NBER WORKING PAPER SERIES

# BARTIK INSTRUMENTS:
# WHAT, WHEN, WHY, AND HOW

Paul Goldsmith-Pinkham
Isaac Sorkin
Henry Swift

Working Paper 24408
http://www.nber.org/papers/w24408

NATIONAL BUREAU OF ECONOMIC RESEARCH
1050 Massachusetts Avenue
Cambridge, MA 02138
March 2018, Revised June 2019

Thanks to the editor (Thomas Lemieux), anonymous referees, Isaiah Andrews, David Autor, Tim Bartik, Paul Beaudry, Kirill Borusyak, Jediphi Cabal, Arun Chandrasekhar, Gabriel Chodorow-Reich, Damon Clark, Richard Crump, Rebecca Diamond, Mark Duggan, Matt Gentzkow, Andrew Goodman-Bacon, David Green, Gordon Hanson, Caroline Hoxby, Peter Hull, Guido Imbens, Xavier Jaravel, Pat Kline, Magne Mogstad, Maxim Pinkovskiy, Luigi Pistaferri, Giovanni Righi, Ben Sand, Pedro Sant'Anna, Juan Carlos Suarez Serrato, Jan Stuhler, Melanie Wallskog, Kenneth West, Wilbert van der Klaauw, Eric Zwick, and numerous seminar participants for helpful comments. Thanks to Maya Bidanda, Jacob Conway, and Victoria de Quadros for research assistance. Thanks to David Card for sharing code, and Rodrigo Adao, Kirill Borusyak, Peter Hull, Xavier Jaravel, Michal Kolesar, and Eduardo Morales for sharing data. Swift was supported by the National Science Foundation Graduate Research Fellowship. Part of the work on this paper was completed while Goldsmith-Pinkham was employed by the Federal Reserve Bank of New York. The views expressed are those of the authors and do not necessarily reflect those of the Federal Reserve Bank of New York, the Federal Reserve Board, or the National Bureau of Economic Research. All errors are our own, please let us know about them.

Bartik Instruments: What, When, Why, and How
Paul Goldsmith-Pinkham, Isaac Sorkin, and Henry Swift
NBER Working Paper No. 24408
March 2018, Revised June 2019
JEL No. C1,C18,C2,J0,J2

**ABSTRACT**

The Bartik instrument is formed by interacting local industry shares and national industry growth rates. We show that the typical use of a Bartik instrument assumes a pooled exposure research design, where the shares measure differential exposure to common shocks, and identification is based on exogeneity of the shares. Next, we show how the Bartik instrument weights each of the exposure designs. Finally, we discuss how to assess the plausibility of the research design. We illustrate our results through three applications: estimating the elasticity of labor supply, estimating local labor market effects of Chinese imports, and estimating the elasticity of substitution between immigrants and natives.

Paul Goldsmith-Pinkham
Yale University School of Management
165 Whitney Ave
New Haven, CT 06511
paulgp@gmail.com

Isaac Sorkin
Department of Economics
Stanford University
579 Serra Mall
Stanford, CA 94305
and NBER
sorkin@stanford.edu

Henry Swift
henry.swift@gmail.com

The Bartik instrument is named after Bartik (1991), and popularized in Blanchard and Katz (1992).[1] These papers define the instrument as the local employment growth rate predicted by interacting local industry employment shares with national industry employment growth rates. The Bartik approach and its formally identical variants have since been used across many fields in economics, including labor, public, development, macroeconomics, international trade, and finance.

In our exposition, we focus on the canonical setting of estimating the labor supply elasticity, but our results apply more broadly wherever Bartik-like instruments are used. For simplicity, consider the cross-sectional structural equation linking wage growth to employment growth

$$y_l = \rho + \beta_0 x_l + \epsilon_l,$$

where $y_l$ is wage growth in location $l$ between two time periods, $x_l$ is the employment growth rate, $\rho$ is a constant, and $\epsilon_l$ is a structural error term that is correlated with $x_l$. Our parameter of interest is $\beta_0$, the inverse elasticity of labor supply. We use the Bartik instrument to estimate $\beta_0$.

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 the share of location $l$'s employment in industry $k$, and $g_{lk}$ is the growth rate of industry $k$ in location $l$. The second is that we can decompose the industry-growth rates as

$$g_{lk} = g_k + \tilde{g}_{lk},$$

where $g_k$ is the industry growth rate and $\tilde{g}_{lk}$ is the idiosyncratic industry-location growth rate. The Bartik instrument is the inner product of the industry-location shares and the industry component of the growth rates; formally, $B_l = \sum_k z_{lk} g_k$.

Because the Bartik instrument combines two accounting identities, it is always possible to construct it. It is not plausible, however, that the Bartik instrument always provides a valid identification strategy. In this paper, we open the black box of the Bartik instrument by formalizing its structure and unpacking the variation that the instrument uses. Our

[1] The intellectual history of the Bartik instrument is complicated. The earliest use of a shift-share type decomposition we have found is Perloff (1957, Table 6), which shows that industrial structure predicts the *level* of income. Freeman (1980) is one of the earliest uses of a shift-share decomposition interpreted as an instrument: it uses the change in industry composition (rather than differential growth rates of industries) as an instrument for labor demand. What is distinctive about Bartik (1991) is that the book not only treats it as an instrument, but also, in the appendix, explicitly discusses the logic in terms of the national component of the growth rates.

goal is to enable researchers to use familiar tools to distinguish between situations where the Bartik instrument would and would not be valid.

In this paper, we discuss the Bartik instruments' identification as coming from the shares. The basis of this view is a numerical equivalence result: we show that the two-stage least squares (TSLS) estimator with the Bartik instrument (the Bartik estimator) is numerically equivalent to a generalized method of moments (GMM) estimator with the local industry shares as instruments and a weight matrix constructed from the national growth rates. We interpret this result as saying that using the Bartik instrument is "equivalent" to using local industry shares as instruments, and so the exogeneity condition should be stated in terms of the shares. In contrast, Borusyak, Hull, and Jaravel (2018) emphasize that under some assumptions the consistency of the estimator can also come from the shocks,[2] and they also provide a motivating numerical equivalence result. How can researchers tell which identification assumption they are making? We argue that a researcher likely has the shares identification assumption in mind if they (i) describe their research design as reflecting differential exposure to common shocks, (ii) emphasize a two-industry example, and/or (iii) emphasize shocks to specific industries as central to their research design.

Once we think about the shares as the instruments, the implied empirical strategy is an exposure research design, where the industry shares measure the differential exposure to the common shock. Because the shares are typically equilibrium objects and likely co-determined with the level of the outcome of interest, it can be hard to assume that the shares are uncorrelated with the levels of the outcome. But this assumption is not necessary for the empirical strategy to be valid. Instead, the strategy asks whether differential exposure to common shocks leads to differential *changes* in the outcome. For example, in the canonical setting, the outcome is wage *growth*, in the China shock setting the outcome is *change* in manufacturing employment, and in the immigrant enclave setting it is *changes* in the residual log wage gap between immigrants and natives. Hence, the empirical strategy can be valid even if the shares are correlated with the levels of the outcomes.

