Design-Based Identification with Formula Instruments: A Review

Kirill Borusyak — UC Berkeley
Peter Hull — Brown
Xavier Jaravel — LSE*

October 2023

Abstract

Many studies in economics use instruments or treatments which combine a set of exogenous shocks with other predetermined variables by a known formula. Examples include shift-share instruments and measures of social or spatial spillovers. We review recent econometric tools for this setting, which leverage the assignment process of the exogenous shocks and the structure of the formula for identification. We compare this design-based approach with conventional estimation strategies based on conditional unconfoundedness, and contrast it with alternative strategies that leverage a model for unobservables.

*Contact: k.borusyak@berkeley.edu, peter_hull@brown.edu, and x.jaravel@lse.ac.uk. Borusyak acknowledges financial support from the European Research Council (ERC) under the European Union's Horizon 2020 research and innovation programme (grant agreement Number 949995).

1 Introduction

Many studies in economics use instruments or treatments that combine a set of observed shocks with other predetermined variables via a known formula. A leading example is shift-share instruments, which sum or average a common set of shocks $(g_k)_{k=1}^{K}$, varying at a different "level" than observations i, with weights s_{ik} ("shares") reflecting heterogenous shock exposure: $z_i = \sum_k s_{ik} g_k$.[1] Other examples include treatments capturing the transmission of shocks along social or geographic networks, and variables based on complex formulas for policy eligibility.[2]

This paper reviews a recent econometric literature which shows how causal effects or structural coefficients in linear models can be estimated in such settings when the shocks in the formula are exogenous with a known "design." Exogeneity here means that the shocks are conditionally independent of the outcome error, allowing all other components of the formula to be endogenous. A known design means some specification of the assignment process from which the vector of observed shocks is drawn. These assumptions jointly formalize the notion of the shocks being "as-good-as-randomly assigned." Borusyak et al. (2022, henceforth BHJ) first develop this approach to shift-share identification and consistency, while Adão et al. (2019) study inference in this setting. Borusyak and Hull (2022, henceforth BH) extend the approach to general formula instruments. Our review unifies the BHJ and BH frameworks and draws connections to alternative approaches to causal inference with formula instruments.

The key problem addressed by BHJ and BH is that exogeneity of the shocks by itself is not enough to avoid omitted variable bias (OVB). This is easily seen with shift-share instruments: suppose the shocks are drawn completely at random (with, say, a positive expectation) but observations vary in the sum of exposure shares to all shocks, $S_i = \sum_k s_{ik}$. The instrument will then be positively correlated with the sum of shares, since $\mathbb{E}[z_i \mid s_{i1}, \dots s_{iK}] = \mathbb{E}[g_k] S_i$, and thus may also correlate with the errors. Randomizing the shocks does not imply instrument validity when shock exposure is not exogenous.

BHJ and BH then show how leveraging the formula allows one to avoid OVB when the shock design is known. In the previous example, the shift-share formula along with the random assignment of the shocks imply that all systematic variation in the instrument is captured by the sum of shares; controlling for this sum eliminates OVB. Identification in the general case, for arbitrary formulas and designs, follows from simple adjustments based on the *expected instrument*: the average value of the formula across counterfactual sets of shocks, drawn from the specified assignment process. Specifically, OVB is avoided by either adding the expected instrument as a control or by using a *recentered instrument* which subtracts the expected instrument from the original formula.

[1] Shift-share (or "Bartik") instruments were originally developed by Bartik (1991) and Blanchard and Katz (1992) to estimate labor demand elasticities. More recent applications study topics in trade (e.g. Autor et al. (2013); Hummels et al. (2014)), immigration (e.g. Card (2009); Peri et al. (2016)), finance (e.g. Greenstone et al. (2014); Xu (2022)), public economics (e.g. Saiz (2010); Diamond (2016)), and macroeconomics (e.g. Oberfield and Raval (2021); Nakamura and Steinsson (2014); Jaravel (2019)).

[2] Settings with such treatments and instruments include Miguel and Kremer (2004), who study spillovers from deworming shocks across students; Donaldson and Hornbeck (2016), who study market access effects from new railroad construction; and Currie and Gruber (1996), who study the effects of Medicaid eligibility.

Controlling for or recentering by the expected instrument is generally necessary for identification with formula instruments, absent auxiliary assumptions on the exogeneity of shock exposure.

How can the shock assignment process be specified in practice? This requirement is trivial when the shocks are generated from a true experiment, as shock counterfactuals then follow from the randomization protocol.[3] In observational data, it is useful to distinguish two cases. With shift-share instruments, only the conditional mean of the shocks needs to be specified. BHJ note that this task, while nontrivial, is no harder than the usual one of selecting controls in a shock-level analysis. In the general case, shock design can follow from the exchangeability of shocks conditional on observables, such that appropriate permutations of shocks can serve as valid shock counterfactuals. Exchangeability of the running variable around a policy threshold, as in the local randomization approach to regression discontinuity designs (e.g., Lee (2008)), scientific knowledge (e.g., geological models when using earthquakes as shocks), and other forms of institutional knowledge can also serve as sources of shock design. Such strategies build on a long tradition in the analysis of randomized experiments, going back to Neyman (1923).

We next review results on the consistency of IV estimators based on (recentered) formula instruments, and on inference using them. BHJ and BH show that shift-share IV estimators converge to the true parameter given a large number of exogenous shocks with sufficiently dispersed average exposure across observations. With shift-share instruments, valid asymptotic inference can be conducted under similar assumptions based on the results of Adão et al. (2019) or via an equivalent shock-level IV regression proposed by BHJ. In settings with a nonlinear formula or few shocks, randomization inference can be a natural (finite-sample) alternative.

The baseline setting we consider involves a constant-effect causal or structural model with a scalar treatment in a cross-section of observations and a known shock design. We discuss, however, various extensions of the design-based approach: to settings with heterogeneous treatment effects, multiple treatments and instruments, shock designs with estimated parameters, and panel data. We also review the results of Borusyak and Hull (2021a) on the most efficient formula instrument constructions for a given treatment variable.

Lastly, we compare design-based identification with two alternative strategies that do not leverage the structure of the formula for identification. First, we draw a connection to conventional identification methods based on a conditional unconfoundedness assumption. Most closely related is the E-estimator of Robins et al. (1992), which uses as an instrument the difference between a treatment and an estimate of its expectation conditional on the covariates. The formula instrument setting is more general, involving instruments constructed from a common set of exogenous shocks that may vary at a different "level" than the observations. Unlike conventional settings, the potentially confounding relationship between the instrument and shock exposure cannot be learned from the data since only one set of shocks is observed. Instead, OVB can be purged by specifying or estimating the assignment process for shocks and using the formula. Second, we contrast the

[3] Note, however, that even in randomized experiments the assumption of exogenous shocks is not trivial: it requires the included treatment variable to capture all mechanisms through which the common shocks affect observations (an implicit exclusion restriction).

design-based approach to formula instruments with strategies that do not directly use the formula but instead model the unobservables (via, e.g., a parallel trends assumption) as in the Goldsmith-Pinkham et al. (2020) approach to shift-share instruments. Such approaches may be especially useful when there are too few shocks or when it is difficult to credibly specify the shock design. We note, however, that the underlying assumptions can be inappropriate when observations are exposed to common unobserved shocks in the same way as they are exposed to the observed shocks in the formula instrument.

The rest of this paper is structured as follows. The next section outlines the basic setting and assumptions. Section 3 reviews results on identification, both in general and in special cases such as shift-share instruments. Section 4 reviews results on consistency and inference. Section 5 discusses extensions of the main results. Section 6 contrasts the design-based approach with alternative methods. Section 7 concludes with open questions in this recent literature.

2 Setting

Consider a simple causal model for an outcome y_i and treatment x_i across units $i = 1, \ldots, N$:

$$y_i = \beta x_i + \varepsilon_i. \tag{1}$$

Here β is the parameter of interest (i.e., the treatment effect) and ε_i is the unobserved error (i.e., unit i's potential outcome when x_i is set to zero). For initial notational simplicity, we assume y_i and x_i (and thus ε_i) have been de-meaned across the units. To estimate β, we suppose a researcher has constructed a candidate "formula" instrument

$$z_i = f_i(\boldsymbol{s}, \boldsymbol{g}), \tag{2}$$

for a set of known functions $f_1(\cdot), \ldots, f_N(\cdot)$, a vector of observed shocks $\boldsymbol{g} \in \mathbb{R}^K$, and some observed data $\boldsymbol{s}$. The shocks are assumed to be exogenous conditional on $\boldsymbol{s}$ and possibly some other observed $\boldsymbol{q}$; formally, with $\boldsymbol{\varepsilon} = (\varepsilon_i)_{i=1}^N$ and $\boldsymbol{w} = (\boldsymbol{s}, \boldsymbol{q})$, we have:

Assumption 1. *(Shock Exogeneity)* $\boldsymbol{\varepsilon} \perp\!\!\!\perp \boldsymbol{g} \mid \boldsymbol{w}$.

Exogeneity, here formalized by the conditional independence of $\boldsymbol{\varepsilon}$ and $\boldsymbol{g}$ given $\boldsymbol{w}$, is satisfied in a four-stage data-generating process where first $\boldsymbol{w}$ is determined, second $\boldsymbol{g}$ and $\boldsymbol{\varepsilon}$ are drawn independently (but possibly in a way that depends on $\boldsymbol{w}$), third $\boldsymbol{x} = (x_i)_{i=1}^N$ is determined in an unspecified way, and fourth $\boldsymbol{y} = (y_i)_{i=1}^N$ is determined by equation (1). We discuss the mapping of this Assumption 1 onto several empirical contexts in Section 3.2; below we discuss a weaker mean-independence condition which suffices in some important settings. Exogeneity can be understood as combining two distinct assumptions: exclusion of the shocks from the causal model and conditionally independent shock assignment. The working paper of BH (Borusyak and Hull 2021b, Appendix C.1) presents a more general potential outcomes model which isolates these two assumptions.

Three examples of candidate instrument constructions illustrate the generality of this setup. In the first example of linear shift-share IV, the instrument is an average (or sum) of the shocks g_k with weights s_{ik} reflecting heterogeneous shock exposure:

Example 1. *(Linear Shift-Share)* $\boldsymbol{s}$ *is an* $N \times K$ *matrix of* $s_{ik} \geq 0$ *and* $z_i = \sum_{k=1}^{K} s_{ik} g_k$.

For instance, supply shocks in China g_k across industries k could be combined with industry employment shares s_{ik} across U.S. commuting zones i to study local employment outcomes, as in Autor et al. (2013). A key feature of such z_i is the linearity of $f_i(\boldsymbol{s}, \boldsymbol{g})$ in $\boldsymbol{g}$; as we discuss below, this simplifies identification, consistency, and inference. This instrument construction is also "anonymous," in that $f_i(\boldsymbol{s}, \boldsymbol{g}) = f(\boldsymbol{s}_i, \boldsymbol{g})$ for a common $f(\cdot)$ across i and $\boldsymbol{s}_i = (s_{ik})_{k=1}^{K}$. Often (though not always) $\sum_k s_{ik} = 1$, in which case z_i is a convex average of the shocks.[4]

Our second example relaxes linearity of $f_i(\boldsymbol{s}, \boldsymbol{g})$ while maintaining the anonymous construction:

Example 2. *(Nonlinear Shift-Share)* $z_i = f(\boldsymbol{s}_i, \boldsymbol{g})$ *for nonlinear* $f(\cdot)$ *and vectors* $\boldsymbol{s}_i$.

Here a concrete setting comes from Boustan et al. (2013), who instrument changes in the income Gini coefficient of municipalities i with a predicted change in the area's Gini coefficient $f(\boldsymbol{s}_i, \boldsymbol{g})$ based on national shocks g_k to the average income of worker groups (defined by percentiles of initial income) k with baseline regional shares of these groups s_{ik}. Another popular construction in this class, seen for instance in Berman et al. (2015), comes from taking the natural log of a linear shift-share variable: i.e., $z_i = \log\left(\sum_k s_{ik} g_k\right)$.

