Novel Shift-Share Instruments and Their Applications

Benjamin Ferri*

September 26, 2022

Click here for most recent version.

Abstract

Shift-Share (Bartik) instruments are among the most important tools for causal identification in economics. In this paper, I crystallize main ideas underlying Shift-Share instruments - their core structure, distinctive claim to validity as instruments, history, uses, and wealth of varieties. I argue that the essence of the Shift-Share approach is to decompose the endogenous explanatory variable into an accounting identity with multiple component parts; preserve that which is most exogenous in the accounting identity, and neutralize that which is most endogenous. Following this framework, I show clearly how several variants in the literature are related. I then develop formulas for several new variants. Particularly, I show how to develop Shift-Share instruments for distribution summaries beyond the mean - the variance, skew, absolute deviation around a central point, and Gini coefficient. As an empirical application that highlights the themes of the paper, I measure the effect of earnings inequality on rates of single parenting in the U.S., comparing results using each of various alternative instruments for the Gini coefficient.

Keywords: shift-share, bartik, instrumental variables, panel data, labor demand and supply, earnings inequality, single parenting

JEL Classification: C23, C26, D31, J20, R12, R23

*Boston College, Department of Economics, benjferri@gmail.com. I am extremely grateful for the generous support and insights of Professors Donald Cox, Christopher Baum, Hanno Foerster, Shakeeb Khan, and Arthur Lewbel.

1 Introduction

The Shift-Share approach is a powerful and flexible framework for developing instrumental variables for causal identification. Many papers have surveyed and examined the validity of particular Shift-Share instruments, as given objects in the literature.[1] This paper rather examines the process by which one develops or arrives at Shift-Share instruments to begin with. In other words, this paper crystallizes the essential features of Shift-Share instruments, the purpose of each feature, and the scope of how each can be extended.

The upcoming Section 2 forms a foundation for the remainder of the paper: It defines the core elements of classical Shift-Share instruments, and develops a unified system of notation into which I will translate disparate papers' models for the purpose of comparison. In following sections, I discuss the history of Shift-Share instruments in the literature, ranging from standard to more exotic variants. By closely comparing many distinct variants, I illustrate both that which is most essential about Shift-Share instruments, and the wealth of different ways in which they may be adapted.

The primary contribution of this paper may be the simple framework it proposes for how to understand Shift-Share instruments. The essence of the Shift-Share approach is to decompose the endogenous explanatory variable as an accounting identity with multiple component parts; preserve that which is most exogenous in the accounting identity, and neutralize that which is most endogenous. Endogeneity is neutralized via delocalizations over space and time. That is, the more endogenous component - the Shift vector - is replaced with nonlocal averages (a strong delocalization); and the more exogenous component - the Share vector - is lagged (a weaker delocalization).

As an additional contribution, I develop several new varieties of Shift-Share instruments, particularly for explanatory variables that are distribution summaries other than means. That is, I develop general formulas for Shift-Share instruments for variances, skews, mean absolute deviations around central points, and Gini coefficients. These formulas strongly emphasize the importance of deriving the instrument from an accounting identity of the

[1] See Baum-Snow and Ferreira (2015), Goldsmith-Pinkham, Sorkin and Swift (2020).

explanatory variable - a central theme of the paper. As an application, I measure the elasticity of single parenting with respect to earnings inequality, using multiple alternative instruments for the Gini coefficient of earnings. Empirically, I find that instruments using only either lags or a Shift Delocalization (but not both) do correct bias, but only part way. This underlines the importance of using both, as in the full Shift-Share instrument.

2 Notation and Core Principles

I begin by defining some standardized notation, which although technical in nature, is critical for understanding the topic at hand. This paper evaluates economic and econometric models from many different papers, each of which has its own completely distinct system of notation. By translating these out of their native notation into a common shared system, I hope to enable far greater insight into their substantive similarities and differences.

Shift-Share instruments are panel data objects. As such, they involve a time dimension, which I denote as t, and cross sectional units, which I denote as z. These cross sectional units z traditionally are geographical units, such as Metropolitan Statistical Areas (MSAs), meant to capture local labor markets. However, they needn't necessarily be geographical units. I call these units z Localities - not in a concrete sense, but rather in an abstract sense of "local" as opposed to "global".

I denote the typical endogenous explanatory variable of interest as $\mathrm{X_{t,z}}$. For example, $\mathrm{X_{t,z}}$ may be the average wage in Locality z (such as z = *Boston*) in time period t (such as t = *2010*). Generally, any average is (at least implicitly) an average over individuals, i.

$$\mathrm{X_{t,z} = N_{t,z}^{-1} \sum_{i \in t,z} X_i}$$

where $\mathrm{X_i}$ the value of the variable per individual (such as an individual person's wage), and $\mathrm{N_{t,z}}$ is the count of individuals in Locality z in time period t. The appropriate kind of count $\mathrm{N_{t,z}}$ may depend on the variable $\mathrm{X_i}$. In particular, it is often the case that we should count only employed individuals, such as when X_i is only defined for those who are employed (wages may be such a case). To emphasize this, I use $NE_{t,z}$ to denote counts of employed persons per se.

The Shift-Share approach begins with a decomposition over industrial, occupational, or similar categories, which I denote as o. For simplicity, I refer to any such categorization as Industries in an abstract sense. The essential feature of such a categorization - whether industrial, occupational, or otherwise - is that the Shares they define should be less sensitive to endogeneity than $X_{t,z}$ is. For example, each employed person works in a particular industry, o. Industrial Shares are $NE_{o,t,z}/NE_{t,z}$,[2] that is, the fraction of the employed who are employed in industry o. It may be that in z = *Detroit* in t = *1990*, there are $NE_{t,z}$ = 600,000 total workers, and $NE_{o,t,z}$ = 120,000 workers in o = *Manufacturing*. The Share would then be 120,000/600,000 = 20%. That about 20% of Detroit's employment opportunities are in Manufacturing would be considered less sensitive to endogeneity than (for example) the average wage $X_{t,z}$ in Detroit is, because the former is a result of geographical and historical factors that cannot be easily adjusted.

A "share" is simply a fraction; therefore other kinds of "shares" may be confused for the Shares of a Shift-Share instrument. A salient case in point is that Industries' shares of Localities are not the same as Localities' shares of Industries. Many authors have understood Industries' shares of Localities, that is,

$$NE_{o,t,z}/NE_{t,z}$$

to be the proper Shares of a Shift-Share instrument. However, Localities' shares of Industries have sometimes been used in a similar way.

$$NE_{o,t,z}/NE_{o,t}$$

The numerator is the same, but the denominator is different. Returning to the example from

[2]Shares may also be $N_{o,t,z}/N_{t,z}$, if X_i is defined for the whole population rather than just the employed.

the previous paragraph, $\mathrm{NE}_{o,t,z} = 120{,}000$ is still the number of Manufacturing workers in Detroit. But now the denominator $\mathrm{N}_{o,t}$ is the national count of Manufacturing workers - let's say 3 million - rather than the local count of all workers ($\mathrm{NE}_{t,z} = 600{,}000$). Although numerically different, Localities' shares of Industries may be considered insensitive to endogenetiy for much the same reasons that Industries' shares of Localities are, and therefore may feature similarly in some instruments.

Shares of Localities - i.e., Industries' shares of Localities, rather than Localities' shares of Industries - possess a unique feature that make them the most natural choice. This is that for *any* variable $\mathrm{X}_{t,z}$ that is an average, the following is true as an accounting identity.

Classical Accounting Identity:

$$\mathrm{X}_{t,z} = \sum_{o} \mathrm{X}_{o,t,z} \cdot \mathrm{N}_{o,t,z}/\mathrm{N}_{t,z}$$

To see why this is so, one must simply unpack definitions:

$$\mathrm{X}_{t,z} = \mathrm{N}_{t,z}^{-1} \sum_{i \in t,z} \mathrm{X}_i \;\Leftrightarrow\; \sum_{i \in t,z} \mathrm{X}_i = \mathrm{X}_{t,z} \cdot \mathrm{N}_{t,z}$$

$$\mathrm{X}_{o,t,z} = \mathrm{N}_{o,t,z}^{-1} \sum_{i \in o,t,z} \mathrm{X}_i \;\Leftrightarrow\; \sum_{i \in o,t,z} \mathrm{X}_i = \mathrm{X}_{o,t,z} \cdot \mathrm{N}_{o,t,z}$$

And pivotally, it is a truism that:[3]

$$\sum_{i \in t,z} \mathrm{X}_i = \sum_{o} \sum_{i \in o,t,z} \mathrm{X}_i \tag{1}$$

[3] The sum of the whole is equal to the sum of (mutually exclusive and collectively exhaustive) subset sums.

Plugging the definition equations above into (1), we have,

$$\mathrm{X_{t,z} \cdot N_{t,z} = \sum_o X_{o,t,z} \cdot N_{o,t,z}}$$

And because $\mathrm{N_{t,z}}$ is constant with respect to o, it can be rewritten (divide both sides by $\mathrm{N_{t,z}}$) inside the summation, as the denominator under $\mathrm{N_{o,t,z}}$.

To arrive at a Shift-Share instrument, one alters the Shift-Share accounting identity in order to preserve that which is most exogenous in $\mathrm{X_{t,z}}$, and remove that which is most sensitive to endogeneity.

Classical Accounting Identity:

$$\mathrm{X_{t,z} = \sum_o \overbrace{X_{o,t,z}}^{Shift} \cdot \overbrace{N_{o,t,z}/N_{t,z}}^{Share}}$$

The primary alteration is to delocalize the Shift over Localities z. That is, replace the Shift $\mathrm{X_{o,t,z}}$ with the Delocalized Shift, $\mathrm{X_{o,t}}$.

Delocalized Shift Instrument:

$$\mathrm{\tilde{X}_{t,z} = \sum_o \overbrace{X_{o,t}}^{Shift} \cdot \overbrace{N_{o,t,z}/N_{t,z}}^{Share}}$$

If $\mathrm{X_{o,t,z}}$ is the average wage of Manufacturing workers in Detroit, $\mathrm{X_{o,t}}$ is the average wage of Manufacturing workers nationally - a *delocalized* average.[4]

The Shift Delocalization is both the strongest and the most distinctive aspect of Shift-Share instruments' claim to exogeneity. As discussed earlier in this section, Industries o should be such that the Shares they define are less sensitive to endogeneity than $\mathrm{X_{t,z}}$ is.

