Bartik Instruments: What, When, Why, and How*

Paul Goldsmith-Pinkham      Isaac Sorkin      Henry Swift

This version: July 21, 2017.

Abstract

We demystify and formalize the Bartik instrument. The Bartik instrument is formed by interacting local industry shares and national industry growth rates. We show that the Bartik instrument is equivalent to using local industry shares as instruments. Hence, the identifying assumption is best stated in terms of these shares, with the national industry growth rates only affecting instrument relevance. This insight applies to a variety of “Bartik-like” instruments. We consider three tests of this identifying assumption and implement them in the context of the canonical application of estimating the inverse elasticity of labor supply.

*Goldsmith-Pinkham: Federal Reserve Bank of New York. Email: paulgp@gmail.com. Sorkin: Stanford University and NBER. Email: sorkin@stanford.edu. Swift: Unaffiliated. Email: henryswift@gmail.com. The views expressed are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of New York or the Federal Reserve Board. All errors are our own.
The Bartik instrument is an instrument for labor demand, named after Bartik (1991), and popularized in Blanchard and Katz (1992). These papers define the instrument as the local employment growth rate predicted by interacting local industry employment shares with national industry employment growth rates. This instrument and its variants have since been used across many fields in economics, including labor, public, development, macroeconomics, international trade, and finance. Indeed, as we discuss at the end of the introduction, numerous instruments beyond those labelled as Bartik instruments per se have the same formal structure as Bartik instruments.

Our goal is to demystify and formalize the Bartik instrument. In our exposition, we focus on the canonical setting of employment growth, but our results apply more broadly wherever Bartik-like instruments are used. For simplicity, consider the cross-sectional regression of wage growth on employment growth

\[ y_l = \alpha + \beta x_l + \epsilon_l, \]

where \( y_l \) is wage growth in location \( l \), \( x_l \) is the employment growth rate, and \( \epsilon_l \) is an error term. \( \beta \) is our estimand of interest, the inverse elasticity of labor supply. The Bartik instrument combines two accounting identities. The first is that employment growth is the inner product of industry shares and local industry growth rates

\[ x_l = \sum_k z_{lk} g_{lk}, \]

where \( z_{lk} \) is location \( l \) share of industry \( k \), and \( g_{lk} \) is industry \( k \) growth rate in location \( l \). The second is that we can decompose the industry-growth rates as follows \( g_{lk} = g_k + g_l + \tilde{g}_{lk} \). The Bartik instrument is the inner product of the industry-location shares and the common industry component of the growth rates; formally, \( B_l = \sum_k z_{lk} g_k \).

We show that the Bartik instrument is equivalent to using local industry shares as instruments, with variation in the common industry component of growth only contributing to instrument relevance. More precisely, using the Bartik instrument in two-stage least squares (TSLS) is numerically equivalent to a generalized methods of moments (GMM) estimator with the local industry shares as instruments and a weight matrix constructed from the national industry growth rates. This result means that the identifying assumption for the Bartik instrument is best stated in terms of local industry composition.

We argue that researchers should conduct three tests for the validity of the instrument. First, because the instrument is implicitly local industry shares, researchers should consider the relationship between local industry shares and observable characteristics. To the extent

---

1 Arguably the Bartik instrument pre-dates Bartik (1991) and is also sometimes known as a shift-share instrument. See Freeman (1980) for an earlier use.
that these observable characteristics are correlated with the error term, then this suggests scope for bias. Controlling for observables addresses these concerns. Second, when Bartik is used in a panel setting, researchers should test for parallel pre-trends. This test cannot be done in an off-the-shelf way because Bartik allows for a new shock in every period. We show how to partial out the effect of past values of the Bartik instrument to test for parallel pre-trends. Third, researchers should leverage the observation that the industry shares are the instruments and conduct an overidentification test. However, because the industry shares are—at least in our application—weak instruments, we reduce the dimensionality of the instruments using principal component analysis to run the test. While the spirit of the first two tests are a standard part of the applied microeconomics toolkit, they are not always used with the Bartik instrument. Failure to pass these checks should be interpreted as evidence of identification concerns related to the industry shares, and not an incorrect choice of national growth rates.

For our empirical application, we consider the Bartik instrument in the canonical setting of estimating the inverse elasticity of labor supply. We write down a simple economic model, relate our econometric results to economic primitives, and, using Census data, implement our proposed tests. We have three main findings. First, industry shares are correlated with many observables, including education. Controlling for these observables attenuates estimates by almost a fifth. Second, we find quantitatively important evidence of pre-trends. Third, the overidentification test rejects the null of exogeneity. The presence of these identification concerns in this canonical application emphasizes the importance of subjecting the Bartik instrument to specification checks in other applications as well.

We next turn to the question of why researchers use Bartik, rather than using the industry shares as instruments. There are two reasons. First, there are some settings where Bartik is used as a proxy because it is not possible to estimate a first-stage (i.e., we observe the $z_{lk}$, but not $x_l$ or $g_{lk}$). Second, using the set of local industry shares as instruments leads to an overidentified estimation setting, which can cause problems of bias in finite samples, especially because the shares are “weak” instruments by themselves. In contrast, the construction of the Bartik instrument converts the high-dimensional industry shares to a single instrument, reducing an overidentified setting to a just identified one. Nevertheless, this suggests that in some applications there might be efficiency gains from using the identifying assumption differently than the Bartik instrument does.

Finally, we provide simulation evidence on the finite sample properties of the Bartik instrument and our tests. In our simulations, the bias of estimates of $\beta$ are affected by the number of locations, whereas the number of industries appears to have little effect. Additionally, our tests have good finite sample properties.
mate the inverse labor supply elasticity, a much broader set of instruments are Bartik-like instruments. We define a Bartik-like instrument as one where the researcher exploits the inner product structure of the endogenous variable to construct an instrument. This encompasses at least six instruments, which are not always labelled as Bartik instruments per se. First, the “China shock” instrument of Autor, Dorn, and Hanson (2013) interacts local industry composition with the growth of Chinese imports to European countries. Second, the “immigrant enclave” instrument introduced by Altonji and Card (1991) interacts initial immigrant composition of a place with immigration flows from origin countries. Third, researchers, such as Greenstone, Mas, and Nguyen (2015), interact pre-existing bank lending shares with changes in bank lending volumes to instrument for credit supply. Fourth, the simulated instrument of Currie and Gruber (1996b) interacts changes in local laws with national population characteristics. Fifth, Acemoglu and Linn (2004) interact age-group spending patterns with demographic changes to instrument for market size. Finally, the “judge” instrument interacts a case’s assigned judge with the leniency of the judge (e.g. Kling (2006)). We discuss these examples in greater detail in Appendix C.

We note three limitations to our analysis. First, we assume locations are independent and so ignore the possibility of spatial spillovers or correlation. Second, we assume that there are no dynamics in the response to shocks. Third, we assume that the number of locations grows large, but the number of industries and time periods are fixed.

Roadmap: Section 1 introduces our notation and shows that the Bartik instrument is equivalent to using industry shares as instruments. Section 2 relates our econometric results to economic primitives, and shows how to probe the identifying assumptions in the canonical application of estimating the inverse elasticity of labor supply. Section 3 discusses why it might be desirable to use Bartik rather than the industry shares as instruments, and presents practical guidance. Section 4 presents simulation evidence. Section 5 concludes by summarizing the results of the paper in a checklist of recommendations for practice.

1 Equivalence between Bartik and industry shares

We first show that using Bartik is equivalent to using industry shares as instruments. This result implies that the identifying assumption is best stated in terms of industry shares.

We begin this section by setting up the most general case: panel data with $K$ industries, $T$ time periods, and controls. Through a series of special cases, we then build up to the main result that Bartik is (numerically) equivalent to using local industry shares as instruments. To focus on identification issues, throughout this section we discuss infeasible Bartik, where

---

2See Ruist, Stuhler, and Jaeger (2017) for discussion of dynamics in the context of immigration.
3See Borusyak and Jaravel (2017) for discussion of the setting where the number of industries grows.
we assume that we know the common national component of industry growth rates. In Section 3 we discuss estimation issues, which includes estimation of the growth rates.

1.1 Full panel setup

We begin by setting up the general panel data case with \( K \) industries and \( T \) time periods. This set-up most closely matches that used in empirical work. It allows for the inclusion of both location and time fixed effects as well as other controls.

We are interested in a regression of wage growth on employment growth and a set of controls:

\[
y_{lt} = W_{lt}\beta_0 + x_{lt}\beta + \epsilon_{lt}.
\]

We consider \( \{\{x_{lt}, W_{lt}, \epsilon_{lt}\}_t\}_{t=1}^T \) to be independent and identically distributed with \( L \) going to infinity and \( T \) fixed. In the canonical setting, \( l \) indexes a location, \( t \) a time period, \( y_{lt} \) is wage growth, \( W_{lt} \) is a vector of controls which could include location and time fixed effects, and \( x_{lt} \) is employment growth. Then \( \beta \) is the inverse elasticity of labor supply. We assume that \( \mathbb{E}[x_{lt}\epsilon_{lt}|W_{lt}] \neq 0 \), so the OLS estimator for \( \beta \) is biased and we need an instrument.

The Bartik instrument exploits the inner product structure of employment growth. Specifically, employment growth is the inner product of industry shares and industry-location growth rates:

\[
x_{lt} = Z_{lt}'G_{lt} = \sum_{k=1}^{K} z_{lkt}g_{kt},
\]

where \( Z_{lt} \) is a \( K \times 1 \) vector of industry-location-time period shares, \( \sum_{k=1}^{K} z_{lkt} = 1 \), and \( G_{lt} \) is a \( K \times 1 \) vector of industry-location-time period growth rates where the \( k^{th} \) entry is \( g_{kt} \). Note that \( K \) is fixed. We decompose the industry-location-period growth rate into industry-period, location-period and idiosyncratic industry-location-period components:

\[
g_{lkt} = g_{kt} + g_{lt} + \tilde{g}_{lkt}.
\]

Without loss of generality, we assume that the location-period and idiosyncratic industry-location-period components \( (g_{lt} \) and \( \tilde{g}_{lkt} \)) are mean zero random variables. For reasons that we discuss further in Section 3 we fix industry shares to an initial time period, so that the Bartik instrument is the inner product of the initial industry-location shares and the industry-period growth rates:

\[
B_{lt} = Z_{l0}'G_{lt} = \sum_{k} z_{lkt}g_{kt},
\]

\[
4
\]
where $G_t$ is a $K \times 1$ vector of the industry growth rates in period $t$ (the $k^{th}$ entry is $g_{kt}$), and $Z_{l0}'$ is the vector of industry shares in location $l$. Define $Z_{L0}$ to be the $K \times L$ matrix of industry-location shares.

Hence, we have a standard two-stage least squares set-up where the first-stage is a regression of employment growth on the set of controls and the Bartik instrument:

$$x_{lt} = W_{lt}' \tau + B_{lt}' \gamma + \eta_{lt}.$$  

The two assumptions for the validity of the TSLS estimator are that the Bartik instrument is exogenous, $\mathbb{E} (B_{lt} \epsilon_{lt} | W_{lt}) = 0$, and that it is relevant, $\text{Cov} (B_{lt}, x_{lt} | W_{lt}) \neq 0$.

1.2 Equivalence in three special cases

We build up to the result that the Bartik instrument is equivalent to using industry shares as instruments through three special cases which each cleanly illustrate one aspect of the general result.

Two industries and one time period

With two industries and one time period, it is most transparent to see that the Bartik instrument is equivalent to industry shares as instruments. To see this, write the Bartik instrument explicitly:

$$B_l = z_{l1} g_1 + z_{l2} g_2,$$

where $g_1$ and $g_2$ are the industry components of growth. Since the shares sum to one, with only two industries, we can write the second industry share in terms of the first, $z_{l2} = 1 - z_{l1}$, and simplify the Bartik instrument to depend only on the first industry share:

$$B_l = g_2 + (g_1 - g_2) z_{l1}.$$  

Because the only term on the right hand side with a location subscript is the first industry share, the cross-sectional variation in the instrument comes from the first industry share. Substitute into the first-stage:

$$x_l = \gamma_0 + \gamma B_l + \eta_l = \gamma_0 + \gamma g_2 + \gamma (g_1 - g_2) z_{l1} + \eta_l.$$  

This equation shows that the difference between using the first industry share and Bartik as the instrument is to rescale the first stage coefficients by the difference in the growth
rates between the two industries \((1/g_1 - g_2)\). But whether we use the Bartik instrument or the first industry share as an instrument, the predicted employment growth (and hence the estimate of the inverse elasticity of labor supply) would be the same. Hence, with two industries, using the Bartik instrument in TSLS is numerically identical to using \(z_{l1}\) and \(z_{l2}\) as an instrument.