How does one build the credibility of such an exposure design? The central identification worry is that the industry shares predict outcomes through channels other than those posited by the researcher. One way to assess this possibility is to look at correlates of the shares. If these correlates suggest other channels through which the shares affect outcomes in the relevant period, then we might be skeptical of the identifying assumption. Second, in some settings there is a pre-period, as in a standard difference-in-differences design. In this case, we can test for parallel pre-trends. Given that the design exploits level differences in the shares, by exploring trends in changes we can assess the plausibility of the assumption that the common shock caused the change in the changes, or whether there

[2] Adao, Kolesar, and Morales (2018) discuss inferential issues in this set-up.

were pre-existing differences in the changes.

There is a third way to explore the validity of the research design, based on the observation that the Bartik instrument is a particular way of combining many instruments. Under the null of constant effects, a researcher can consider alternative estimators which combine multiple instruments or run overidentification tests. One interpretation of the divergence between estimators and the failure of overidentification tests is that the null of constant effects is unreasonable, and to instead interpret these tests as pointing to the presence of treatment effect heterogeneity, rather than failure of exogeneity. With multiple unordered instruments, it is difficult to allow unrestricted heterogeneity of the form discussed by Imbens and Angrist (1994). Instead, we follow Borusyak, Hull, and Jaravel (2018) and Adao, Kolesar, and Morales (2018) and consider a restricted form of linear heterogeneity where there are constant effects within each location. We highlight that even if each instrument separately places convex weights on each location's parameter, the Bartik instrument can generate negative weights on each instrument, raising the possibility of negative weights on location-specific treatment effects. In this case, the Bartik estimator would not have a LATE-like interpretation as a weighted average of treatment effects. To the extent that researchers wish to embrace a treatment effect heterogeneity interpretation of the Bartik instrument, they should be comfortable with the patterns of underlying heterogeneity. We develop a visual diagnostic to aid researchers in this task.

How does the Bartik instrument combine the exposure designs? We build on Rotemberg (1983) and decompose the Bartik estimator into a weighted sum of the just-identified instrumental variable estimators that use each industry share ($z_{lk}$) as a separate instrument. The weights, which we refer to as Rotemberg weights, are simple to compute and sum to 1. They depend on the covariance between the $k^{th}$ instrument's fitted value of the endogenous variable and the endogenous variable itself. The weights are a scaled version of the Andrews, Gentzkow, and Shapiro (2017) sensitivity-to-misspecification parameter, and tell us how sensitive the overidentified estimate of $\beta_0$ is to misspecification (i.e., endogeneity) in any instrument. Heuristically, they also tell us which exposure design gets more weight in the overall estimate, and thus which of these identifying assumptions is most worth testing. If the high-weight designs—where it is concrete what comparisons the researcher is doing—pass basic specification tests, then researchers should feel reassured about the overall empirical strategy.

In many contexts where researchers use Bartik instruments, it is used in the reduced-form, whereas in our analysis we discuss the instrumental variables setting. We note that the insights of this paper still apply when Bartik is used in the reduced-form. Specifically, the relevant moment condition (exclusion restriction) is still the same. Moreover, it is still possible to compute the Rotemberg weights.

We note two limitations to our analysis. First, we assume locations are independent and so ignore the possibility of spatial spillovers or correlation.[3] Second, we assume that the data consist of a series of steady states.[4]

To summarize, we view our contribution as explaining identification in the context of Bartik instruments in two ways. First, our GMM result shows that Bartik is numerically equivalent to using industry shares as instruments. Hence, we argue that the typical identifying assumption is best stated in terms of industry shares, rather than growth rates. Second, we build on Andrews, Gentzkow, and Shapiro (2017) to provide tools to measure the "identifying variation," and formalize how to use Rotemberg weights to highlight the subset of instruments to which the estimated parameter is most sensitive to endogeneity.

**Applications:** We illustrate our results through three applications. In our first application, we look at the canonical example of estimating the inverse elasticity of labor supply in US Census data using decadal differences from 1980-2010 and instrumenting for labor demand with the Bartik instrument. We first show that the national growth rates explain less than one percent of the variance of the Rotemberg weights. Hence, the growth rates are a poor guide to understanding what variation in the data is driving estimates. Second, the weights are skewed, with over fourty percent of the weight on the top five industries. In the particular, the oil and gas extraction industry receives the largest weight. Hence, a concrete example of the comparisons being made by the estimator is between changes in employment growth and wage growth in places with more and less oil and gas extraction. Third, industry shares, including oil and gas extraction, are correlated with many observables, including the immigrant share, which potentially predicts innovations in labor supply. Fourth, alternative estimators deliver substantively different point estimates and overidentification tests reject the null of exogeneity. Fifth, consistent with the overidentification tests rejecting, we find substantial visual dispersion in the estimates from each individual instrument. Moreover, some of outlying point estimates receive negative Rotemberg weights, which suggests that—under the treatment effect heterogeneity interpretation—some of the underlying effects receive negative weight so that there is unlikely to be a LATE-like interpretation of the parameter estimate.

In our second application, we estimate the effect of Chinese imports on manufacturing employment in the United States (using the China shock of Autor, Dorn, and Hanson (2013)). We first show that the growth rates of imports from China to other high-income countries explain about twenty percent of the variance in the Rotemberg weights. Hence, the growth rates are a poor guide to what variation in the data is driving estimates. Sec-

[3]Monte, Redding, and Rossi-Hansberg (2018) document the presence and economic importance of spatial spillovers through changes in commuting patterns in response to local labor demand shocks.

[4]See Jaeger, Ruist, and Stuhler (2018) for discussion of out-of-steady-state dynamics in the context of immigration.

ond, the two highest weight industries are electronic computers and games & toys. Hence, a concrete example of the comparisons being made by the estimator is comparing outcomes in locations with high and low shares of the electronic computers industry. Third, the industries that get the most weight tend have larger shares in more educated areas. Fourth, we examine pre-trends among the industries with high Rotemberg weights and find that the comparisons implied by the industries (i.e., places with more and less of the industry) exhibits substantial pre-trends, potentially explaining part of the large effects in the 2000s (when the China shock was largest). Fifth, alternative estimators deliver substantively different point estimates and overidentification tests reject the null of exogeneity. Sixth, the underlying point estimates are visually less dispersed than in the canonical example, and the point estimates receiving negative weight are less varied, suggesting that overall negative weights are less likely.