Our third example of network spillovers illustrates the generality of the setting by relaxing both linearity and anonymity:

Example 3. *(Network Spillovers)* $K = N$, $\boldsymbol{s} = (s_{ik})_{i,k=1}^{N}$ *is an adjacency matrix, and* $z_i = f_i(\boldsymbol{s}, \boldsymbol{g})$.

In Carvalho et al. (2021), for instance, a natural disaster generates shocks g_k to firms k which propagate across the firm-to-firm supplier network. Here $f_i(\boldsymbol{s}, \boldsymbol{g})$ returns the network distance between i and the nearest shocked firm, which depends on the full adjacency matrix $\boldsymbol{s}$. Of course, there are also network spillover applications which satisfy linearity and anonymity. For example, in Miguel and Kremer (2004), health shocks from a randomized deworming intervention propagate across a network of students captured by $\boldsymbol{s}$. Here $f_i(\boldsymbol{s}, \boldsymbol{g}) = \boldsymbol{s}_i' \boldsymbol{g}$ simply counts the number of dewormed neighbors of student i, showing this example of network spillovers can also be viewed as a linear shift-share variable with varying sum of exposure shares $\sum_k s_{ik}$.

BH discuss other examples of formula instruments, including market access instruments used to study the effects of transportation upgrades (as in Donaldson and Hornbeck (2016)) and policy eligibility instruments used to study the effects of complex eligibility changes (as in Currie and Gruber (1996)). The class of formula instruments is quite broad, containing any z_i which can be computed from a set of observed shocks $\boldsymbol{g}$ and other observed data $\boldsymbol{s}$.

[4] The restriction that $s_{ik} \geq 0$ is not essential. We impose it because it is satisfied in most applications while it simplifies the results below. In particular, the conditions for consistent estimation in Proposition 1 need to be adjusted to accommodate negative exposure shares.

Remark 1. The setup allows $x_i = z_i$, in which case β captures the reduced-form effect of the formula instrument itself. The exclusion restriction in this case requires the formula $f_i(\cdot)$ to be correct, in the sense of capturing all channels by which the shocks affect the outcome. In the more general IV case, Assumption 1 allows the consideration of any formula $f_i(\cdot)$ as long as x_i captures all shock effects. While strong, this exclusion restriction is standard.

Remark 2. Formula instruments are often derived from the structure of x_i by removing or replacing some endogenous components. For example, canonical shift-share instruments were derived in settings where $x_i = \sum_k s_{ik}\tilde{x}_{ik}$ for observation-specific shocks $\tilde{x}_{ik}$ (e.g. local industry growth rates) which were replaced with national averages when constructing z_i. Similarly, in the Boustan et al. (2013) example, z_i replaces the local income shocks at national income percentiles with the corresponding national changes in their predictions of local Gini coefficients. We return to this relationship between x_i and z_i when discussing instrument relevance in Section 4.1 and asymptotic efficiency in Section 5.

Remark 3. The setup does not assume the (y_i, x_i) are independently and identically distributed (*iid*), as when sampled from some population, or make other assumptions which restrict the dependence of ε_i across units. This permits the units to have common exposure to shocks (both observed and unobserved). This approach is consistent with the design-based tradition of conditioning on the set of potential outcomes, though we do not require such conditioning. It further allows the N units to represent a population—for example, all regions of a country (Abadie et al. 2020)—where random sampling assumptions are inappropriate.

In addition to shock exogeneity, we assume knowledge about the shock "assignment process"—i.e., restrictions on the distribution of $\boldsymbol{g}$ given $\boldsymbol{w}$. The most demanding version of such design knowledge is complete specification of this distribution:

Assumption 2. *(Known Design) The distribution of $\boldsymbol{g}$ given $\boldsymbol{w}$, denoted $G(\boldsymbol{g} \mid \boldsymbol{w})$, is known.*

In randomized controlled trials (RCTs), where $\boldsymbol{g}$ is drawn from according to an experimental protocol that may depend on $\boldsymbol{q}$ (e.g. Miguel and Kremer (2004)), this assumption holds trivially as $G(\boldsymbol{g} \mid \boldsymbol{w})$ is then given by the protocol. For instance, the g_k may be independent Bernoulli random variables with known strata-specific means, with strata indicators included in $\boldsymbol{q}$.

Outside of true experiments, Assumption 2 may be satisfied by appropriate exchangeability assumptions which specify the permutations of shocks that were as likely to have occurred. For example, assuming the g_k are *iid* across k conditional on $\boldsymbol{s}$ implies permutations of $\boldsymbol{g}$ are equally likely to arise. Hence $G(\boldsymbol{g} \mid \boldsymbol{w})$ is known to be uniform over permutations $\Pi(\boldsymbol{g}) = \{\pi(\boldsymbol{g}) \mid \pi \in \Pi_K\}$, with Π_K denoting the set of permutation operators $\pi(\cdot)$ on vectors of length K, and with $\boldsymbol{q} = \Pi(\boldsymbol{g})$. Similar design knowledge follows under weaker shock exchangeability conditions, such as when the g_k are *iid* within, but not across, a set of known clusters and $\boldsymbol{q}$ contains the class of within-cluster permutations of $\boldsymbol{g}$. BH, for example, consider a setting where the shocks are the indicators for whether each planned interregional railway line has opened before a certain date; the treatment is regional market access created by those lines. The timing of line opening is assumed random within groups of planned lines with similar characteristics—specifically, the number of regions a line connects. Thus, expected market access is computed by permuting the opening dates of planned lines that connect the same number of regions. Exchangeability could also be assumed local to some policy threshold, as in the local randomization view of regression discontinuity (Lee 2008; Cattaneo et al. 2015). Scientific or other institutional knowledge (e.g. geological models when using earthquakes as shocks) may also yield specifications of $G(\boldsymbol{g} \mid \boldsymbol{w})$, as we discuss more below.

An important relaxation of Assumptions 1 and 2 follows when $f_i(\boldsymbol{s}, \boldsymbol{g})$ is linear in the shocks, as in Example 1. Specifically, they can be jointly weakened to a restriction on the conditional mean of the shocks:

Assumption 3. *(Conditionally Linear Shock Means) There exists unknown* θ *such that* $\mathbb{E}[g_k \mid \boldsymbol{\varepsilon}, \boldsymbol{w}] = \boldsymbol{q}_k'\theta$ *for known* $\boldsymbol{q} = (\boldsymbol{q}_k)_{k=1}^K$ *included in* $\boldsymbol{w}$ *and for all* k.

This restriction can be seen as combining a shock mean-independence condition, of $\mathbb{E}[g_k \mid \boldsymbol{\varepsilon}, \boldsymbol{w}] = \mathbb{E}[g_k \mid \boldsymbol{w}]$, with a linearity restriction of $\mathbb{E}[g_k \mid \boldsymbol{w}] = \boldsymbol{q}_k'\theta$ (for a θ which implicitly can depend on w). The former condition weakens Assumption 1 to allow higher moments of the shocks to depend on the unobserved $\boldsymbol{\varepsilon}$ while the latter restriction weakens Assumption 2 to a partial parameterization of $G(\boldsymbol{g} \mid \boldsymbol{w})$. For example, in the shift-share application of Autor et al. (2013), the latter condition allows the mean of industry-specific supply shocks in China to systematically vary across broader sectors (e.g., consumer electronics, apparel, food, etc.). This would be captured by a set of sectoral dummies $\boldsymbol{q}_k$, with the aim of leveraging residual shock variation within each sector (e.g., between TVs and computers). Below we further discuss and illustrate the usefulness of Assumption 3 in the linear shift-share setting, while also showing how it can be used with linear approximations of nonlinear shift-share instruments. In Section 5, we discuss other incomplete specifications for $G(\boldsymbol{g} \mid \boldsymbol{w})$.

3 Identification

3.1 Expected and Recentered Instruments

We first consider IV identification of β: i.e., whether the parameter can be learned from the first-stage and reduced-form moments.[5] Formally, we consider the key orthogonality condition $\mathbb{E}\left[\frac{1}{N}\sum_i z_i\varepsilon_i\right] = 0$. Identification follows under this condition as $\beta = \mathbb{E}\left[\frac{1}{N}\sum_i z_i y_i\right] / \mathbb{E}\left[\frac{1}{N}\sum_i z_i x_i\right]$ so long as $\mathbb{E}\left[\frac{1}{N}\sum_i z_i x_i\right] \neq 0$: a relevance condition we return to below.[6]

[5] Identification in this setting is less straightforwardly defined than it would be under the usual assumption of *iid* data (see Remark 3). Following Goldsmith-Pinkham and Imbens (2013), we approach the problem in two steps: first whether certain moment conditions hold (in this section) and then whether the moments can be asymptotically learned from the data (in our consistency results in Section 4.1). Appropriately for non-*iid* data, we consider full-sample moments like $\mathbb{E}\left[\frac{1}{N}\sum_i z_i\varepsilon_i\right]$ which generalize moments like $\mathbb{E}[z_i\varepsilon_i]$ studied in *iid* data.

[6] Expectations here are taken with respect to the joint distribution of $(\boldsymbol{g}, \boldsymbol{x}, \boldsymbol{w}, \boldsymbol{\varepsilon})$. The framework allows, but doesn't require, $(\boldsymbol{w}, \boldsymbol{\varepsilon})$ to be considered fixed, in line with the design-based tradition of focusing on randomness in the shocks and induced variation in the treatment.

A key insight of BH is that shock exogeneity is not, by itself, enough for formula instrument orthogonality. To see this, define $\mu_i(\boldsymbol{w}) \equiv \mathbb{E}\left[f_i(\boldsymbol{s}, \boldsymbol{g}) \mid \boldsymbol{w}\right]$ as the expectation of the candidate instrument of unit i across the distribution of shocks while conditioning on the predetermined data in $\boldsymbol{s}$ and $\boldsymbol{q}$. BH call this object, which we write as μ_i for brevity, the *expected instrument.* They show that under Assumption 1:

$$\mathbb{E}\left[\frac{1}{N}\sum_i z_i \varepsilon_i\right] = \mathbb{E}\left[\frac{1}{N}\sum_i \mu_i \varepsilon_i\right]. \tag{3}$$

In other words, shock exogeneity does not make the cross-sectional covariance between z_i and ε_i zero, unless the expected instrument is itself orthogonal to ε_i. Adjusting for μ_i—the key confounder in this setting—is therefore necessary for IV identification, absent additional restrictions on the exogeneity of shock exposure. The proof of (3) follows by repeated use of the law of iterated expectations: for any i, $\mathbb{E}\left[z_i \varepsilon_i\right] = \mathbb{E}\left[\mathbb{E}\left[f_i(\boldsymbol{g}, \boldsymbol{w})\varepsilon_i \mid \boldsymbol{w}\right]\right] = \mathbb{E}\left[\mu_i \mathbb{E}\left[\varepsilon_i \mid \boldsymbol{w}\right]\right] = \mathbb{E}\left[\mu_i \varepsilon_i\right]$, using Assumption 1 in the second equality. When z_i is linear in the shocks, as with linear shift-share instruments, the same result holds under Assumption 3.

Two simple adjustments for μ_i follow: recentering and controlling. Note first that equation (3) immediately implies that the *recentered instrument* $\tilde{z}_i = z_i - \mu_i$ satisfies orthogonality:

$$\mathbb{E}\left[\frac{1}{N}\sum_i \tilde{z}_i \varepsilon_i\right] = 0. \tag{4}$$

Thus, as long as it also satisfies relevance, $\mathbb{E}\left[\frac{1}{N}\sum_i \tilde{z}_i x_i\right] \neq 0$, using $\tilde{z}_i$ as an instrument identifies β. This follows because $\mathbb{E}\left[\tilde{z}_i \mid \boldsymbol{w}\right] = 0$ by construction. Hence $\tilde{z}_i$ cannot be correlated with any function of $\boldsymbol{w}$ which, through its unrestricted dependence with $\boldsymbol{\varepsilon}$, could violate orthogonality.