**K industries and one time period**

Next, we show that with \(K\) industries as instruments in a generalized method of moments (GMM) set-up with a specific weight matrix, Bartik is identical to using the set of industry shares as instruments.

To prove this result, we introduce some additional notation. Let \(G\) be the \(K \times 1\) vector of industry growth rates, let \(Z_L\) be the \(K \times L\) matrix of industry shares, let \(Y_L\) be the \(L \times 1\) vector of outcomes, let \(X_L\) be the \(L \times 1\) vector of endogenous variables, let \(B_L = (G'Z_L)'\) be the \(L \times 1\) vector of Bartik instruments, and let \(\Omega\) be an arbitrary \(K \times K\) matrix. Let a bar over a vector denote the mean of the vector times the all ones vector of the appropriate dimension.\(^4\) We define the Bartik and the GMM estimator using industry shares as instruments:

\[
\hat{\beta}_{L,Bartik} = \frac{(B_L - \bar{B}_L)'Y_L}{(B_L - \bar{B}_L)'X_L}; \quad \text{and} \quad \hat{\beta}_{L,GMM} = \frac{X_L'(Z_L - \bar{Z}_L)'\Omega(Z_L - \bar{Z}_L)Y_L}{X_L'(Z_L - \bar{Z}_L)'\Omega(Z_L - \bar{Z}_L)X_L}.
\]

The following proposition says that Bartik and GMM are equivalent for a particular choice of weight matrix.

**Proposition 1.1.** If \(\Omega = (G - \bar{G})(G - \bar{G})'\), then \(\hat{\beta}_{L,GMM} = \hat{\beta}_{L,Bartik}\).

**Proof.**

\[
\hat{\beta}_{L,GMM} = \frac{X_L'(Z_L - \bar{Z}_L)'(G - \bar{G})(G - \bar{G})'(Z_L - \bar{Z}_L)Y_L}{X_L'(Z_L - \bar{Z}_L)'(G - \bar{G})(G - \bar{G})'(Z_L - \bar{Z}_L)X_L} = \frac{X_L'(B_L - \bar{B}_L)(B_L - \bar{B}_L)'Y_L}{X_L'(B_L - \bar{B}_L)(B_L - \bar{B}_L)'X_L} = \hat{\beta}_{L,Bartik},
\]

where \(B_L - \bar{B}_L = (Z_L - \bar{Z}_L)'(G - \bar{G})\) because \((Z_L - \bar{Z}_L)'G = Z_L'G - \bar{Z}_L'G = B_L - \bar{B}_L\), \((Z_L - \bar{Z}_L)'\bar{G} = 0\) (an \(L \times 1\) vector of zeros), and \(X_L'(B_L - \bar{B}_L)\) is a scalar and so cancels. □

Hence, with \(K > 2\) industries, the Bartik instrument and industry shares as instruments

\(^4\)In the case of \(Z_L\), which is a matrix, these are the column means (i.e., the mean share of industry \(k\) across locations).
are numerically equivalent for a particular choice of weight matrix. This result means that
the Bartik instrument is a way of reducing the number of instruments. Specifically, Bartik
chooses a particular way to weight the $K$ moment conditions given by the exogeneity of
industry shares. Hence, the decision of what growth rates to use is a decision about how to
weight a set of moment conditions.

Two industries and two time periods
In a panel with two time periods, if we allow for time-varying coefficients on the time-
invariant industry shares, then Bartik is equivalent to using industry shares as instruments.
Recall that the industry shares are time-invariant because we exclusively use the initial in-
dustry shares.

To see this result, we focus on two industries, and define the infeasible Bartik instrument
so that it varies over time:

$$B_{lt} = g_{1t}z_{l10} + g_{2t}z_{l20} = g_{2t} + (g_{1t} - g_{2t})z_{l10},$$

where $g_{1t}$ and $g_{2t}$ are the industry-by-time growth rate for industry 1 and 2. Because we
fix the shares to an initial time-period, denoted by $z_{lk0}$, the time variation in $B_{lt}$ comes from
the difference between $g_{1t}$ and $g_{2t}$.

To see the relationship between the cross-sectional and the panel estimating equations,
specialize our general panel setup to have the vector of controls consist solely of location
and time fixed effects. Then the first-stage is

$$x_{lt} = \tau_l + \tau_t + B_{lt} \gamma + \eta_{lt}.$$ 

Now substitute in the Bartik instrument and rearrange the first stage:

$$x_{lt} = \tau_l + \left(\tau_t + g_{2t} \gamma\right) + z_{l10} \Delta g_{t} \gamma + \eta_{lt},$$

where $\Delta g_{t} = g_{1t} - g_{2t}$. This first-stage is more complicated than in the cross-sectional case
because there is a time-varying growth rate multiplying the time-invariant industry share.

To recover the equivalence between Bartik and using shares as instruments in the panel
setting, interact industry shares with time effects. Once we allow for time-varying coeffi-
ciants on the industry shares, this approach is equivalent to Bartik:

\[ x_{it} = \tau_i + \tilde{\tau}_t + z_{i10} \Delta_{g,it} + \eta_{it} \]  
(Bartik)

\[ x_{it} = \tau_i + \tilde{\tau}_t + z_{i10} \sum_{s \neq 0} \tilde{\gamma}_s (t = s) + \eta_{it}. \]  
(Industry Shares)

Hence, using Bartik in a panel is analogous to a difference-in-difference estimation framework, with a continuous treatment and non-parametrically estimated exposure. Here, the size of the policy is measured by the dispersion in national industry growth rates, \( \Delta_{g,it} \), and the exposure to the policy is given by \( z_{i10} \). And, because the industry shares are time invariant, \( \tilde{\gamma}_t \) can only be estimated relative to a base period.

### 1.3 Equivalence with \( K \) industries and \( T \) time periods

We now have the ideas necessary to show the equivalence between Bartik and industry shares as instruments in the general setup of Section 1.1. With \( K \) industries and \( T \) time periods (but no controls), the equivalence simply involves creating \( K \times T \) instruments (industry shares interacted with time periods). Then an identical GMM result holds as we proved in the cross-section with \( K \) industries. Extending the result to include controls—such as location and time fixed effects—is notationally cumbersome so we leave the formal details to Appendix A.

In the more general set-up, we also have a limiting result that under standard regularity conditions and if Bartik is a valid instrument, then Bartik, TSLS with industry shares as instruments, and our GMM estimator are all consistent for the true parameter:

\[ \lim_{L \to \infty} \hat{\beta}_{L,GMM} = \lim_{L \to \infty} \hat{\beta}_{L,Bartik} = \lim_{L \to \infty} \hat{\beta}_{L,TSLS} = \beta. \]

### 2 An empirical example

So far we have emphasized that Bartik is equivalent to using industry shares (interacted with time) as instruments and hence that the identifying assumption is best viewed in terms of the industry shares. We now present an empirical example to make these ideas concrete. We begin with an economic model that maps our econometric conditions into economic primitives in the canonical setting of estimating the inverse elasticity of labor supply. We then use the model and our econometric results to guide an empirical investigation of the plausibility of using industry shares as instruments.
2.1 An economic model

We consider \( L \) independent locations indexed by \( l \). Labor is homogeneous so that the wage in location \( l \) in period \( t \) is \( w_{lt} \). The labor supply curve in location \( l \) in period \( t \) is:

\[
\ln N_{lt}^S = \sigma_{lt} + \theta \ln w_{lt}.
\] (2.1)

Here, \( N_{lt}^S \) is the quantity of labor supplied and \( \sigma_{lt} \) is a location-period-specific shifter of the level of labor supply. The local labor supply elasticity, \( \theta \), is the parameter of interest and is common across industries and locations.

The demand curve for industry \( k \) in location \( l \) at time \( t \) is given by

\[
\ln N_{lkt}^D = T_{lk} \alpha_{lkt} - \phi \ln w_{lt}.
\] (2.2)

Here, \( N_{lkt}^D \) is the quantity of labor demanded, \( T_{lk} \) is a fixed factor that generates persistent differences in industry composition, \( \alpha_{lkt} \) is the time-varying industry-location level of labor demand, and \( \phi \) is the common elasticity of local labor demand. Letting \( \alpha_{lt} = \ln \left( \sum_k \exp \{ T_{lk} \alpha_{lkt} \} \right) \) be the aggregated location-specific shifter of labor demand, the location-level demand curve is:

\[
\ln N_{lt}^D = \alpha_{lt} - \phi \ln w_{lt}.
\] (2.3)

The equilibrium condition in market \( l \) in period \( t \) is a labor market clearing condition:

\[
N_{lt} = N_{lt}^S = \sum_k N_{lkt}^D = N_{lt}^D.
\]

We let \( \tilde{x}_t = \ln x_t \) and \( dx_t \) be the per-period change in \( x_t \).

To construct the infeasible Bartik instrument, write the change in log employment in an industry-location, and then label the components of this decomposition in the same notation as the previous section:

\[
d \tilde{N}_{lkt} = d \alpha_{kt} - \left( \frac{\phi}{\theta + \phi} d \alpha_{lt} - \frac{\phi}{\theta + \phi} d \sigma_{lt} \right) + T_{lk} d \alpha_{lkt} - d \alpha_{kt}.
\]

Define \( z_{lk0} = \exp(T_{lk} \alpha_{lk0}) / \sum_{k'} \exp(T_{lk'} \alpha_{lk'0}) \) to be the industry shares in period 0. Then the infeasible Bartik instrument that isolates the industry component of the innovations to demand shocks is \( B_{lt} = \sum_k z_{lk0} d \alpha_{kt} \).

In differences and with only two time periods, the equation we are interested in esti-
mating is:

\[ (d\bar{\omega}_{lt+1} - d\bar{\omega}_{lt}) = (\tau_{t+1} - \tau_t) + \beta(d\bar{N}_{lt+1} - d\bar{N}_{lt}) + (\epsilon_{lt+1} - \epsilon_{lt}) \]  \hspace{1cm} (2.4)

where we have differenced out a location fixed effect, \( \epsilon_{lt} \) is an additive error term and the goal is to recover the inverse labor supply elasticity \( \beta = \frac{1}{\phi} \). Traditional OLS estimation of equation (2.4) is subject to concerns of endogeneity and hence the Bartik instrument may provide a way to estimate \( \beta \) consistently.

### 2.2 The model’s empirical analogue

It is instructive to compare the population expressions for \( \hat{\beta}_{OLS} \) and \( \hat{\beta}_{Bartik} \):

\[
\hat{\beta}_{OLS} = \frac{\theta}{(\theta + \phi)^2} \cdot \frac{Var(da_{lt+1} - da_{lt}) - \phi \frac{\theta}{(\theta + \phi)^2} \cdot Var(d\sigma_{lt+1} - d\sigma_{lt}) + \phi \frac{\theta}{(\theta + \phi)^2} \cdot Cov(d\sigma_{lt+1} - d\alpha_{lt}, d\sigma_{lt+1} - d\sigma_{lt})}{\theta \cdot \frac{\phi}{(\theta + \phi)^2} \cdot Var(da_{lt+1} - da_{lt}) + \phi \frac{\theta}{(\theta + \phi)^2} \cdot Var(d\sigma_{lt+1} - d\sigma_{lt}) + \phi \frac{\theta}{(\theta + \phi)^2} \cdot Cov(d\sigma_{lt+1} - d\alpha_{lt}, d\sigma_{lt+1} - d\sigma_{lt})}
\]

\[
\hat{\beta}_{Bartik} = \frac{Cov[da_{lt+1} - da_{lt}, \sum_k z_{lk0} (d\alpha_{kt+1} - d\alpha_{kt})]}{\theta \cdot Cov[da_{lt+1} - da_{lt}, \sum_k z_{lk0} (d\alpha_{kt+1} - d\alpha_{kt})] + \sum_k z_{lk0} Cov[d\sigma_{lt+1} - d\sigma_{lt}, \sum_k z_{lk0} (d\alpha_{kt+1} - d\alpha_{kt})].}
\]

We see that for \( \hat{\beta}_{OLS} \) to be consistent, an important sufficient condition is that there are no changes in supply shocks, or \( Var(d\sigma_{lt+1} - d\sigma_{lt}) = 0 \). In contrast, for \( \hat{\beta}_{Bartik} \) to be consistent, industry composition must not be related to innovations in supply shocks, or \( Cov[da_{lt+1} - da_{lt}, \sum_k z_{lk0} (d\alpha_{kt+1} - d\alpha_{kt})] = 0 \). Bartik is invalid if the innovations in the supply shocks are predicted by industry composition. For example, Bartik would not be valid if \( d\sigma_{lt+1} - d\sigma_{lt} = \bar{d}\sigma_{lt+1} - d\sigma_{lt} + \sum_k z_{lk0} (d\sigma_{kt+1} - d\sigma_{kt}) \). The relevance condition is that \( Cov[da_{lt+1} - da_{lt}, \sum_k z_{lk0} (d\alpha_{kt+1} - d\alpha_{kt})] \neq 0 \). A necessary condition for instrument relevance is that there is variation in the innovations to demand shocks between at least two industries.