In our third application, we estimate the inverse elasticity of substitution between immigrants and natives in 2000 (following the empirical strategy of Card (2009)). Here, the relevant shares are the share of migrants from an origin country that live in a particular location in the base year, and the shocks are the immigrant inflows. First, we find that for high school equivalent workers, the Rotemberg weights are almost completely explained by the immigrant inflows. For the college equivalent workers, the explanatory power of the inflows is higher than in our other two examples. Hence, the growth rates (the shocks) are a good guide to the variation in the data that drives estimates. Second, for high school equivalent workers, the share of Mexican immigrants in a city in 1980 gets almost half the weight in the estimator, a possibility that Card (2009, pg. 9) acknowledges. Hence, for high school equivalent workers, a concrete example of the comparison the estimator is making is between places with more and fewer Mexican immigrants in 1980. For college equivalent workers, the highest weight instrument is the Philippines, and so the comparison is between places with higher and lower Philippines share. Third, among the covariates used by Card (2009), we do not find any systematic patterns of correlations with the immigrant shares. Fourth, unlike in our other examples, most overidentification tests fail to reject and we do not find differences among estimators. Fifth, we find no evidence of pre-trends for the high school equivalent workers. In contrast, we find statistically and economically significant pre-trends for the estimates involving the college equivalent workers, consistent with the concerns emphasized by Jaeger, Ruist, and Stuhler (2018).

Besides these three examples, a much broader set of instruments is Bartik-like. We define a Bartik-like instrument as one that uses the inner product structure of the endogenous variable to construct an instrument. In Appendix A, we discuss two additional examples. First, researchers, such as Greenstone, Mas, and Nguyen (Forthcoming), interact pre-existing bank lending shares with changes in bank lending volumes to instrument for credit supply. Second, Acemoglu and Linn (2004) interact age-group spending patterns with demographic changes to instrument for market size.

**Literature:** A vast literature of papers uses Bartik-like instruments, and many of these discuss the identifying assumptions in ways that are close to the benchmark results in this paper. For example, Baum-Snow and Ferreira (2015, pg. 50) survey the literature and state that the "validity [of the Bartik instrument]...relies on the assertion that neither industry composition nor unobserved variables correlated with it directly predict the outcomes of interest conditional on controls." Similarly, Beaudry, Green, and Sand (2012) provide a careful discussion of identifying assumptions in the context of an economic model. Given the vast diversity of ways in which Bartik instruments are discussed and understood in the literature, we can only claim novelty for the formalism along this dimension.

Beyond the vast literature of papers using Bartik-like instruments, this paper is also related to a growing literature that comments on specific papers (or literatures) that use Bartik-like instruments. This literature includes at least three papers: Christian and Barrett (2017), which comments on Nunn and Qian (2014), Jaeger, Joyce, and Kaestner (Forthcoming), which comments on Kearney and Levine (2015), and Jaeger, Ruist, and Stuhler (2018), which comments on the use of the immigrant enclave instrument. Relative to this literature, our goal is to develop a formal econometric understanding of the Bartik instrument and provide methods to increase transparency in its use.

# 1 Equivalence between Bartik IV and GMM with industry shares

We first show that the Bartik instrument is numerically equivalent to using industry shares as instruments, which we use to argue that the identification condition 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. To focus on identification issues, we discuss infeasible Bartik, where we assume that we know the common national component of industry growth rates. Section 2 discusses asymptotics.

## 1.1 Full panel setup

We begin by setting up the general panel data case with $K$ industries and $T$ time periods. This setup 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 the following structural equation:

$$y_{lt} = D_{lt}\rho + x_{lt}\beta_0 + \epsilon_{lt}. \tag{1.1}$$

In the canonical setting, $l$ indexes a location, $t$ a time period, $y_{lt}$ is wage growth, $D_{lt}$ is a vector of $Q$ controls which could include location and time fixed effects, $x_{lt}$ is employment growth and $\epsilon_{lt}$ is a structural error term. The parameter of interest is $\beta_0$. We assume that the ordinary least squares (OLS) estimator for $\beta_0$ is biased and we need an instrument to estimate $\beta_0$.

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

$$x_{lt} = Z_{lt} G_{lt} = \sum_{k=1}^{K} z_{lkt} g_{lkt},$$

where $Z_{lt}$ is a $1 \times K$ vector of industry-location-time period shares, and $G_{lt}$ is a $K \times 1$ vector of industry-location-time period growth rates where the $k^{th}$ entry is $g_{lkt}$. We decompose the industry-location-period growth rate into an industry-period and an idiosyncratic industry-location-period components:

$$g_{lkt} = g_{kt} + \tilde{g}_{lkt}.$$

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:[5]

$$B_{lt} = Z_{l0} G_t = \sum_k z_{lk0} g_{kt},$$

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 $1 \times K$ vector of industry shares in location $l$. Hence, we have a standard two-stage least squares setup where the first-stage is a regression of employment growth on the controls and the Bartik instrument:

$$x_{lt} = D_{lt}\tau + B_{lt}\gamma + \eta_{lt},$$

and the structural equation is given by (1.1).

Let $y_l = (y_{l1}, \ldots, y_{lT})$, $x_l = (x_{l1}, \ldots, x_{lT})$, $Z_l = (Z_{l1}, \ldots, Z_{lT})$, $G_l = (G_{l1}, \ldots, G_{lT})$, $D_l = (D_{l1}, \ldots, D_{lT})$, and $\epsilon_l = (\epsilon_{l1}, \ldots, \epsilon_{lT})$. We assume that the data

$$\{y_l, D_l, G_l, Z_l, Z_{l0}\}_{l=1}^{L}$$

[5] If $\epsilon_{lt}$ are correlated with growth rates, and the $\epsilon_{lt}$ are serially correlated, then future shares will be endogenous. This potential for serial correlation motivates fixing industry shares to some initial period. Beaudry, Green, and Sand (2018, pg. 18-19) discuss Bartik instruments and advocate updating the shares under the assumption that the error term is not serially correlated.

is drawn i.i.d. across $l$.

We assume that $D_{lt}$ is strictly exogeneous, and focus on estimating $\beta_0$ using residual regression. Define $Y_L = (y_1, \ldots, y_L)$, $X_L = (x_1, \ldots, x_L)$, $D_L = (D_1, \ldots, D_L)$ and $\epsilon_L = (\epsilon_1, \ldots, \epsilon_L)$. Let $M_D = I_L - D_L(D_L' D_L)^{-1} D_L'$ denote the annhilator matrix for $D$, the $L \times Q$ matrix of controls, where $I_L$ is the $L \times L$ identity matrix. We define $X_L^{\perp} \equiv M_D X_L$ and $Y_L^{\perp} \equiv M_D Y_L$ to be the residualized $X_L$ and $Y_L$ such that $M_D(Y_L - X_L\beta_0) = M_D(D_L\rho + \epsilon_L) = M_D\epsilon_L$, since $M_D D_L = 0$. Finally, define $\epsilon_L^{\perp} \equiv M_D \epsilon_L$.

## 1.2 Equivalence in three special cases

We build up to the general result that the Bartik instrument is numerically equivalent to using industry shares as instruments for a particular weight matrix in GMM through three special cases. Each of these special cases also illustrates a research design implicit in using a Bartik instrument and suggests a specification test.