The second adjustment is perhaps more familiar: controlling for the confounder μ_i—or, more generally, for a vector of functions $\boldsymbol{r}_i \equiv \boldsymbol{r}_i(\boldsymbol{w})$ that linearly span μ_i—in an IV regression that uses the unadjusted z_i as the instrument. This approach works because controlling for such a $\boldsymbol{r}_i$ implicitly recenters z_i by residualizing it on μ_i. Formally, by the Frisch-Waugh-Lovell theorem, the controlled IV regression is equivalent to an uncontrolled IV regression of $y_i^\perp$ on $x_i^\perp$ where $\perp$ denotes the residuals from an in-sample projection on $\boldsymbol{r}_i$. The orthogonality condition of this regression,

$$\mathbb{E}\left[\frac{1}{N}\sum_i z_i \varepsilon_i^\perp\right] = 0, \tag{5}$$

holds because $\mathbb{E}\left[\frac{1}{N}\sum_i \tilde{z}_i \varepsilon_i^\perp\right] = 0$ by analogy to equation (3) and because $\mathbb{E}\left[\frac{1}{N}\sum_i \left(z_i - \tilde{z}_i\right) \varepsilon_i^\perp\right] = 0$ by the orthogonality between fitted values and residuals of the in-sample projection.[7] The $\boldsymbol{r}_i$

[7]Similarly, if the recentered instrument is used, any set of functions of $\boldsymbol{w}$ can be controlled for, whether or not this set includes μ_i. For this reason, BH argue recentering is the key step for identification while controlling for $\boldsymbol{r}_i$ is a convenient implementation. We note that these results go through even if the number of controls is large. As is well known (e.g., Freedman (2008)), including many controls can create bias with heterogeneous effects (which we consider in Section 5) or otherwise reduce estimation efficiency and affect validity of inference techniques.

controls remove some variation from ε_i, which may improve efficiency in large samples.

Under Assumption 2, the recentering and controlling procedures are both straightforward to implement because μ_i is easily computed. In some cases the expectation $\mathbb{E}[z_i \mid \boldsymbol{w}]$ can be computed analytically or (as we show in the next section) is known to be linear in some observed $\boldsymbol{r}_i$. Otherwise, μ_i can be approximated by a simple simulation procedure:

1. For simulations $j = 1, \dots, J$, redraw counterfactual shocks $\boldsymbol{g}^{(j)}$ from $G(\cdot \mid \boldsymbol{w})$ (e.g. by following the RCT protocol or taking appropriate permutations of observed shocks);
2. Compute $z_i^{(j)} = f_i(\boldsymbol{s}, \boldsymbol{g}^{(j)})$;
3. For each observation, take the average across simulations: $\mu_i \approx \frac{1}{J}\sum_{j=1}^{J} z_i^{(j)}$.

For identification, it is enough to approximate μ_i by *any* number of simulations J so long as draws are independent of $\boldsymbol{g}$ and $\boldsymbol{\varepsilon}$ conditionally on $\boldsymbol{w}$: $\mathbb{E}\left[\frac{1}{N}\sum_i (z_i - \frac{1}{J}\sum_j f_i(\boldsymbol{s}, \boldsymbol{g}^{(j)}))\varepsilon_i\right] = 0$ by the law of iterated expectations, since $\mathbb{E}[z_i \mid \boldsymbol{w}, \boldsymbol{\varepsilon}] = \mathbb{E}\left[f_i(\boldsymbol{s}, \boldsymbol{g}^{(j)}) \mid \boldsymbol{w}, \boldsymbol{\varepsilon}\right]$.[8]

Remark 4. Should recentering or controlling be used in practice? BH note that the answer may depend on whether the shocks arise from an RCT or a natural experiment. In RCTs, Assumption 2 is satisfied trivially. Thus, the researcher can first recenter the instrument to avoid OVB and then choose the controls solely based on efficiency considerations. Controlling for μ_i, as opposed to any other $\boldsymbol{r}_i$, is only preferable if μ_i is believed to be strongly correlated with ε_i. In a natural experiment, however, the controlling strategy has an additional appeal relative to recentering, as the former allows the researcher to incorporate multiple guesses for μ_i, arising from different specifications of $G(\boldsymbol{g} \mid \boldsymbol{w})$. If these candidate μ_i are included as controls while using z_i as the instrument, identification follows if any one candidate is correctly specified.

Remark 5. A related approach to avoid bias from non-random exposure is developed by Aronow and Samii (2017) in the context of discrete network treatments x_i and discrete shocks g_k. While their estimator is based on inverse probability weighting, recentering is a regression adjustment and does not require an additional overlap condition.

3.2 Linear Shift-Share IV

A particularly important case of formula instruments is linear shift-share instruments, with $z_i = \sum_k s_{ik} g_k$, where identification can be achieved under weaker conditions than in the general case. We consider three cases of increasing complexity.

Case 1: Complete shares, no controls We first suppose the exposure shares are "complete," in the sense that $\sum_k s_{ik} = 1$ for all i, and that Assumption 3 holds with no additional controls such that $\boldsymbol{q}_k = 1$ and $\mathbb{E}[g_k] = \theta$ for all k. The latter condition requires each shock g_k to have the same mean θ, regardless of the realizations of the unobservables $\boldsymbol{\varepsilon}$ and exposure shares $\boldsymbol{s}$.

[8] BH further show that the number of simulations is not important for consistency of the estimator. The proof of their Proposition 1 shows the number of simulations generally affects the estimator's efficiency, however.

In this case, instrument orthogonality holds without any recentering as long as the estimating equation includes an intercept. This follows because the expected instrument is constant, $\mu_i = \sum_k s_{ik}\mathbb{E}[g_k \mid \boldsymbol{w}] = \theta \sum_k s_{ik} = \theta$, and is thus absorbed by the intercept. The weaker mean-independent restriction of Assumption 3 is furthermore enough because of the linear construction of z_i. Intuitively, when all the shocks are exogenous and have the same expectation, there is no reason for a weighted average of those shocks z_i to be systematically higher or lower for units with any particular shares $\boldsymbol{s}_i$ and thus with higher or lower ε_i.

For further intuition, it is instructive to rewrite the ($\boldsymbol{r}_i$-controlled) orthogonality condition at the shock "level," following BHJ:

$$0 = \mathbb{E}\left[\frac{1}{N}\sum_i z_i \varepsilon_i^{\perp}\right] = \mathbb{E}\left[\frac{1}{N}\sum_i \sum_k s_{ik} g_k \varepsilon_i^{\perp}\right] = \mathbb{E}\left[\sum_k s_k g_k \bar{\varepsilon}_k^{\perp}\right], \tag{6}$$

where $s_k = \frac{1}{N}\sum_i s_{ik}$ and $\bar{\varepsilon}_k^{\perp} = \frac{\sum_i s_{ik}\varepsilon_i^{\perp}}{\sum_i s_{ik}}$ with $\varepsilon_i^{\perp}$ again denoting the residualized error. This shows that the orthogonality of z_i and $\varepsilon_i^{\perp}$ is equivalent to a weighted "shock-level" orthogonality condition, which considers the relationship between g_k and unobserved $\bar{\varepsilon}_k^{\perp}$. The s_k weights capture the "importance" of each shock, in terms of its average exposure s_{ik}, while $\bar{\varepsilon}_k^{\perp}$ are exposure-weighted averages of $\varepsilon_i^{\perp}$. Since $\sum_k s_k \bar{\varepsilon}_k^{\perp} = \frac{1}{N}\sum_i \sum_k s_{ik}\varepsilon_i^{\perp} = \frac{1}{N}\sum_i \varepsilon_i^{\perp} = 0$, with $\boldsymbol{r}_i$ including a constant, equation (6) shows that instrument orthogonality holds when the weighted covariance of g_k and $\bar{\varepsilon}_k^{\perp}$ is zero.

As a concrete example, consider the setting of Aghion et al. (2022, henceforth AABJ) who study the impact of lowering the cost of automation technologies on labor demand in France in the 2000s. Here y_i is the change in employment of firm i, x_i is the change in the firm's stock of automation technologies, and $z_i = \sum_k s_{ik} g_k$ is the predicted change in the cost of importing these technologies. This prediction is based on shocks g_k to the supply of imported technologies across source countries and technology categories, such that k indexes country-technology pairs (e.g., robots from Italy). Exposure shares s_{ik} are computed as the proportion of imports from cell k in firm i's total imports of automation technologies in a previous period (which sum to one).

In this example, Assumption 3 requires cell-level supply shocks to have the same expectation regardless of all firm-level unobservables. The representation (6) provides further intuition: the shift-share instrument would violate orthogonality if the country-technology pairs that have preexisting relationships with French firms with growing employment (for reasons other than automation) receive systematically different supply shocks. For instance, orthogonality fails if better managed French firms (with better employment trends) established ties with Chinese suppliers in the 1990s in anticipation of the growing supply from China in the 2000s. We next relax the assumption of equal expected shocks to accommodate situations like this.

Case 2: Complete shares with controls We now turn to linear shift-share instruments with general shock-level controls $\boldsymbol{q}_k$ in Assumption 3. In the AABJ application, for example, it appears important to allow for productivity shocks that differ systematically across supplier countries (e.g., China experiences positive productivity shocks in the 2000s) and for firms connected to certain countries to have systematically different unobservables. Similarly, one may allow productivity shocks to vary across technology categories (e.g., robots experience faster productivity growth than textile machines), with firms connected to those technologies being systematically different. The shock-level controls $\boldsymbol{q}_k$ here could then include supplier country and technology fixed effects.

With shock-level controls, the expected instrument is linear in exposure-weighted averages of these controls:

$$\mu_i = \sum_k s_{ik} \boldsymbol{q}_k' \theta = \boldsymbol{Q}_i' \theta, \qquad \boldsymbol{Q}_i = \sum_k s_{ik} \boldsymbol{q}_k. \tag{7}$$

Hence, we can view z_i as combining systematic (and potentially confounding) variation in $\mu_i = \boldsymbol{Q}_i'\theta$ with idiosyncratic variation from certain shocks drawn above or below their expectation by chance:

$$z_i = \boldsymbol{Q}_i' \theta + \sum_k s_{ik} \left(g_k - \boldsymbol{q}_k' \theta \right). \tag{8}$$

Including $\boldsymbol{Q}_i$ in the control vector $\boldsymbol{r}_i$ removes the systematic variation, making instrument orthogonality hold. In the AABJ example this would mean controlling for the total exposure share of a firm to each supplying country, as well as the total exposure share to each technology.

An alternative approach in this case would be to first residualize the g_k shocks on the $\boldsymbol{q}_k$ controls, then use a modified shift-share instrument $\sum_k s_{ik}(g_k - \boldsymbol{q}_k'\hat{\theta})$, possibly with no additional controls. This can be viewed as a recentering approach, although with an estimated parameter $\hat{\theta}$ affecting shock expectations (we allow for shock assignment process specified with parameters in Section 5). In the AABJ case this can be implemented by residualizing cell-level supply shocks on country and technology fixed effects and constructing the shift-share instruments from the estimated residuals.

Case 3: Incomplete shares In certain shift-share applications, the exposure shares are "incomplete" in that they do not sum to one: i.e. $S_i = \sum_k s_{ik}$ varies across observations i. This is the case in the "China shock" study of Autor et al. (2013, henceforth ADH) where i denotes commuting zones and k indexes manufacturing industries. ADH estimate the response of labor market outcomes (e.g. manufacturing employment) across commuting zones to exposure to industry-specific supply shocks in China. Here g_k captures the supply shock in China, while the shares s_{ik} are lagged industry employment shares in each commuting zone. The shares are measured relative to total employment (i.e., both in manufacturing and in other sectors) and so (unlike in AABJ) add up to the lagged total manufacturing share of region i: $S_i < 1$.