The condition for Bartik to be consistent is weaker than for OLS, since the variance of the innovations to the supply shocks enters into the location-level component of growth (\( g_{lt} \)) and Bartik removes these (but not their correlation with demand shocks). The observation that the Bartik estimator does not include the variance of the innovations to the supply shocks helps explain why Bartik tends to produce results that “look like” a demand shock.

In this model, any given industry share would be a valid instrument. The exclusion restriction is that the industry share does not predict innovations to supply shocks: \( Cov(d\sigma_{lt+1} - d\sigma_{lt}, z_{lk0}) = 0 \). The relevance condition is that \( Cov[da_{lt+1} - da_{lt}, z_{lk0}] \neq 0 \), which says that the industry share is correlated with the innovations in the demand shocks.

While the exclusion restriction is not directly testable, we describe three simple ways to test auxiliary implications of this assumption. We then apply these tests while estimating the inverse labor supply elasticity using the Bartik instrument.
Empirical Test 1: Correlates of industry composition  First, it is helpful to explore the relationship between industry composition and location characteristics \((W_{lt})\) that may be correlated with innovations to supply shocks \((d\sigma_{lt} - d\sigma_{lt-1})\). This provides descriptive evidence both of where the variation comes from, and the types of mechanisms that would be problematic for the exclusion restriction. Since \(Z_{l0}\) is high-dimensional, this can be done in two ways: we can simply regress \(B_{lt}\) on \(W_{lt}\); or we can decompose \(Z_{l0}\) using a technique like principal components analysis (PCA) and examine the correlation between the components of \(Z_{l0}\) and \(W_{lt}\). We prefer using PCA because it uses information only in the industry shares.

Empirical Test 2: Pre-trends  Second, we can use the insight from Section 1.2 that Bartik is analagous to difference-in-differences, and examine pre-trends. However, since Bartik allows for a new shock in every period, if the instrument is serially correlated, parallel pre-trends may not hold. We develop a simple procedure to test for parallel pre-trends in the face of a time-varying instrument, and show that there is evidence of pre-trends using the Bartik instrument.

It is consistent with the validity of the Bartik instrument to find evidence of pre-trends. In a traditional difference-in-differences setting, testing for pre-trends amounts to regressing future values of Bartik on current values of earnings growth

\[
y_{lt-1} = \pi^\text{noresid}_0 + \pi^\text{noresid}_1 B_{lt} + u^\text{noresid}_{lt-1},
\]

and traditionally, we would interpret a significant coefficient on future values of the instrument, \(\pi^\text{noresid}_1\), as evidence of pre-trends.

However, even if the Bartik instrument is valid, this test might find evidence of pre-trends. Consider the relationship between current values of the outcome and future values of the instrument:

\[
\text{Cov}(y_{lt-1}, B_{lt}) = \text{Cov}(a_0 + \beta x_1 + \beta \gamma B_{lt-1} + \beta \eta_1 + e_1, B_{lt}).
\]

If the Bartik instrument is correlated through time, then the future values of the instrument can predict current outcomes through the persistent component of the Bartik instrument. Hence, this naive test for pre-trends might reject the null even if the Bartik instrument is valid.

To address this issue, we ask whether the residuals from the second stage in the current period can be predicted by the values of Bartik in a future period. That is, we remove the part of wage growth that we would predict from the Bartik instrument and then examine the relationship between the residualized wage growth and future values of Bartik. For-
mally, we run the reduced form regression of the current outcome on the current value of
the (constructed) Bartik instrument
\[ y_{lt} = \alpha_0 + \beta\gamma\hat{B}_{lt} + \epsilon_{lt} + \beta\eta_{lt}, \] (2.6)
compute wage growth residuals that take out the part of wage growth predicted by
the current value of the Bartik instrument:
\[ \tilde{y}_{l1} = y_{l1} - \hat{\beta}\gamma\hat{B}_{l1}, \] (2.7)
and then regress the residualized current period wage growth on the future value of the
Bartik instrument:
\[ \tilde{y}_{l-1} = \pi_{0}^{\text{resid}} + \pi_{1}^{\text{resid}}\hat{B}_{lt} + \epsilon_{l-1}^{\text{resid}}. \] (2.8)
We interpret the statistical significance of the coefficient on the future value of the Bartik
instrument, \( \pi_{1}^{\text{resid}} \), as our test of pre-trends. This test has power against alternatives in
which Bartik is endogenous in a time-varying way: e.g., in period 1, the error term contains
a some function of the first \( \frac{K}{2} \) industries, and in period 2, the error term contains some
function of the last \( \frac{K}{2} \) industries. In Section 4 we conduct a Monte Carlo analysis of this test
and show that it has good finite sample properties in the sense that under the null of no
endogeneity, it rejects at the 5% level about 5% of the time.

**Empirical Test 3: Overidentification** Because the Bartik instrument implicitly leverages
a large number of instruments, the Bartik instrument is overidentified. Hence, we can
use an overidentification test to test the null hypothesis that all of the industry shares are
uncorrelated with innovations to supply shocks. Conceptually, the test asks whether the
instruments are correlated with the error term beyond what would be expected by chance.

The simplest way to implement an overidentification test would be to test the overiden-
tifying restrictions in terms of industry shares. Monte Carlo simulations, however, showed
that the Sargan-Hansen test was poorly behaved when using \( K \times T \) separate instruments,
since the industry shares are all individually quite weak as instruments. A simple way to
reduce the dimensionality of the instruments as well as improve their individual power is
to use a technique like principal components analysis (PCA) and use a subset of the most
predictive components as instruments. We show in Section 4 that in simulations, these PCA
components work well in the Sargan-Hansen test.
2.3 Dataset

We use U.S. Census Bureau data. In particular, we use the 5% sample of IPUMS (Ruggles et al. [2015]) for 1980, 1990 and 2000 and we pool the 2009-2011 ACSs for 2010. We look at PUMAs and 3-digit IND1990 industries. In the notation given above, our $y$ variable is earnings growth, and $x$ is employment growth. We use people aged 18 and older who report usually working at least 30 hours per week in the previous year. We fix industry shares at the 1980 values, and then construct the Bartik instrument using 1980 to 1990, 1990 to 2000 and 2000 to 2010 leave-one-out growth rates. To construct the industry growth rates, we weight by employment.

2.4 Instrument relevance

The relevance of the Bartik instrument relies on cross-industry variation in national growth rates. If there is no cross-industry variation, then the Bartik instrument is constant and the instrument is not relevant.

To quantify the size of the industry component, we decompose the variance of the industry-location growth rates into the industry, location and idiosyncratic industry-location components. Table [1] provides evidence that there is indeed a common industry component to growth rates. At the 3-digit level and using PUMAs we find that the industry component explains about 10% of the variance of growth rates.

2.5 Correlates of 1980 industry composition

Table [2] shows that 1980 industry composition is related to observable characteristics of places. In the table, we take the first principal component of the matrix of 1980 industry shares, which explains about 10% of the variance of industry shares, and regress it on a variety of observable characteristics measured in 1980.[8] The table shows that this component of the cross-sectional variation is tightly linked to observable characteristics: the $R^2$ of the regression is 0.77, and many observables are significantly related to industry composition. For example, the share of immigrants (or native share) is related to industry composition. Hence, if there are innovation to supply shocks that vary systematically with initial immigrant share (as predicted by the immigrant enclave instrument), then this would pose a problem for the validity of the instrument.

---

[7] There are 228 non-missing 3-digit IND1990 industries in 1980. There are 543 PUMAs. We have looked at other combinations of industry and geography detail, and results are broadly similar. These results are available upon request.

[8] Appendix Table A1 shows the variance share of the first 10 principal components. In Appendix Table A2 we report an analogous exercise where we regress the Bartik instrument constructed using various growth rates on 1980 characteristics.
2.6 Pre-trends

Column (1) of Table 3 shows that there is evidence of pre-trends. Namely, we pool wage growth from 1980 to 1990 and 1990 to 2000, and regress it on Bartik constructed one period forward (as well location and time fixed effects). Column (1) shows that we can predict past wage growth using future values of the instrument.

Some of this relationship might be mechanical: the correlation between adjacent period values of the Bartik instrument is 0.4. Hence, what might drive the correlation is that the future value of the instrument is correlated with the past good shocks that led to wage growth.

Columns (3) and (4) of Table 3 show that after addressing the mechanical reason for correlation, we still find evidence of pre-trends. Column (3) residualizes without controls, while column (4) includes the controls from Table 2 (interacted with time). Two aspects of the table are quantitatively notable. First, controlling for observable characteristics of the location does reduce the magnitude of the implied pre-trends. Second, even after this residualization, the size of the coefficient is economically large and statistically significant. For example, the reduced-form for the effect of Bartik on wage growth for the specification in column (4) is 0.35\(^9\). Hence, the coefficient in column (4) of 0.04 is large.

2.7 Parameter estimates

Despite the fact that we find evidence of pre-trends, we consider OLS and IV estimates with and without controlling for the 1980 covariates we discussed in Section 2.5.

Table 4 reports the results of these various exercises. The table makes two main points. First, the IV estimates are bigger than the OLS estimates. Second, the Bartik results are sensitive to the inclusion of controls. Adding controls moves the parameter estimate from 1.01 to 0.83 and these are statistically distinguishable.

We can use the model in Section 2.1 to interpret the difference between OLS and IV and the movements in IV when we add controls. The most compact explanation for why \( \hat{\beta}_{\text{OLS}} < \hat{\beta}_{\text{Bartik}} \) is that there are innovations in the supply shocks, since such innovations bias \( \hat{\beta}_{\text{OLS}} \) down. To interpret the effect of controls, suppose for simplicity that after including controls we get a consistent estimate of \( \beta \). Then before controls we have that \( \text{Cov}(d\eta_{lt+1} - d\eta_{lt}, \sum_k z_{ltk}(d\alpha_{kt} - d\alpha_{kt-1})) < 0 \). That is, observable characteristics of a place, such as the immigrant share, are correlated with negative innovations to the supply shocks.

\(^9\)This multiplies the coefficient in columns (4) and (6) in Table 4.
2.8 Overidentification test

In order to reduce the high-dimensionality of our instrument, we use the first four principal components of the 1980 industry share interacted with time dummies as our instrument set. We then have 8 (4 × 2) instruments for 2 endogenous variables (growth rates in each period after subtracting off one period for the location fixed effects) and report the Hansen J-statistic. The null hypothesis of the test is that all instruments are exogenous.

The bottom row of Table 5 reports the p-value of the Hansen J-statistic. We reject the null of the exogeneity of the complete instrument set with a p-value of 0.00. Hence, we conclude that the set of observable characteristics is not sufficient to control for all confounds.

3 Estimation

So far we have established that using the Bartik instrument is equivalent to using industry shares as instruments. Thus, what is distinctive about the Bartik instrument is not as an instrument per se, but as an estimation approach. Put differently, to call it the Bartik instrument is a slight misnomer, and in fact it could be called the Bartik approach.

This raises the basic question why researchers use Bartik as an instrument, rather than the industry shares as an instrument. In this section we offer two answers. First, as an estimation approach, it deals with problems arising from the fact that each individual industry share as an instrument is weak and so would result in poor performance of traditional TSLS estimators. Second, there are settings where the researcher is only interested in the reduced-form effect of an instrument, and is not able to observe the endogenous variable.

We also consider two additional estimation issues: first, we discuss feasible Bartik and the desirability of using a leave-one-out estimator; and, second, we show why a researcher should not update the shares in a panel setting.

---

10 The first four components explain about 20% of the variance of the industry-location shares. See Appendix Table A1. The p-values of our test statistic are identical when using anywhere from the first 2 to the first 10 principal components.

11 Relative to the parameter estimates in the previous section, our baseline parameter estimates move around because we are not using identical instruments. We continue to find, however, that parameter estimates attenuate when we control for observable characteristics.