### Two industries and one time period

With two industries whose shares sum to one within each location and one time period, the Bartik instrument is identical to using one of the industry shares as an instrument. To see this, expand the Bartik instrument:

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

where $g_1$ and $g_2$ are the industry components of growth. Since the shares sum to one, 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:

$$x_l = \gamma_0 + \gamma B_l + \eta_l = \underbrace{\gamma_0 + \gamma g_2}_{\text{constant}} + \underbrace{\gamma(g_1 - g_2)}_{\text{coefficient}} z_{l1} + \eta_l.$$

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 $z_{l1}$ (or $z_{l2}$) as an instrument.

What is the research design inherent in this special case? Here, $z_{l1}$ measures exposure to the policy that affects industry 1, and $g_1 - g_2$ is the size of the policy. The outcome is $y_l$, which is the *change* in outcomes between two periods. Hence, in this special case the empirical strategy asks about the effects of levels of $z_{l1}$ on changes in $y_l$. The identification concern is whether $z_{l1}$ is correlated with *changes* in the outcome, and not *levels* of the outcome. As we discuss more below in Test 1 in Section 5, studying covariates of $z_{l1}$ is helpful in making clear the types of concerns one might have. Concretely, while $z_{l1}$ might be correlated with many covariates that predict the level of the outcome, this correlation is not necessarily a problem for the research design. Instead, the central question a researcher should have in mind is whether these correlates predict changes in the outcome in the relevant period.

**Two industries and two time periods**

In a panel with two time periods, if we interact the time-invariant industry shares with time, then Bartik is equivalent to a special case of using industry shares as instruments. To see this result, we again specialize to two industries, and define the 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 panel estimating equations, restrict our 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:

$$x_{lt} = \tau_l + \underbrace{(\tau_t + g_{2t}\gamma)}_{\equiv \tilde{\tau}_t} + z_{l10} \underbrace{(g_{1t} - g_{2t})\gamma}_{\equiv \tilde{\gamma}_t} + \eta_{lt}. \tag{1.2}$$

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, write $g_{1t} - g_{2t} = (g_{11} - g_{21}) + (\Delta g_1 - \Delta g_2)\mathbb{1}(t=2)$, where $\Delta g_1 = g_{12} - g_{11}$, $\Delta g_2 = g_{22} - g_{21}$, and $\mathbb{1}$ is the indicator function. Then, rewrite the first stage as

$$x_{lt} = \underbrace{\tau_l + z_{l10}(g_{11} - g_{21})\gamma}_{\equiv \tilde{\tau}_l} + \underbrace{(\tau_t + g_{2t}\gamma)}_{\equiv \tilde{\tau}_t} + z_{l10}\mathbb{1}(t=2)\underbrace{(\Delta g_1 - \Delta g_2)\gamma}_{\equiv \tilde{\gamma}_t} + \eta_{lt}. \tag{1.3}$$

We can now see the equivalence between Bartik and using the shares as instruments:

$$x_{lt} = \tilde{\tau}_l + \tilde{\tau}_t + z_{l10}\mathbb{1}(t=2)(\Delta g_1 - \Delta g_2)\gamma + \eta_{lt} \qquad \text{(Bartik)}$$
$$x_{lt} = \tilde{\tau}_t + \tilde{\tau}_t + z_{l10}\mathbb{1}(t=2)\tilde{\gamma} + \eta_{lt}. \qquad \text{(Industry Shares)}$$

In this case, again $\tilde{\gamma} = \gamma/(\Delta g_1 - \Delta g_2)$. If we view $z_{l10}$ as the effect of exposure to a policy, then $\tilde{\gamma}$ captures the "unscaled" effect on $x_{lt}$, while $\gamma$ is rescaled by the size of the policy, where the size of the policy is the dispersion in national industry growth rates, $\Delta g_{1t} - \Delta g_{2t}$.

What is the research design inherent in this special case? Viewing the growth rates as a measure of policy size and the industry shares as measures of exposure emphasizes a useful connection to difference-in-differences. In some settings, there are more than two time periods, and there is a pre-period before a policy takes effect. In this case, if we write the reduced form in terms of the Bartik instrument, or, alternatively, in terms of industry shares interacted with time, we see a natural comparison to a standard difference-in-difference estimation procedure:

$$y_{lt} = \tau_l + \tau_t + \gamma\beta B_{lt} + \underbrace{\tilde{\epsilon}}_{\eta_{lt}\beta + \epsilon_{lt}}$$
$$y_{lt} = \tau_l + \tau_t + \sum_{s \neq t_0} \mathbb{1}(s=t)\underbrace{\tilde{\gamma}_s}_{\gamma\Delta_{gs}}\beta z_{l10} + \underbrace{\tilde{\epsilon}}_{\eta_{lt}\beta + \epsilon_{lt}}.$$

Relative to the single change case discussed before, having an extra period provides additional tools to assess the plausibility of using the level of $z_{l1}$ to predict the change in the outcome. In this case, the testable implication of parallel pre-trends is that $\gamma_t\beta = 0$ for $t < t_0$, where $t_0$ demarcates the pre-period. Intuitively, a researcher is asking whether in the pre-period, the level of $z_{l1}$ predicts changes in the outcome. Failing to find a pre-trend gives credence to a research design where the researcher assumes that $z_{l1}$ is relevant for predicting the change in period $t_0$ and afterwards. We return to this point in Test 2 in Section 5.

**K industries and one time period**

Finally, we show that with $K$ industries as instruments in a generalized method of moments (GMM) estimator setup with a specific weight matrix, the Bartik estimator is identical to using the set of industry shares as instruments.

To show this result, recall that $G$ is the $K \times 1$ vector of industry growth rates, $Z$ is the $L \times K$ matrix of industry shares, $Y$ is the $L \times 1$ vector of outcomes, $X$ is the $L \times 1$ vector of endogenous variables, and $B = ZG$ is the $L \times 1$ vector of Bartik instruments. Let $W$ be an arbitrary $K \times K$ matrix.

We define the Bartik and the GMM estimator using industry shares as instruments:

$$\hat{\beta}_{Bartik} = \frac{B'Y^{\perp}}{B'X^{\perp}}; \text{ and } \hat{\beta}_{GMM} = \frac{X^{\perp\prime}ZWZ'Y^{\perp}}{X^{\perp\prime}ZWZ'X^{\perp}}.$$

PROPOSITION 1.1. *If $W = GG'$, then $\hat{\beta}_{GMM} = \hat{\beta}_{Bartik}$.*

**Proof.** See appendix B. □

Proposition 1.1 says the Bartik instrument and industry shares as instruments are numerically equivalent for a particular choice of weight matrix.

