\beta} = \sum_k \omega_k \frac{\hat{\rho}_k}{\hat{\pi}_k} = \frac{\sum_k (\omega_k / \hat{\pi}_k^2) \hat{\rho}_k \hat{\pi}_k}{\sum_k (\omega_k / \hat{\pi}_k^2) \hat{\pi}_k^2},$$

which is the slope from a regression of $\hat{\rho}_k$ on $\hat{\pi}_k$, with no intercept and with weights $\omega_k / \hat{\pi}_k^2$ (which are not necessarily convex since Rotemberg weights can take negative values).

[23]See Proposition 3 in Goldsmith-Pinkham et al. (2020) for the definition of Rotemberg weights and Section IV.B for the adjustments needed when the shares add up to one.