Journal of Economic Perspectives—Volume 39, Number 1—Winter 2025—Pages 181–204

A Practical Guide to Shift-Share Instruments

Kirill Borusyak, Peter Hull, and Xavier Jaravel

Many economic studies consider units that are exposed differently to a common set of shocks. Consider, for example, the influential Autor, Dorn, and Hanson (2013) 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 hope 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

■ *Kirill Borusyak is Assistant Professor of Agricultural and Resource Economics, University of California Berkeley, Berkeley, California. Peter Hull is Professor of Economics, Brown University, Providence, Rhode Island. Xavier Jaravel is Professor of Economics, London School of Economics, London, United Kingdom. Their email addresses are k.borusyak@berkeley.edu, peter_hull@brown.edu, and x.jaravel@lse.ac.uk.*

For supplementary materials such as appendices, datasets, and author disclosure statements, see the article page at https://doi.org/10.1257/jep.20231370.

the last ten years (Goldsmith-Pinkham 2024). Today, around one-eighth 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 (for example, industries) than the unit of analysis (for example, local labor markets). Are they still useful then? In contrast, the shares vary across units but are usually predetermined (for example, 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, Hull, and Jaravel (2022) and Adão, Kolesár, and Morales (2019), leverages many exogenous shifts while making no assumption on the exogeneity of the shares. The second path, proposed by Goldsmith-Pinkham, Sorkin, and Swift (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, Hull, and Jaravel (2024) for a more technical review.

Online 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; 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, Dorn, and Hanson (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:

$$y_i = \beta x_i + \gamma' \mathbf{w}_i + \varepsilon_i, \tag{1}$$

where $\mathbf{w}_i$ 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*: that is, a non-zero correlation between x_i and ε_i. In the Autor, Dorn, and Hanson (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 that 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, Dorn, and Hanson (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

Table 1
Shift-Share Instrument Examples

					Instrument (z_i)	
Study	Unit (i)	Outcome (y_i)	Treatment (x_i)	Level of shift variation (k)	Share (s_{ik})	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	$\mathbf{1}\{k \text{ is friend of } i\}$	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, Dorn, and Hanson (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, Janvry, and Sadoulet (2015)	Individual	Takeup of insurance	% of friends selected for an information session*	Individual	$\mathbf{1}\{k \text{ is friend of } i\}$/# of friends i has	Dummy of information session
Jaravel (2019)	Product category	Inflation and innovation	Δ Quantity demanded	Sociodemographic group	Sales of i to group k/Total sales of i	Population change
Greenstone, Mas, and Nguyen (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. (2024)	Local labor market	Wage	Shift-share exposure to the intervention*	Residential neighborhood	$\text{Commuters}_{ik}/\text{Employment}_i$	Dummy of public works intervention
Mohnen (2025)	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

Note: 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, and so forth. Asterisks (*) indicate ordinary least squares regressions, in which the treatment itself is the shift-share with shares s_{ik} and shifts g_k.

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 $x_i = (X_{i1} - X_{i0})/X_{i0}$ 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:

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

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, for example, 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 $s_{ik} = X_{ik0}/X_{i0}$ 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 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. Online 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 (for example, Cai, Janvry, and Sadoulet 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*; that is, 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 $\text{cov}[z_i, \varepsilon_i] = 0$). 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 versus 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 that 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 versus 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 versus native workers in labor demand, β. Here the model (1) relates changes in the relative wages of migrants versus 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

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

Mariel Boatlift studied by Card (1990). One might be willing to assume that regions that 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 versus native labor: that is, that $\text{cov}[\varepsilon_i, s_{i\kappa}] = 0$. 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, Sorkin, and Swift (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 bolstered by constructing the instrument with shares that 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 (for example, 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, Dorn, and Hanson (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, Dorn, and Hanson (2013) and Acemoglu and Restrepo (2020) use essentially the same instruments (lagged employment shares) for different treatments.