Even when Assumption 3 holds with no controls ($\boldsymbol{q}_k = 1$), the incompleteness of shares can make instrument orthogonality fail. Intuitively, z_i leverages variation in S_i which may be correlated with the error term even when shocks are fully random. More formally, we again consider the expected instrument:

$$\mu_i = \sum_k s_{ik} \mathbb{E}\left[g_k\right] = \theta \sum_k s_{ik} = \theta S_i.$$

Instrument orthogonality generally fails when $\theta \neq 0$ and when S_i is cross-sectionally correlated with

ε_i. For example, in the ADH setting, industry-level China shocks g_k are positive in expectation such that z_i is systematically higher in regions where manufacturing as a whole is a bigger share of the local economy. Because manufacturing employment is on a downward trend for reasons other than trade with China (e.g., structural transformation), this can be a source of negative bias.[9]

Controlling for the sum of shares S_i isolates the idiosyncratic variation in z_i which yields instrument orthogonality.[10] Constructing the shift-share instrument from de-meaned shocks is again an alternative recentering solution. The more general case with incomplete shares and other controls $\boldsymbol{q}_k$ follows similarly. Here $\mu_i = \boldsymbol{Q}_i'\theta$ for $\boldsymbol{Q}_i = \sum_k s_{ik}\boldsymbol{q}_k$, but this μ_i is not an exposure-weighted average of $\boldsymbol{q}_k$ since the weights do not add up to one. Still, as before, including $\boldsymbol{Q}_i$ in $\boldsymbol{r}_i$ would isolate the idiosyncratic variation in the instrument.[11]

3.3 Linear vs. Nonlinear Formulas

We now contrast identification with linear and nonlinear formula instruments, using the nonlinear shift-share example of $z_i = f(\boldsymbol{s}_i, \boldsymbol{g})$ for concreteness. As with the linear case, a common motivation for such instruments is that the endogenous treatment of interest can be written as $x_i = f(\boldsymbol{s}_i, \tilde{\boldsymbol{x}}_i)$ for some (potentially unobserved) $\tilde{\boldsymbol{x}}_i = (\tilde{x}_{ik})_{k=1}^K$. An intuitive instrument in this case replaces the $\tilde{x}_{ik}$ with an exogenous shifter g_k. Concrete examples of such cases include the predicted change in a regional Gini coefficient in Boustan et al. (2013), the predicted share of migrants in Basso and Peri (2015), and the predicted foreign demand instrument of Berman et al. (2015).

Nonlinearity of $f(\cdot)$ complicates the structure of the expected instrument relative to the linear case. Even with complete shares and fully random shocks (i.e., *iid* g_k conditionally on $\boldsymbol{w}$)—a case where μ_i is constant for linear shift-share instruments—a nonlinear shift-share instrument generally has a μ_i which depends on $\boldsymbol{s}_i$ in complex ways. Outside this case, μ_i further depends on the heteroskedasticity and mutual dependence between the shocks—making the mean-independence condition (Assumption 3) insufficient.

Borusyak and Hull (2021b, Appendix D.4) make these points concrete by considering an instrument which could be called a "shift-share in logs" (see, e.g., Berman et al. (2015), Berthou et al. (2019), and Costa et al. (2019)). Suppose $x_i = \log(X_{i1}/X_{i0})$ measures the growth rate of a

[9]In practice, ADH include a control which is highly correlated with S_i: total manufacturing employment share from a later period. BHJ show their estimates of manufacturing employment effects are similar when controlling for S_i itself.

[10]As BHJ note, another way to arrive at this solution is to imagine a "missing" shock $g_0 = 0$ with exposure share $s_{i0} = 1 - S_i$: e.g., the non-manufacturing sector in ADH which is not subject to the China shock. With this shock added, there are no incomplete shares. However, $\mathbb{E}[g_0] = 0$ while $\mathbb{E}[g_k] = \theta$ for $k \neq 0$. Thus, unless $\theta = 0$ (e.g., the expected supply shock in China in each manufacturing industry is zero), the indicator $\mathbf{1}[k > 0]$ should be included in $\boldsymbol{q}_k$, and correspondingly the share-weighted exposure to it, which equals S_i, should be controlled for.

[11]For example, ADH study two periods (the 1990s and 2000s) with systematically higher shocks to manufacturing industries in the 2000s. With period fixed effects in $\boldsymbol{q}_k$, the corresponding $\boldsymbol{Q}_i$ includes the total manufacturing share interacted with period fixed effects. BHJ show that adding this control is empirically relevant in the ADH context. The intuition is that regions with a higher total manufacturing share always have systematically higher z_i, but especially in the 2000s. If manufacturing-heavy regions face stronger declines in employment in the 2000s for reasons other than trade with China (again, perhaps because of structural change), the shift-share IV estimator is biased even if period dummies and S_i are separately controlled for.

regional variable which can be represented as a sum of local industry components, $X_{it} = \sum_k \tilde{X}_{ikt}$ for $t \in \{0, 1\}$. Then x_i can be rewritten as a nonlinear function of initial shares $s_{ik} = \frac{\tilde{X}_{ik0}}{X_{i0}}$ and regional growth rates $\widetilde{x}_{ik} = \frac{\tilde{X}_{ik1}}{\tilde{X}_{ik0}}$, as $x_i = \log\left(\sum_k s_{ik} \widetilde{x}_{ik}\right)$. Suppose the $\tilde{x}_{ik}$ are endogenous but we observe an industry characteristic G_{kt} with plausibly exogenous growth rates $g_k = \frac{G_{k1}}{G_{k0}}$ that predict the $\widetilde{x}_{ik}$. A natural nonlinear shift-share instrument is then $z_i = \log(\sum_k s_{ik} g_k)$.

The log transformation makes μ_i heterogeneous across regions, even with fully random g_k and complete shares. In particular, by Jensen's inequality, $\mathbb{E}\left[\log\left(\sum_k s_{ik} g_k\right) \mid \boldsymbol{s}\right]$ is the lowest for regions where $\sum_k s_{ik} g_k$ has little variance conditional on $\boldsymbol{s}$. Thus, regions with very dispersed industries (i.e., where $\sum_k s_{ik}^2 \approx 0$) will tend to have lower μ_i and z_i. Such regions may also have systematically different unobservables ε_i. For example, a more dispersed local economy may not reap the benefits from returns to scale leading to higher regional economic growth, which would generate a downward bias in an IV estimator taking growth as outcome.

One solution to this problem, following the general approach in Section 3.1, is to compute μ_i from a specification of shock counterfactuals and either recenter by or control for it. An alternative solution is to take a first-order approximation of $f(\cdot)$ around some fixed vector of shocks to return to the linear case and obviate the need for a full specification of shock counterfactuals. For example, taking a log-linear approximation around $g_k = 1$ in the preceding example gives $\breve{z}_i = \sum_k s_{ik} \log g_k$, with $s_{ik} = \frac{\partial f(\boldsymbol{s}_i, 1)}{\partial \log g_k}$. This is a linear shift-share instrument, with logged shocks and exposure shares which need not sum to one. As an approximation to z_i, the linear instrument might predict x_i less well and thus be less efficient. However, unlike z_i, the orthogonality of $\breve{z}_i$ depends only on correct specification of $\mathbb{E}\left[\log g_k \mid \boldsymbol{\varepsilon}, \boldsymbol{w}\right]$ (i.e., Assumption 3), making identification more robust.

3.4 Non-Anonymous Constructions

We now consider a case where the formula used to build the instrument is non-anonymous: $z_i = f_i(\boldsymbol{s}, \boldsymbol{g})$, i.e. where the exposure of unit i to the shocks cannot be naturally summarized by some observation-specific vector $\boldsymbol{s}_i$.[12] For example, one may study the propagation of shocks across a network where the units i represent nodes and $\boldsymbol{g}$ captures shocks that are as-good-as-randomly assigned to nodes.

In such cases, avoiding bias may seem particularly challenging. Consider for example the setting of Carvalho et al. (2021), who study the effects of network distance between a focal firm i and the nearest firm located in the geographic area of the 2011 Tohoku earthquake. Exposure to the earthquake shock is given by the entire firm-to-firm supplier network, with network distance yielding a complex non-anonymous construction.

In this setting, it is not clear how to pick the right set of controls to span μ_i without leveraging the formula and some knowledge of the shock assignment process. However, the general simulation procedure from Section 3 yields μ_i given specification of shock counterfactuals: for

[12] Note that it is technically always possible to write $z_i = f_i(\boldsymbol{s}, \boldsymbol{g})$ anonymously, for example as $z_i = \tilde{f}(\boldsymbol{a}_i, \boldsymbol{g})$ where $\boldsymbol{a}_i$ includes copies of $\boldsymbol{s}$ and an observation indicator such that $\tilde{f}(\cdot)$ returns $f_i(\boldsymbol{s}, \boldsymbol{g})$. Such constructions are unnatural in most settings with non-anonymous constructions, however.

example, by drawing on geological models of earthquake probabilities across regions to redraw earthquake realizations. This example demonstrates how the recentering logic and the specification of counterfactual shocks help address the specific challenges with non-anonymous (and non-linear) constructions.

4 Consistency and Inference

4.1 Consistency with Many Shocks

We next consider consistency of IV estimators which recenter by or control for μ_i. Even if the orthogonality condition (4) holds and $\tilde{z}_i$ is relevant, additional conditions are generally required for consistent estimation because of the potentially complex dependencies in the data. Indeed, $\tilde{z}_i$ captures common exposure of the observations to the shocks in $\boldsymbol{g}$ and may thus be correlated across i in non-standard ways. Similarly, the ε_i may exhibit non-standard dependencies from their common dependence on unobserved shocks. Consistency may nevertheless be guaranteed as long as $\boldsymbol{g}$ includes a large number of sufficiently-independent shocks regardless of the dependence structure of unobservables.

Formally, we consider the recentered IV estimator

$$\hat{\beta} = \frac{\sum_i y_i \tilde{z}_i}{\sum_i x_i \tilde{z}_i} = \beta + \frac{\sum_i \varepsilon_i \tilde{z}_i}{\sum_i x_i \tilde{z}_i}. \tag{9}$$

Here we assume that y_i and x_i (and correspondingly ε_i) have been residualized on some regression controls $\boldsymbol{r}_i \equiv \boldsymbol{r}_i(w)$, dropping the previous $\perp$ notation. This definition of $\hat{\beta}$ therefore nests estimators which control for functions of $\boldsymbol{w}$ that span μ_i (making recentering unnecessary) or include any predetermined controls after recentering z_i.

BHJ and BH study the convergence of $\hat{\beta}$ along a sequence of distributions P_N for the complete data $\left\{(y_i, x_i, z_i)_{i=1}^N, \boldsymbol{s}, \boldsymbol{g}, \boldsymbol{q}\right\}$ where the number of shocks $K_N = \dim(\boldsymbol{g})$ can vary with the number of observations N. In particular, they consider consistency: $\hat{\beta} \xrightarrow{p} \beta$ as $N \to \infty$. Here we suppress the N subscript for brevity.

As with the identification discussion in Section 3.2, it is instructive to first study consistency in the special case of shift-share instruments. Without loss of generality, assume $s_k = \frac{1}{N}\sum_i s_{ik} > 0$ for every k since, if $s_{ik} = 0$ for all i, shock g_k can be dropped. Consider:

Assumption 4. *(a) Many uncorrelated shocks:* $\mathbb{E}\left[\sum_{k=1}^K s_k^2\right] \to 0$ *and* $\operatorname{Cov}\left[g_k, g_{k'} \mid \boldsymbol{\varepsilon}, \boldsymbol{w}\right] = 0$ *for all* $k \neq k'$*;*

(b) First-stage: $\frac{1}{N}\sum_i x_i \tilde{z}_i \xrightarrow{p} \pi \neq 0$*;*

(c) Regularity: $\operatorname{Var}\left[g_k \mid \boldsymbol{\varepsilon}, \boldsymbol{w}\right] \leq B_g$ *and* $\mathbb{E}\left[\bar{\varepsilon}_k^2 \mid \boldsymbol{w}\right] \leq B_\varepsilon$ *uniformly over* N*, for* $\bar{\varepsilon}_k = \frac{1}{N}\sum_i s_{ik}\varepsilon_i / s_k$.[13]

We then have:

[13] All conditions here and in Assumption 5 below should be understood as applying almost surely with respect to $\boldsymbol{w}$.