[4]Alternatively to the national average $\mathrm{X_{o,t}}$, some authors have used the average over all Localities excluding z, $\mathrm{X_{o,t,\neg z}}$. For large cross sections, these are essentially the same, because each individual Locality plays only a very small role in the national average.

The Shares' complement - the local Shift vector $X_{o,t,z}$ - is by contrast the part of $X_{t,z}$ *most* sensitive to endogeneity. Replacing $X_{o,t,z}$ with $X_{o,t}$ annihilates this main part of $X_{t,z}$'s endogeneity, because national averages $X_{o,t}$ cannot be a function of anything at the level of isolated Localities z. As an example, suppose we are interested in the effects of wages on hours of work supplied. The average wage of Manufacturing workers in Detroit, $X_{o,t,z}$, is sensitive to reverse causality: an abundance of able workers in Detroit may depress their wages. However, the national average wage of Manufacturing workers, $X_{o,t}$, is not sensitive to the local overabundance of workers in Detroit. It represents rather the portion of $X_{o,t,z}$ that is independent of Detroit's local conditions.

Because the main part of $X_{t,z}$'s endogeneity is in the local Shifts, the Shift Delocalization removes this main part. However, there remains a question of whether some endogeneity still lingers in the Shares. For this reason, a secondary alteration to the accounting identity is made: Lag the Shares, such as to a base period τ.

Classical Accounting Identity:

$$X_{t,z} = \sum_{o} \overbrace{X_{o,t,z}}^{\text{Shift}} \cdot \overbrace{N_{o,t,z}/N_{t,z}}^{\text{Share}}$$

Delocalized Shift Instrument:

$$\tilde{X}_{t,z} = \sum_{o} \overbrace{X_{o,t}}^{\text{Shift}} \cdot \overbrace{N_{o,t,z}/N_{t,z}}^{\text{Share}}$$

Shift-Share Instrument:

$$\tilde{X}_{t,z} = \sum_{o} \overbrace{X_{o,t}}^{\text{Shift}} \cdot \overbrace{N_{o,\tau,z}/N_{\tau,z}}^{\text{Share}}$$

The purpose for lagging the Shares is similar to that for delocalizing the Shifts. A lag is a kind of delocalization, only over time rather than over the cross section. However, it is a weaker kind of delocalization, as the lagged Share retains a connection to the Locality z, while the (Delocalized) Shift does not at all. The mildness of the remedy is apropos to the mildness of the problem. That is, the Shares would already be considered mostly or plausibly exogenous, even in time period t; the lag is just an added guarantee.

Alternatively, it may be questioned whether it is even necessary for the Shares to be exogenous. Goldsmith-Pinkham, Sorkin and Swift (2020) discuss that there are two designs by which Shift-Share instruments may be viewed to be exogenous. That is, researchers may view the exogeneity either as "coming from the Shares," or as "coming from the shocks [Shifts]." I would phrase this differently in that, as discussed in the previous paragraphs, the Delocalized Shifts are definitely exogenous. The question rather is whether this is sufficient: Does the product of an exogenous (Delocalized Shift) vector and an endogenous (Share) vector come out as exogenous like the Shift, or as endogenous like the Share? Borusyak, Hull and Jaravel (2022) present a scenario by which the former is true: The product is exogenous by force of the Shifts' exogeneity alone. But Adão, Kolesár and Morales (2019) show this to be tenuous, and as Goldsmith-Pinkham, Sorkin and Swift (2020) also discuss, most researchers have conceived of Shift-Share instruments as relying on both the (Delocalized) Shifts and the (Lagged) Shares to be exogenous.

In addition to delocalizing the Shifts and lagging the Shares, there may be a zeroth or third step, which is meant to account for effects of unobservables. The Shift Delocalization and Share Lag are powerful measures that make the instrument causally prior with respect to any variable that is observed in the estimation, including the outcome variable. However there is little they can do with respect to unobserved factors that might be correlated with both the instrument and the outcome. For such unobservables, the Shift-Share approach relies on standard methods of time differencing, fixed effects and other controls. As a zeroth step, $\mathrm{X}_{\mathrm{t,z}}$ may be defined as a time difference, such as a growth rate; or as a third step, both the estimating equation and the instrument may be time differenced. This will have the implication of canceling out time-invariant unobservables, though not necessarily time-invariant unobservables. Alternatively as a third step, the instrument may be accompanied by fixed effects or other controls meant absorb the effects of unobservables.

3 Classical Variants

Bartik (1991) and the closely related Blanchard and Katz (1992) are broadly credited with introducing Shift-Share instruments to modern economics literature. Both investigate the effects of job growth on wages, employment and unemployment rates as outcomes, in Metropolitan Statistical Areas (MSAs) as labor markets.[5] In other words, they ask how local job growth is absorbed: does it pull locals into employment out from unemployment, or out from non-participation (not looking for a job at all), or does it rather pull migrant workers from elsewhere; and how are local wages impacted. Although these are multiple outcomes of interest, the authors are interested in the effects on each outcome of only one main causal factor, that is, job growth. Therefore a single instrumental variable for job growth is sufficient to identify all the effects in question. Bartik (1991) develops such an instrument, and Blanchard and Katz (1992) import it directly from the former.

The concept of a Shift-Share decomposition has earlier precedent; but Bartik (1991) may have been the first to apply it to create an instrumental variable.[6] As a general framework, the Shift-Share instrument is essentially an altered version of the endogenous explanatory variable itself. Therefore it is applicable in a great variety of settings, leading to widespread adoption. Nonetheless, there may be ambiguity as to what "Shift-Share" is actually supposed to mean. The below quote from Bartik (1991) may be helpful in clarifying his terminology:

> A shift-share analysis decomposes MSA growth into three components: a national growth component, which calculates what growth would have occurred if all industries in the MSA had grown at the all-industry national average; a share component, which calculates what extra growth would have occurred if each industry in the MSA had grown at that industry's national average; and a shift

[5] Job (growth) refers to (change over time in) the absolute number of employed workers - unlike employment or unemployment *rates*, which are divided by a local population count even in level terms.

[6] See Goldsmith-Pinkham, Sorkin and Swift (2020).

component, which calculates the extra growth that occurs because industries grow at different rates locally than they do nationally.

In my notation, this can be translated as follows:

$$X_{t,z} = \sum_{o}(\overbrace{X_t}^{\text{“National Component”}} + \overbrace{X_{o,t} - X_t}^{\text{“Share Component”}} + \overbrace{X_{o,t,z} - X_{o,t}}^{\text{“Shift Component”}}) \cdot NE_{o,t,z}/NE_{t,z}$$

$$X_{t,z} = (NE_{t,z} - NE_{t-1,z})/NE_{t-1,z}$$

$$X_t = (NE_t - NE_{t-1})/NE_{t-1}$$

$$X_{o,t} = (NE_{o,t} - NE_{o,t-1})/NE_{o,t-1}$$

$$X_{o,t,z} = (NE_{o,t,z} - NE_{o,t-1,z})/NE_{o,t-1,z}$$

where each NE is a count of employed persons (see Section 2). The above is an accounting identity, as

$$X_t + X_{o,t} - X_t + X_{o,t,z} - X_{o,t} = X_{o,t,z}$$

and

$$\sum_{o}(X_{o,t,z}) \cdot NE_{o,t,z}/NE_{t,z} = X_{t,z}$$

Is the Classical Accounting Identity (see Section 2).To arrive at an instrument, Bartik *removes* the “Shift Component” - because this is the component most sensitive to endogeneity - and also lags the Shares $NE_{o,t,z}/NE_{t,z}$ to a base period τ:[7]

$$\tilde{X}_{t,z} = \sum_{o}(\overbrace{X_t}^{\text{“National Component”}} + \overbrace{X_{o,t} - X_t}^{\text{“Share Component”}}) \cdot NE_{o,\tau,z}/NE_{\tau,z}$$

[7]The formula actually used by Bartik (1991) may be different. It appears that he lags Localities' Shares of Industries rather than Industries' Shares of Localities (see Section 2), but does not provide a rationale for this choice. Localities' Shares of Industries can be part of the same accounting identity in this case, only because the variable X itself contains NE terms, which can be exchanged with those of the natural Shares. Rather, the interpretation I present here is applicable to any variable X, and is truer to how most researchers have understood Bartik instruments subsequently.

$$= \sum_{o} (X_{o,t}) \cdot NE_{o,\tau,z}/NE_{\tau,z}$$

I posit that there is some redundancy in Bartik's terminology. Although Bartik refers to,

$$(X_{o,t} - X_t) \cdot NE_{o,t,z}/NE_{t,z}$$

as the "Share Component," the Share itself is surely just,

$$NE_{o,t,z}/NE_{t,z}$$

That which Bartik calls the "Shift Component,"

$$(X_{o,t,z} - X_{o,t}) \cdot NE_{o,t,z}/NE_{t,z}$$

also contains the Share. I call rather $X_{o,t,z}$ the Shift, and $X_{o,t}$ the Delocalized Shift. In either case, the steps for converting the accounting identity into the instrumental variable are equivalent, only said in a different way. That is, *removal* of what Bartik calls the "Shift Component" is equivalent to *delocalization* of what I call the Shift. In either case, the accounting identity can be written as,

$$X_{t,z} = \sum_{o} X_{o,t,z} \cdot NE_{o,t,z}/NE_{t,z}$$

and the instrument is,

$$\tilde{X}_{t,z} = \sum_{o} X_{o,t} \cdot NE_{o,\tau,z}/NE_{\tau,z}$$

The main step in converting $X_{t,z}$ to $\tilde{X}_{t,z}$ is the Shift Delocalization, $X_{o,t,z} \longrightarrow X_{o,t}$, and the secondary step is the Share Lag, $NE_{o,t,z}/NE_{t,z} \longrightarrow NE_{o,\tau,z}/NE_{\tau,z}$.

Bound and Holzer (2000), also closely related to Bartik (1991) and Blanchard and Katz (1992), study in particular how the impacts of local job growth fall differently across racial and demographic groups. Their Shift-Share instrument is also the same,[8] only with a very slight context alteration. Where Bartik and Blanchard and Katz view job growth in terms of the total number of people working, Bound and Holzer (2000) view it rather in terms of the total number of hours worked. In other words, where Bartik would view each individual as either 1 (a worker) or 0 (not a worker), Bound and Holzer would view a part-time worker as effectively 1/2.