The role of the shifts is secondary with the exogenous shares strategy: Goldsmith-Pinkham, Sorkin, and Swift (2020) show that the shifts affect the weights in their representation of shift-share as pooled one-at-a-time share-instrument

[2] To see this link, note that we can formalize the above example as a setting with only one non-zero immigration shift: that is, $g_\kappa \neq 0$ 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.

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 (for example, 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 an idealized shift-level experiment. Any compelling instrumental-variable design begins with thinking about what endogeneity bias is being addressed: that is, 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 (for example, 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 that 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 (1) specifying some control variables and (2) describing how observed shifts proxy for the ideal ones.

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 versus 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: for example, 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, for example, because they are too few or endogenous	You would not want to use shares directly as instruments, for example, because they are "generic" (capturing the unit's exposure to many types of shocks)

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 shifts 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, Hull, and Jaravel (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; for example, 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 versus 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, Dorn, and Hanson (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 online 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, Hull, and Jaravel (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

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

[4] In online 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 online 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, Hull, and Jaravel (2022), consider the labor supply setting and imagine that subsidies now occur in two periods. Supposing regions vary in labor market flexibility, the reallocation of employment towards industries with larger subsidies is stronger in

matter for instrument power; lagging shares 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 (Jaeger, Ruist, and Stuhler 2018). Simply lagging the shares does not help with this problem. We discuss the problems arising in panels and possible solutions in online 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

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.

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

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 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, Kolesár, and Morales (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, Kolesár, and Morales (2019) provide a variance estimator that 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 (for example, by group of industries). Second, Borusyak, Hull, and Jaravel (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, because 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; for example, not only standard clustering but also spatial clustering and serial

[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}.

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 (for example, population weights in a regional analysis) yields similar results.

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. Online Appendix A.1 provides an additional example leveraging time-series variation in the shifts.

Shift-Share in a Randomized Trial. Franklin et al. (2024) 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 baseline-period commuting, Franklin et al. (2024) 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, Dorn, and Hanson (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, for example, 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 United States or if US consumers substitute to Chinese goods in industries where US productivity is lagging.

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 predetermined industrial composition of each area. In practice, productivity changes are unobserved and, as a proxy, Autor, Dorn, and Hanson (2013) use the observed growth of imports from China in industry k in eight high-income countries (excluding the United States). 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, because 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, for example, 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. Because Autor, Dorn, and Hanson (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, Dorn, and Hanson (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 (for example, 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,

[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.

average wages, and investment in new technologies (for example, computers) in a pre-period. Using the data from Autor, Dorn, and Hanson (2013), Borusyak, Hull, and Jaravel (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, Dorn, and Hanson (2013) and Borusyak, Hull, and Jaravel (2022) find that most of these tests pass.

Using this shift-share strategy, Autor, Dorn, and Hanson (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, Hull, and Jaravel (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 (for example, import competition in Autor, Dorn, and Hanson 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 is 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 (1) labor demand shocks that attract migrants are transitory, and (2) 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—as with 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, Sorkin, and Swift (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, Sorkin, and Swift (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, as 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 since the pre-period or there were no shocks of the same nature as g_k in the past (Jaeger, Ruist, and Stuhler 2018).

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, because 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). Because 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 (that is, when the overidentification test "passes").[10] An alternative graph is proposed by Goldsmith-Pinkham,

[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, Sorkin, and Swift (2020) recommend choosing the Rotemberg weights that correspond to the demeaned shifts.

[10] Formally, online 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.

Sorkin, and Swift (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.

Because 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—for example, 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, Imbens, and Krueger 1999), limited information maximum likelihood, the heteroskedasticity-robust Fuller estimator (Hausman et al. 2012), and modified bias-corrected two-stage least squares (Kolesár 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 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, Torgovitsky, and Walters 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 reanalysis by Goldsmith-Pinkham, Sorkin, and Swift (2020). The goal is to estimate the (inverse) elasticity of substitution between immigrant workers and native workers in labor demand, that is, the relationship between the

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

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 United States from 1990 to 2000, normalized by the national stock of migrants from k already in the United States 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, as the migrant shares are measured before these shocks are realized. In contrast, persistent regional labor demand factors (for example, 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, as 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, Sorkin, and Swift (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, Sorkin, and Swift (2020) document that the top country is the Philippines, receiving 15 percent of the total weight.