What is the research design inherent in this special case? Under the shares interpretation that we discuss further below, if there is a shock in a single period, then this research design pools many different exposure designs. In Section 3, we show the way that Bartik pools these designs. The tools for building the credibility of any given share are the same as in the single instrument case. Moreover, the many instruments provide the researcher with the opportunity to test whether the parameter estimates from all of these instruments are the same using overidentification tests. Alternatively, if these parameters are not similar, the researcher might be interested in trying to characterize this heterogeneity. In Test 3 in Section 5, we discuss overidentification tests. In Section 4, we discuss heterogeneity.

REMARK 1.1. *When $\sum_{k=1}^{K} z_{lk} = 1$, there are $K-1$ instruments and not K instruments. In practice, any of the K industries can be dropped by subtracting off that industry's growth rate from the G vector, and the Bartik instrument will maintain its numerical equivalence from Proposition 1.1. To see the intuition behind this, suppose that $\sum_k z_{lk} = 1 \forall l$. Consider the first stage regression:*

$$x_l = \gamma_0 + \gamma_1 B_l + \eta_l.$$

*Now add and subtract $\gamma_1 \sum_k z_{lk} g_j$ from the right hand side:*

$$x_l = \underbrace{\gamma_0 + \gamma_1 \sum_k z_{lk} g_j}_{\gamma_0 + \gamma_1 g_j} + \underbrace{\gamma_1 \sum_k z_{lk}(g_k - g_j)}_{B_l - g_j} + \eta_l. \tag{1.4}$$

*This expression generalizes our result from the two industry and one time period example. It says that normalizing the growth rates by a constant $g_j$ changes the first-stage intercept and does not affect the slope estimate. Hence, the first-stage prediction is unaffected.*

### 1.3 Summary

With $K$ industries and $T$ time periods, the numerical equivalence 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 is notationally cumbersome so we leave the formal details to Appendix C. We now turn to discussing how these finite sample results map into identification conditions.

## 2 Asymptotic consistency and identifying assumptions

We now consider consistency of the TSLS estimator that uses the Bartik instrument. In the previous section, we established a finite sample equivalence result between the TSLS estimator using the Bartik instrument, and the GMM estimator using industry shares as instruments and a weight matrix defined by the industry growth rates. Here, we use this equivalence to show that a sufficient condition for consistency is strict exogeneity of the shares.

To fix ideas, consider the difference between the TSLS estimator and the parameter of interest:

$$\hat{\beta} - \beta_0 = \frac{\sum_{t=1}^{T} \sum_{k=1}^{K} g_{kt} \sum_{l=1}^{L} z_{lk0} \epsilon_{lt}^{\perp}}{\sum_{t=1}^{T} \sum_{k=1}^{K} g_{kt} \sum_{l=1}^{L} z_{lk0} x_{lt}^{\perp}}. \tag{2.1}$$

Broadly, conditions for the consistency of $\hat{\beta}$ can be stated either in terms of the *shares*, the $z_{lk0}$, or the *shocks*, the $g_{kt}$. In this paper, we consider a setting where we observe increasingly larger samples of locations, but a fixed number of time periods and industries (fixed $T$ and $K$). As we show below, in this setting it is natural to state conditions for consistency in terms of the shares.

A natural extension of this setup studied by Borusyak, Hull, and Jaravel (2018) considers a situation where we not only observe increasingly larger samples of locations, but also of industries. They show that while a sufficient condition for consistency of $\hat{\beta}$ is exogeneity of the shares, it is not necessary. With many industries, it is possible to use the exogeneity of the shocks, e.g. $g_{kt}$, instead.

In this section, we first state the sufficient conditions in our setting, highlighting the relevance and exogeneity assumptions. We then discuss when these exogeneous shares assumptions are reasonable, and how they contrast to the exogenous shocks assumptions.

## 2.1 Identifying assumptions

Consider the following sums, which denote the empirical moments and first derivatives, respectively:

$$m^1_{L,kt} = L^{-1}\sum_{l=1}^{L} z_{lk0}\epsilon^{\perp}_{lt}, \text{ and} \qquad m^2_{L,kt} = L^{-1}\sum_{l=1}^{L} z_{lk0}x^{\perp}_{lt}.$$

Our consistency condition (2.1) is thus

$$\hat{\beta} - \beta_0 = \frac{\sum_{t=1}^{T}\sum_{k=1}^{K} g_{kt}m^1_{L,kt}}{\sum_{t=1}^{T}\sum_{k=1}^{K} g_{kt}m^2_{L,kt}}, \tag{2.2}$$

which combines $KT$ moment conditions.

Note that two assumptions must hold for consistency. First, the denominator must converge to a non-zero term. Intuitively, for this assumption to hold, there must be an industry and time period when the industry share has predictive power for $x^{\perp}_{lt}$, and the growth rates $g_{kt}$ cannot weight the covariances in such a way that they exactly cancel. This first condition holds under the following low-level assumption:

ASSUMPTION 1 (Relevance). $\mathbb{E}[z_{lk0}x^{\perp}_{l}] = C_k$, *and* $\sum_k g_k C_k = C \neq 0$, *where* $g_k = (g_{k1}, \ldots, g_{kT})$.

Note that these $C_k$ are unscaled versions of the first-stage coefficients.

The second necessary assumption for consistency is that the numerator must converge to zero. This assumption is the exclusion restriction, and to hold generically, the industry share must be uncorrelated with the structural error term, *after controlling for $D_{lt}$*, for industries that have non-zero growth rates.[6] The following identifying assumption ensures that the numerator converges to zero:

ASSUMPTION 2 (Strict Exogeneity). $\mathbb{E}[\epsilon^{\perp}_{l}|z_{lk0}] = 0$ *for all $k$ where $g_k \neq 0$.*

This assumption is standard in empirical settings that use exposure designs. For example, this assumption is made in difference-in-differences designs that use location fixed effects.

[6] Even if $m^1_{L,kt}$ converges to a non-zero term, then the numerator could still converge to zero non-generically if the $g_{kt}$ are such that these biases cancel out exactly. For fixed $K$ and $T$, this case is unlikely to hold in practice. When $K$ increases, Borusyak, Hull, and Jaravel (2018) show that this can hold generically. We discuss this point further below.

It is now straightforward to show consistency.

PROPOSITION 2.1. *Given Assumptions 1 and 2, and standard regularity conditions,*

$$\operatorname*{plim}_{L\to\infty} \hat{\beta} - \beta_0 = \frac{\sum_{t=1}^{T}\sum_{k=1}^{K} g_{kt}\mathbb{E}[z_{lk0}\epsilon_{lt}^{\perp}]}{\sum_{t=1}^{T}\sum_{k=1}^{K} g_{kt}\mathbb{E}[z_{lk0}x_{lt}^{\perp}]} = 0.$$

As a result, and the Bartik TSLS IV estimator is consistent.[7]

These results have two implications: first, under our sampling process, strict exogeneity of the industry shares is necessary for the Bartik estimator to be generically consistent. This assumption is standard in many difference-in-differences settings. Second, it highlights that the Bartik estimator uses a particular weighting of these moment conditions; other weightings would imply other estimators.