12 An alternative interpretation of this result is that it is a rejection of the null of constant treatment effects and suggests the presence of treatment effect heterogeneity. Appendix B discusses what monotonicity means in this canonical setting and why it is unlikely to hold.
3.1 Why use Bartik rather than industry shares as instruments?

3.1.1 Bartik as dimension reduction

Our first explanation for why researchers use Bartik rather than industry shares as an instrument is that it is a form of dimension reduction. In the two-stage set-up, the generic first-stage estimation problem is to estimate the expectation of endogenous variable given the vector of instruments (\( \mathbb{E}[x_l|Z_l] \)), using a linear predictor:

\[
\mathbb{E}^*(x_l|Z_l) = Z_l'\pi. \tag{3.1}
\]

If a researcher estimates this \( \pi \) using least squares, then estimation may be very noisy and can lead to bias in the two-stage estimator.

Bartik uses the inner product structure of the endogenous variable to do dimension reduction. Given this structure, the first-stage estimation problem is to estimate the expectation of the industry-location growth rates given industry composition:

\[
\mathbb{E}[x_l|Z_l] = \mathbb{E}[Z_l'G_l|Z_l] = Z_l'E[G_l|Z_l].
\]

This estimation problem is also very high-dimensional. The Bartik instrument avoids it by using the national mean of the industry growth rates as the estimator for the expectation of the industry-location growth rates given industry composition: \( \hat{\mathbb{E}}[G_l|Z_l] = \mathbb{E}[G_l] = \bar{g}_k \). This reduces the \( K \) industry shares to an exactly identified instrument, which has desirable finite sample properties.

The Bartik approximation has a tight link to the first stage coefficient. When the Bartik assumption of \( \mathbb{E}[G_l|Z_l] = \mathbb{E}[G_l] \) is exact, then the first stage coefficient is 1.

**Proposition 3.1.** Suppose that \( G \) and \( Z \) are independent which implies that \( \mathbb{E}[G_l|Z_l] = \mathbb{E}[G_l] \). Then the expectation of the first stage coefficient is one.

**Proof.** See Appendix [A].

To see this more intuitively, rewrite the first-stage in terms of our decomposition of the \( g_{ik} \):

\[
x_l = B_l + \bar{g}_l + Z_l'\tilde{g}_{lk}.
\]

This equation says that employment growth is equal to the Bartik instrument and an error term. The error term contains two components: the location component of growth rates and weighted idiosyncratic industry-location growth rates. The first stage coefficient is one.
when the Bartik instrument is uncorrelated with the error term, which is implied by the Bartik approximation of $E[G_l|Z_l]$ holding exactly.

Substantively, this Bartik approximation assumes that the expectation of the growth rate of a particular industry does not depend on a location’s industry composition. However, it would be consistent with the identifying assumption of Bartik to allow for richer dependence. To see this, note that we can allow the idiosyncratic industry-location component of the growth rates to have rich dependence on industry composition. Formally, write $\tilde{g}_{lk} = f_k(Z_l) + \tilde{g}_{lk}$, where $f_k: [0,1]^K \to \mathbb{R}$ is an arbitrary industry-specific function that maps combinations of industry shares to growth rates. Then we expand our initial decomposition of the industry-location growth rates to include this component:

$$g_{lk} = g_k + f_k(Z_l) + g_l + \tilde{g}_{lk}.$$

If Bartik is a valid instrument, then the instrument would be more powerful if we estimated $f_k$ in the first stage. This $f_k$ function could capture any number of economic mechanisms. For example, we could allow for inter-industry spillovers through input-output linkages. Similarly, we could allow for the effects of a shock to depend on levels, i.e., it might be that there is curvature in the location-industry production function and a given shock has a larger (or smaller) effect in locations where the industry is more prominent.

### 3.1.2 Bartik as a proxy

Another reason to use the Bartik instrument rather than industry shares as instruments is that, in some cases, Bartik is used as a proxy in a setting where it would not be feasible to use the shares as instruments. This happens when the industry shares ($z_{lk}$) are observed, but the industry-location growth rates ($g_{lk}$) are unobserved. In this case, it is not possible to construct the endogenous variable, $x_l$, and hence not possible to run the first-stage regression. Nevertheless, because $\hat{g}_k$ is observed, it is possible construct the Bartik instrument and use it as a proxy. For example, in Appendix C we discuss how the “China Shock” instrument of Autor, Dorn, and Hanson (2013) has this structure. In this case, it is possible to estimate the reduced form effect of $B_l$ on the outcome of interest, $y_l$.

---

13Since this process constructs a generated instrument to be used in a just-identified two-stage estimation procedure, there is no impact on the asymptotic distribution of the estimator. (See Section 6.1.2 in Wooldridge [2002], pg. 117.) This implies that under standard asymptotics, any improvement in estimating $E[x_l|Z_l]$ by estimating $E[G_l|Z_l]$ more efficiently will only improve results in finite samples. In fact, it’s worth noting that the Bartik weight matrix $\Omega$ is of reduced rank (rank one), rather than full rank ($K - 1$). The asymptotically efficient weight matrix would be full rank.
3.2 Feasible Bartik and leave-one-out

In practice, researchers do not observe the true industry component of employment growth, \( g_k \), and so have to estimate it. Our results provide a formal justification for why—starting with Autor and Duggan (2003)—the literature has often used a leave-one-out estimator of the industry growth rates to construct the Bartik instrument. The motivation for doing leave-one-out is fundamentally a finite sample estimation issue and not an identification issue.

To see this, define the feasible estimator of the industry growth rates as the national mean of the industry-location growth rates in that industry.

\[
\hat{g}_k = \frac{1}{L} \sum_{l=1}^{L} g_{lk} = g_k + \frac{1}{L} \sum_{l=1}^{L} [g_l + \hat{g}_{lk}].
\]

The key point in this expression is that the location and idiosyncratic industry-location terms are, in expectation, mean zero and so as the number of locations grow the national mean converges to the infeasible value. Formally, under standard regularity conditions \( \text{plim}_{L \to \infty} \hat{g}_k = g_k \). Not surprisingly, this result means that as the number of locations grow the feasible Bartik instrument converges to the infeasible Bartik instrument and so in the limit the estimation error disappears. To write this more formally, let \( \hat{G} = (\hat{g}_1, \cdots, \hat{g}_K)' \) be the vector of estimated national growth rates. Then the constructed Bartik instrument is:

\[
\hat{B}_l = Z'_l \hat{G}.
\]

In finite samples, not doing leave-one-out can be a source of bias. In our notation, define the leave-one-out Bartik instrument as the feasible Bartik instrument less the own-location growth rate,

\[
\hat{B}_{lo} = \frac{L}{L-1} Z'_l \hat{G} - \frac{1}{L-1} x_l.
\]

The reason to use Bartik rather than OLS is the concern that the parts of \( g_{lk} \) that are not the industry component are correlated with the error term.\(^{15}\) Using simple means to compute \( \hat{g}_k \) adds these components of \( g_{lk} \) into estimation and thus reintroduces the bias that an instrument is designed to address. Because these components are multiplied by \( \frac{1}{L} \), we expect bias in estimation when \( L \) is not sufficiently large. The leave-one-out mean estimator solves this problem directly, and is analogous to the jackknife estimator in JIVE (Angrist,\(^{16}\))

---

\(^{14}\)In practice, sometimes the \( \hat{g}_k \) is computed by weighting by the \( z_{lk} \). Doing this weighting complicates the algebra without added insight.

\(^{15}\)To see this, return to the two-industry cross-sectional case. If \( E[x_i e_i] \neq 0 \) and \( E[B_i e_i] = 0 \), then \( E[z_{11} \hat{g}_{11} + z_{12} \hat{g}_{12} + \hat{g}_1 e_1 | x_1 ] \neq 0 \).
The need to estimate $g_k$ in feasible Bartik is related to the question of how finely to disaggregate the level of industries. We show two simple results relating to this decision, which suggest that the main consideration in choosing the level of aggregation should be ensuring that the industry growth rates are well-estimated. Later, in our simulation results in Section 4, we show that feasible Bartik has substantially worse behavior than infeasible Bartik when it is more difficult to estimate the vector of $g_k$.

Our two results on industry aggregation are as follows. First, if Bartik is valid at a finer level of industry aggregation, then it is also valid at a coarser level. Second, the converse is not necessarily true, but relies on knife-edge conditions.

We now introduce notation to define what we mean by a “coarse” and “fine” industry partition.

**Definition 3.1.** Let $Z^1_k$ and $Z^2_k$ be $K_1 \times 1$ and $K_2 \times 1$ vectors defining industry shares at different levels of aggregation ($K_1 < K_2$). $K_1$ is our “coarse” partition and $K_2$ is our “fine” partition. We assume that there exists a surjective function $f$ such that for each $k \in [1, \ldots, K_2]$, $f(k) \in [1, \ldots, K_1]$. Moreover, $\sum_{k \text{ s.t. } f(k) = j} z^2_{lk} = z^1_{lj}$.

We now show that if the fine partition is a valid instrument, then the coarse partition is as well.

**Proposition 3.2.** If $\mathbb{E}[z^2_{lk}e_l] = 0 \forall k \in [1, \ldots, K_2]$, then $\mathbb{E}[z^1_{f(k)l}e_l] = 0 \forall f(k) \in [1, \ldots, K_1]$.

**Proof.** Note that $\mathbb{E}[z^1_{lk}e_l] = \mathbb{E}[\sum_{j \text{ s.t. } f(j) = k} z^2_{lj}e_l] = \sum_{j \text{ s.t. } f(j) = k} \mathbb{E}[z^2_{lj}e_l] = 0$. 

However, it is not necessarily the case that if the coarse partition is a valid instrument that the fine partition is as well. Formally, the reason is that $\mathbb{E}[z^1_{lk}e_l] = 0$, $\forall k \in K_1$, does not necessarily imply that $\mathbb{E}[z^2_{lk}e_l] = 0 \forall k \in K_2$. The knife-edge case where this fails is if a particular set of covariances sums to zero across all $k \in [1, \ldots, K_1]$.

We interpret the knife-edgeness of this result as suggesting that in general estimation concerns, rather than identification concerns, should guide the choice of fineness of industry bins.

### 3.3 Updating shares in the panel

In many panel application, it is unclear which time period of industry shares to use to construct the Bartik instrument. We argue that in most settings, the industry shares should...
be fixed at the initial period in order to avoid endogeneity. To see the need for fixing industry shares, consider how industry shares update as a function of the growth rates. Let the initial period shares be $z_{lk0}$ and next period’s shares be $z_{lk1}$. For simplicity consider the two-industry case. Next period’s shares depend on this period’s growth rates ($g_{lk0}$) as follows:

$$z_{l11} = \frac{(1 + g_{l10})z_{l10}}{(1 + g_{l10})z_{l10} + (1 + g_{l20})z_{l20}}; \quad z_{l21} = \frac{(1 + g_{l20})z_{l20}}{(1 + g_{l10})z_{l10} + (1 + g_{l20})z_{l20}}.$$

If there is serial correlation in the error term, then the updated industry shares may be correlated with the error term, invalidating the identification assumption. Formally, suppose that $\epsilon_{lt} = \sum_k \tilde{g}_{lk,t-1} + \omega_{lt}$, where $\omega_{lt}$ is noise, then the current industry shares, $z_{lklt}$, may be correlated with the second-stage error term, $\epsilon_{lt}$. This potential for serial correlation motivates fixing industry shares to some initial period.

This adjustment does not come without a cost, however. For a panel with $T$ time periods with industry shares fixed at $t = 0$, the correlation between $Z_{l0}$ and $x_{lt}$ will typically be lower than that of $Z_{lt}$ and $x_{lt}$. Unfortunately, there is no way to identify “how early” to fix the shares—this choice will be a function of the serial correlation of the endogeneous component of $\epsilon_{lt}$, which is unobservable.

## 4 Simulations

Our simulations are broadly designed to mimic properties of industry growth rates and industry shares in the U.S. Specifically, we report simulations with 88 and 228 industries, where this reflects the number of non-missing IND1990 2 and 3 digit industries in 1980. Similarly, we use 50, 150 and 800 locations to correspond (loosely speaking) to states, MSAs and PUMAs. Note that (possibly unlike in the data) we assume that the location are independent. We also use the empirical distribution of industry shares at the various levels of aggregation. Finally, we anchor the properties of the industry-location growth rates to the U.S. data.\(^\underline{18}\)

In all simulations, we begin by randomly drawing industry shares for each location (with replacement) from the empirical distribution of industry shares, as well as $g_{lt}$, $\tilde{g}_{lk}$ and $g_{l}$ terms, which we assume to be normally distributed. We use these terms to construct an $x_{lt}$, and then draw a random error term $\epsilon_l$ for each location with mean zero and variance 0.0006. To create endogeneity in OLS, we add the $g_l$ term to $\epsilon_l$. In all simulations, the true value of $\beta$ is assumed to be 2.