Although only slightly different from Bartik (1991), the case of Bound and Holzer (2000) provides a simple illustration of how the Shift-Share instrument may adapt while retaining the same structure and properties. Because the endogenous explanatory variable of interest $\mathrm{X}_{\mathrm{t,z}}$ for Bound and Holzer is the (change over time in) total hours worked, the Shift $\mathrm{X}_{\mathrm{o,t,z}}$ in their accounting identity is also the (change over time in) total hours worked - by industry o.

Classical Accounting Identity:

$$\mathrm{X}_{\mathrm{t,z}} = \sum_{\mathrm{o}} \overbrace{\mathrm{X}_{\mathrm{o,t,z}}}^{\text{Shift}} \cdot \overbrace{\mathrm{NE}_{\mathrm{o,t,z}} \Big/ \mathrm{NE}_{\mathrm{t,z}}}^{\text{Share}}$$

Shift-Share Instrument:

$$\tilde{\mathrm{X}}_{\mathrm{t,z}} = \sum_{\mathrm{o}} \overbrace{\mathrm{X}_{\mathrm{o,t}}}^{\text{Shift}} \cdot \overbrace{\mathrm{NE}_{\mathrm{o,\tau,z}} \Big/ \mathrm{NE}_{\mathrm{\tau,z}}}^{\text{Share}}$$

Both the accounting identity and instrument are the same as in the original case. Only the content of the explanatory variable $\mathrm{X}_{\mathrm{t,z}}$ has changed - from a measure of (change over time in) total number of workers, to a measure of (change over time in) total number of hours worked.

[8] That is, they are the same under my interpretation in which Shares are Industries' Shares of Localities. Bound and Holzer (2000) appear indeed to view Shares as Industries' Shares of Localities, although they do not clarify this explicitly.

The classical accounting identity and instrument are applicable for any endogenous explanatory variable $X_{t,z}$ that can be decomposed over an industrial, occupation, or similar categorization o.[9] Rather than job growth, Diamond (2016) for example studies wages as explanatory variables $X_{t,z}$. Therefore, the Delocalized Shift $X_{o,t}$ for Diamond refers to industrial average wages - whereas for Bound and Holzer (2000) and others it referred to industrial average job growth.

4 Essential and Adjustable Features

It is worth emphasizing that Shift-Share instruments, even in the classical setup, contain multiple features that may be adjusted without straying from the essential properties and purpose. This section discusses some of these features, partly in order to clarify which aspects of classical structure are most essential.

4.1 Lagging of Shares

One of the core properties of Shift-Share instruments is that the Share vector retains some information that is distinctive to its cross sectional unit or Locality, z. By contrast, the Shift vector is delocalized over the cross section for the sake of exogeneity, severing its connection to z. But the precise manner in which Share component retains its connection to z may be adjusted. Typically it is lagged to a historical base period; but it may be lagged in other ways. Moreover, because the Share vector is meant to retain that which is most exogenous about the regressor per z, it is conceivable that the Shares needn't be lagged at all.

Most commonly, and including in Bartik (1991), the Share vector is lagged to a historical base period τ. That is, the Share component of the regressor $X_{t,z}$ is replaced as $(z, t) \longrightarrow (z, \tau)$. I call this a Frozen Lag: All t are lagged to a fixed τ, regardless of the distance in time between t and τ. In Bartik (1991), τ is 1970, while $t \in [1971, 1986]$ for the panel itself.

[9] The essential feature of such a categorization is that the Share vector it defines should be less sensitive to endogeneity than is the explanatory variable itself.

Critically, this lagged Share vector (of industrial shares) is assumed to be causally prior to (exogenous with respect to) future job growth. The main argument for this exogeneity is that industrial shares are deep characteristics of localities (MSAs) z, arising from forces outside the model, such as geography, technology, and historical accident. The lag provides an auxiliary argument, applicable but not conclusive (on its own) in any context: Anything in the past is likely to be causally to prior to anything in the present or future.

I argue that, although less common in Shift-Share instruments than Frozen Lags are, Updating Lags are usually better. In contrast to Frozen Lags, which lag all t to a common base period τ, lagging $(z, t) \longrightarrow (z, \tau)$, Updating Lags rather maintain a common distance in time between each t and its lag, $(z, t) \longrightarrow (z, t-1)$. As discussed in the previous paragraph, the purpose of either kind of lag is to help insure that the Shares are causally prior to the dependent variable in the present (t), although this is not the main argument for the Shares' exogeneity. Either kind of lag accomplishes this purpose equally, as either is prior in time to the present (t). However, Frozen Lags have the downside that because the base period τ is a different distance in time from each t in the panel, it is likely to be far more relevant for the earlier time periods - those closer to τ - than it is for the later time periods. The magnitude of this problem is likely to depend on the span of time under study.

In addition to Frozen or Updating Lags, Shares may be averaged over all lags and leads, or not lagged at all. Averaging over all lags and leads may be sensible particularly if the number of time periods is large, as in Nunn and Qian (2014). With a large number of time periods, that is, the role of any one time period's values in determining the averages for z is small. Hence these averages can be assumed to reflect intrinsic characteristics of z, causal first movers rather than endogenous reactions. However, perhaps most thematically important to consider is the option of not lagging the Shares at all. As discussed earlier in this section, the lag (of whichever kind) is in any case not the primary argument for the Shares' exogeneity. Rather, the primary argument is that Shares are deep characteristics of localities z, arising from forces outside the model - such as geography, technology and historical accident. This primary argument may be applicable in the present (t) as well as in any lag.

4.2 Differencing and Fixed Effects

Although Shift-Share instruments provide a powerful framework for bypassing reverse causality, it should be noted that they do not offer anything novel with respect to omitted variable bias. Rather, Shift-Share instruments rely on standard methods for unobservables: differencing, fixed effects, and other controls. Differencing and fixed effects are methods meant to account for the sum total of all unobserved effects. However, neither of these is perfect, and the potential for unobserved information to play a confounding role is ultimately unavoidable in any setting.

The stereotypical Shift-Share instrument is accompanied by a differencing method. This can be understood as follows. Suppose the model is of a form,

$$\mathrm{Y_{t,z}} = \beta \cdot \mathrm{X_{t,z}} + \delta_{\mathrm{t}} + \delta_{\mathrm{z}} + \delta_{\mathrm{t,z}} \tag{2}$$

where $\mathrm{X_{t,z}}$ is the endogenous regressor of interest, and δ are unobserved factors. As we know, the endogenous regressor $\mathrm{X_{t,z}}$ can be written as a product of Shift and Share components (accounting identity),

$$\mathrm{X_{t,z}} = \sum_{\mathrm{o}} \overbrace{\mathrm{X_{o,t,z}}}^{\text{Shift}} \cdot \overbrace{\mathrm{N_{o,t,z}} \big/ \mathrm{N_{t,z}}}^{\text{Share}} \tag{3}$$

The Shift-Share instrument for $\mathrm{X_{t,z}}$,

$$\tilde{\mathrm{X}}_{\mathrm{t,z}} = \sum_{\mathrm{o}} \overbrace{\mathrm{X_{o,t}}}^{\text{Shift}} \cdot \overbrace{\mathrm{N_{o,\tau,z}} \big/ \mathrm{N_{\tau,z}}}^{\text{Share}} \tag{4}$$

corrects for reverse causality, but not necessarily for bias arising from the unobserved factors

δ.

Stereotypically, a time difference of (2) is taken to eliminate the time-invariant unobserved component, δ_z. The differenced equation is,

$$(Y_{t,z} - Y_{\tau,z}) = \beta \cdot (X_{t,z} - X_{\tau,z}) + (\delta_t - \delta_\tau) + (\delta_{t,z} - \delta_{\tau,z}) \quad (5)$$

with

$$\tilde{X}_{t,z} - \tilde{X}_{\tau,z} = \sum_o \overbrace{(X_{o,t} - X_{o,\tau})}^{\text{Shift}} \cdot \overbrace{N_{o,\tau,z} \big/ N_{\tau,z}}^{\text{Share}}$$

straightforwardly serving as the instrument for $(X_{t,z} - X_{\tau,z})$.[10] Estimation of (5) can recover β without fear of confounding from δ_z, because the time difference has eliminated δ_z. Or, rather than taking the time difference, an alternative method to accomplish the same purpose is to impose fixed effects $\hat{\hat{\delta}}_z$ to absorb δ_z in the estimation of (2). In either case, time period fixed effects $\hat{\hat{\delta}}_t$ are imposed as well to absorb δ_t.

A weakness of either differencing or fixed effects $\hat{\hat{\delta}}_z$ is that, although both nullify time-invariant unobservables δ_z completely, neither does anything whatsoever to nullify time-varying unobservables $\delta_{t,z}$. There is typically no reason to assume that all potentially confounding unobservables would be time-invariant, and hence time-varying unobservables are the most important blind spot of many Shift-Share instruments. Alternatively, time-varying regional effects $\hat{\hat{\delta}}_{t,Z}$ may be imposed to absorb some time-varying unobservables, as well as some time-invariant unobservables. (Regions Z are closely related clusters of Localities z.)[11] The optimal balance of such effects should absorb the most important potentially confounding unobserved factors of both types, that is, time-varying and time-invariant.

5 Novel Variants

Autor, Dorn and Hanson (2013) study the effects of import competition on labor market

[10] Often $(X_{t,z} - X_{\tau,z})$ is written rather as the original $X_{t,z}$, and $(X_{o,t} - X_{o,\tau})$ as the original $X_{o,t}$.

[11] Time-varying effects at the z level $\hat{\hat{\delta}}_{t,z}$ would absorb the entire panel, leaving nothing to be explained by any regressor(s) of interest.

outcomes for US workers, that is, unemployment, labor force participation, and wages. The endogenous explanatory variable of interest is a change in labor market exposure to Chinese imports. Total China-to-US import volume is $\mathrm{I_t}$. The basic idea is that these imports $\mathrm{I_t}$ are in competition with the output of US workers: The competitive impact of $\mathrm{I_t}$ on US workers is spread out equally, so that the impact per individual US worker is $\mathrm{I_t/NE_t}$, where $\mathrm{NE_t}$ is the total number of workers in the US. The change in exposure can be written as

$$\mathrm{X_{t,z} = (I_t - I_{t-1})/NE_{t-1}} \tag{6}$$

for any US Commuting Zone z in which there are any workers. This $\mathrm{X_{t,z}}$ is constant over z, assuming there are any workers in z, because the impact is assumed to be spread equally over workers nation-wide.