Proposition 1. *In the shift-share case where* $\tilde{z}_i = \sum_{k=1}^{K} s_{ik}\tilde{g}_k$ *with* $s_{ik} \geq 0$ *and* $\tilde{g}_k = g_k - \boldsymbol{q}_k'\theta$*, Assumptions 3 and 4 imply* $\hat{\beta} \xrightarrow{p} \beta$*.*

Proof. Follows from Proposition 3 of BHJ applied to the demeaned shocks $\tilde{\boldsymbol{g}} = (\tilde{g}_k)_{k=1}^{K}$. □

Assumption 4(a) ensures that the law of large numbers applies to $\frac{1}{N}\sum_i \varepsilon_i \tilde{z}_i$ regardless of the mutual correlations of ε_i. It requires that the number of shocks grows with the sample size (since $\sum_{k=1}^{K} s_k^2 \geq 1/K$), that shocks are mutually uncorrelated, and average exposure to them is dispersed in the sense of the Herfindahl index converging to zero.[14] Just like in conventional regressions, mutual uncorrelatedness of the shocks can easily be relaxed, e.g. by having many uncorrelated clusters of shocks or other forms of weak dependence between them, with the Herfindahl condition appropriately strengthened (see Assumptions 5 and 6 in BHJ). In a panel context the large number of shocks can arise either from many shocks in each cross-section or from a long time series, a point we return to in Section 5.

Assumption 4(b) requires the first-stage covariance $\frac{1}{N}\sum_i \tilde{z}_i x_i$ to converge to some non-zero constant. BHJ provide low-level conditions sufficient for this using a model that often underlies the usage of shift-share instruments in the first place (and which holds trivially in reduced-form studies where $x_i = z_i$). Specifically, they assume that the treatment can naturally be decomposed into k-specific components each strongly affected by the respective shock: $x_i = \sum_k s_{ik}\tilde{x}_{ik}$ with $\tilde{x}_{ik} = \pi_{ik} g_k + u_{ik}$, $\pi_{ik} \geq \bar{\pi} > 0$ and $\mathbb{E}\left[\boldsymbol{g} \mid \boldsymbol{w}, \boldsymbol{\varepsilon}, \boldsymbol{u}\right] = \mathbb{E}\left[\boldsymbol{g} \mid \boldsymbol{w}\right]$ for $\boldsymbol{u} = (u_{ik})_{i,k}$. With this model and additional regularity conditions, Assumption 4(b) follows if $\mathbb{E}\left[\frac{1}{N}\sum_i \left(\sum_k s_{ik}^2\right)\right]$ is bounded from below. This last condition requires an average observation to be effectively exposed to only a small number of shocks, in the sense of a non-vanishing Herfindahl index of exposure shares, such that the law of large numbers does *not* apply to individual $\tilde{z}_i$ and the variance of the instrument does not converge to zero. It is instructive to compare this condition with Assumption 4(a) which requires the *average* exposure to be dispersed. Both conditions can hold when most observations are exposed to a small set of shocks, but different ones for different observations. For example, they hold in the ADH setting when most local labor markets specialize in a small number of manufacturing industries — the identities of which vary across markets.

Propositions 1–6 in BH generalize Proposition 1 and the sufficient conditions for a non-vanishing first stage to nonlinear and non-anonymous formula instruments. While the assumptions become more technically restrictive, the main intuition extends: dispersed average exposure to sufficiently independent shocks makes $\frac{1}{N}\sum_i \varepsilon_i \tilde{z}_i$ converge to zero regardless of the mutual correlation in the errors, while concentrated individual shock exposure is key to the first-stage. We reproduce one of the sets of sufficient conditions for convergence of $\frac{1}{N}\sum_i \tilde{z}_i \varepsilon_i$, applicable to continuous shocks:

Assumption 5. *(a) Many independent shocks: the components of* $\boldsymbol{g}$ *are jointly independent conditionally on* $\boldsymbol{w}$*, and* $\sum_k \mathbb{E}\left[\left(\frac{\partial \bar{f}(\boldsymbol{g},\boldsymbol{w})}{\partial g_k}\right)^2\right] \to 0$ *for* $\bar{f}(\boldsymbol{g},\boldsymbol{w}) = \frac{1}{N}\sum_i \left(f_i(\boldsymbol{s},\boldsymbol{g}) - \mu_i(\boldsymbol{w})\right)$*;*

[14] The literal interpretation of $\sum_k s_k^2$ as a Herfindahl index and the result of $\sum_{k=1}^{K} s_k^2 > 1/K$ both technically require complete shares, $\sum_k s_{ik} = 1$, which implies $\sum_k s_k = 1$. However, in typical incomplete share cases where $\sum_k s_k$ is between zero and one and bounded away from zero, the same intuition applies.

(b) Each $f_i(\boldsymbol{s}, \boldsymbol{g})$ is weakly monotone in $\boldsymbol{g}$;

(c) Each g_k has a Gaussian distribution;

(d) Regularity: $\mathbb{E}\left[\varepsilon_i^2 \mid \boldsymbol{w}\right] \leq U_\varepsilon$ uniformly; $\mathbb{E}\left[\left|\frac{\partial \bar{f}(\boldsymbol{g},\boldsymbol{w})}{\partial g_k}\right| \mid \boldsymbol{w}\right] < \infty$ for all k; $\operatorname{Var}\left[g_k \mid \boldsymbol{w}\right] \in [L_\sigma, U_\sigma]$ for fixed constants $0 < L_\sigma < U_\sigma < \infty$.

Proposition 2. *Assumptions 1 and 5 imply* $\operatorname{Var}\left[\frac{1}{N}\sum_i \tilde{z}_i \varepsilon_i\right] \to 0$.

Proof. Follows from Propositions 1 and 4(a) of BH. □

BH also establish a similar result for Bernoulli-distributed shocks, where $\frac{\partial \bar{f}(\boldsymbol{g},\boldsymbol{w})}{\partial g_k}$ is replaced by discrete derivatives, as well as analogous low-level conditions for convergence of the first stage.

4.2 Asymptotic Inference with Shift-Share Variables

We now consider statistical inference on β. Inference is challenging in this setting, even with a large number of shocks, again because the observations of z_i (ε_i) are jointly exposed to the observed (unobserved) shocks and may therefore exhibit complex mutual correlations. Adão et al. (2019, henceforth AKM) illustrate this problem in the linear shift-share setting, via a Monte Carlo simulation based on the specifications in Autor et al. (2013). They show that conventional heteroskedasticity-robust and state-clustered standard errors, with nominal 5% significance level, lead to rejections of the true null of $\beta = 0$ in around half of their placebo samples with randomly generated shocks.

To address this issue, AKM derive new design-based central limit theorems and variance estimators for $\hat{\beta}$ in the linear shift-share context, which are asymptotically valid regardless of the correlation structure of the errors—what BHJ label "exposure-robust." The variance estimators can be motivated by first writing

$$\sqrt{N}(\hat{\beta} - \beta) = \frac{\frac{1}{\sqrt{N}}\sum_i \varepsilon_i \tilde{z}_i}{\frac{1}{N}\sum_i x_i \tilde{z}_i} = \left(\frac{1}{\sqrt{N}}\sum_i \varepsilon_i \tilde{z}_i\right)\pi^{-1}(1 + o_p(1)), \tag{10}$$

using Assumption 4(b). For simplicity, suppose Assumption 3 holds with $\theta = 0$, such that $\mathbb{E}\left[g_k \mid \boldsymbol{\varepsilon}, \boldsymbol{w}\right] = 0$ and $\tilde{z}_i = z_i$. Then:

$$\begin{aligned}\operatorname{Var}\left[\hat{\beta}\right] &\approx \frac{1}{\pi^2}\operatorname{Var}\left[\frac{1}{N}\sum_i \varepsilon_i z_i\right] \\ &= \frac{1}{\pi^2}\operatorname{Var}\left[\frac{1}{N}\sum_k g_k\left(\sum_i s_{ik}\varepsilon_i\right)\right] \\ &= \frac{1}{\pi^2}\sum_k \mathbb{E}\left[g_k^2\left(\frac{1}{N}\sum_i s_{ik}\varepsilon_i\right)^2\right]\end{aligned} \tag{11}$$

using Assumption 4(a) in the final line. This suggests a feasible estimator of $\frac{1}{\hat{\pi}^2}\sum_k g_k^2\left(\frac{1}{N}\sum_i s_{ik}\hat{\varepsilon}_i\right)^2$ where $\hat{\pi} = \frac{1}{N}\sum_i \tilde{z}_i x_i$ and $\hat{\varepsilon}_i = y_i - \hat{\beta}x_i$. AKM prove the validity of similar variance estimators in the general case of $\theta \neq 0$ and when Assumption 4(a) is weakened to allow for clustered shocks.[15]

BHJ propose a convenient way to obtain exposure-robust standard errors based on this asymptotic approximation, which come from standard formulas applied to an equivalent IV regression estimated at the "level" of the shocks. To see this result, it is first useful to note (following Proposition 1 in BHJ) that $\hat{\beta}$ equals the second-stage coefficient from a s_k-weighted shock-level IV regression, which uses the shocks g_k as the instrument to estimate[16]

$$\bar{y}_k = \beta \bar{x}_k + \bar{\varepsilon}_k. \tag{12}$$

This result follows from the definition of z_i:

$$\hat{\beta} = \frac{\sum_i \frac{1}{N} \left(\sum_k s_{ik} g_k\right) y_i}{\sum_i \frac{1}{N} \left(\sum_k s_{ik} g_k\right) x_i} = \frac{\sum_k g_k \left(\frac{1}{N} \sum_i s_{ik} y_i\right)}{\sum_k g_k \left(\frac{1}{N} \sum_i s_{ik} x_i\right)} = \frac{\sum_k s_k g_k \bar{y}_k}{\sum_k s_k g_k \bar{x}_k}. \tag{13}$$

Moreover, if $\boldsymbol{Q}_i = \sum_k s_{ik} \boldsymbol{q}_k$ has been included in the controls $\boldsymbol{r}_i$, the same coefficient is obtained from a richer shock-level IV specification:

$$\bar{y}_k = \beta \bar{x}_k + \boldsymbol{q}_k' \delta + \bar{\varepsilon}_k. \tag{14}$$

The *ssaggregate* packages in Stata and R automate the translation of linear shift-share regressions to the shock level.

Under regularity conditions (see BHJ, Proposition 5), the conventional heteroskedasticity-robust standard error for $\hat{\beta}$ in (14) yields asymptotically valid confidence intervals for β. While this derivation follows under the assumption of conditionally uncorrelated shocks (Assumption 4(a)), BHJ note the same shock-level regression can be used to conduct asymptotically valid inference when shocks are clustered, again using conventional clustered standard errors. Indeed, one convenient feature of their approach is the flexibility of accommodating other relaxations of Assumption 4(a), such as serial correlation (with conventional heteroskedastic-and-autocorrelation-consistent standard errors) or two-way clustering.

Remark 6. While BHJ only consider the case where $\boldsymbol{Q}_i$ has been initially controlled for, equation (14) produces valid estimates (and, we conjecture, standard errors) regardless of this, consistent with the recentering logic. Indeed, by the Frisch-Waugh-Lovell theorem, $\hat{\beta}$ from (14) can be obtained by residualizing the shocks on $\boldsymbol{q}_k$ (with weights s_k) and constructing the shift-share instrument from the residualized shocks.

Remark 7. A key assumption underlying the validity of the BHJ approach to exposure-robust inference is that only two types of controls are included in $\boldsymbol{r}_i$. First, all sources of shock-level

[15] AKM also consider other extensions, including to confidence intervals which impose a null hypothesis on β and to settings with heterogeneous treatment effects (but assuming that each observation is asymptotically exposed to only one shock).

[16] Note that (12) does not include the intercept, which is unusual, but (14) below shows that including it does not change the estimate so long as the shares are complete or the sum of shares is controlled for.

confounding have to be captured by controls with a shift-share structure (i.e., $\boldsymbol{Q}_i$). Second, all other controls should not be asymptotically correlated with the instrument; these controls may however increase the asymptotic efficiency of the estimator. BHJ show that the standard errors are asymptotically conservative under a much weaker assumption, which allows for controls in $\boldsymbol{r}_i$ of the form $\sum_k s_{ik}\boldsymbol{p}_k + u_i$, where $\boldsymbol{p}_k$ are unobserved shock-level confounders and u_i is noise. AKM provide an alternative variance estimator which is asymptotically exact in the case of controls with an "approximate" shift-share structure. Asymptotically exact shift-share inference with general control vectors (which are necessary for identification and do not have this form) remains an open problem.