Goldsmith-Pinkham, Sorkin, and Swift (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 percent (because of supply shocks), the wage of migrants relative to native workers falls by 4 percent for high school graduates, and by 7 percent for college graduates. This implies that migrant and native workers are more substitutable for low-skill groups. Goldsmith-Pinkham, Sorkin, and Swift (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 online Appendix Figure 1, we report the visual instrumental variables graph. Because 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 (2025) studies the impact of the retirement rate of older generations on labor market outcomes for younger generations in the United States, conducting the analysis at the commuting zone level. Specifically, the author relates ten-year differences in labor market outcomes for the young (unemployment rate, share working in high-skilled jobs, and so forth) to retirement rates over ten years in the commuting zone. The specification further includes start-of-period regional controls: employment share of manufacturing and routine occupations, unemployment rate, and so forth. 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 ten-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 ten-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 ten-year national retirement rates by age.

Shift-share estimates suggest that the retirement slowdown in the United States 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"; for example, 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.

■ *We thank Matilde Bombardini, David Dorn, Michael Gmeiner, Paul Goldsmith-Pinkham, David McKenzie, Paul Mohnen, Jonathan Parker, Nina Pavcnik, Timothy Taylor, and Heidi Williams for helpful suggestions.*

References

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

Adão, Rodrigo, Costas Arkolakis, and Federico Esposito. 2023. "General Equilibrium Effects in Space: Theory and Measurement." NBER Working Paper 25544.

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

Aghion, Philippe, Céline Antonin, Simon Bunel, and Xavier Jaravel. 2022. "Modern Manufacturing Capital, Labor Demand, and Product Market Dynamics: Evidence from France." Programme on Innovation and Diffusion Working Paper 044.

Angrist, Joshua D., Guido W. Imbens, and Alan B. Krueger. 1999. "Jackknife Instrumental Variables Estimation." *Journal of Applied Econometrics* 14 (1): 57–67.

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

Autor, David H., David Dorn, and Gordon H. Hanson. 2013. "The China Syndrome: Local Labor Market Impacts of Import Competition in the United States." *American Economic Review* 103 (6): 2121–68.

Autor, David H., and Mark G. Duggan. 2003. "The Rise in the Disability Rolls and the Decline in Unemployment." *Quarterly Journal of Economics* 118 (1): 157–205.

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

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

Blanchard, Olivier Jean, and Lawrence F. Katz. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity* 23 (1): 1–75.

Bombardini, Matilde, and Bingjing Li. 2020. "Trade, Pollution and Mortality in China." *Journal of International Economics* 125: 103321.

Borusyak, Kirill, and Peter Hull. 2021. "Non-Random Exposure to Exogenous Shocks: Theory and Applications." NBER Working Paper 27845.

Borusyak, Kirill, and Peter Hull. 2023. "Non-Random Exposure to Exogenous Shocks." *Econometrica* 91 (6): 2155–85.

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

Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2024. "Design-Based Identification with Formula Instruments: A Review." *Econometrics Journal*. https://doi.org/10.1093/ectj/utae003.

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2024. "Revisiting Event-Study Designs: Robust and Efficient Estimation." *Review of Economic Studies* 91 (6): 3253–85.

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

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

Card, David. 1990. "The Impact of the Mariel Boatlift on the Miami Labor Market." *ILR Review* 43 (2): 245–57.

Card, David. 2009. "Immigration and Inequality." *American Economic Review* 99 (2), 1–21.

Christian, Paul, and Christopher B. Barrett. 2024. "Spurious Regressions and Panel IV Estimation: Revisiting the Causes of Conflict." *Economic Journal* 134 (659): 1069–99.

Clots-Figueras, Irma. 2011. "Women in Politics: Evidence from the Indian States." *Journal of Public Economics* 95 (7–8): 664–90.