\(^{18}\)The top panel of Table A4 reports the empirical variances of industry employment growth with states and PUMAs and 2 and 3 digit industries, while the bottom panel reports the overall variances of industry-location growth rates.
Table 6 reports the results of 3 regressions. In Column (1), we report OLS and see substantial bias. In Columns (2)-(4), we report results from using feasible Bartik, where we estimate Bartik using the leave-one-out empirical means for each industry. In Column (2), we report the average $\hat{\beta}$, in Column (3) we report the median of the estimated $\hat{\beta}$ and in Column (4), we report the average first-stage F-statistic. In Columns (5)-(7), we repeat the exercise, but construct infeasible Bartik using the true value of $g_k$ in the place of $\hat{g}_k$.

Table 6 makes four points. First, variation in industry growth rates is important for power. As we move down sub-blocks of the table and increase the variance of industry growth rates ($\sigma^2_k$), Bartik does better in the sense that we have more first-stage power and the medians and means converge. Second, increasing the number of locations is very important for the performance of Bartik. With 50 locations, on average Bartik is not a powerful instrument. With 150 locations Bartik is marginal at conventional levels of first-stage power. Only once we get to 800 locations does feasible Bartik perform well. Third, with feasible Bartik and a small number of locations, the fact that the industry growth rates are poorly estimated means that the estimator is poorly behaved. We can see this point by comparing the feasible and infeasible columns, where moving towards the infeasible column does marginally increases first-stage power and makes the median of the distribution better behaved. Fourth, over this range, increasing the number of industries does not make Bartik more effective (despite treating industries as independent).

Table 7 shows that our test for pre-trends is correctly sized. We make the Bartik instrument correlated between two time periods by setting the correlation in the industry growth rates between two time periods to 0.5. Columns (1) and (2) show that we can predict past values of wage growth using future values of the Bartik instrument—specifically, column (2) shows that we reject the null of “no pre-trends” at the 5% level at least 15% of the time. We then residualize. Columns (3) and (4) show that after residualization we reject the null of “no pre-trends” at the 5% level, typically around 5% of the time.

Table 8 shows evidence on the finite sample performance of the Sargan-Hansen tests. Column (1) shows rejection rates at the 5% under the null when we first do dimension reduction to four instruments using PCA. The rejection rates range from 5% to 8.4%, which implies that in samples of these sizes and for empirically plausible datasets, the test is correctly sized. In contrast, column (2) shows that when we use the industry shares directly as instruments the overidentification test is poorly behaved, with the rejection rates ranging from 0% to 35%.

$^{19}$Table A5 shows analogous results using infeasible Bartik.
5 Summary: recommendations for practice

This paper aims to demystify and formalize the Bartik instrument, and develop recommendations for practice. We summarize this paper by these recommendations:

- The instrument is the industry shares so the exogeneity condition is best stated in terms of industry shares.
- Consider correlates of industry shares, and consider controlling for these correlates.
- Look at pre-trends after partialling out the direct effects of previous values of Bartik.
- Overidentification test: reduce the dimensionality of industry shares using PCA, and then use two or more components as instruments and run a Sargan-Hansen test.
- Use the initial period shares – do not update.
- Use leave-one-out means.
References


Table 1: Instrument relevance: variance decomposition of industry-location growth rates

<table>
<thead>
<tr>
<th></th>
<th>Industry</th>
<th>Location</th>
<th>Residual</th>
</tr>
</thead>
<tbody>
<tr>
<td>1990</td>
<td>0.11</td>
<td>0.02</td>
<td>0.87</td>
</tr>
<tr>
<td>2000</td>
<td>0.11</td>
<td>0.02</td>
<td>0.87</td>
</tr>
<tr>
<td>2010</td>
<td>0.07</td>
<td>0.01</td>
<td>0.92</td>
</tr>
</tbody>
</table>

Notes: This table reports the results of a variance decomposition of the industry-location growth rates. The variance shares come from a regression of the form $g_{lk} = g_l + g_k + \tilde{g}_{lk}$, and the variance shares correspond to the covariance of the right hand side variables with the left hand side, divided by the variance of the left hand side variable.
Table 2: Relationship between industry shares and characteristics

<table>
<thead>
<tr>
<th>1980 characteristics</th>
<th>First principal component of 1980 industry shares</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>-0.32</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
</tr>
<tr>
<td>White</td>
<td>-0.26</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
</tr>
<tr>
<td>Native Born</td>
<td>-0.75</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
</tr>
<tr>
<td>12th Grade Only</td>
<td>0.14</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
</tr>
<tr>
<td>Some College</td>
<td>0.54</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
</tr>
<tr>
<td>Veteran</td>
<td>0.47</td>
</tr>
<tr>
<td></td>
<td>(0.08)</td>
</tr>
<tr>
<td># of Children</td>
<td>-0.13</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
</tr>
</tbody>
</table>

| R²                   | 0.77                                             |
| F                    | 201.53                                           |
| p                    | 0.00                                             |

*Notes:* This table reports a regression of the first principal component of 1980 industry shares on 1980 characteristics. Each characteristic is standardized to have unit standard deviation. The first principal component also has unit standard deviation. Standard errors in parentheses.
Table 3: Test for pre-trends

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>Residualized</th>
<th></th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
</tr>
<tr>
<td>Bartik (1980 shares)</td>
<td>0.11</td>
<td>0.02</td>
<td>0.12</td>
<td>0.04</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.04)</td>
<td>(0.01)</td>
<td>(0.01)</td>
</tr>
<tr>
<td>Year and Puma FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Controls</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1086</td>
<td>1086</td>
<td>1086</td>
<td>1086</td>
</tr>
</tbody>
</table>

Notes: This table reports regression results of wage growth on future values of the Bartik instrument (i.e., one observation is 1980 to 1990 wage growth, on the value of the Bartik instrument constructed using 1990 to 2000 employment growth). Column (1) shows the basic relationship, while column (2) shows the relationship after adding controls from Table 2 interacted with time. Column (3) uses wage growth that is residualized from the Bartik instrument following equation (2.7). Column (4) adds the controls. Standard errors are bootstrapped to take into account the estimation error in the residualization and are in parentheses.
Table 4: OLS and IV estimates

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>First Stage</th>
<th>Second Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td>Δ Emp</td>
<td>0.44 (0.03)</td>
<td>0.42 (0.03)</td>
<td>1.01 (0.08)</td>
</tr>
<tr>
<td>Bartik (1980 shares)</td>
<td>0.39 (0.03)</td>
<td>0.42 (0.03)</td>
<td></td>
</tr>
<tr>
<td>Controls</td>
<td>No</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Year and Puma FE</td>
<td>Yes</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1,629</td>
<td>1,629</td>
<td>1,629</td>
</tr>
<tr>
<td>Partial R²</td>
<td>0.88</td>
<td>0.91</td>
<td>0.36</td>
</tr>
<tr>
<td>Test of equality (p-value)</td>
<td>0.08</td>
<td>0.16</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Notes: This table reports OLS and TSLS estimates of the inverse elasticity of labor supply. The regression are at the PUMA level with 3 digit industry. The odd-numbered columns show estimates without controls and the even-numbered columns show estimates with controls. Column (1) and (2) show the OLS estimates. Columns (3) and (4) show the first stage. Columns (5) and (6) show the TSLS estimates. The $R^2$ is partial after absorbing location and time fixed effects. The p-value for the equality of coefficients compares the adjacent columns with and without controls. The controls are the 1980 characteristics (interacted with time) displayed in Table 2. See Table A3 for the coefficients on the controls. Standard errors are in parentheses.
Table 5: Overidentification tests

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>4 PCA Components</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Δ Wage</td>
<td>Δ Wage</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
</tr>
<tr>
<td>Δ Emp</td>
<td>0.44</td>
<td>0.42</td>
</tr>
<tr>
<td></td>
<td>(0.03)</td>
<td>(0.03)</td>
</tr>
<tr>
<td>Year and Puma FE</td>
<td>Yes</td>
<td>Yes</td>
</tr>
<tr>
<td>Controls</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1,629</td>
<td>1,629</td>
</tr>
<tr>
<td>Partial $R^2$</td>
<td>0.88</td>
<td>0.91</td>
</tr>
<tr>
<td>J-Statistic</td>
<td>97.62</td>
<td>76.81</td>
</tr>
<tr>
<td>(p-value)</td>
<td>0.00</td>
<td>0.00</td>
</tr>
</tbody>
</table>

Notes: This table reports OLS and TSLS estimates of the inverse elasticity of labor supply. Relative to Table 4, we have four instruments per time period. The four instruments are the first four principal components of 1980 industry shares. The controls in columns (2) and (4) are the 1980 characteristics (interacted with time) displayed in Table 2. Standard errors are in parentheses.
Table 6: Simulation results

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>Feasible Bartik</th>
<th>Infeasible Bartik</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>$\hat{E}(\hat{\beta})$</td>
<td>$\hat{E}(\hat{\beta})$</td>
<td>$\hat{E}(\hat{\beta})$</td>
</tr>
<tr>
<td>$\sigma^2_k = 0.0026$, $\sigma^2_l = 0.001$, $\sigma^2_{lk} = 0.0023$</td>
<td>$K = 88, L = 50$</td>
<td>2.84</td>
<td>0.32</td>
</tr>
<tr>
<td></td>
<td>$K = 88, L = 150$</td>
<td>2.83</td>
<td>1.55</td>
</tr>
<tr>
<td></td>
<td>$K = 88, L = 800$</td>
<td>2.84</td>
<td>1.95</td>
</tr>
<tr>
<td></td>
<td>$K = 228, L = 50$</td>
<td>2.87</td>
<td>0.37</td>
</tr>
<tr>
<td></td>
<td>$K = 228, L = 150$</td>
<td>2.86</td>
<td>4.27</td>
</tr>
<tr>
<td></td>
<td>$K = 228, L = 800$</td>
<td>2.86</td>
<td>1.95</td>
</tr>
<tr>
<td>$\sigma^2_k = 0.0038$, $\sigma^2_l = 0.001$, $\sigma^2_{lk} = 0.01$</td>
<td>$K = 88, L = 50$</td>
<td>2.61</td>
<td>7.24</td>
</tr>
<tr>
<td></td>
<td>$K = 88, L = 150$</td>
<td>2.58</td>
<td>1.49</td>
</tr>
<tr>
<td></td>
<td>$K = 88, L = 800$</td>
<td>2.60</td>
<td>1.98</td>
</tr>
<tr>
<td></td>
<td>$K = 228, L = 50$</td>
<td>2.65</td>
<td>1.62</td>
</tr>
<tr>
<td></td>
<td>$K = 228, L = 150$</td>
<td>2.67</td>
<td>2.41</td>
</tr>
<tr>
<td></td>
<td>$K = 228, L = 800$</td>
<td>2.67</td>
<td>1.96</td>
</tr>
<tr>
<td>$\sigma^2_k = 0.0046$, $\sigma^2_l = 0.001$, $\sigma^2_{lk} = 0$</td>
<td>$K = 88, L = 50$</td>
<td>2.90</td>
<td>1.40</td>
</tr>
<tr>
<td></td>
<td>$K = 88, L = 150$</td>
<td>2.91</td>
<td>1.86</td>
</tr>
<tr>
<td></td>
<td>$K = 88, L = 800$</td>
<td>2.91</td>
<td>1.98</td>
</tr>
<tr>
<td></td>
<td>$K = 228, L = 50$</td>
<td>2.91</td>
<td>2.17</td>
</tr>
<tr>
<td></td>
<td>$K = 228, L = 150$</td>
<td>2.93</td>
<td>1.70</td>
</tr>
<tr>
<td></td>
<td>$K = 228, L = 800$</td>
<td>2.92</td>
<td>1.97</td>
</tr>
</tbody>
</table>

Notes: This table reports simulation results. The true $\beta = 2$. In the feasible Bartik columns, the $g_k$ is estimated. In the infeasible column we use the true $g_k$. $\sigma^2_k$ is the variance of the industry growth rates, $\sigma^2_l$ is the variance of the location growth rates, and $\sigma^2_{lk}$ is the variance of the industry-location growth rates. See Section 4 for details.
Table 7: Simulation results: Feasible Bartik, pre-trends test