Instead of (6), which is an oversimplification, the authors suppose rather that the competitive impact of Chinese imports *in industry o* is spread equally nation-wide over all US workers *in industry o*.

$$\mathrm{X_{o,t,z} = (I_{o,t} - I_{o,t-1})/NE_{o,t-1}} \tag{7}$$

Alhough specific by industry o, $\mathrm{X_{o,t,z}}$ is still constant over z, for the same reason that equation (6)'s $\mathrm{X_{t,z}}$ would be constant over z. However, the unconditional average impact per worker in z now depends on the industrial makeup of z, because each industry-specific impact $\mathrm{X_{o,t,z}}$ applies only to the workers in industry o:

$$\mathrm{X_{t,z}} = \sum_{\mathrm{o}} \overbrace{\mathrm{X_{o,t,z}}}^{\text{Shift}} \cdot \overbrace{\mathrm{NE_{o,t,z}} \big/ \mathrm{NE_{t,z}}}^{\text{Share}} \tag{8}$$

$$\mathrm{X_{o,t,z} = (I_{o,t} - I_{o,t-1})/NE_{o,t-1}}$$

Equation (8) happens to have (apparently) the same form as the Classical Accounting Identity (see Section 2). And, as in the classical setting, Autor, Dorn and Hanson (2013) use lagged shares in their instrument, $\mathrm{NE_{o,t-1,z}} \big/ \mathrm{NE_{t-1,z}}$ in place of $\mathrm{NE_{o,t,z}} \big/ \mathrm{NE_{t,z}}$. Yet, the Shift $X_{o,t,z}$ is unlike that in the classical setting, because it is already constant over z. Typically, a Shift Delocalization involves replacing local Industry averages $X_{o,t,z}$ with national Industry averages $X_{o,t}$. However, as discussed in the previous paragraph, it is already the case that $X_{o,t,z} = X_{o,t}$, due to the way in which the exposure variable itself is defined. But Autor, Dorn and Hanson (2013) go further than this, delocalizing their Shift on a higher level. They replace *nation*-wide industry averages $X_{o,t}$ with *world*-wide averages $X'_{o,t}$.

$$\tilde{X}_{t,z} = \sum_o \overbrace{X'_{o,t}}^{\text{Shift}} \cdot \overbrace{NE_{o,t-1,z} \Big/ NE_{t-1,z}}^{\text{Share}} \tag{9}$$

$$X'_{o,t} = (I'_{o,t} - I'_{o,t-1})/NE_{o,t-1}$$

I'_o is export volume from China industry o *world*-wide, to other high income countries besides the US.

Nunn and Qian (2014) study the effect of food aid on armed conflict in countries receiving the aid. Although food aid might be expected to cool tensions between opposing factions, often rather the opposite is observed, as armed theft of the aid ignites further conflict. There is also a strong channel for reverse causality, however: Countries experiencing conflict may be more likely to receive aid.

The endogenous variable of interest $X_{t,z}$ for Nunn and Qian is the the amount of food aid (wheat) that country z receives from the US in year t. An accounting identity for this variable is,

$$X_{t,z} = \overbrace{I_{t,z}}^{\text{Share}} \cdot \overbrace{X_{t-1}}^{\text{Push}} \tag{10}$$

where $I_{t,z}$ is a binary variable equal to one if country z is selected to receive wheat aid in year t, and X_{t-1} is the quantity of wheat produced in the US in the previous year (scaled by the number of countries to receive the aid). This quantity is lagged because it takes a one-year cycle for the wheat to transition from production to distribution. Their instrument is,

$$\tilde{X}_{t,z} = \overbrace{\tau^{-1} \sum_{\tau} I_{\tau,z}}^{\text{Share}} \cdot \overbrace{X_{t-1}}^{\text{Push}} \tag{11}$$

accompanied by region specific time trends and country fixed effects. The summation index τ covers all years in the sample, and $\tau = 36$ is the number of years τ. Observations are 125 non-OECD countries, over the 36 years. The outcome variable is a binary indicator of whether the country is in a state of conflict, defined as experiencing more than 25 battle deaths in the year.

Nunn and Qian's instrument is similar to a classical Shift-Share instrument in the Share, except that they use an unusual kind of share, and lag it in an unusual way. Rather than an Updating or Frozen Lag (see Section 4.1), they take the average over all lags and leads. This is appropriate both because the number of time periods is large, and because of the unusual binary nature of the Share. Because the number of time periods is large, the effect of any one time period's value in determining the average is small. Therefore the time averaged Share can be viewed as a deep characteristic of the country z (conditional on country fixed effects), rather than an endogenous reaction to present period conditions.

Because X_{t-1} is already constant over countries z, it cannot be delocalized as a Shift would be. I therefore refer to it as something else, a Push. I define a Push as a component of the accounting identity that is not altered between the accounting identity and the instrument. Many Shift-Share instruments contain such components. But would I not refer to Nunn and Qian's instrument as a Shift-Share instrument per se, because it does not contain a Shift that is delocalized for the purpose of enhancing exogeneity.

Card (2001) studies the effects of immigrant inflows on occupation specific wages, employment and unemployment rates of natives. A larger pool of workers competing for the same set of jobs j may depress wages by diminishing laborers' bargaining power against employers. However, this effect is difficult to isolate because, for example, cities with higher wages may attract more immigrant workers.

For each occupation group j, Card's endogenous explanatory variable of interest $X_{j,t,z}$

is the inflow of immigrant workers of group j, $\mathrm{NI}_{j,t,z}$, to each Metropolitan Statistical Area (MSA) z in the US. This inflow can be expressed as the sum of inflows from each of various countries of origin, o.

$$\mathrm{X}_{j,t,z} = \mathrm{NI}_{j,t,z} = \sum_{o} \mathrm{NI}_{j,o,t,z}$$

This can be multiplied by $\mathrm{NI}_{o,t,z}/\mathrm{NI}_{o,t,z}$ and $\mathrm{NI}_{o,t}/\mathrm{NI}_{o,t}$ to yield the following accounting identity:

$$\mathrm{X}_{t,z} = \sum_{o} \overbrace{\mathrm{NI}_{j,o,t,z}/\mathrm{NI}_{o,t,z}}^{\text{Shift}} \cdot \overbrace{\mathrm{NI}_{o,t,z}\big/\mathrm{NI}_{o,t}}^{\text{Share}} \cdot \overbrace{\mathrm{NI}_{o,t}}^{\text{Push}} \tag{12}$$

$\mathrm{NI}_{o,t}$ is the total inflow to the US of immigrant workers from country of origin o. $\mathrm{NI}_{o,t,z}/\mathrm{NI}_{o,t}$ are the fraction of immigrant workers from country of origin o who accrue to each MSA z. Abstractly, $\mathrm{NI}_{o,t,z}/\mathrm{NI}_{o,t}$ are Localities' Shares of Industries (see Section 2), with countries of origin o as abstract Industries. $\mathrm{NI}_{j,o,t,z}/\mathrm{NI}_{o,t,z}$ are the fraction of immigrant workers from country of origin o to MSA z who are workers in occupation group j.

To arrive at an instrument, Card delocalizes the Shifts $\mathrm{NI}_{j,o,t,z}/\mathrm{NI}_{o,t,z}$ over z, and lags the Shares $\mathrm{NI}_{o,t,z}/\mathrm{NI}_{o,t}$ to a historical base period τ.[12]

$$\tilde{\mathrm{X}}_{t,z} = \sum_{o} \underbrace{\overbrace{\mathrm{NI}_{j,o,t}/\mathrm{NI}_{o,t}}^{\text{Shift}}}_{\tau_{gj}} \cdot \underbrace{\overbrace{\mathrm{NI}_{o,\tau,z}\big/\mathrm{NI}_{o,\tau}}^{\text{Share}}}_{\lambda_{gc}} \cdot \underbrace{\overbrace{\mathrm{NI}_{o,t}}^{\text{Push}}}_{\mathrm{M}_g}$$

Similarly as in a classical Shift-Share instrument, the Shares represent that which is most exogenous about $\mathrm{X}_{t,z}$ per Locality z. Newly arriving immigrants tend to move to enclaves established by earlier immigrants from the same source country, not for endogenous labor demand reasons, but rather due to cultural and family ties. The lag provides an additional guarantee of this exogeneity.

Also like in a classical Shift-Share instrument, the local Shifts would be highly endogenous, and therefore are delocalized over the cross section. These local Shifts $\mathrm{NI}_{j,o,t,z}/\mathrm{NI}_{o,t,z}$ give the occupational makeup of immigrant workers entering each city. Occupational makeup

[12] τ_{gj}, λ_{gc}, M_g are the notation actually used by Card (2001) for these objects.

is driven heavily by the local economy's labor demand, which is endogenous in this setting. The Delocalized Shifts $\mathrm{NI_{j,o,t}/NI_{o,t}}$ replace the occupational makeup entering each city with the average occupational makeup of immigrants entering the US as a whole, by country of origin o.

Boustan et al. (2013) study the effects of income inequality on public taxation and expenditure. On the one hand, a highly unequal local population may have little in common with one another, and therefore little that they can agree on in the way of public programs. On the other hand, median voter theory suggests that a more unequal distribution will yield a more covetous median voter, and therefore a larger appetite for public programs. The authors show that the latter effect dominates. But to do so, they must get around the reverse causal effect of public programs on inequality.