### 2.2 When are these assumptions plausible?

The exogenous shares assumption discussed in the last section might seem implausible because shares are equilibrium objects likely co-determined with the level of the outcome of interest. But this reasoning does not reflect the assumption that is typically being made. Instead, the assumption is about exogeneity *conditional on observables*, which typically include location fixed effects.[8] Taking typical controls into account, the assumption is that the shares are exogenous to *changes* in the error term (i.e., changes in the outcome variable), rather than *levels* of the outcome variable.

The plausibility of the substantive restrictions implied by this identifying assumption might be more intuitive in a setting with two industries and a differential exposure design, which we discussed in Section 1. In this setting, the identifying assumption is that the differential effect of higher exposure of one industry (compared to another) only affects the change in the outcome ($y_{lt}$) through the endogeneous variable of interest, and not through any potential confounding channel. This assumption is standard in difference-in-differences. In the shares view, the identifying assumption underlying the Bartik setting is simply this differential exposure design applied to each industry separately.

[7] This result is a straightforward application of standard theorems for GMM. Let $\mathbf{m}_L^1 = (m_{L,11}^1, \ldots, m_{L,KT}^1)$ denote our $KT$ empirical moment conditions, and $\operatorname{plim}_{L\to\infty} \mathbf{m}_L^1 = \mathbf{m}^1$. Then, it is straightforward to see that for $\mathbf{m}_L^2 = (m_{L,11}^2, \ldots, m_{L,KT}^2)$, $\operatorname{plim}_{L\to\infty} \mathbf{m}_L^2 = d\mathbf{m}_L^1/d\beta$. Hence, the empirical analog to $d\mathbf{m}_L^1/d\beta$ naturally fits into the standard GMM framework, with our weight matrix defined as $GG'$, where $G = (g_{11}, \ldots, g_{KT})'$. It is also possible to see why assuming $G$ is known is not a restrictive condition. So long as an estimated $G$ converges to a fixed non-stochastic $G$, as in standard GMM analysis, the results follow.

[8] In some cases, the application does not include fixed effects, but instead uses first differences instead. An example of this would be Autor, Dorn, and Hanson (2013), where the outcome of interest is the change in employment, rather than the level of employment.

This type of identification assumption is natural to make when the shares create differential exposure to a common economic or policy shock (or sets of shocks). In these cases, the most natural description of the identification comes from highlighting a few key industries which best illustrate the exposure design. In Section 3, we show how to do this. While natural to make, this type of assumption may not always be satisfied. For example, areas with high vs. low exposure may have other features that predict change in the outcome through channels other than the endogeneous variable, violating the exclusion restriction.

In cases when the assumption of exogeneous shares is not plausible, consistency of the estimator can instead come from many exogeneous shocks. As proved in Borusyak, Hull, and Jaravel (2018), exogeneous independent shocks to many industries leads the Bartik estimator to be consistent, even when the shares are not exogeneous. The core intuition to this result can be seen in Equation 2.2. In cases when the shares are not exogeneous, $m^1_{L,kt}$ does not converge to zero. As a result, the weighted sum of the industry shocks and the shares are non-zero. With many exogeneous and independent shocks, however, Borusyak, Hull, and Jaravel (2018) show that the estimator is still consistent. The reason is that the random shocks are uncorrelated with the bias from the shares, and the presence of many shocks causes this bias to average out (see also Kolesar et al. (2015)).

For the shocks assumption to be plausible, there needs to be a large number of different industries, each experiencing random growth. Importantly, the asymptotic thought experiment involves adding additional industries, rather than more finely partitioning existing industries. This idea is most clearly seen in the numerical equivalence result from Borusyak, Hull, and Jaravel (2018), which shows that the Bartik IV regression can be rewritten as an industry-level regression. If an industry-level regression seems unnatural for the economic question at hand, then the exogenous shocks assumption is unlikely to hold.

How can researchers tell which assumption they have in mind? When a researcher explains identification using a two-industry example, she is emphasizing differential exposure, which underlies the shares assumption. Hence, she has in mind an identification assumption stated in terms of the shares. Indeed, because the shocks assumption is fundamentally about a large number of industries, the logic of how this identification works is not captured by the two-industry example. Similarly, if a researcher emphasizes the performance of a particular industry (or a small handful of industries), then this reasoning also suggests that she has in mind an identifying assumption in terms of the shares. In contrast, when having a large number of industries is central to how the researcher thinks about identification, then it is likely that she has in mind the shocks assumption.

While a best case scenario for a researcher using a Bartik instrument is for both the exogenous shares and shocks assumptions to hold, in practice, this coincidence seems unlikely. Typically, a researcher will only have one identification strategy at their disposal.

We encourage researchers to pick one or the other, be clear about why, and then defend the relevant assumptions in their setting.

# 3 Opening the black box of the Bartik estimator

The previous sections showed that under standard panel asymptotics, the Bartik instrument is equivalent to using industry shares as instruments. Thus, the Bartik estimator combines many instruments using a specific weight matrix.

Empirical work using a single instrument is transparent because there is a small number of covariances that enter the estimator. With many instruments, it is less intuitive how the estimator combines the different instruments. This lack of intuition underlies much of the empirical work using Bartik instruments, where it is hard to explain what variation in the data drives estimates, and can often feel like a black box.

In this section, we show how to open the black box of the Bartik estimator. First, we decompose the Bartik estimator into a weighted combination of just-identified estimates based on each instrument. This decomposition increases the transparency of the estimator because the weights highlight the industries whose variation in the data drives the overall Bartik estimate. Building on Andrews, Gentzkow, and Shapiro (2017) (AGS), we show that these weights can be interpreted as sensitivity-to-misspecification elasticities. High-weight instruments are more sensitive to misspecification, and hence are the instruments that are most important for researchers to defend.

## 3.1 Decomposing the Bartik estimator

We first present a finite sample decomposition of the linear overidentified GMM estimator due to Rotemberg (1983).[9] For expositional simplicity, we use a single cross-section, though it is straightforward to extend results to a panel with $T$ time periods.

PROPOSITION 3.1. *We can write:*

$$\hat{\beta}_{Bartik} = \sum_k \hat{\alpha}_k \hat{\beta}_k,$$

*where*

$$\hat{\beta}_k = (Z_k' X^{\perp})^{-1} Z_k' Y^{\perp} \text{ and } \hat{\alpha}_k = \frac{g_k Z_k' X^{\perp}}{\sum_{k'} g_{k'} Z_{k'}' X^{\perp}},$$

*so that* $\sum_k \hat{\alpha}_k = 1$.

**Proof.** See appendix B. □

[9] Andrews (Forthcoming, Section 3.1) reports this decomposition for constant-effect linear instrumental variables.