<table>
<thead>
<tr>
<th></th>
<th>$\hat{E}(\hat{\gamma}_{\text{noresid}})$</th>
<th>Reject</th>
<th>$\hat{E}(\hat{\gamma}_{\text{resid}})$</th>
<th>Reject</th>
<th>$\hat{E}(\hat{\beta}\hat{\gamma})$</th>
</tr>
</thead>
<tbody>
<tr>
<td><strong>(1)</strong></td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$\sigma_k^2 = 0.0026, \sigma_l^2 = 0.001, \sigma_{lk}^2 = 0.0023$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K = 88, L = 50$</td>
<td>-0.49</td>
<td>0.07</td>
<td>-0.79</td>
<td>0.09</td>
<td>-0.64</td>
</tr>
<tr>
<td>$K = 88, L = 150$</td>
<td>1.35</td>
<td>0.15</td>
<td>-0.13</td>
<td>0.06</td>
<td>1.26</td>
</tr>
<tr>
<td>$K = 88, L = 800$</td>
<td>1.88</td>
<td>0.62</td>
<td>-0.04</td>
<td>0.06</td>
<td>1.87</td>
</tr>
<tr>
<td>$K = 228, L = 50$</td>
<td>-0.02</td>
<td>0.08</td>
<td>-0.63</td>
<td>0.08</td>
<td>-0.02</td>
</tr>
<tr>
<td>$K = 228, L = 150$</td>
<td>1.53</td>
<td>0.17</td>
<td>-0.08</td>
<td>0.06</td>
<td>1.43</td>
</tr>
<tr>
<td>$K = 228, L = 800$</td>
<td>1.83</td>
<td>0.57</td>
<td>-0.06</td>
<td>0.06</td>
<td>1.88</td>
</tr>
<tr>
<td>$\sigma_k^2 = 0.0038, \sigma_l^2 = 0.001, \sigma_{lk}^2 = 0.01$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K = 88, L = 50$</td>
<td>0.59</td>
<td>0.08</td>
<td>-0.24</td>
<td>0.06</td>
<td>0.26</td>
</tr>
<tr>
<td>$K = 88, L = 150$</td>
<td>1.53</td>
<td>0.19</td>
<td>0.01</td>
<td>0.06</td>
<td>1.41</td>
</tr>
<tr>
<td>$K = 88, L = 800$</td>
<td>1.90</td>
<td>0.68</td>
<td>-0.01</td>
<td>0.04</td>
<td>1.91</td>
</tr>
<tr>
<td>$K = 228, L = 50$</td>
<td>0.21</td>
<td>0.08</td>
<td>-0.37</td>
<td>0.08</td>
<td>0.04</td>
</tr>
<tr>
<td>$K = 228, L = 150$</td>
<td>1.56</td>
<td>0.17</td>
<td>-0.05</td>
<td>0.05</td>
<td>1.48</td>
</tr>
<tr>
<td>$K = 228, L = 800$</td>
<td>1.92</td>
<td>0.68</td>
<td>0.01</td>
<td>0.05</td>
<td>1.98</td>
</tr>
<tr>
<td>$\sigma_k^2 = 0.0046, \sigma_l^2 = 0.001, \sigma_{lk}^2 = 0$</td>
<td></td>
<td></td>
<td></td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K = 88, L = 50$</td>
<td>1.13</td>
<td>0.13</td>
<td>-0.20</td>
<td>0.08</td>
<td>1.06</td>
</tr>
<tr>
<td>$K = 88, L = 150$</td>
<td>1.65</td>
<td>0.26</td>
<td>-0.14</td>
<td>0.05</td>
<td>1.66</td>
</tr>
<tr>
<td>$K = 88, L = 800$</td>
<td>1.97</td>
<td>0.82</td>
<td>-0.03</td>
<td>0.06</td>
<td>1.93</td>
</tr>
<tr>
<td>$K = 228, L = 50$</td>
<td>1.27</td>
<td>0.11</td>
<td>-0.14</td>
<td>0.06</td>
<td>1.18</td>
</tr>
<tr>
<td>$K = 228, L = 150$</td>
<td>1.54</td>
<td>0.21</td>
<td>-0.14</td>
<td>0.06</td>
<td>1.58</td>
</tr>
<tr>
<td>$K = 228, L = 800$</td>
<td>1.91</td>
<td>0.78</td>
<td>0.01</td>
<td>0.06</td>
<td>1.94</td>
</tr>
</tbody>
</table>

Notes: This table reports Monte Carlo simulations of the test for parallel pre-trends. The true $\gamma = 1$ and the true $\beta = 2$ so that the true $\beta\gamma = 2$. Column (1) reports the results of the test for pre-trends without residualizing (equation (2.5)). Column (2) reports the probability of rejecting that $\hat{\gamma}_{\text{noresid}} = 0$ at the 5% level of significance. Column (3) reports the results of the test for pre-trends after residualizing the earnings growth (equation (2.8)), while Column (4) reports the probability of rejecting that $\hat{\gamma}_{\text{resid}} = 0$ at the the 5% level of significance. Finally, Column (5) reports the results of the pooled reduced-form regression of earnings growth on Bartik (Equation 2.6). $K$ is the number of industries and $L$ is the number of locations. See Section 4 for details.
<table>
<thead>
<tr>
<th>Case</th>
<th>Reject (4 PCA)</th>
<th>Reject (Industry shares)</th>
</tr>
</thead>
<tbody>
<tr>
<td>$\sigma_k^2 = 0.0026$, $\sigma_l^2 = 0.001$, $\sigma_{lk}^2 = 0.0023$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K = 88, L = 150$</td>
<td>0.07</td>
<td>0.00</td>
</tr>
<tr>
<td>$K = 88, L = 800$</td>
<td>0.07</td>
<td>0.26</td>
</tr>
<tr>
<td>$K = 228, L = 800$</td>
<td>0.07</td>
<td>N/A</td>
</tr>
<tr>
<td>$\sigma_k^2 = 0.0038$, $\sigma_l^2 = 0.001$, $\sigma_{lk}^2 = 0.01$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K = 88, L = 150$</td>
<td>0.07</td>
<td>0.00</td>
</tr>
<tr>
<td>$K = 88, L = 800$</td>
<td>0.06</td>
<td>0.12</td>
</tr>
<tr>
<td>$K = 228, L = 800$</td>
<td>0.07</td>
<td>N/A</td>
</tr>
<tr>
<td>$\sigma_k^2 = 0.0046$, $\sigma_l^2 = 0.001$, $\sigma_{lk}^2 = 0$</td>
<td></td>
<td></td>
</tr>
<tr>
<td>$K = 88, L = 150$</td>
<td>0.05</td>
<td>0.00</td>
</tr>
<tr>
<td>$K = 88, L = 800$</td>
<td>0.07</td>
<td>0.35</td>
</tr>
<tr>
<td>$K = 228, L = 800$</td>
<td>0.08</td>
<td>N/A</td>
</tr>
</tbody>
</table>

Notes: This table reports Monte Carlo simulations of the overidentification tests. The columns report the rejection rates of Sargan-Hansen overidentification tests at the 5% level of significance under the null that the industry shares are exogenous. Column (1) takes the first 4 components from principal component analysis, while column (2) uses the industry shares as instruments directly. See Section 4 for details.
A Omitted proofs

Equivalence with $K$ industries, $T$ locations, and controls

The two stage least squares system of equations is:

$$y_{lt} = W_{lt} \alpha + x_{lt} \beta + \epsilon_{lt}$$  \hspace{1cm} (A1)

$$x_{lt} = W_{lt} \tau + B_{lt} \gamma + \eta_{lt}$$ \hspace{1cm} (A2)

where $W_{lt}$ is a $1 \times L$ vector of controls. Typically in a panel context, $W_{lt}$ will include location and year fixed effects, while in the cross-sectional regression, this will simply include a constant. It may also include a variety of other variables. Let $n = L \times T$, the number of location-years. For simplicity, let $Y$ denote the $n \times 1$ stacked vector of $y_{lt}$, $W$ denote the $n \times L$ stacked vector of $W_{lt}$ controls, $X$ denote the $n \times 1$ stacked vector of $x_{lt}$, $G$ the stacked $K \times T$ vector of the $g_{kt}$, and $B$ denote the stacked vector of $B_{lt}$. Denote $P_W = W(W'W)^{-1}W'$ as the $n \times n$ projection matrix of $W$, and $M_W = I_n - P_W$ as the annhilator matrix. Then, because this is an exactly identified instrumental variable our estimator is

$$\hat{\beta}_{\text{Bartik}} = \frac{(M_W B)' Y}{(M_W B)' X}.$$ \hspace{1cm} (A3)

We now consider the alternative approach of using industry shares as instruments. The two-equation system is:

$$y_{lt} = W_{lt} \alpha + x_{lt} \beta + \epsilon_{lt}$$ \hspace{1cm} (A4)

$$x_{lt} = W_{lt} \tau + Z_{lt} \gamma_t + \eta_{lt},$$ \hspace{1cm} (A5)

where $Z_{lt}$ is a $1 \times K$ row vector of industry shares, and $\gamma_t$ is a $K \times 1$ vector, and, reflecting the lessons of previous section, the $t$ subscript allows the effect of a given industry share to be time-varying. In matrix notation, we write

$$Y = W\alpha + X\beta + \epsilon$$ \hspace{1cm} (A6)

$$X = W\tau + Z\Gamma + \eta,$$ \hspace{1cm} (A7)

where $\Gamma$ is a stacked $1 \times (T \times K)$ row vector such that

$$\Gamma = [\gamma_1 \cdots \gamma_T],$$ \hspace{1cm} (A8)
and \( \tilde{Z} \) is a stacked \( n \times (T \times K) \) matrix such that

\[
\tilde{Z} = \begin{bmatrix} Z \odot 1_{t=1} & \cdots & Z \odot 1_{t=T} \end{bmatrix},
\]

(A9)

where \( 1_{t=t'} \) is an \( n \times K \) indicator matrix equal to one if the \( n \)th observation is in period \( t' \), and zero otherwise. \( \odot \) indicates the Hadamard product, or pointwise product of the two matrices. Let \( Z^\perp = M_W \tilde{Z} \) and \( P_{Z^\perp} = \tilde{Z}^\perp (\tilde{Z}^{\perp \prime} \tilde{Z}^{\perp})^{-1} \tilde{Z}^{\perp \prime} \). Then, the TSLS estimator is

\[
\hat{\beta}_{TSLS} = \frac{X'P_{Z^\perp}Y}{X'P_{Z^\perp}X},
\]

(A10)

Alternatively, using the \( Z \) as instruments, the GMM estimator is:

\[
\hat{\beta}_{GMM} = \frac{X'M_W \tilde{Z} \Omega \tilde{Z}'M_W Y}{X'M_W \tilde{Z} \Omega \tilde{Z}'M_W X'},
\]

where \( \Omega \) is a \((K \times T) \times (K \times T)\) weight matrix.

**Proposition A.1.** If \( \Omega = G'G \), then \( \hat{\beta}_{GMM} = \hat{\beta}_{Bartik} \).

**Proof.** Start with the Bartik estimator,

\[
\hat{\beta}_{Bartik} = \frac{(M_W B)'Y}{(M_W B)'X}
\]

(A12)

\[
= \frac{B'M_W Y}{B'M_W X}
\]

(A13)

\[
= \frac{G\tilde{Z}'M_W Y}{G\tilde{Z}'M_W X}
\]

(A14)

\[
= \frac{X'M_W \tilde{Z}G'G\tilde{Z}'M_W Y}{X'M_W \tilde{Z}G'G\tilde{Z}'M_W X'}
\]

(A15)

where the second equality is algebra, the third equality follows from the definition of \( B \), and the fourth equality follows because \( X'M_W \tilde{Z}G' \) is a scalar. By inspection, if \( \Omega = G'G \), then \( \hat{\beta}_{GMM} = \hat{\beta}_{Bartik} \).

**Proposition 3.1**

**Proof.** Note that we can write \( G_i = G + \tilde{G}_i \) where, for example, \( \tilde{G}_i \) is a \( K \times 1 \) vector made up of \( \tilde{g}_{ik} \). Similarly, \( B_i = Z_i'G \). Hence, we can write:

\[
Var(X_i) = Var(Z_i'G_i) = Var(Z_i'G + Z_i'\tilde{G})
\]

(A16)

\[
\quad = Var(Z_i'G) + 2Cov(Z_i'G, Z_i'\tilde{G}) + Var(Z_i'\tilde{G}),
\]

(A17)
In the limit as the number of locations grows large, the first stage coefficient is:

\[
\gamma = \frac{\text{Cov}(\hat{B}_l, X_l)}{\text{Var}(\hat{B}_l)} = 1 + \frac{\text{Cov}(Z'_l G, Z'_l \hat{G})}{\text{Var}(Z'_l G)}.
\] (A18)

Hence, whether the first stage coefficient is 1 depends on the properties of \(\text{Cov}(Z'_l G, Z'_l \hat{G})\).