Boustan et al. (2013) build an instrument for the Gini coefficient of income, a traditional measure of inequality. With income data for individuals i, the Gini coefficient in any given Locality (school district) z in time period t would be,

$$\mathrm{X_{t,z}} = \frac{\sum_{i \in \mathrm{t,z}} \sum_{\mathrm{i} \in \mathrm{t,z}} \left| \mathrm{Y_{i,t,z}} - \mathrm{Y}_{i,\mathrm{t,z}} \right|}{2\mathrm{N_{t,z}} \sum_{i \in \mathrm{t,z}} \mathrm{Y}_{i,\mathrm{t,z}}} \tag{13}$$

where $\mathrm{Y_{i,t,z}}$ or $\mathrm{Y}_{i,\mathrm{t,z}}$ is the income of any given individual i or i. For data reasons, the authors use rather a discretized Gini, that is, an approximation of the Gini using discrete bins h. Concretely, whatever bin h that $\mathrm{Y_{i,t,z}}$ falls into, $\mathrm{Y_{i,t,z}}$ is discretized rather to be equal to $\mathrm{Y_{h,t}}$, a midpoint value for the bin.[13]

$$\mathrm{X_{t,z}} = \frac{\sum_{h} \sum_{\mathrm{h}} \overbrace{\mathrm{P_{h,t,z}} \mathrm{P}_{h,\mathrm{t,z}}}^{\text{Share}} \cdot \overbrace{\left| \mathrm{Y_{h,t}} - \mathrm{Y}_{h,\mathrm{t}} \right|}^{\text{Push}}}{2 \sum_{h} \underbrace{\mathrm{P}_{h,\mathrm{t,z}}}_{\text{Share}} \cdot \underbrace{\mathrm{Y}_{h,\mathrm{t}}}_{\text{Push}}}, \quad \mathrm{P_{h,t,z}} = \mathrm{N_{h,t,z}/N_{t,z}} \tag{14}$$

[13] For example, if the cutoffs for bin a bin h are \$30,000 on the low end and \$34,000 on the high end, an individual i's income $\mathrm{Y_{i,t,z}} =$ \$30,500 may be discretized as $\mathrm{Y}_{h,\mathrm{t}} =$ \$32,000.

$P_{h,t,z} = N_{h,t,z}/N_{t,z}$, which I call Bin Shares, are the fractions of the local z population who fall into each bin h. The cutoffs of the bins h are shared in common over the nation, so the Bin Shares vary by Locality z. For example, more unequal Localities z will tend to have higher $P_{h,t,z}$ values for the highest and lowest bins h. To arrive at an instrument, the authors lag the Bin Shares to a historical base period τ.

$$\tilde{X}_{t,z} = \frac{\sum_{h} \sum_{\mathrm{h}} \overbrace{P_{\mathrm{h},\tau,z} P_{h,\tau,z}}^{\text{Share}} \cdot \overbrace{\left| Y_{\mathrm{h},t} - Y_{h,t} \right|}^{\text{Push}}}{2 \sum_{h} \underbrace{P_{h,\tau,z}}_{\text{Share}} \cdot \underbrace{Y_{h,t}}_{\text{Push}}} \tag{15}$$

They also use a time difference, plus linear trend effects and other controls, to absorb unobservables. Because there is no Shift vector that is delocalized over the cross section in moving from the accounting identity to the instrument, I would not call it a Shift-Share instrument per se - although it may nonetheless be a valid instrument. Using this example specifically, I discuss what I view as the potential importance of this distinction in depth in Section 7.

6 New Creations

The abstracted Shift-Share instrument is a flexible framework for constructing potential instrumental variables. I posit that the core idea of the Shift-Share approach is to decompose the endogenous explanatory variable as an accounting identity that contains a more endogenous factor (the Shift vector) and a more exogenous factor (the Share vector); replace the Shift vector with an analogue that is completely exogenous, because it is delocalized over the cross section; and lag the Share vector to enhance its already solid claim to exogeneity. The Shift Delocalization and Share Lag - in conjunction with differencing methods and controls for unobservables - provide Shift-Share instruments' claim to exogeneity. Conditional on these, starting from an accounting identity provides a guarantee that the instrument should be relevant.

In Section 2, I presented a general formula for classical Shift-Share instruments, applicable to any endogenous explanatory variable $X_{t,z}$ which is a mean over individuals who can be rearranged into subgroups to form Shares. In this section, I derive analogous formulas for other distribution summaries, that is, besides the mean. As with the formula for the mean, basing the instrument on an accounting identity is a guide to relevance. That is, it provides the most natural approximation of the endogenous explanatory variable as a function of the exogenous factors.[14]

6.1 Variance and Skew

In Section 2, I considered any endogenous explanatory variable $X_{t,z}$ that is a mean over individuals i. Following from the definitions of means and from the fact that,

$$\sum_{i \in t,z} X_i = \sum_{o} \sum_{i \in o,t,z} X_i \tag{16}$$

I showed that $X_{t,z}$ can be expressed as the Classical Accounting Identity,

$$X_{t,z} = \sum_{o} \overbrace{X_{o,t,z}}^{\text{Shift}} \cdot \overbrace{N_{o,t,z}/N_{t,z}}^{\text{Share}}$$

I now consider any endogenous explanatory variable $V_{t,z}$ that is a variance over individuals i.

$$V_{t,z} = N_{t,z}^{-1} \sum_{i \in t,z} (X_i - X_{t,z})^2 \tag{17}$$

[14] A first stage consisting of all the lagged Shares as separate instruments, for example, may yield a higher overall fit - but this would most likely be a case of overfitting.

For readability, I suppress t subscripts throughout the remainder of Section 6.

$$\mathrm{V_z = N_z^{-1} \sum_{i \in z} (X_i - X_z)^2 \;\Leftrightarrow\; \sum_{i \in z} (X_i - X_z)^2 = N_z \cdot V_z} \tag{18}$$

In addition to (16), it is helpful to apply the identity,

$$\mathrm{(X_i - X_z)^2 = (X_i - X_{o,z})^2 + ((X_i - X_z)^2 - (X_i - X_{o,z})^2)} \tag{19}$$

(19) beneficially decomposes the moment into within-category and across-category components.

$$\begin{aligned} \mathrm{V_z} &= \mathrm{N_z^{-1} \sum_{o} \sum_{i \in o,z} (X_i - X_z)^2} \\ &= \mathrm{N_z^{-1} \sum_{o} \Big\{ \sum_{i \in o,z} \underbrace{(X_i - X_{o,z})^2}_{within-o} + \sum_{i \in o,z} \underbrace{((X_i - X_z)^2 - (X_i - X_{o,z})^2)}_{across-o} \Big\}} \end{aligned} \tag{20}$$

The within-category component is straightforward, as by definition,

$$\mathrm{V_{o,z} = N_{o,z}^{-1} \sum_{i \in o,z} (X_i - X_z)^2 \;\Leftrightarrow\; \sum_{i \in o,z} (X_i - X_{o,z})^2 = N_{o,z} \cdot V_{o,z}} \tag{21}$$

The across-category component simplifies after expansion and collection of terms:

$$
\begin{aligned}
&\sum_{\mathrm{i}\in \mathrm{o,z}} ((\mathrm{X_i} - \mathrm{X_z})^2 - (\mathrm{X_i} - \mathrm{X_{o,z}})^2) \\
&= \sum_{\mathrm{i}\in \mathrm{o,z}} (\mathrm{X_i^2} - 2\mathrm{X_i X_z} + \mathrm{X_z^2} - \mathrm{X_i^2} + 2\mathrm{X_i X_{o,z}} - \mathrm{X_{o,z}^2}) \\
&= 2(\mathrm{X_{o,z}} - \mathrm{X_z}) \sum_{\mathrm{i}\in \mathrm{o,z}} (\mathrm{X_i}) - (\mathrm{X_{o,z}^2} - \mathrm{X_z^2}) \sum_{\mathrm{i}\in \mathrm{o,z}} (1) \\
&= 2(\mathrm{X_{o,z}} - \mathrm{X_z}) \cdot \mathrm{N_{o,z}} \cdot \mathrm{X_{o,z}} - (\mathrm{X_{o,z}^2} - \mathrm{X_z^2})\mathrm{N_{o,z}} \\
&= 2(\mathrm{X_{o,z}^2} - \mathrm{X_{o,z}X_z}) \cdot \mathrm{N_{o,z}} - (\mathrm{X_{o,z}^2} - \mathrm{X_z^2}) \cdot \mathrm{N_{o,z}} \\
&= (\mathrm{X_{o,z}^2} - 2\mathrm{X_{o,z}X_z} + \mathrm{X_z^2}) \cdot \mathrm{N_{o,z}} \\
&= \mathrm{N_{o,z}} \cdot (\mathrm{X_{o,z}} - \mathrm{X_z})^2
\end{aligned}
\tag{22}
$$

Combining the above,

$$
\mathrm{V_z} = \sum_{\mathrm{o}} \overbrace{\Big\{ \underbrace{\mathrm{V_{o,z}}}_{\substack{\text{within–o}\\ \text{variance}}} + \underbrace{(\mathrm{X_{o,z}} - \mathrm{X_z})^2}_{\substack{\text{across–o}\\ \text{variance}}} \Big\}}^{\text{Shift}} \cdot \overbrace{\mathrm{N_{o,z}/N_z}}^{\text{Share}}
\tag{23}
$$

Equation (23) is the analogue of the Classical Accounting Identity for variances rather than means. With time period subscripts, this is:

$$
\mathrm{V_{t,z}} = \sum_{\mathrm{o}} \overbrace{\Big\{ \underbrace{\mathrm{V_{o,t,z}}}_{\substack{\text{within–o}\\ \text{variance}}} + \underbrace{(\mathrm{X_{o,t,z}} - \mathrm{X_{t,z}})^2}_{\substack{\text{across–o}\\ \text{variance}}} \Big\}}^{\text{Shift}} \cdot \overbrace{\mathrm{N_{o,t,z}/N_{t,z}}}^{\text{Share}}
\tag{24}
$$

To arrive at the Shift-Share instrument, delocalize the Shift vector over z, and lag the Share vector:

$$\tilde{V}_{t,z} = \sum_o \overbrace{\left\{ \underbrace{V_{o,t}}_{\substack{\text{within–o} \\ \text{variance}}} + \underbrace{(X_{o,t} - \tilde{X}_{t,z})^2}_{\substack{\text{across–o} \\ \text{variance}}} \right\}}^{\text{Shift}} \cdot \overbrace{N_{o,\tau,z}/N_{\tau,z}}^{\text{Share}} \tag{25}$$

where $\tilde{X}_{t,z}$ is the Shift-Share instrument for the mean.