Proposition 3.1 has two implications. First, mirroring our results from Section 2, the validity of each just-identified $\hat{\beta}_k$ depends on the exogeneity of a given $Z_k$. Second, for some $k$, $\hat{\alpha}_k$ can be negative. Under the constant effects assumption we have maintained so far, these negative weights do not pose a conceptual problem. In Section 4, we introduce a restricted form of treatment effect heterogeneity and revisit the implications of the negative Rotemberg weights.

In Appendix D, we discuss how to interpret the Rotemberg weights in terms of sensitivity-to-misspecification following work by Conley, Hansen, and Rossi (2012) and Andrews, Gentzkow, and Shapiro (2017). The basic intuition is that if any particular instrument is misspecified, then $\alpha_k$ tells us how much that misspecification translates into the overall bias of the estimator. For example, if $\alpha_k$ is small, then bias in the $k^{th}$ instrument does not affect the overall bias in the estimator very much. We also show that this measure is different than simply dropping instruments and seeing how estimates change, since dropping an instrument combines sensitivity-to-misspecification (i.e., $\alpha_k$) as well as the relative misspecification of different instruments (i.e., how far $\hat{\beta}_k$ diverges from $\hat{\beta}$).

We recommend researchers report the instruments associated with the largest values of $\alpha_k$ for two reasons: First, reporting the instruments with the largest $\alpha_k$ provides a more concrete way to describe the empirical strategy. Second, to the extent that the researcher is concerned about misspecification, these are the instruments that are most worth probing.

### 3.2 Normalization

When the industry shares sum to one within a location, the instruments are linearly dependent and so we can write each instrument as a function of the remaining $K-1$ instruments. This fact has a couple implications. First, following Remark 1.1, we can drop any industry through normalization by subtracting off $g_j$ from all the growth rates, and leave our point estimates unchanged. Second, the fact that we can drop any one industry means that the Rotemberg weights are not invariant to the choice of which industry to drop. To take an extreme example, suppose industry $j$ has the largest weight. Then, by dropping industry $j$ through normalization, a researcher could make industry $j$ have a weight of zero, but the Bartik estimate would remain the same.

To address this issue, in applications where the industry shares sum to one, we report Rotemberg weights that come from demeaning the (unweighted) industry growth rates. In Appendix E, we show that this normalization is the average of the $K$ possible normalizations of dropping each of the industries.

To understand the intuition for why the normalization matters, return to the two industry example: suppose we think that Bartik is biased in this case. Does the bias arise from the fact that the industry 1 share is correlated with the error term, or that the industry 2 share is correlated with the error term? Conceptually, it is not meaningful to distinguish between these two possibilities, because industry 1 and 2 shares are exactly negatively correlated. Hence, saying the bias is correlated with industry 1 is the same as saying the bias is correlated with industry 2. In this case, our normalization assigns weight 0.5 to each industry.

### 3.3 Aggregation

Below, we consider applications with panel data and multiple time periods. As a result, the underlying instruments are industry shares interacted with time fixed effects. Rather than reporting results at the level of $\alpha_{k,t}$, we report $\alpha_k = \sum_t \alpha_{k,t}$. It is typically easier to think about the variation coming from a cross-sectional difference, rather than the variation coming from a cross-sectional difference in a particular time period. When aggregating to the $k$th industry, we report $\hat{\beta}_k$, which comes from using $B_{lkt} = z_{lk0} g_{kt}$, the Bartik instrument built from just the $k^{th}$ industry, as the instrument.

To interpret such an aggregated $\alpha$ in terms of the underlying misspecification, suppose that $\tilde{\beta}_{kt} = \tilde{\beta}_k$ for all $t$. Then,

$$\tilde{\beta} = \sum_k \alpha_k \sum_t \frac{\alpha_{kt}}{\alpha_k} \tilde{\beta}_{kt} = \sum_k \alpha_k \tilde{\beta}_k \sum_t \frac{\alpha_{kt}}{\alpha_k} = \sum_k \alpha_k \tilde{\beta}_k.$$

These equations say that the $\alpha_k$ measures the sensitivity-to-misspecification where we assume that the endogeneity associated with the $k^{th}$ industry is constant across time.

## 4 Heterogeneous effects

In previous sections, we showed that the Bartik estimator combines many instruments with a specific weight matrix. A key assumption was that of constant effects. In many contexts, a researcher might prefer to think that there are heterogeneous effects that vary across locations or time. For example, in the canonical labor supply elasticity application that we discuss below, some locations might have more elastic labor supply than others.

In this section, we discuss a heterogeneous effects interpretation of the Bartik instrument. Because the Bartik instrument combines multiple unordered instruments, it is difficult to allow unrestricted heterogeneity of the form discussed in Imbens and Angrist (1994) and ensure interpretable estimates.[10] Specifically, assuming monotonicity as in Imbens and Angrist (1994) is not sufficient to ensure estimates reflect non-negative weights on the un-

[10]To see why industry shares are unordered instruments, note that increasing the share of an industry can increase the predicted growth rates in some locations and decrease it in others depending on which industry share decreases to offset.

derlying heterogeneity. For further lucid discussion of these issues, see Kirkeboen, Leuven, and Mogstad (2016), among others. Instead, we impose a restricted form of linear heterogeneity and then state assumptions to ensure interpretable just-identified estimates. We also emphasize that even if each just-identified IV estimate produces a convex combination of heterogeneous effects, the overall Bartik instrument can produce negative weights if there are negative Rotemberg weights.

### 4.1 Setup with restricted heterogeneity

We follow Borusyak, Hull, and Jaravel (2018) and expand our model to include location specific coefficients.[11] Formally, consider the structural model:

$$y_l^{\perp} = x_l^{\perp} \beta_l + \epsilon_l^{\perp}, \tag{4.1}$$

where now $\beta_l$ replaces $\beta_0$.[12] We also assume the following linear relationship between $z_{lk}$ and $x_l$:

$$x_l^{\perp} = z_{lk}^{\perp} \pi_{lk} + u_{lk}^{\perp}, \tag{4.2}$$

where $\pi_{lk}$ is the location-industry specific first-stage coefficient and $u_{lk}^{\perp}$ is the location-industry specific error. We assume that $\beta_l$ is a random variable with well-defined moments.

Relative to Imbens and Angrist (1994), this setup is restricted because it assumes constant linear effects within a location over the whole support of $x_l^{\perp}$. One substantive restriction it imposes is that identically sized shocks have identical effects regardless of the level of employment in the location.

We now impose assumptions which are sufficient to ensure that in this linear model the weights on the $\beta_l$ are all weakly positive. In this sense, they are analogous to monotonicity assumptions in non-parametric models:

ASSUMPTION 3. *1. For each $k$, $\pi_{lk}$ is (weakly) the same sign for all $l$.*

*2. $\mathbb{E}[z_{lk}^{\perp} u_{lk}^{\perp} \beta_l] = 0$.*

We now state the result that the just-identified IV estimates represents a convex combination of the $\beta_l$:

PROPOSITION 4.1. *Suppose that equations (4.1) and (4.2) are true, and Assumption 3 holds, then*

[11] Adao, Kolesar, and Morales (2018) include location-industry coefficients. For simplicity, we maintain location specific coefficients.