One sufficient condition for \(\text{Cov}(Z'_l G, Z'_l g_{lk}) = 0\) is that \(\mathbb{E}[G_l | Z_l] = \mathbb{E}[G_l]\).

\[
\text{Cov}(Z'_l G, Z'_l \hat{G}_l) = \mathbb{E}[Z'_l G Z'_l \hat{G}_l] - \mathbb{E}[Z'_l G] \mathbb{E}[Z'_l \hat{G}_l]
\] (A19)
\[
= \mathbb{E}[Z'_l G (G_l - G)] - \mathbb{E}[Z'_l G] \mathbb{E}[Z'_l (G_l - G)]
\] (A20)
\[
= \mathbb{E}[Z'_l G (G_l - G)' Z_l] - \mathbb{E}[Z'_l G] \mathbb{E}[Z'_l (G_l - G)]
\] (A21)
\[
= \mathbb{E}[Z'_l \mathbb{E}[G (G_l - G)' | Z_l] Z_l] - \mathbb{E}[Z'_l \mathbb{E}[G | Z_l]] \mathbb{E}[Z'_l \mathbb{E}[(G_l - G) | Z_l]]
\] (A22)
\[
= \mathbb{E}[Z'_l \mathbb{E}[G (G_l - G)' | Z_l] - \mathbb{E}[Z'_l] \mathbb{E}[G] \mathbb{E}[Z'_l] \mathbb{E}[(G_l - G)]]
\] (A23)
\[
= 0.
\] (A24)

The first line is the definition of covariance, the second line is the definition of \(g_{lk}\), the third line takes the transpose of a scalar, the fourth line is the law of iterated expectations, the fifth line is the assumption that \(G\) and \(Z\) are independent, and the sixth follows from the fact that \(\mathbb{E}[G_l - G] = 0\) and \(\text{Cov}(G, G_l - G) = 0\). \(\square\)
B Monotonicity

Throughout our results, we maintain the assumption of constant treatment effects. Here we briefly consider an extension to heterogeneous treatment effects, and discuss why the monotonicity condition seems unlikely to hold in this setting and hence why it makes sense to restrict attention to constant treatment effects.

For simplicity of exposition, we focus on the cross-sectional case with two industries and a discrete set of possible values for employment growth. The equation we are now interested in estimating is:

$$y_l = \alpha + \beta_l x_l + \epsilon_l,$$

where the difference from the main text is that now the parameter of interest, the inverse elasticity of labor supply, has an $l$ subscript. The definition of $x_l$ is as follows (we omit $g_l$ for simplicities’ sake):

$$x_l \equiv z_{l1}(g_1 + g_{l1}^{0.25}) + z_{l2}(g_2 + g_{l2}^{0.25}),$$

where $z_{l1}$ only takes on two values: $z_{l1} \in \{0.25, 0.75\}$, and $g_{l1}^{0.25}, g_{l2}^{0.25} \in \{g, \bar{g}\}$. Hence, we can write the endogenous variable in terms of values of the treatment:

$$x_l(z_{l1} = 0.25) = \left[0.25(g_1 + g_{l1}^{0.25}) + 0.75(g_2 + g_{l2}^{0.25})\right]$$

and

$$x_l(z_{l1} = 0.75) = \left[0.75(g_1 + g_{l1}^{0.75}) + 0.25(g_2 + g_{l2}^{0.75})\right],$$

and thus there are sixteen unique combinations of $(z_{l1}, g_{l1}^{0.25}, g_{l2}^{0.25})$.

In this setting, monotonicity means that one of the following conditions holds:

$$x_l(0.25) \geq x_l(0.75) \forall l, \text{ or } x_l(0.75) \geq x_l(0.25) \forall l.$$  

Consider the difference between these two values of the endogenous variable:

$$x_l(z_{l1} = 0.75) - x_l(z_{l1} = 0.25)$$

$$= \left[0.75(g_1 + g_{l1}^{0.75}) + 0.25(g_2 + g_{l2}^{0.75})\right] - \left[0.25(g_1 + g_{l1}^{0.25}) + 0.75(g_2 + g_{l2}^{0.25})\right]$$

$$= 0.5(g_1 - g_2) + 0.75(g_{l1}^{0.75} - g_{l2}^{0.25}) + 0.25(g_{l2}^{0.75} - g_{l1}^{0.25}).$$

To see what restrictions monotonicity imposes, let us focus on the case where $x_l(0.75) \geq$
\( x_l(0.25) \forall l \). In this case, note that \( g_1 - g_2 \geq 0 \). The sharpest inequality imposed by monotonicity comes from setting \( g_{0.75}^l - g_{0.25}^l = g_{0.75}^2 - g_{0.25}^2 = \bar{\delta} - \bar{\delta} \). This case implies that \( g_1 - g_2 \geq 2(\bar{\delta} - \bar{\delta}) \). Intuitively, monotonicity requires that the variation in the idiosyncratic industry-location component of the growth rates is “small” relative to the variation in the common industry component. While the evidence in Table 1 suggests the presence of a common industry component, it does not suggest that it is the dominant source of variation. Hence, we view monotonicity as unlikely to hold, which motivates why we focus on the constant treatment effects setting.

\footnote{To derive this, set \( g_{0.75}^l = g_{0.25}^2 \) and \( g_{0.75}^2 = g_{0.25}^l \).}
C Instruments encompassed by our structure

We now discuss six other instruments that our encompassed by our framework. This list cannot be exhaustive, but illustrates the widespread applicability of our results.

C.1 China shock

Autor, Dorn, and Hanson (2013) are interested in understanding the effect of the increase in imports from China on local labor market outcomes. In our notation, the authors’ endogenous variable of interest, $x_l$, is the commuting zone-specific increase in worker exposure to imports from China. This can written in inner product form, where $Z_l$ is commuting-zone specific industry composition (lagged by several years). Ideally, the authors would observe $g_{lk}$, or commuting-zone and industry specific growth in import competition from China. That is, ideally it would be possible to map the rise of the import of each good from China to the local labor market. Instead, they observe $g_k$, or the national growth in imports.

The problem of not observing $g_{lk}$ is that $g_k$ might contain import growth that reflects supply shocks in location $l$. To get around the problem that they do not observe the $g_{lk}$ (and thus cannot do a leave-one-out estimator of $\hat{g}_{lk}$), they use growth in imports to other high income countries from China.

C.2 Immigrant enclave instrument

Altonji and Card (1991) are interested in the effects of immigration on native wages, but are concerned that the correlation between immigrant inflows and local economic conditions may confound their estimates. To fit our notation, let $x_l$ denote the number of newly arriving immigrants in location $l$ in a given interval. Let $k$ denote one of $K$ countries of origin and let $z_{lk}$ denote the share of people arriving from origin country $k$ living in location $l$. Hence, $\sum_{l=1}^{L} z_{lk} = 1$, $\forall k$. In contrast, in the industry-location setting it is the sum over $k$ that sums to one. Let $g_k$ denote the number of people arriving from origin $k$. The instrument comes from lagging the $z_{lk}$. Once we lag $z$, say $z_{lk0}$ in some initial period, then let $i_{lk}$ be the number of immigrants from origin country $k$ arriving in destination $l$. Then define $g_{lk} = \frac{i_{lk}}{z_{lk0}}$ to be the hypothetical flow of immigrants from $k$ that would have to have occurred to have generated the extent of flows; this allows us to write $x_l = \sum_k z_{lk0} g_{lk}$. Then rather than using the $g_{lk}$ that makes this an identity, the researcher uses $g_k = \sum_l i_{lk} = \sum_l g_{lk} z_{lk0}$. (This is analogous to in the industry-location setting weighting the $g_{lk}$ by the $z_{lk}$ to compute the $g_k$, rather than equal-weighting across locations). This emphasizes that researchers should be using leave-one-out to compute the $g_k$. 

38
C.3 Bank lending relationships

Greenstone, Mas, and Nguyen (2015) are interested in the effects of changes in bank lending on economic activity during the Great Recession. They observe county-level outcomes and loan origination by bank to each county. In our notation, let \( x_l \) be credit growth in a county, let \( z_{lk} \) be the share of loan origination in county \( l \) from bank \( k \) in some initial period, and let \( g_{lk} \) be the growth in loan origination in county \( l \) by bank \( k \) over some period. Then

\[
x_l = \sum_k z_{lk} g_{lk}.
\]

The most straightforward Bartik estimator would compute \( \hat{g}_{lk} - \frac{1}{L-1} \sum_{l' \neq l} g_{lk} \). However, Greenstone, Mas, and Nguyen (2015) are concerned that there is spatial correlation in the economic shocks and so leave-one-out is not enough to remove mechanical correlations. One approach would be to instead leave out regions. Instead, they pursue a generalization of this approach and regress:

\[
g_{lk} = g_l + g_k + \epsilon_{lk}, \tag{A1}
\]

where the \( g_l \) and \( g_k \) are indicator variables for location and bank. Then the \( g_l \) captures the change in bank lending that is common to a county, while \( g_k \) captures the change in bank lending that is common to a bank. To construct their instrument, they use \( \hat{B}_l = \sum_k z_{lk} \hat{g}_k \), where the \( \hat{g}_k \) comes from equation (A1).

C.4 Simulated instruments

Currie and Gruber (1996a) are interested in the effect of Medicaid eligibility on infant mortality and low birth weight. To see this instrument as a Bartik-like instrument, let \( x_l \) be the share of the population in state \( l \) eligible for Medicaid, and let \( k \) index one of \( K \) discrete types, where the types are defined by eligibility criterion of the policy. \( Z_l \) is the Medicaid eligibility rules in state \( l \) for each type. Finally, \( G_l \) is the share of the population in each state that is of each type. In this case, for \( \hat{E}[G|Z] \), Currie and Gruber use the national shares of each type of person. (Note that their estimation is in a panel with state and time fixed effects, so in practice they exploit changes in \( Z \)).

There are two differences from the Bartik setting. First, future values of \( Z \) are not mechanically related to past values of \( G \), and so they update the weights by allowing for changes in policy to generate variation in the instrument over time. Second, the \( Z \)'s are not shares and do not sum to one. This means that even if the \( G \) were constant—i.e, \( g_{lk} = \frac{1}{|K|} \)—then the instrument could still have variation and be relevant.
C.5 Market size and demography

Acemoglu and Linn (2004) are interested in the effects of market size on innovation. Naturally, the concern is that the size of the market reflects both supply and demand factors: a good drug will increase consumption of that drug. To construct an instrument, their basic observation is that there is an age structure to demand for different types of pharmaceuticals and there are large shifts in the age structure in the U.S. in any sample. They use this observation to construct an instrument for the change in market size.

In our notation, \( z_{lk} \) is the share of spending on drug category \( l \) that comes from age group \( k \). Hence, \( \sum_k z_{lk} = 1 \). Then \( g_{lk} \) is the growth in spending of age group \( k \) on drug category \( l \). Hence, \( x_l = \sum_k z_{lk} g_{lk} \). To construct an instrument, they use the fact that there are large shifts in the age distribution. Hence, they estimate \( \hat{g}_k \) as the increase in the number of people in age group \( k \), and sometimes as the total income (people times incomes) in age group \( k \). This is similar to the “China shock” setting where for both conceptual and data limitation issues \( g_{lk} \) is fundamentally unobserved and so the researcher constructs \( \hat{g}_k \) using other information.

C.6 The “judge” instrument

Kling (2006) is interested in the impact of incarceration length on outcomes. He is concerned that incarceration length reflects characteristics of the person (or the crime). One of his research designs relies on the random assignment of people to judges as an instrument for incarceration length.