Remark 8. The asymptotic inference results of AKM and BHJ are also useful to conduct regression-based falsification tests of shock orthogonality, and for verifying the strength of the instrument. Falsification tests for orthogonality can be conducted by regressing any proxy for the unobserved error (such as a pre-trend in the outcome) on the shift-share instrument while controlling for $\boldsymbol{Q}_i$. For asymptotically valid tests, the coefficient on the instrument can be computed by the shock-level regression (14) with exposure-robust standard errors. Similarly, a valid first-stage F-statistic can be computed by translating the regression of x_i on z_i and $\boldsymbol{Q}_i$ to the shock level.

Remark 9. Note that reduced-form regressions on z_i are still IV regressions at the shock level, which instrument $\bar{z}_k$ with g_k controlling for $\boldsymbol{q}_k$.

4.3 Randomization Inference

The asymptotic results of AKM and BHJ are specific to linear shift-share variables with a large number of shocks. We now discuss an alternative randomization inference (RI) approach, which BH propose for other formula instruments or settings with only a small number of shocks. This approach may be natural when Assumption 2 is made, because specified shock counterfactuals immediately yield confidence intervals that are exact in finite samples of observations and shocks—albeit only under the constant-effect model (1).[17]

Formally, BH propose to construct confidence intervals for β using the sample covariance between $\tilde{z}_i$ and the residual as a randomization test statistic:

$$T = \frac{1}{N}\sum_i \left(f_i(\boldsymbol{s},\boldsymbol{g}) - \mu_i\right)\left(y_i - bx_i\right), \tag{15}$$

where b is a candidate parameter value. Under the null hypothesis of $\beta = b$ and Assumption 1, $y_i - bx_i = \varepsilon_i$, and the distribution of T conditional on $\boldsymbol{\varepsilon}$ and $\boldsymbol{w}$ is implied by the shock assignment process $G(\boldsymbol{g} \mid \boldsymbol{w})$. One may simulate this distribution under Assumption 2, by redrawing the shocks and recomputing T. The null is rejected if the original value of T is far in the tails of the simulated

[17]RI methods are typically used to test the "sharp null" of zero treatment effects for all observations; valid inference under weaker nulls of (say) no average effect is a more challenging problem in general. See, e.g., Chung and Romano (2013) for an approach motivated by this challenge.

distribution.[18] The confidence interval for β is then obtained by inversion of such tests, i.e. by collecting all b that are not rejected. These intervals have correct size, both conditionally on $(\boldsymbol{\varepsilon}, \boldsymbol{w})$ and unconditionally.

While any statistic $T(\boldsymbol{g}, \boldsymbol{y} - b\boldsymbol{x}, \boldsymbol{w})$ can in principle be used in a similar procedure, the choice of (15) is motivated by its close connection to the IV estimator $\hat{\beta}$. Specifically, Borusyak and Hull (2021b) show that $\hat{\beta}$ is the Hodges-Lehman estimator corresponding to T (Hodges and Lehmann 1963; Rosenbaum 2002) and provide guarantees for the consistency of the randomization tests when $\hat{\beta}$ is consistent.

Remark 10. RI-based confidence intervals can be useful even when few shocks are observed in the data, making the asymptotic approach inapplicable even with linear formulas. Consider for example the single earthquake studied by Carvalho et al. (2021). Even without spillovers, it is not possible to consistently estimate the effect of earthquakes if only one region is shocked. Yet randomization inference remains informative in this case, as it can produce confidence intervals of finite length. For example, if the true effect is zero, it is unlikely that unobserved shocks hit exactly the same region where the earthquake randomly happened and the randomization test could reject $\beta = 0$.

Remark 11. As with asymptotic shift-share inference, RI-based confidence intervals can be used for falsification tests of the identifying assumptions. A falsification test of Assumption 1 is obtained by checking that the sample covariance between $\tilde{z}_i$ and a proxy for ε_i is close to zero (or, more precisely, is not in the tail of the distribution of this covariance across counterfactual shocks). Similarly, to test the correct specification of the shock assignment process (Assumption 2) and the expected instrument implied by it, one can check that $\tilde{z}_i$ is not correlated with any prespecified functions of $\boldsymbol{w}$. While the assumption of constant effects required for the validity of RI-based confidence intervals can be restrictive in practice, it is not a problem for falsification tests where the true effect should be zero for all i.[19]

5 Extensions

Identification with heterogeneous effects Borusyak and Hull (2021b, Appendix C.1) and BHJ (Appendix A.1) show that classic results on linear IV identification in the presence of heterogeneous treatment effects (e.g., Imbens and Angrist 1994) extend to formula instruments, provided μ_i is recentered by or controlled for and a version of the usual first-stage monotonicity condition holds. Specifically, they show that when an exclusion restriction holds but the outcome model features nonlinear and heterogeneous effects of the treatment (as opposed to the linear specification in (1)), the recentered IV estimand is a convex weighted average of the marginal effects of x_i on y_i. For example, in the reduced-form case with linear effect heterogeneity, i.e. $y_i = \beta_i z_i + \varepsilon_i$, recentered IV identifies $\mathbb{E}\left[\frac{1}{N}\sum_i \omega_i \beta_i\right]$ where the weights ω_i are proportional to $\operatorname{Var}[z_i \mid \boldsymbol{w}]$. This conditional

[18] Note that, while we recenter T in equation (15), this is not necessary since the term $\frac{1}{N}\sum_i \mu_i(y_i - bx_i)$ shifts the observed and re-randomized T equally.

[19] Testing the first-stage relationship between x_i and $\tilde{z}_i$ is similarly straightforward, though as noted by Imbens and Rosenbaum (2005) there is no weak instrument problem for RI-based confidence intervals.

variance can be computed by a simulation analogously to μ_i, allowing researchers to study the implied weights of the estimand. When $\mathrm{Var}\left[z_i \mid \boldsymbol{w}\right]$ is bounded away from zero, a recentered IV estimator which inversely weights by it identifies the unweighted average causal effect $\mathbb{E}\left[\frac{1}{N}\sum_i \beta_i\right]$.

In the IV case, Borusyak and Hull (2021b) generalize Imbens and Angrist (1994) by allowing the regression of x_i on $\tilde{z}_i$ to be non-causal, in the sense that the first-stage coefficient on $\tilde{z}_i$ does not capture the causal effect of shocks $\boldsymbol{g}$ on x_i. This extension is useful for formula instruments, since $\tilde{z}_i$ is a constructed variable that does not usually correspond to a real economic object. Thus, it is more appropriate to specify the potential values of x_i in terms of the primitive shocks $\boldsymbol{g}$. BH show that $\hat{\beta}$ retains its interpretation as a convexly-weighted average of causal effects under an appropriate monotonicity assumption, even if $\tilde{z}_i$ does not capture all of the ways how the shocks affect the treatment. In the context of linear shift-share instruments, BHJ consider the case where may use exposure shares s_{ik} are misspecified, such that they do not correlate imperfectly with the true shares through which shocks g_k affect x_i. Nevertheless, a causally interpretable IV estimand is guaranteed as long as x_i is correctly specified, the s_{ik} are non-negative, the shocks are weakly positively correlated, and the true effects of shocks on each treatment are monotone.

These results can be contrasted to those of Blandhol et al. (2022), who show the importance of including controls flexibly for IV regressions to be "weakly causal" (i.e.,to identify a convex average of heterogeneous effects). Here design knowledge yields the single control μ_i needed for the IV estimand to be weakly causal; other $\boldsymbol{r}_i(\boldsymbol{w})$ can again increase precision while not affecting this estimand.

Multiple treatments and instruments The identification results of Section 3 extend immediately to cases with multiple formula instruments $f_{\ell i}(\boldsymbol{s}, \boldsymbol{g})$, $\ell = 1, \ldots, L$, for either a single treatment x_i or multiple treatments x_{pi}, $p = 1, \ldots, P \leq L$. As long as each expected instrument $\mathbb{E}\left[f_{\ell i}(\boldsymbol{s}, \boldsymbol{g}) \mid \boldsymbol{w}\right]$ is adjusted for via recentering or controlling, the corresponding orthogonality conditions hold—such that parameters $(\beta_1, \ldots, \beta_P)$ are identified under a standard rank condition.

The case of multiple treatments is particularly useful in network settings where it allows including a node's own shock along with a treatment capturing network spillovers; e.g, student i's own deworming status and the number of her neighbors who have been dewormed, in the Miguel and Kremer (2004) setting. Notably, the same set of exogenous shocks allows for identification of the effects of both treatments, via different exposure mappings.[20]

The case of multiple instruments is also relevant when the researcher has access to multiple sets of exogenous shocks. BHJ illustrate this scenario in the ADH China shock setting. While ADH measure industry supply shocks as average Chinese import growth across eight non-U.S. countries, in principle the shocks from each individual country may be considered as-good-as-randomly assigned, leading to multiple shift-share instruments for the regional exposure to import competition from China and thus overidentification of the parameter of interest. BHJ show how to extend the shock-level representation of the shift-share IV estimator when the exposure shares

[20]Constant effects play a bigger role with multiple treatments, however. See Bhuller and Sigstad (2023) and Goldsmith-Pinkham et al. (2022) for discussions in the context of IV and reduced-form regressions.

used to construct the instruments are the same but the shocks differ. This allows them to obtain not only identification but also asymptotic exposure-robust inference.

Other necessary controls The framework so far may seem unable to accommodate certain controls that may be deemed necessary for identification. Specifically, the identification results in Section 3.1 only allow for controls $\boldsymbol{r}_i$ which are functions of $\boldsymbol{w}$. The purpose of these controls for identification is to absorb confounding variation in μ_i; beyond that, they only play a role in estimation efficiency. In the linear shift-share case of $z_i = \sum_k s_{ik} g_k$, for example, this means all necessary controls must have a shift-share structure of the form $\sum_k s_{ik} \boldsymbol{q}_k$.

One could imagine, however, other necessary controls which remove confounding variation in ε_i. Such controls can be accommodated by augmenting the causal model (1) to have $\varepsilon_i = \boldsymbol{c}_i' \lambda + \eta_i$ for an observed $\boldsymbol{c}_i$ (which is not assumed conditionally independent of g) and an error η_i defined to be orthogonal to $\boldsymbol{c}_i$ (without loss of generality, since λ is not interpreted causally). Then

$$y_i = \beta x_i + \boldsymbol{c}_i' \lambda + \eta_i \tag{16}$$

and the identification question shifts to the orthogonality of z_i and η_i. Modifying Assumption 1 to the conditional independence of $\boldsymbol{\eta} = (\eta_i)_{i=1}^N$ and $\boldsymbol{g}$ given $\boldsymbol{w}$ (and thus allowing violations of $\boldsymbol{g} \perp\!\!\!\perp \boldsymbol{\varepsilon} \mid \boldsymbol{w}$ via $\boldsymbol{c}$) and maintaining Assumption 2 and appropriate instrument relevance, we obtain identification in a recentered IV regression of y_i on x_i which instruments by $\tilde{z}_i$ and controls for $\boldsymbol{c}_i$, because orthogonality conditions $\mathbb{E}\left[\frac{1}{N}\sum_i \tilde{z}_i \eta_i\right] = \mathbb{E}\left[\frac{1}{N}\sum_i \boldsymbol{c}_i \eta_i\right] = 0$ both hold. The same result holds for linear instruments under Assumption 3 modified to specify that the conditional mean of g_k given $(\boldsymbol{\eta}, \boldsymbol{w})$ is linear in $\boldsymbol{q}_k$. Alternatively, as before, μ_i can be controlled for along with $\boldsymbol{c}_i$ with the original z_i used as the instrument.[21]

To make these ideas concrete, consider again the setting of Autor et al. (2013). Imagine that regions more exposed to supply shocks in China also have higher (lagged) local unemployment rates, due to some location-specific factors (e.g., local regulations). The local unemployment rate does not have a shift-share structure, as there is residual variation in unemployment rates across locations with the same industry composition. However, one can directly control for the (lagged) local unemployment rate for identification.

Panel data While we have denoted observations by i having cross-sectional data in mind, the results of Sections 2–4 extend directly to panel data where observations $i = (\ell, t)$ correspond to units ℓ in periods t and shocks can also vary over time. We make two remarks specific to panel data.