The above equations illustrate the importance of deriving the instrument from an accounting identity, rather than from intuition alone. Intuitively, one might construct an instrument for the variance using only the within category variance,

$$\tilde{V}_{t,z} = \sum_o \overbrace{V_{o,t}}^{\text{Shift}} \cdot \overbrace{N_{o,\tau,z}/N_{\tau,z}}^{\text{Share}} \tag{26}$$

because this may seem most analogous to the Shift-Share instrument for the mean,

$$\tilde{X}_{t,z} = \sum_o \overbrace{X_{o,t}}^{\text{Shift}} \cdot \overbrace{N_{o,\tau,z}/N_{\tau,z}}^{\text{Share}} \tag{27}$$

Similarly as (27) is a Share-weighted average of the category (Industry) o's national averages, (26) would be a Share-weighted average of the Industries' national variances. Indeed, one might conceive of constructing an instrument for any kind of moment in this way. But the accounting identity (24) demonstrates that there is additional term, the across-category variance, that should be included as well in order to arrive at the most natural prediction of $V_{t,z}$ as a function of the national industry distributions.

To derive the instrument for the skew,

$$W_z = N_z^{-1} \sum_{i \in z} (X_i - X_z)^3 \tag{28}$$

it is again helpful to decompose into within and across category components,

$$(X_i - X_z)^3 = (X_i - X_{o,z})^3 + ((X_i - X_z)^3 - (X_i - X_{o,z})^3) \tag{29}$$

The across category component now simplifies as,

$$
\begin{aligned}
&\sum_{i \in o,z} ((X_i - X_z)^3 - (X_i - X_{o,z})^3) \\
= &\sum_{i \in o,z} (X_i^3 - 3X_i^2 X_z + 3X_i X_z^2 - X_z^3 - X_i^3 + 3X_i^2 X_{o,z} - 3X_i X_{o,z}^2 + X_{o,z}^3) \\
&= 3(X_{o,z} - X_z) \sum_{i \in o,z} (X_i^2) - 3(X_{o,z}^2 - X_z^2) \sum_{i \in o,z} (X_i) + (X_{o,z}^3 - X_z^3) \sum_{i \in o,z} (1) \\
&= 3(X_{o,z} - X_z) \cdot N_{o,z} \cdot (V_{o,z} + X_{o,z}^2) - 3(X_{o,z}^2 - X_z^2) \cdot N_{o,z} \cdot X_{o,z} + (X_{o,z}^3 - X_z^3) \cdot N_{o,z} \\
&= (X_{o,z}^3 - 3X_{o,z}^2 X_z + X_{o,z} X_z^2 - X_z^3) \cdot N_{o,z} + 3(X_{o,z} - X_z) \cdot N_{o,z} \cdot V_{o,z} \\
&= N_{o,z} \cdot ((X_{o,z} - X_z)^3 + 3(X_{o,z} - X_z) \cdot V_{o,z})
\end{aligned}
\tag{30}
$$

That is, the across category component itself now contains two sub-components.

$$
W_z = \sum_o \overbrace{\Big\{ \underbrace{W_{o,z}}_{\substack{\text{within–o} \\ \text{skew}}} + \underbrace{(X_{o,z} - X_z)^3}_{\substack{\text{across–o–mean} \\ \text{skew}}} + \underbrace{3(X_{o,z} - X_z) V_{o,z}}_{\substack{\text{across–o–variance} \\ \text{skew}}} \Big\}}^{\text{Shift}} \cdot \overbrace{N_{o,z}/N_z}^{\text{Share}} \tag{31}
$$

The instrument follows as,

$$
\tilde{W}_{t,z} = \sum_o \overbrace{\Big\{ \underbrace{W_{o,t}}_{\substack{\text{within–o} \\ \text{skew}}} + \underbrace{(X_{o,t} - \tilde{X}_{t,z})^3}_{\substack{\text{across–o–mean} \\ \text{skew}}} + \underbrace{3(X_{o,t} - \tilde{X}_{t,z}) V_{o,t}}_{\substack{\text{across–o–variance} \\ \text{skew}}} \Big\}}^{\text{Shift}} \cdot \overbrace{N_{o,\tau,z}/N_{\tau,z}}^{\text{Share}} \tag{32}
$$

This process will solve similarly for any higher order n moment,

$$
N_z^{-1} \sum_{i \in z} (X_i - X_z)^n
$$

thus allowing the researcher in principle to construct Shift-Share instruments for the entire distribution beyond the mean. I leave to the reader to derive the general n formula.

6.2 Mean Absolute Deviation

A mean absolute deviation around a central point X is defined as,

$$\mathrm{M_z} = \mathrm{N_z^{-1}} \sum_{\mathrm{i} \in \mathrm{z}} |\mathrm{X_i} - \mathrm{X}| \tag{33}$$

First, it is helpful to split the piecewise function into its two parts, which I call the Mean Inferior Deviation $\lfloor \mathrm{M_z}$ and Mean Superior Deviation $\lceil \mathrm{M_z}$.

$$\begin{aligned} \mathrm{M_z} &= \lfloor \mathrm{X_z} + \lceil \mathrm{X_z} \\ \lfloor \mathrm{M_z} &= \mathrm{N_z^{-1}} \sum_{\mathrm{i} \in \mathrm{z}} (\mathrm{X} - \mathrm{X_i}) \cdot I[\mathrm{X_i} \leq \mathrm{X}] \\ \lceil \mathrm{M_z} &= \mathrm{N_z^{-1}} \sum_{\mathrm{i} \in \mathrm{z}} (\mathrm{X_i} - \mathrm{X}) \cdot I[\mathrm{X_i} > \mathrm{X}] \end{aligned} \tag{34}$$

Applying (16),

$$\lceil \mathrm{M_z} = \mathrm{N_z^{-1}} \sum_{\mathrm{o}} \left\{ \sum_{\mathrm{i} \in \mathrm{o,z}} (\mathrm{X_i} - \mathrm{X}) \cdot I[\mathrm{X_i} > \mathrm{X}] \right\} = \mathrm{N_z^{-1}} \sum_{\mathrm{o}} \left\{ (\lceil \mathrm{X_{o,z}} - \mathrm{X}) \cdot \lceil \mathrm{N_{o,z}} \right\} \tag{35}$$

$\lceil \mathrm{X_{o,z}}$ is defined as the mean value of $\mathrm{X_i}$ in z restricted both to category o, and to the condition of being greater than the central point X; and $\lceil \mathrm{N_{o,z}}$ is the count satisfying the same. $\lfloor \mathrm{X_{o,z}}$ and $\lfloor \mathrm{N_{o,z}}$ are similar, for the condition of being less than the central point.

$$\lceil \mathrm{M_z} = \sum_{\mathrm{o}} \overbrace{\left(\lceil \mathrm{X_{o,z}} - \mathrm{X}\right)}^{\text{Shift}} \cdot \overbrace{\lceil \mathrm{N_{o,z}} / \mathrm{N_z}}^{\text{Share}}, \quad \lfloor \mathrm{M_z} = \sum_{\mathrm{o}} \overbrace{\left(\lfloor \mathrm{X_{o,z}} - \mathrm{X}\right)}^{\text{Shift}} \cdot \overbrace{\lfloor \mathrm{N_{o,z}} / \mathrm{N_z}}^{\text{Share}} \tag{36}$$

The instrument follows from delocalizing the Shifts over z and lagging the Shares,

$$\begin{aligned}
\lceil\tilde{\mathrm{M}}_{\mathrm{t,z}} &= \sum_{\mathrm{o}} \overbrace{\left(\lceil \mathrm{X}_{\mathrm{o,t}} - \mathrm{X}_{\mathrm{t}}\right)}^{\text{Shift}} \cdot \overbrace{\lceil \mathrm{N}_{\mathrm{o,t}}/\mathrm{N}_{\mathrm{t}}}^{\text{Share}} \\
\lfloor\tilde{\mathrm{M}}_{\mathrm{t,z}} &= \sum_{\mathrm{o}} \overbrace{\left(\lfloor \mathrm{X}_{\mathrm{o,t}} - \mathrm{X}_{\mathrm{t}}\right)}^{\text{Shift}} \cdot \overbrace{\lfloor \mathrm{N}_{\mathrm{o,t}}/\mathrm{N}_{\mathrm{t}}}^{\text{Share}}
\end{aligned} \tag{37}$$

$$\tilde{\mathrm{M}}_{\mathrm{t,z}} = \lfloor\tilde{\mathrm{M}}_{\mathrm{t,z}} + \lceil\tilde{\mathrm{M}}_{\mathrm{t,z}}$$

Although $\lceil \mathrm{X}_{\mathrm{o,t}}$, $\lceil \mathrm{N}_{\mathrm{o,t}}$, $\lfloor \mathrm{X}_{\mathrm{o,t}}$, $\lfloor \mathrm{N}_{\mathrm{o,t}}$ are unusual mathematical objects, they follow directly from the accounting identity as the natural components for the instrument, are straightforward to construct from national Industry o data.

6.3 Bin Shares and Gini

Bin Shares are objects that can flexibly map the shape of a distribution. Bin Shares are,

$$\mathrm{P}_{\mathrm{h,z}} = \mathrm{N}_{\mathrm{h,z}}/\mathrm{N}_{\mathrm{z}}$$

where $\mathrm{N}_{\mathrm{h,z}}$ is the count of persons i in Locality z who fall into bin h; and for i to fall into bin h means that the variable X_{i} falls between an upper and lower cutoff particular to h. Table 1 provides a concrete example of Bin Shares. Here, the variable X_{i} is each individual i's self reported yearly earnings, Y_{i}, and the bin cutoffs are defined by deciles of the national earnings distribution.

The right-most column of Table 1 reports the Bin Shares for z = *Boston* (the Commuting Zone) in 2010. If the earnings distribution in Boston were a perfect match to the national earnings distribution, then all of Boston's Bin Shares would be equal to 0.10, because the bins are deciles of the national distribution. Instead, because Boston is wealthy, it has above par shares in the three highest bins, especially the top bin. The Cutoff column gives the lower cutoff for each national decile bin h. Individual earnings are reported to the nearest

\$1,000, yielding clean bin cutoffs. $Y_{h,t}$ is the national average within each bin, which may serve as a (discrete approximation) representative value for any Y_i within each.