Dauth, Wolfgang, Sebastian Findeisen, and Jens Suedekum. 2014. "The Rise of the East and the Far East: German Labor Markets and Trade Integration." *Journal of the European Economic Association* 12 (6): 1643–75.

de Chaisemartin, Clément, and Xavier D'Haultfoeuille. 2020. "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects." *American Economic Review* 110 (9): 2964–96.

de Chaisemartin, Clément, and Ziteng Lei. 2023. "More Robust Estimators for Instrumental-Variable Panel Designs, with an Application to the Effect of Imports from China on US Employment." http://dx.doi.org/10.2139/ssrn.3802200.

Franklin, Simon, Clément Imbert, Girum Abebe, and Carolina Mejia-Mantilla. 2024. "Urban Public Works in Spatial Equilibrium: Experimental Evidence from Ethiopia." *American Economic Review* 114 (5): 1382–1414.

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

Freeman, Richard B. 1980. "An Empirical Analysis of the Fixed Coefficient 'Manpower Requirements' Model, 1960–1970." *Journal of Human Resources* 15 (2): 176–99.

Goldsmith-Pinkham, Paul. 2024. "Tracking the Credibility Revolution across Fields." https://doi.org/10.48550/arXiv.2405.20604.

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. "Bartik Instruments: What, When, Why, and How." *American Economic Review* 110 (8): 2586–2624.

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

Hausman, Jerry A., Whitney K. Newey, Tiemen Woutersen, John C. Chao, and Norman R. Swanson. 2012. "Instrumental Variable Estimation with Heteroskedasticity and Many Instruments: Instrumental Variable Estimation." *Quantitative Economics* 3 (2): 211–55.

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

Imbens, Guido W., and Joshua D. Angrist. 1994. "Identification and Estimation of Local Average Treatment Effects." *Econometrica* 62 (2): 467–75.

Jaeger, David A., Joakim Ruist, and Jan Stuhler. 2018. "Shift-Share Instruments and the Impact of Immigration." NBER Working Paper 24285.

Jaravel, Xavier. 2019. "The Unequal Gains from Product Innovations: Evidences from the US Retail Sector." *Quarterly Journal of Economics* 134 (2): 715–83.

Kolesár, Michal, Raj Chetty, John Friedman, Edward Glaeser, and Guido W. Imbens. 2015. "Identification and Inference with Many Invalid Instruments." *Journal of Business and Economic Statistics* 33 (4): 474–84.

Miguel, Edward, and Michael Kremer. 2004. "Worms: Identifying Impacts on Education and Health in the Presence of Treatment Externalities." *Econometrica* 72 (1), 159–217.

Mogstad, Magne, Alexander Torgovitsky, and Christopher R. Walters. 2021. "The Causal Interpretation of Two-Stage Least Squares with Multiple Instrumental Variables." *American Economic Review* 111 (11): 3663–98.

Mohnen, Paul. 2025. "The Impact of the Retirement Slowdown on the US Youth Labor Market." *Journal of Labor Economics* 43 (1). https://doi.org/10.1086/725874.

Nunn, Nathan, and Nancy Qian. 2014. "US Food Aid and Civil Conflict." *American Economic Review* 104 (6): 1630–66.

Rotemberg, Julio J. 1983. "Instrument Variable Estimation of Misspecified Models." MIT Sloan Working Paper 1508-83.

Terry, Stephen J., Thomas Chaney, Konrad B. Burchardi, Lisa Tarquinio, and Tarek A. Hassan. 2023. "Immigration, Innovation, and Growth." https://sciencespo.hal.science/hal-03869993v1.

Topalova, Petia. 2010. "Factor Immobility and Regional Impacts of Trade Liberalization: Evidence on Poverty from India." *American Economic Journal: Applied Economics* 2 (4): 1–41.

Wooldridge, Jeffrey M. 2002. *Econometric Analysis of Cross Section and Panel Data.* MIT Press.

Xu, Chenzi. 2022. "Reshaping Global Trade: The Immediate and Long-Run Effects of Bank Failures." *Quarterly Journal of Economics* 137 (4): 2107–61.