First, including unit fixed effects does not generally alleviate the need to adjust the formula instrument, unless the corresponding expected instrument is time-invariant. Consider the instrument $z_{\ell t} = f_{\ell t}(\boldsymbol{s}_t, \boldsymbol{g}_t)$ with $\boldsymbol{g}_t \perp\!\!\!\perp \boldsymbol{\varepsilon}_t \mid \boldsymbol{w}_t$ (here objects like $\boldsymbol{v}_t$, with just a t subscript, refer to the vector of $v_{\ell t}$). Then $\mu_{\ell t} = \mathbb{E}[z_{\ell t} \mid \boldsymbol{w}_t]$ is absorbed by the unit fixed effects under the very restrictive

[21] Consistency and inference with such controls may be more complicated (see Remark 7).

conditions that the $f_{\ell t}(\cdot)$ mapping, the value of $\boldsymbol{w}_t$, and the shock assignment process $G(\boldsymbol{g}_t \mid \boldsymbol{w}_t)$ are all time-invariant. In linear and nonlinear shift-share settings, for example, even if exposure shares are measured in a fixed pre-period for the entire panel, the shocks often vary systematically over time. Footnote 11 above illustrates this issue in the ADH setting where industry-level China shocks are higher in the 2000s than in the 1990s, making $\mu_{\ell t}$ time-varying.

Second, an advantage of panel data is that the number of cross-sectional shocks (and observations) needs not be large; instead, a long time series of shocks g_t with weak serial dependence may also suffice for consistency of $\hat{\beta}$. BHJ formalize this insight for linear shift-share settings. Nunn and Qian (2014) is an empirical example of constructing shift-share instruments from purely time-series exogenous shocks (specifically, in total U.S. foreign aid arising from wheat production in the U.S.) with no cross-sectional variation.

Estimated shock assignment processes Assumption 2 requires complete knowledge of the shock assignment process, conditionally on $\boldsymbol{w}$. Appendix C.5 of Borusyak and Hull (2021b) shows that this assumption can be relaxed with the assignment process specified up to a vector of parameters, $G(\boldsymbol{g} \mid \boldsymbol{w}; \theta)$. If these parameters can be consistently estimated by $\hat{\theta}$ from shock-level data (e.g., by maximum likelihood) and some regularity conditions hold, adjusting the instrument by the estimated $\hat{\mu}_i = \int f_i(\boldsymbol{s}, \boldsymbol{g}) dG(\boldsymbol{g} \mid \boldsymbol{w}; \hat{\theta})$ leads to consistent estimation of β. This generalizes Assumption 3 (which allowed the conditional shock mean to be parameterized) to nonlinear formula instruments. This approach is especially attractive when the shocks are binary (and mutually independent conditionally on $\boldsymbol{w}$), as a model the conditional mean, e.g. via logit or probit, yields the entire conditional distribution.

Inference on β will typically be affected by the first-step estimation of θ. Appendix C.5 of Borusyak and Hull (2021b) shows how the baseline BH randomization inference procedure can be extended to build conservative finite-sample confidence intervals, drawing on the approach of Berger and Boos (1994). It further notes that there are cases in which exact confidence intervals can be constructed by using sufficient statistics that obviate the need to estimate θ. In some settings valid asymptotic inference may also be possible.

Most efficient recentered instruments The results of Sections 2–4 take the instrument construction $f_i(\cdot)$ as given, but different constructions are likely to yield estimates with different levels of precision in large samples. Borusyak and Hull (2021a) study the question of optimal instrument construction (under constant effects), building on Chamberlain (1987, 1992). In a class of regular recentered IV estimators, the asymptotically most efficient one can be described in a three step process: the best prediction of $\boldsymbol{x}$ from $(\boldsymbol{g}, \boldsymbol{w})$ is constructed, recentered using the shock assignment process, and then adjusted for "heteroskedasticity" (formally, weighted by the inverse of $\mathbb{E}[\boldsymbol{\varepsilon}\boldsymbol{\varepsilon}' \mid \boldsymbol{w}]$ which also depends on $\mathbb{E}[\boldsymbol{\varepsilon} \mid \boldsymbol{w}]$). The prediction and recentering steps can be seen to justify the empirical practice, noted in Remark 2, of forming instruments from the structure of x_i by removing or replacing some endogenous components. The heteroskedasticity adjustment step is infeasible in non-*iid* data but it is often ignored in practice even in settings where it would be feasible (see, e.g., Coussens and Spiess (2021)).

6 Comparisons with Alternative Approaches

The design-based approach leverages knowledge of the formula construction $f_i(\cdot)$ to correct for potential confounding. A natural question is whether identification is possible without such knowledge, potentially under alternative assumptions. We answer this question by comparing the design-based approach to two alternative strategies for identification in linear models: those based on conventional assumptions of conditional treatment unconfoundedness, and those based on alternative restrictions on outcome unobservables (such as parallel trend assumptions).

For both comparisons we use the fact that, under Assumption 1, equation (1) can be written as a partially linear model in the spirit of Robinson (1988). Specifically, with $h_i(\boldsymbol{w}) = \mathbb{E}[\varepsilon_i \mid \boldsymbol{w}]$,

$$y_i = \beta x_i + h_i(\boldsymbol{w}) + \eta_i, \qquad \mathbb{E}[\eta_i \mid \boldsymbol{g}, \boldsymbol{w}] = 0. \tag{17}$$

One approach to identification in partially linear models, advanced by Robins et al. (1992) and in line with a long literature on propensity score methods, adjusts x_i for its association with potential confounders in $\boldsymbol{w}$ before estimating β. A different approach, considered by Goldsmith-Pinkham et al. (2020) for linear shift-share instruments, restricts the form of $h_i(\boldsymbol{w})$ in order to directly estimate (17). This section connects the design-based approach to formula instruments to the first strategy—showing that explicit use of the formula is essential in this case—before contrasting it with the alternative restrictions implied by the second strategy.

6.1 Conventional Unconfoundedness

Partially linear models are typically considered in *iid* data where $x_i = g_i$ is assumed to be unconfounded given observation-specific $\boldsymbol{w}_i$. That is, with $h(\boldsymbol{w}_i) = \mathbb{E}[\varepsilon_i \mid \boldsymbol{w}_i]$,

$$y_i = \beta x_i + h(\boldsymbol{w}_i) + \eta_i, \qquad \mathbb{E}[\eta_i \mid x_i, \boldsymbol{w}_i] = 0. \tag{18}$$

Based on a suggestion by Newey (1990), Robins et al. (1992) consider estimation of β when a researcher has "sharper" information on the relationship between the observed x_i and $\boldsymbol{w}_i$ than on the form of unobservables as captured by $h(\boldsymbol{w}_i)$. This motivates their "E-estimator":

$$\hat{\beta}_E = \frac{\sum_i y_i \left(x_i - \mathbb{E}[x_i \mid \boldsymbol{w}_i]\right)}{\sum_i x_i \left(x_i - \mathbb{E}[x_i \mid \boldsymbol{w}_i]\right)}. \tag{19}$$

The recentered IV estimator of BH can be seen to generalize the E-estimator to the setting where $x_i = f_i(\boldsymbol{s}, \boldsymbol{g})$ is constructed by a formula from common shocks. The $\mathbb{E}[x_i \mid \boldsymbol{w}_i]$ term becomes μ_i in this case. A further generalization is obtained by letting $\tilde{z}_i = z_i - \mu_i$ instrument for a different treatment x_i, now with $z_i = f_i(\boldsymbol{s}, \boldsymbol{g})$ and $\mu_i = \mathbb{E}[z_i \mid \boldsymbol{w}_i]$.[22]

A central difference in the BHJ/BH settings, where the treatment or instrument is constructed from common shocks, is that estimation of $\mu_i = \mathbb{E}[f_i(\boldsymbol{s}, \boldsymbol{g}) \mid \boldsymbol{w}]$ via some first-stage procedure is not possible without using the formula. In the conventional setting Robins et al. (1992) study, $\mathbb{E}[x_i \mid \boldsymbol{w}_i]$ may be learned non-parametrically from the conditional distribution of observation-specific shocks g_i given the confounders $\boldsymbol{w}_i$. But when all observations i are exposed to the *same* vector of shocks $\boldsymbol{g}$, and only one realization of this vector is observed, such non-parametric estimation is impossible. Indeed, in conventional settings there are many observations with similar $\boldsymbol{w}_i$ but different values of x_i, such that $\mathbb{E}[x_i \mid \boldsymbol{w}_i]$ can be learned. In contrast, for the linear shift-share instruments there is no cross-sectional variation in $\sum_k s_{ik} g_k$ conditionally on $\boldsymbol{s}_i$, while variation across possible realizations of $\boldsymbol{g}$ vectors is not observed. This problem is distinct from challenges related to high-dimensional confounding, which can be solved by modern machine learning methods (e.g. Chernozhukov et al. (2018)).[23]

Leveraging the knowledge contained in the formula, however, allows the researcher to circumvent the common shocks issue. This is achieved in two steps: by first specifying or estimating the assignment process for shocks, similar to the design-based approach of Robins et al. (1992), and then "translating" it to the level at which observations and treatment are observed. The shock-level equivalence results of BHJ make this translation especially clear in the linear shift-share case.

6.2 Outcome Model Restrictions

An alternative approach to estimating β in the partially linear model (17) restricts the unobserved outcome error ε_i, without restricting the assignment process of conditionally unconfounded shocks. Specifically, one posits a model for $h_i(\boldsymbol{w})$ and jointly estimates it along with β. Specification of the shock design plays no role in this strategy. In fact, the strategy is coherent when the shocks are considered non-random, in which case Assumptions 1 and 2 hold trivially. More generally, consider a linear outcome model replacing Assumptions 1 and 2 (or 3):

Assumption 6. *(Outcome Model) There exists unknown γ such that $\mathbb{E}[\varepsilon_i \mid \boldsymbol{g}, \boldsymbol{w}] = \boldsymbol{q}_i'\gamma$ for observed $\boldsymbol{q}_i$ included in $\boldsymbol{w}$ and for all i.*

Assumption 6 requires a weaker version of Assumption 1, $\mathbb{E}[\varepsilon_i \mid \boldsymbol{g}, \boldsymbol{w}] = \mathbb{E}[\varepsilon_i \mid \boldsymbol{w}]$, which is trivially satisfied when the shocks are non-random.

A concrete example of Assumption 6, in the linear shift-share setting, comes from Goldsmith-Pinkham et al. (2020). Conditioning on the shocks $\boldsymbol{g}$, they assume $\boldsymbol{\varepsilon}$ is mean-independent of the

[22]When z_i is binary, μ_i can be seen to generalize the propensity score of Rosenbaum and Rubin (1983) to the formula instrument setting. Propensity scores are typically defined in settings with randomly sampled data and a conditionally unconfounded treatment, and used in weighting or matching estimators rather than with regression. Appendix C.1 of Borusyak and Hull (2021b) shows how such estimators can be adapted to formula instruments.

[23]Consider, for instance, a shift-share instrument z_i where $\mathbb{E}[g_k \mid \boldsymbol{s}] = 0$ for all but a small set of $k \in \mathcal{K}$. The expected instrument $\mu(\boldsymbol{s}_i)$ is then sparse in $\boldsymbol{s}_i$, suggesting one might estimate it via LASSO methods. However, this intuition is misleading: the observed z_i are based on a single draw of each g_k which may be far from zero even for $k \in \mathcal{K}$ giving no hope to identify $\mathcal{K}$ or $\mu(\cdot)$ even as $N \to \infty$.

shares $\boldsymbol{s}$ after partialling out a vector of controls. Then $z_i = \sum_k s_{ik} g_k$ satisfies orthogonality. Justification for the key mean-independence assumption is similar to that underlying conventional "parallel trends" restrictions in difference-in-differences estimation and related strategies for causal inference in panel data. Indeed, shift-share regressions are often specified in first-differences, in which case $\mathbb{E}[\varepsilon_i \mid \boldsymbol{s}] = 0$ amounts to an assumption that observations with different exposure shares would have been on similar outcome trends if not for a change in x_i. Assumption 6 relaxes this condition by allowing for time-varying controls $\boldsymbol{q}_i$ and stochastic shocks.