Table 1: (Earnings Decile) Bin Shares for (z = Boston) in 2010

Decile (h)	Cutoff	$Y_{h,t}$ (\$$10^3$)	$P_{h,\ t=2010,\ z=Boston}$
1	0	0	0.0970
2	6.200	3.176	0.0859
3	13.00	9.933	0.0764
4	20.00	17.12	0.0816
5	26.00	23.63	0.0589
6	34.00	30.16	0.0835
7	42.00	38.23	0.0938
8	55.00	48.90	0.1157
9	76.00	65.16	0.1257
10	-	130.8	0.1814

Bin shares $P_{h,t,z} = N_{h,t,z}/N_{t,z}$ map the shape of a distribution. Here, bin cutoffs are defined by deciles of the national earnings distribution; the Cutoff column gives the upper cutoff of each. $Y_{h,t}$ is the national average within each bin. Data are from the American Community Survey via IPUMS USA.

Boustan et al. (2013) use Bin Shares as the Shares in their instrument for the Gini coefficient of income. However, as Bin Shares are simply an approximate mapping of the local income distribution, they are similar in terms of exogeneity to the average (or any other summary) of the local income distribution, which would often be considered endogenous even in the lag. As an alternative, I the cast the Bin Shares themselves as endogenous variables, for which we can construct classical Shift-Share instruments. Applying the Bin Share, $P_{h,t,z} = N_{h,t,z}/N_{t,z}$, as $X_{t,z}$ in the Classical Accounting Identity, we have as an accounting identity,

$$\mathrm{P_{h,t,z}} = \sum_{\mathrm{o}} \overbrace{\mathrm{N_{h,o,t,z}} \Big/ \mathrm{N_{o,t,z}}}^{\text{Shift}} \cdot \overbrace{\mathrm{N_{o,t,z}} \Big/ \mathrm{N_{t,z}}}^{\text{Share}} \tag{38}$$

The Shift-Share instrument for the Bin Share then follows from delocalizing the Shift vector over Localities z, and lagging the Industrial Share vector.

$$\tilde{\mathrm{P}}_{\mathrm{h,t,z}} = \sum_{\mathrm{o}} \overbrace{\mathrm{N_{h,o,t}} \Big/ \mathrm{N_{o,t}}}^{\text{Shift}} \cdot \overbrace{\mathrm{N_{o,\tau,z}} \Big/ \mathrm{N_{\tau,z}}}^{\text{Share}} \tag{39}$$

These can then be applied in the place of the endogenous Bin Shares in the discretized Gini, yielding a Shift-Share instrument for the Gini:

$$\frac{\sum_{h} \sum_{\mathrm{h}} \overbrace{\tilde{\mathrm{P}}_{\mathrm{h,t,z}} \tilde{\mathrm{P}}_{h,\mathrm{t,z}}}^{\text{Shift·Share}} \cdot \overbrace{\left| \mathrm{Y_{h,t}} - \mathrm{Y}_{h,\mathrm{t}} \right|}^{\text{Push}}}{2 \sum_{h} \underbrace{\tilde{\mathrm{P}}_{h,\mathrm{t,z}}}_{\text{Shift·Share}} \cdot \underbrace{\mathrm{Y}_{h,\mathrm{t}}}_{\text{Push}}} \tag{40}$$

Although Boustan et al. (2013)'s instrument may be validly exogenous, (40) provides arguably stronger guarantees of exogeneity, more like that of a classical Shift-Share instrument. The upcoming section tests multiple related instruments for the Gini in practice.

7 Application: Inequality and Single Parenting

There are many reasons that we might expect earnings inequality to result in higher rates of single parenting. However, there has been little direct work on this topic in economics, despite its importance.[15] Gould and Paserman (2003) articulate a theory as to why inequality would result in lower rates of marriage: In more unequal markets, women hold out longer in

[15] Chetty et al. (2014) find single parenting to be the strongest and most robust single predictor of socioeconomic immobility.

hopes of securing more desirable matches. Kearney and Levine (2014) discuss several other theories pertaining to single parenting directly, emphasizing rather hope*less*ness: Underprivileged women in more unequal markets may have little hope of securing decent marriage matches (or decent careers for themselves), anyway. Therefore, the burden of single parenting makes relatively little difference to their life prospects.

In examining the relationship between inequality and single parenting propensity (or marriage rates), there is often an elephant in the room. Although both Gould and Paserman (2003) and Kearney and Levine (2014) use rich sets of control variables to test for alternative explanatory factors, neither clearly addresses the question of direct reverse-causal effects of single parenting (or marriage rates) on inequality. Indeed, Chetty et al. (2014) find single parenting to be the strongest and most robust single predictor of socioeconomic immobility: Although not causal, this is highly suggestive. Single parenting may exacerbate inequality, either directly - by distorting women's career prospects, or distorting men's career incentives - or via societal trauma more generally.

As an application in this paper, I test various alternative instruments for the Gini coefficient of earnings to estimate the effect of earnings inequality on rates of single parenting. The Gini is one of many plausible (but perhaps the single most well-known) measures of inequality. Because my goal is only to examine the implications of using different instruments, I use a (causal) "reduced-form" approach. That is, rather than deriving a model of equilibrium earnings and single parenting from assumed utility functions (i.e. a "structural" approach), I begin rather from the more downstream assumption that single parenting's equilibrium response to earnings can be written in a general log-linear form,

$$log\ (\ \mathrm{P}_{\mathrm{t,z}}^{SingleParent}\ )\ =\ \beta^{Inequality}\ \cdot\ log\ (\ gini\ [\ \mathrm{Y}_{\mathrm{i}\ \in\ \mathrm{t,z}}\ ]\ )\ +\ \{\ \beta^{x}\ \cdot\ \mathrm{x}_{\mathrm{t,z}}\ \} \tag{41}$$

where $\mathrm{P}_{\mathrm{t,z}}^{SingleParent}$ is the fraction of adults in Commuting Zone z who are single parents, and $\mathrm{x}_{\mathrm{t,z}}$ are controls. $\beta^{Inequality}$ is the causal elasticity of single parenting with respect to inequality: It encompasses the *sum* effect of all mechanisms through which inequality would motivate single parenting - including both the "hope" of Gould and Paserman (2003), and the "hopelessness" of Kearney and Levine (2014).

Although equation (41) represents all effects of inequality on single parenting propensity, it does not represent any reverse effects - that is, of single parenting on inequality.[16] For this reason, an ordinary regression estimation of (41) would be confounded by simultaneity bias, and fail to recover $\beta^{Inequality}$. The purpose of instrumenting is to restrict the information about the Gini that enters into the estimation of (41) to only variation that cannot be driven by single parenting. This removes any reverse-casual effects, hence identifying $\beta^{Inequality}$.

As discussed throughout this paper, Shift-Share instruments use delocalizations over space and time. These delocalizations can make the instruments causally prior (exogenous) with respect to almost any outcome. The Shift-Share instrument restricts to information from previous time periods (Shares), and from other cross sectional units (Localities) beyond z - typically the national or global average over all such units. Therefore, the instrument is valid so long as the outcome cannot affect either Shares in previous time periods, or national averages beyond z in the current time period. These conditions can usually be supposed to hold, given that Shares are chosen to represent deep characteristics of Localities z.

Like Shift-Share instruments, the instrument of Boustan et al. (2013), which I call the Lagged Bin Share Gini, can be considered valid in a wide variety of settings. In Boustan et al. (2013), the outcome of interest is local government taxation and expenditure. Enamorado et al. (2016) use the same instrument, but for a very different outcome variable - violent crime rates. In either case, the explanatory variable of interest is the local Gini coefficient of income, as a measure of inequality. Although the reason for inequality to drive taxation is unrelated to the reason for inequality to drive crime, the instrument is valid for the same reason in either setting: *present* values of the outcome variable cannot drive *past* values of the Shares that are used to construct the instrument.

I apply the Lagged Bin Share Gini (and my alternative Shift-Share Gini) in yet another setting, that is, with single parenting rates as the outcome variable. As in the settings of

[16] The sum of these reverse effects would follow an analogous, additional equation (unwritten).

local taxation and of crime rates, the instrument is valid for the same reason: *present* period single parenting rates cannot drive *past* values of the Shares. This raises the question: What advantage do Shift-Share instruments have over simply using lags as instruments? That is, why not use lagged values of the Gini itself as instrument for the Gini in the present?

Indeed, it is generally plausible to assume that present values of *any* variable do not affect past values of any other variable, and as such, the use of simple lags as instruments is widespread. The exception would be if agents are "forward looking" (able to predict the future), and also able to adjust the lagged explanatory variable accordingly (in response to the future). The advantage of Shift-Share instruments over simple lag instruments comes from the split between Shift and Share: Only the Shares are lagged, while the Shifts are delocalized in a stronger way. Shares are chosen to represent deep characteristics - things that cannot be readily adjusted in response to future expectations. The remaining, shallower or more adjustable components of the explanatory variable can thus be resolved into the Shift, and hence neutralized by the stronger delocalization.

I argue that the Lagged Bin Share Gini is essentially similar to a simple lag instrument. Although a lag in itself may yield a valid instrument, a true Shift-Share instrument takes the additional step of more strongly delocalizing a component (the Shift) of the explanatory variable that is most vulnerable to endogeneity. The Lagged Share Gini instrument (with a Frozen Lag τ) is,

$$\tilde{\mathrm{X}}_{\mathrm{t,z}} = \frac{\sum_{h}\sum_{\mathrm{h}} \overbrace{\mathrm{P}_{\mathrm{h},\tau,\mathrm{z}}\mathrm{P}_{h,\tau,\mathrm{z}}}^{\text{Share}} \cdot \overbrace{\left|\mathrm{Y}_{\mathrm{h,t}} - \mathrm{Y}_{h,\mathrm{t}}\right|}^{\text{Push}}}{2\sum_{h} \underbrace{\mathrm{P}_{h,\tau,\mathrm{z}}}_{\text{Share}} \cdot \underbrace{\mathrm{Y}_{h,\mathrm{t}}}_{\text{Push}}} \tag{42}$$

With Updating Lags rather,[17]

$$\tilde{X}_{t,z} = \frac{\sum_h \sum_{\text{h}} \overbrace{P_{\text{h},t-1,z} P_{h,t-1,z}}^{\text{Share}} \cdot \overbrace{\left| Y_{\text{h},t} - Y_{h,t} \right|}^{\text{Push}}}{2 \sum_h \underbrace{P_{h,t-1,z}}_{\text{Share}} \cdot \underbrace{Y_{h,t}}_{\text{Push}}} \tag{43}$$

For comparison, a simple lag of the Gini is,

$$X_{t-1,z} = \frac{\sum_h \sum_{\text{h}} \overbrace{P_{\text{h},t-1,z} P_{h,t-1,z}}^{\text{Share}} \cdot \overbrace{\left| Y_{\text{h},t-1} - Y_{h,t-1} \right|}^{\text{Push}}}{2 \sum_h \underbrace{P_{h,t-1,z}}_{\text{Share}} \cdot \underbrace{Y_{h,t-1}}_{\text{Push}}} \tag{44}$$

The difference between (43) and (44) is that the latter lags the Push, while the former uses present values for the Push. However, because the Push is invariant over Localities z in either case, there is little scope for this change to yield meaningful differences in terms of exogeneity.[18] In other words, because the Shares in this case contain essentially all of the information in the original Gini that that is distinctive per Locality z, lagging these Shares is similar to lagging the whole Gini. By contrast, a Shift-Share instrument would sequester some of the distinctive information - that which is most sensitive to endogeneity - into a Shift vector which is delocalized, hence yielding a stronger claim of exogeneity than that offered just by a lag.