In our notation, \( z_{lk} \) is an indicator variable for whether person \( l \) is assigned to judge \( k \). Then \( g_{lk} \) is the sentence when person \( l \) is assigned to judge \( k \). That is, \( g_{lk} \) captures potential outcomes. Letting \( x_l \) be the realized sentence length for person \( l \), we can write our endogenous variable in inner product form as follows: \( x_l = \sum_k z_{lk} g_{lk} \). To construct the instrument, the most straightforward approach is to compute the leave-one-out estimator of the \( \hat{g} \): \( \hat{g}_{l \neq i, k} = \frac{1}{1 - \sum_l z_{lk}} \sum_{l' \neq l} \text{s.t. } z_{l'k} = 1 g_{l'k} \), where this notation recognizes that \( g_{lk} \) is only observed when person \( l \) is assigned to judge \( k \). Then the instrument is: \( \hat{B}_l = \sum_k z_{lk} \hat{g}_{l \neq i, k} \).

\[\text{For expositional purposes, we simplify along two dimensions relative to Kling (2006): first, it is cases rather than people that are randomly assigned to judges. Second, there is only random assignment conditional on covariates.}\]
### D Additional tables

Table A1: Principal components of 1980 industry shares

<table>
<thead>
<tr>
<th></th>
<th>Component Share Explained</th>
<th></th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Marginal (1)</td>
<td>Cumulative (2)</td>
</tr>
<tr>
<td>1</td>
<td>0.091</td>
<td>0.091</td>
</tr>
<tr>
<td>2</td>
<td>0.055</td>
<td>0.145</td>
</tr>
<tr>
<td>3</td>
<td>0.036</td>
<td>0.181</td>
</tr>
<tr>
<td>4</td>
<td>0.029</td>
<td>0.210</td>
</tr>
<tr>
<td>5</td>
<td>0.025</td>
<td>0.234</td>
</tr>
<tr>
<td>6</td>
<td>0.021</td>
<td>0.256</td>
</tr>
<tr>
<td>7</td>
<td>0.021</td>
<td>0.277</td>
</tr>
<tr>
<td>8</td>
<td>0.018</td>
<td>0.295</td>
</tr>
<tr>
<td>9</td>
<td>0.016</td>
<td>0.311</td>
</tr>
<tr>
<td>10</td>
<td>0.016</td>
<td>0.327</td>
</tr>
</tbody>
</table>

*Notes:* This table reports the marginal and cumulative shares of the first ten principal components of the 1980 industry shares.
Table A2: Relationship between industry shares and characteristics

<table>
<thead>
<tr>
<th>Characteristic</th>
<th>1990</th>
<th>2000</th>
<th>2010</th>
</tr>
</thead>
<tbody>
<tr>
<td>Male</td>
<td>-0.20</td>
<td>-0.49</td>
<td>0.40</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.03)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>White</td>
<td>-0.12</td>
<td>0.04</td>
<td>-0.20</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.03)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>Native Born</td>
<td>-0.32</td>
<td>-0.20</td>
<td>0.01</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.04)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>12th Grade Only</td>
<td>0.06</td>
<td>0.32</td>
<td>0.00</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.06)</td>
<td>(0.09)</td>
</tr>
<tr>
<td>Some College</td>
<td>0.69</td>
<td>0.62</td>
<td>0.44</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.08)</td>
</tr>
<tr>
<td>Veteran</td>
<td>0.11</td>
<td>0.37</td>
<td>-0.28</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.06)</td>
<td>(0.11)</td>
</tr>
<tr>
<td># of Children</td>
<td>-0.05</td>
<td>0.14</td>
<td>-0.06</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.03)</td>
<td>(0.05)</td>
</tr>
</tbody>
</table>

R²       | 0.64  | 0.54  | 0.25  |
F        | 182.42| 115.66| 16.98 |
p        | 0.00  | 0.00  | 0.00  |

Notes: This table reports a regression of 3 values of the Bartik instrument on 1980 characteristics. Each characteristic is standardized to have unit standard deviation. Standard errors are in parentheses.
Table A3: Coefficients on controls corresponding to Table 4

<table>
<thead>
<tr>
<th></th>
<th>OLS</th>
<th>First Stage</th>
<th>Second Stage</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Δ Wage</td>
<td>Δ Wage</td>
<td>Δ Emp</td>
</tr>
<tr>
<td></td>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
</tr>
<tr>
<td>2000 × Male (1980)</td>
<td>0.17</td>
<td>0.23</td>
<td>0.12</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.10)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>2010 × Male (1980)</td>
<td>0.28</td>
<td>0.01</td>
<td>0.17</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.08)</td>
<td>(0.05)</td>
</tr>
<tr>
<td>2000 × White (1980)</td>
<td>-0.04</td>
<td>0.05</td>
<td>-0.08</td>
</tr>
<tr>
<td></td>
<td>(0.04)</td>
<td>(0.05)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>2010 × White (1980)</td>
<td>-0.02</td>
<td>-0.06</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(0.05)</td>
<td>(0.05)</td>
<td>(0.04)</td>
</tr>
<tr>
<td>2000 × Native Born (1980)</td>
<td>0.26</td>
<td>0.38</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.07)</td>
<td>(0.07)</td>
</tr>
<tr>
<td>2010 × Native Born (1980)</td>
<td>0.37</td>
<td>-0.27</td>
<td>0.43</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.06)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>2000 × 12th Grade Only (1980)</td>
<td>0.05</td>
<td>-0.03</td>
<td>0.02</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.08)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>2010 × 12th Grade Only (1980)</td>
<td>0.05</td>
<td>-0.05</td>
<td>0.08</td>
</tr>
<tr>
<td></td>
<td>(0.07)</td>
<td>(0.08)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>2000 × Some College (1980)</td>
<td>0.07</td>
<td>-0.06</td>
<td>0.11</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.08)</td>
<td>(0.06)</td>
</tr>
<tr>
<td>2010 × Some College (1980)</td>
<td>0.00</td>
<td>-0.19</td>
<td>0.13</td>
</tr>
<tr>
<td>----------------------</td>
<td>-----------------------</td>
<td>-----------------------</td>
<td>-----------------------------</td>
</tr>
<tr>
<td></td>
<td>-0.05</td>
<td>-0.08</td>
<td>0.07</td>
</tr>
<tr>
<td></td>
<td>(0.09)</td>
<td>(0.10)</td>
<td>(0.04)</td>
</tr>
<tr>
<td></td>
<td>-0.48</td>
<td>-0.24</td>
<td>0.09</td>
</tr>
<tr>
<td></td>
<td>(0.16)</td>
<td>(0.13)</td>
<td>(0.07)</td>
</tr>
<tr>
<td></td>
<td>0.11</td>
<td>0.09</td>
<td>-0.00</td>
</tr>
<tr>
<td></td>
<td>(0.06)</td>
<td>(0.09)</td>
<td>(0.04)</td>
</tr>
</tbody>
</table>

<table>
<thead>
<tr>
<th></th>
<th>Year and Puma FE</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
<th>Yes</th>
</tr>
</thead>
<tbody>
<tr>
<td></td>
<td>Controls</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
<td>No</td>
<td>Yes</td>
</tr>
<tr>
<td>Observations</td>
<td>1,629</td>
<td>1,629</td>
<td>1,629</td>
<td>1,629</td>
<td>1,629</td>
<td>1,629</td>
<td></td>
</tr>
</tbody>
</table>

*Notes: See notes to Table 4.*
### Table A4: Growth Summary Statistics results

<table>
<thead>
<tr>
<th></th>
<th>Mean</th>
<th>Variance</th>
</tr>
</thead>
<tbody>
<tr>
<td>Wage Growth</td>
<td>0.0006</td>
<td></td>
</tr>
<tr>
<td><strong>Panel A: Industry:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State 2 Digit Emp. Growth</td>
<td>0.0334</td>
<td>0.0026</td>
</tr>
<tr>
<td>State 3 Digit Emp. Growth</td>
<td>0.0370</td>
<td>0.0038</td>
</tr>
<tr>
<td>Puma 2 Digit Emp. Growth</td>
<td>0.0497</td>
<td>0.0040</td>
</tr>
<tr>
<td>Puma 3 Digit Emp. Growth</td>
<td>0.0443</td>
<td>0.0044</td>
</tr>
<tr>
<td><strong>Panel B: Pooled:</strong></td>
<td></td>
<td></td>
</tr>
<tr>
<td>State 2 Digit Emp. Growth</td>
<td>0.0334</td>
<td>0.0059</td>
</tr>
<tr>
<td>State 3 Digit Emp. Growth</td>
<td>0.0372</td>
<td>0.0141</td>
</tr>
<tr>
<td>Puma 2 Digit Emp. Growth</td>
<td>0.0501</td>
<td>0.0235</td>
</tr>
<tr>
<td>Puma 3 Digit Emp. Growth</td>
<td>0.0472</td>
<td>0.0364</td>
</tr>
</tbody>
</table>

**Notes:** This table reports a variance decomposition of industry and industry-location growth rates. Panel A reports the means and variances of industry growth rates when these are constructed at various levels of aggregation. Panel B reports the means and variance of industry-location growth rates constructed at various levels of aggregation.
Table A5: Simulation results: Infeasible Bartik, Pre-Test

<table>
<thead>
<tr>
<th>( \hat{E}(\hat{\pi}_{\text{noresid}}) )</th>
<th>Reject</th>
<th>( \hat{E}(\hat{\pi}_{\text{resid}}) )</th>
<th>Reject</th>
<th>( \hat{E}(\hat{\beta\gamma}) )</th>
</tr>
</thead>
<tbody>
<tr>
<td>(1)</td>
<td>(2)</td>
<td>(3)</td>
<td>(4)</td>
<td>(5)</td>
</tr>
</tbody>
</table>

\( \sigma_k^2 = 0.0026, \sigma_l^2 = 0.001, \sigma_{lk}^2 = 0.0023 \)
- \( K = 88, L = 50 \)
  - 2.01
  - 0.11
  - -0.00
  - 0.06
  - 2.03
- \( K = 88, L = 150 \)
  - 1.99
  - 0.21
  - 0.02
  - 0.07
  - 1.95
- \( K = 88, L = 800 \)
  - 1.98
  - 0.65
  - -0.02
  - 0.07
  - 1.99
- \( K = 228, L = 50 \)
  - 1.79
  - 0.11
  - -0.14
  - 0.06
  - 1.88
- \( K = 228, L = 150 \)
  - 2.15
  - 0.23
  - 0.04
  - 0.06
  - 2.08
- \( K = 228, L = 800 \)
  - 1.94
  - 0.61
  - -0.04
  - 0.07
  - 2.00

\( \sigma_k^2 = 0.0038, \sigma_l^2 = 0.001, \sigma_{lk}^2 = 0.01 \)
- \( K = 88, L = 50 \)
  - 1.87
  - 0.13
  - -0.07
  - 0.06
  - 2.02
- \( K = 88, L = 150 \)
  - 1.97
  - 0.25
  - 0.03
  - 0.06
  - 1.95
- \( K = 88, L = 800 \)
  - 1.99
  - 0.72
  - 0.00
  - 0.04
  - 2.02
- \( K = 228, L = 50 \)
  - 1.79
  - 0.10
  - -0.11
  - 0.06
  - 1.89
- \( K = 228, L = 150 \)
  - 2.00
  - 0.24
  - -0.04
  - 0.06
  - 2.03
- \( K = 228, L = 800 \)
  - 2.01
  - 0.71
  - 0.01
  - 0.05
  - 1.98

\( \sigma_k^2 = 0.0046, \sigma_l^2 = 0.001, \sigma_{lk}^2 = 0 \)
- \( K = 88, L = 50 \)
  - 1.95
  - 0.2
  - 0.05
  - 0.07
  - 1.88
- \( K = 88, L = 150 \)
  - 1.98
  - 0.33
  - -0.05
  - 0.05
  - 1.98
- \( K = 88, L = 800 \)
  - 2.02
  - 0.84
  - -0.01
  - 0.06
  - 1.99
- \( K = 228, L = 50 \)
  - 2.15
  - 0.19
  - 0.09
  - 0.06
  - 2.05
- \( K = 228, L = 150 \)
  - 1.92
  - 0.28
  - -0.04
  - 0.06
  - 1.96
- \( K = 228, L = 800 \)
  - 1.97
  - 0.80
  - 0.02
  - 0.06
  - 2.00

Notes: The true \( \gamma = 1 \) and the true \( \beta = 2 \) so that the true \( \beta\gamma = 2 \). Column (1) reports the results of the test for pre-trends without residualizing (equation (2.5)). Column (2) reports the probability of rejecting that \( \hat{\pi}_{\text{noresid}} = 0 \) at the 5% level of significance. Column (3) reports the results of the test for pre-trends after residualizing the earnings growth (equation (2.8)), while column (4) reports the probability of rejecting that \( \hat{\pi}_{\text{resid}} = 0 \) at the the 5% level of significance. Finally, column (5) reports the results of the pooled reduced-form regression of earnings growth on Bartik (Equation 2.6). \( K \) is the number of industries and \( L \) is the number of locations. See Section 4 for details.