Assumption 6 is very powerful; this can be seen from the fact that it justifies the orthogonality of *any* formula instrument $f_i(\boldsymbol{s}, \boldsymbol{g})$ when including the same set of controls $\boldsymbol{q}_i$. Indeed, in the Goldsmith-Pinkham et al. (2020) case, any function of $\boldsymbol{s}_i$ (including individual s_{ik}) satisfies instrument orthogonality. In this sense, identification strategies based on such models for ε_i can be said to not leverage the specific formula construction. Restrictions on ε_i may be especially valuable when design-based estimation is imprecise (e.g. because the shocks are too few or too mutually correlated), if credible assumptions on their assignment process are lacking, or if the shocks are unobserved.

The cost of this alternative model-based approach is that Assumption 6 can be very restrictive. In particular, BHJ show that in the shift-share setting the Goldsmith-Pinkham et al. (2020) approach generally fails when there are *any* unobserved shocks ν_k that affect the outcomes through the same (or correlated) exposure shares; e.g., when

$$\varepsilon_i = \sum_k s_{ik} \nu_k + \tilde{\varepsilon}_i,$$

for some idiosyncratic $\tilde{\varepsilon}_i$.[24] Ruling out such ν_k is implausible in applications where g_k are "specific" shocks that are of interest, while other shocks varying at the same level are likely.[25] For instance, in the ADH China shock setting, Assumption 6 is generally violated if there are any unobserved industry shocks affecting the outcome of interest—such as those arising from automation or changing tastes. Indeed, Goldsmith-Pinkham et al. (2020) find evidence against their "exogenous shares" assumption in the ADH setting.

Models for ε_i can more generally be unsatisfying for the same reason that motivates Robins et al. (1992): one may have limited information about the "right" form of $h_i(\boldsymbol{w})$, particularly since ε_i is unobserved. It may be unclear, for example, whether linearity of the unobservables more plausibly holds when the outcome is specified in levels vs. in logs (while both generally cannot hold; see Roth and Sant'Anna (2023)). It may moreover be challenging to pick the appropriate observation-specific features $\boldsymbol{q}_i$ in the kinds of non-*iid* settings where formula instruments are deployed, such as with spatial or network data. The challenge is amplified in the presence of heterogeneous effects where

[24] If $\mathbb{E}[\nu_k \mid \boldsymbol{s}] = 0$ for all k, Assumption 6 technically holds. But the presence of the ν_k shocks may still affect the consistency of the non-recentered shift-share IV estimator; see Appendix A.2 of BHJ.

[25] Conversely, BHJ argue this problem may be less concerning when the shares are "tailored" to the treatment of interest such that they are unlikely to mediate other unobserved shocks. One candidate for this is Card (2009), who uses lagged immigrant enclave shares and builds on an earlier difference-in-difference strategy in Card (1990).

a correct model of untreated potential outcomes may not suffice for the regression to estimate a convex average of treatment effects.[26] The design-based approach helps meet these challenges, as the formula for z_i and specification of the distribution of observed shocks guide the choice of controls via μ_i and, as noted in Section 5, the design-based approach guarantees a causally interpretable estimand even with heterogeneous effects under a first-stage monotonicity condition.

7 Conclusion

The design-based approach to formula instruments can be seen to bring new insights to longstanding identification strategies in economics (shift-share IV, in particular) while also pointing a way forward for novel econometric strategies which leverage more complex instrument constructions. An ever-increasing richness of data, sophistication of economic models, and creativity in discovering plausibly exogenous shocks are likely to yield many new settings where Assumptions 1–2 or 3 credibly hold. A focus on shock design in such settings can allow researchers to avoid extraneous or undesirably strong assumptions on how model unobservables relate to predetermined observables, provided the formula construction is leveraged appropriately for identification and inference.

Several open paths remain in this agenda. First, asymptotic inference results are only available in some special cases. It would be valuable to develop asymptotically valid inference techniques outside the linear shift-share case, and for shift-share instruments with unrestricted control vectors, as well as to extend results on randomization inference with heterogeneous treatment effects (e.g. Chung and Romano (2013)) to formula instruments. Second, it would be interesting to extend recent double/debiased machine learning results (e.g. Chernozhukov et al. (2018)) to shift-share instruments—allowing the vector $\boldsymbol{q}_k$ of shock-level confounders to be high-dimensional. Third, it would be useful to characterize properties of the recentering approach when the shock assignment process is estimated in a flexible way, e.g. by non-parametric estimation of the distribution of $g_k \mid \boldsymbol{q}_k$ when $(g_k, \boldsymbol{q}_k)$ are *iid.* Fourth, the estimand of the recentered IV procedure is not known when the exclusion restriction is violated. For instance, is recentering guaranteed to reduce bias in some situations even if the network spillovers embedded in x_i are misspecified? Finally, open questions remain on how recentering can relax the assumptions of more elaborate econometric methods, such as structural models for differentiated product demand. On these and other issues we expect many interesting developments in the growing literature on formula instruments.

References

Abadie, Alberto, Susan Athey, Guido W. Imbens, and Jeffrey M. Wooldridge. 2020. "Sampling-based vs. Design-based Uncertainty in Regression Analysis." *Econometrica* 88:265–296.

[26] This is now well-known in the context of staggered difference-in-differences designs (e.g., Chaisemartin and D'Haultfœuille (2020), Borusyak et al. (2023)). Restrictions on treatment effects can only be avoided in a restricted class of settings where "imputation" estimators are feasible (e.g., Borusyak et al. (2023), Wooldridge (2021)).

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

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

Aronow, Peter M., and Cyrus Samii. 2017. "Estimating average causal effects under general interference, with application to a social network experiment." *Annals of Applied Statistics* 11:1912–1947.

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* 103:2121–2168.

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

Basso, Gaetano, and Giovanni Peri. 2015. "The Association between Immigration and Labor Market Outcomes in the United States." *Working Paper.*

Berger, Roger L., and Dennis D. Boos. 1994. "P values maximized over a confidence set for the nuisance parameter." *Journal of the American Statistical Association* 89:1012–1016.

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

Berthou, Antoine, John Jong-hyun Chung, Kalina Manova, and Charlotte Sandoz Dit Bragard. 2019. "Productivity, (Mis)allocation and Trade." *Working Paper.*

Bhuller, Manudeep, and Henrik Sigstad. 2023. "2SLS with Multiple Treatments." *Mimeo.*

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

Blandhol, Christine, John Bonney, Magne Mogstad, and Alexander Torgovitsky. 2022. "When is TSLS Actually LATE?" *SSRN Electronic Journal.*

Borusyak, Kirill, and Peter Hull. 2021a. "Efficient Estimation with Non-Random Exposure to Exogenous Shocks." *Mimeo.*

———. 2021b. "Non-Random Exposure to Exogenous Shocks: Theory and Applications." *Mimeo.*

———. 2022. "Non-Random Exposure to Exogenous Shocks." *Working Paper.*

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

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess. 2023. "Revisiting Event Study Designs: Robust and Efficienct Estimation." *Working Paper.*

Boustan, Leah, Fernando Ferreira, Hernan Winkler, and Eric M. Zolt. 2013. "The effect of rising income inequality on taxation and public expenditures: Evidence from U.S. Municipalities and school districts, 1970-2000." *Review of Economics and Statistics* 95:1291–1302.

Card, David. 1990. "The Impact of the Mariel Boatlift on the Miami Labor Market." *ILR Review* 43:245–257.

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

Carvalho, Vasco M., Makoto Nirei, Yukiko U. Saito, and Alireza Tahbaz-Salehi. 2021. "Supply Chain Disruptions: Evidence from the Great East Japan Earthquake." *Quarterly Journal of Economics* 136:1255–1321.

Cattaneo, Matias D., Brigham R. Frandsen, and Rocío Titiunik. 2015. "Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate." *Journal of Causal Inference* 3:1–24.

Chaisemartin, Clément de, and Xavier D'Haultfœuille. 2020. "Two-way fixed effects estimators with heterogeneous treatment effects." *American Economic Review* 110:2964–2996.

Chamberlain, Gary. 1987. "Asymptotic efficiency in estimation with conditional moment restrictions." *Journal of Econometrics* 34:305–334.

———. 1992. "Efficiency Bounds for Semiparametric Regression." *Econometrica* 60:567–596.

Chernozhukov, Victor, Denis Chetverikov, Mert Demirer, Esther Duflo, Christian Hansen, Whitney Newey, and James Robins. 2018. "Double/debiased machine learning for treatment and structural parameters." *Econometrics Journal* 21:C1–C68.

Chung, Eunyi, and Joseph P. Romano. 2013. "Exact and asymptotically robust permutation tests." *Annals of Statistics* 41:484–507.

Costa, Rui, Swati Dhingra, and Stephen Machin. 2019. "Trade and Worker Deskilling." *Working Paper.*

Coussens, Stephen, and Jann Spiess. 2021. "Improving Inference from Simple Instruments Through Compliance Estimation." *Mimeo.*

Currie, Janet, and Jonathan Gruber. 1996. "Health Insurance Eligibility, Utilization of Medical Care, and Child Health." *The Quarterly Journal of Economics* 111:431–466.

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

Donaldson, Dave, and Richard Hornbeck. 2016. "Railroads and American Economic Growth: A "Market Access" Approach." *Quarterly Journal of Economics* 131:799–858.

Freedman, David A. 2008. "On regression adjustments to experimental data" [in en]. *Advances in Applied Mathematics* 40:180–193. Accessed October 23, 2023.

Goldsmith-Pinkham, Paul, Peter Hull, and Michal Kolesár. 2022. "Contamination Bias in Linear Regressions." *Working Paper.*

Goldsmith-Pinkham, Paul, and Guido Imbens. 2013. "Social Networks and the Identification of Peer Effects." *Journal of Business and Economic Statistics* 31:253–264.

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

Greenstone, Michael, Alexandre Mas, and Hoai-Luu Nguyen. 2014. "Do Credit Market Shocks Affect the Real Economy? Quasi-Experimental Evidence from the Great Recession and 'Normal' Economic Times." NBER Working Paper 20704.

Hodges, J.L. Jr., and Erich L Lehmann. 1963. "Estimates of Location Based on Rank Tests." *The Annals of Mathematical Statistics* 34:598–611.

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

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

Imbens, Guido W., and Paul R. Rosenbaum. 2005. "Robust, accurate confidence intervals with a weak instrument: quarter of birth and education." *Journal of the Royal Statistical Society: Series A (Statistics in Society)* 168 (January): 109–126.

Jaravel, Xavier. 2019. "The unequal gains from product innovations: Evidence from the us retail sector." *The Quarterly Journal of Economics* 134:715–783.

Lee, David S. 2008. "Randomized experiments from non-random selection in U.S. House elections." *Journal of Econometrics* 142:675–697.

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

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

Newey, Whitney K. 1990. "Efficient Instrumental Variables Estimation of Nonlinear Models." *Econometrica* 58:809–837.

Neyman, Jerzy. 1923. "On the application of probability theory to agricultural experiments. Essay on principles." *Ann. Agricultural Sciences,* 1–51.

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

Oberfield, Ezra, and Devesh Raval. 2021. "Micro data and macro technology." *Econometrica* 89:703–732.

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

Robins, James M, Steven D Mark, and Whitney K Newey. 1992. "Estimating Exposure Effects by Modelling the Expectation of Exposure Conditional on Confounders." *Biometrics* 48:479–495.

Robinson, P.M. 1988. "Root-N-Consistent Semiparametric Regression." *Econometrica* 56:931–954.

Rosenbaum, Paul R, and Donald B Rubin. 1983. "The Central Role of the Propensity Score in Observational Studies for Causal Effects Paul R. Rosenbaum, Donald B. Rubin." 70:41–55.

Rosenbaum, Paul R. 2002. "Covariance adjustment in randomized experiments and observational studies." *Statistical Science* 17:286–327.

Roth, Jonathan, and Pedro H.C. Sant'Anna. 2023. "When Is Parallel Trends Sensitive to Functional Form?" *Econometrica* 91.

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

Wooldridge, Jeffrey M. 2021. "Two-Way Fixed Effects, the Two-Way Mundlak Regression, and Event Study Estimators." *Working Paper.*

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