Whether a lag in itself is sufficient to yield exogeneity depends on context - particularly, how plausible it is that the explanatory variable responds to *future* values of the outcome

[17] See Section 4.1. Frozen Lags have the clear downside that they make the instrument more relevant in earlier time periods (closer to the lag), and weaker as time goes on. Updating Lags have no clear downside in my opinion.

[18] The Lagged Bin Share instrument may still have an advantage over the simple lag in terms of relevance, although it does not in terms of exogeneity. In other words, the Lagged Bin Share Gini may be better correlated with the present Gini than the lagged Gini is.

of interest. In the setting of Boustan et al. (2013), this question would be whether the local income distribution responds to future changes in local taxation and expenditure policy. Lower earning people for example, if accurately predicting future policy, may select into localities in which future taxation and expenditure are on an upward trajectory. With single parenting rates as the outcome of interest, an analogous confounding scenario may be that wealthier people avoid localities in which single parent homes are on the rise.

Where a lag instrument relies on the explanatory variable to not respond to future changes in the outcome variable, a Shift-Share instrument rather relies on only the Shares to not do so. To accomplish this purpose, it is vital that Shares represent deep characteristics of cross sectional units (Localities), such as can be viewed as causal first movers of the economic system under study. The industrial profile each Locality (the classical Share vector) meets this purpose because it arises from geographical and historical forces that are beyond the scope of the model. For example, shipping relies on access to water, and locations are bound to particular industries, such as Detroit to manufacturing, for historical reasons that cannot be easily adjusted.

Unlike classical industrial Shares, the Shares of the Lagged Bin Share instrument needn't coincide with anything particularly fundamental to localities. These Shares are Bin Shares, which simply map the local income distribution. They are sensitive to anything that alters the local income distribution, including selective migration of higher or lower earning individuals. As such, these Shares are more vulnerable to confoundedness by future expectations than Shares of a typical Shift-Share instrument would be. This is the flip side of lacking a Shift component that is delocalized over the cross section. The local averages that are most vulnerable to endogeneity would be in the Shift rather than the Share, and hence neutralized by the Shift delocalization.

Where the Lagged Bin Share Gini replaces the endogenous Bin Share with a lagged Bin Share, my Shift-Share Gini rather replaces it with a classical Shift-Share instrument for the Bin Share itself. Although it is its own kind of share, the Bin Share does not constitute the Share of a Shift-Share instrument, for two reasons that are flip sides of the same coin. First, there is no corresponding Shift - that is, a component distinctive to each Locality that is delocalized in the instrument. Second, indeed because there is no Shift component to carry the regressor's local information that is most vulnerable to endogeneity, the Bin Share (even lagged) is left vulnerable. My Shift-Share Gini fixes this problem by resolving the Bin Share into an underlying Shift and exogenous Share.

To examine concrete implications of each of the above versions of the Gini instrument, I estimate (44) using each.[19] OLS estimation of (44) should be confounded by reverse causality in so far as single parenting rates have effects on the distribution of earnings. If both the forward and reverse effects are positive - that is, if inequality drives more single parenting, and single parenting also drives more inequality - then it is natural for the OLS estimate to be biased downward in magnitude.[20] Indeed, Table 2 shows that the elasticity estimate reported by OLS is smaller than that of 2SLS using any of the instrument alternatives.

[19]Source data are from the US Census and American Community Survey via IPUMS USA, Ruggles et al. (2020). I aggregate by Commuting Zones (CZs) as Localities z. CZs are defined based on actual commuting patterns in 1990, and hence capture local labor markets; see Tolbert and Sizer (1996) and Autor and Dorn (2013).

[20]Although a mutually positive reverse-causal effect will increase the correlation between the two variables, it will nonetheless decrease the magnitude of the OLS slope coefficient estimate, by increasing the "run" in "slope = rise/run" more than it increases the "rise."

Table 2: Elasticity of Single Parenting w.r.t. Gini of Earnings

	OLS	2SLS			
	Gini	Lagged Share Gini	Lagged Gini	Delocalized Shift Gini	Shift-Share Gini
Gini	0.654***	1.077***	1.105***	1.184***	1.487***
	(0.114)	(0.271)	(0.268)	(0.221)	(0.369)
Gini	0.465***	1.002***	0.997***	1.081***	1.217***
	(0.117)	(0.272)	(0.271)	(0.227)	(0.401)
Mean	-0.280***	-0.226***	-0.209**	-0.165**	-0.151*
	(0.045)	(0.082)	(0.082)	(0.083)	(0.088)

Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. Observations are 722 Commuting Zones by 5 time periods 1980-2019. Dependent variable is log single parenting rate; regressor is log Gini of earnings. Controls are region by year effects, as well log average earnings in the lower panel. The 2SLS columns each use a different instrument for the Gini as stated, and also instrument average earnings with the Shift-Share average.

The results given in Table 2 illustrate two main points concerning the different instruments. First is that the Lagged Bin Share instrument (Column 2) gives numerically very similar results as the simple lag instrument (Column 3). This is an empirical confirmation of my theoretical argument that the Lagged Bin Share instrument is similar to the simple lag in terms of exogeneity. The second main point is that - supposing the Shift-Share instrument (Column 5) is fully exogenous, and hence yielding the most accurate result - then the Lagged Bin Share and lag instruments are getting part of the way there, that is, partially but not fully correcting the bias from OLS.[21] This is indeed what we ought to expect given that the Shift-Share instrument uses both a (Share) lag and a Shift delocalization, and that both of these are essential in correcting bias.

[21] Likewise, the Delocalized Shift instrument (Column 4) corrects bias only part way. This uses the original (non-lagged) Share vector in conjunction with the delocalized Shift vector.

8 Conclusion

This paper may be unique in focusing on the generative process by which a researcher may arrive at Shift-Share instruments. By deriving the instrument directly from an accounting identity of the explanatory variable, one can appreciate both the essential features of Shift-Share instruments, and the scope of their potential varieties. Using my simple approach for understanding Shift-Share instruments as modified accounting identities, I closely compare a wide variety of instruments from the literature. I then also develop general formulas for several new varieties - instruments for variances, skews, mean absolute deviations, and Gini coefficients.

As an empirical application, I measure the effect of earnings inequality on rates of single parenting, using multiple alternative instruments for the Gini coefficient of earnings. The empirical results illustrate core themes from earlier in the paper. That is, Shift-Share instruments both delocalize (replace with nonlocal averages) the more endogenous part of the accounting identity (the Shift vector), and lag the more exogenous part of the accounting identity (the Share vector). Empirically, I show that instruments that do only one or the other - delocalize the Shift, or lag the Share - also correct bias, but only part way. Thus, although each of these steps - delocalizing the Shift, and lagging the Share - provides its own argument of exogeneity, both make meaningful contributions to the exogeneity of the Shift-Share instrument as a whole.

References

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

Autor, David, and David Dorn. 2013. "The Growth of Low Skill Service Jobs and the Polarization of the U.S. Labor Market." *American Economic Review*, 103(5): 1553–1597.

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

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

Baum-Snow, Nathaniel, and Fernando Ferreira. 2015. "Causal Inference in Urban and Regional Economics." *Handbook of Regional and Urban Economics*, 5: 3–68.

Blanchard, Olivier, and Lawrence Katz. 1992. "Regional Evolutions." *Brookings Papers on Economic Activity*, 23(1): 1–76.

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

Bound, John, and Harry Holzer. 2000. "Demand Shifts, Population Adjustments, and Labor Market Outcomes during the 1980s." *Journal of Labor Economics*, 18(1): 20–54.

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

Card, David. 2001. "Immigrant Inflows, Native Outflows, and the Local Labor Market Impacts of Higher Immigration." *Journal of Labor Economics*, 19(1): 22–64.

Chetty, Raj, Nathaniel Hendren, Patrick Kline, and Emmanuel Saez. 2014. "Where is the Land of Opportunity? The Geography of Intergenerational Mobility in the United States." *Quarterly Journal of Economics*, 129(4): 1553–1623.

Diamond, Rebecca. 2016. "The Determinants and Welfare Implications of US Workers' Diverging Location Choices by Skill: 1980-2000." *American Economic Review*, 106(3): 479–524.

Enamorado, Ted, Luis F López-Calva, Carlos Rodríguez-Castelán, and Hernán Winkler. 2016. "Income inequality and violent crime: Evidence from Mexico's drug war." *Journal of Development Economics*, 120: 128–143.

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

Gould, Eric D, and Daniele M Paserman. 2003. "Waiting for Mr. Right: Rising Inequality and Declining Marriage Rates." *Journal of Urban Economics*, 53(2): 257–281.

Kearney, Melissa S, and Phillip B Levine. 2014. "Income Inequality and Early Nonmarital Childbearing." *The Journal of Human Resources*, 49(1): 1–31.

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

Ruggles, Steven, Sarah Flood, Ronald Goeken, Josiah Grover, Erin Meyer, Jose Pacas, and Matthew Sobek. 2020. IPUMS USA: Version 10.0 [dataset]. Minneapolis, MN.

Tolbert, Charles M, and Molly Sizer. 1996. "U.S. Commuting Zones and Labor Market Areas: A 1990 Update."