[12] We focus on a single time period, but these points generalize.

*we can write:*

$$\hat{\beta}_k = \sum_l \omega_{lk}\beta_l + o_p(1), \tag{4.3}$$

*where*

$$\omega_{lk} = \frac{z_{lk}^{\perp,2}\pi_{lk}}{\sum_{l'} z_{l'k}^{\perp,2}\pi_{l'k}} \geq 0 \textit{ and } \sum_l \omega_{lk} = 1.$$

This result explains why in the presence of heterogeneity using different instruments (i.e., $z_{lk}$) would generate different point estimates (i.e., $\hat{\beta}_k$) even without misspecification. Each instrument estimates a parameter that is a different weighted combination of location-specific parameters. Because these parameters differ (i.e., there is heterogeneity), different instruments generate different estimates.

### 4.2 The Bartik estimator with heterogeneity

In this heterogeneous effects interpretation of Bartik, we can combine the Rotemberg weights and the $\omega_{lk}$ to write the Bartik estimate in terms of the location-specific coefficients:

$$\hat{\beta}_{Bartik} = \sum_l \beta_l \sum_k \alpha_k \omega_{lk} + o_p(1). \tag{4.4}$$

When $\sum_k \alpha_k \omega_{lk}$ is non-negative for all $l$, the Bartik estimator thus reflects a convex combination of the $\beta_l$. When are these weights non-negative? In the previous section, we discussed assumptions such that the $\omega_{lk}$ are non-negative. These assumptions, however, do not imply that the $\alpha_k$ are all positive. Thus, negative $\alpha_k$ are possible, which raises the possibility (but does not necessarily imply) non-convex weights on the $\beta_l$, in which case the overall Bartik estimate does not have a LATE-like interpretation as a weighted average of treatment effects.

When are negative weights on the $\beta_l$ likely to arise? We note first that we cannot estimate the $\omega_{lk}$ and hence we cannot directly compute the weights on the $\beta_l$. We can, however, estimate the $\alpha_k$ and the $\beta_k$, and use information in these two estimates to gauge the possibility of negative weights on the $\beta_l$.

If the $\hat{\beta}_k$ are all similar, then the negative weights on the $k$ are unlikely to generate negative weights on the $\beta_l$. The reason is that the similarity of the $\beta_k$ suggests that the $\omega_{lk}$ are similar across $k$, so that each instrument is likely estimating a similar weighted combination of effects. Hence, the negative $\alpha_k$ are likely just subtracting off the same $\beta_l$, with the overall weight on each $\beta_l$ remaining positive.

In contrast, if the $\beta_k$ are very different, then the $\omega_{lk}$ are different across $k$ and each instrument is estimating a different weighted combination of effects. It is then more likely that there are negative weights on the $\beta_l$, as the negative $\alpha_k$ place weight on $\beta_l$ that do not receive positive weight from other instruments. A way to assess the quantitative importance of these negative weights is to split the instruments into those with positive and negative $\alpha_k$ and compare their weighted sums; i.e., to compare $\sum_{k|\alpha_k>0} \hat{\alpha}_k \hat{\beta}_k$ and $\sum_{k|\alpha_k<0} \hat{\alpha}_k \hat{\beta}_k$. If the weighted sum of the instruments with the negative $\alpha_k$ is relatively large, then it is more likely that there are negative weights on the $\beta_l$ that are important in the overall estimate.

# 5 Testing the plausibility of the identifying assumptions

The identifying assumptions necessary for consistency are typically not directly testable. However, it is possible to partially assess their plausibility. We focus on the assumptions from Section 2; in the context of the canonical setting of estimating the inverse elasticity of labor supply, the identifying assumption is that initial industry composition ($Z_{l0}$) does not predict innovations to labor supply ($\epsilon_{lt}$).[13]

## 5.1 Empirical Test 1: Correlates of industry composition

It is helpful to explore the relationship between industry composition and location characteristics that may be correlated with innovations to supply shocks. This relationship provides an empirical description of the variation and the types of mechanisms that may be problematic for the exclusion restriction. In particular, the key question researchers should have in mind is whether the correlates of the levels of the shares predict *changes* in the outcome. For the empirical strategy to be valid, it is fine if the level of the correlates are related to the level of the outcome.

Since we argued in footnote 5 that it is typically desirable to fix industry shares to an initial time period ($Z_{l0}$), we recommend considering the correlation with initial period characteristics, as this reflects the instruments' cross-sectional variation. If $Z_{l0}$ is correlated with potential confounding factors, this can imply that there are omitted variables biasing estimation. Naturally, it is always possible to control for observable confounders, but following the logic of Altonji, Elder, and Taber (2005) and Oster (Forthcoming), movements in point estimates when conditioning on observable confounders suggest the potential importance of unobserved confounders. Looking at industries with the largest Rotemberg weights focuses attention on the instruments where confounding variables are most problematic.

[13] In Appendix F, we write down an economic model which allows us to derive this statement more precisely.

### 5.2 Empirical Test 2: Pre-trends

In some applications, there is a policy change in period $s_0$. As we discussed in Section 1.2, a researcher can use this sharp policy change to implement a difference-in-differences research design. The analogy to difference-in-differences is most straightforward when the shares are fixed over time (emphasizing the point in footnote 5). In this case, the industry shares measure the exposure to the policy change, while the national growth rates proxy for the size of the policy change.[14] In these settings, it is natural to test for pre-trends. We recommend looking at pre-trends in terms of the instruments with the largest Rotemberg weights, as well as looking at pre-trends in terms of the overall Bartik instrument. We suspect that researchers will be more comfortable with the plausibility of their empirical design if parallel pre-trends are satisfied for the instruments to which their estimates are most sensitive to misspecification. For more details on pre-trends tests, see DiNardo and Lee (2011). We additionally present examples below.

### 5.3 Empirical Test 3: Alternative estimators, overidentification tests, and patterns of heterogeneity

So far, we have emphasized that the Bartik estimator combines many moment conditions with a particular weight matrix. In this section, we discuss how researchers can use these moment conditions. Broadly speaking, there are two directions that a researcher can go. Under homogeneous effects, researchers can consider alternative estimators that combine the moment conditions in potentially more efficient ways. Additionally, researchers can use overidentification tests. If alternative estimators yield different estimates and overidentification tests reject, then these findings point to misspecification. In contrast, under heterogeneous effects, each instrument will converge to a different estimate (say, $\beta_k$) as discussed in Section 4. Under this assumption, it is important that the patterns of heterogeneity make sense, and we discuss some ways of assessing this.

**Homogeneous effects:** We begin in a world of homogeneous effects. Because the overidentified TSLS estimator is biased in finite samples, we encourage researchers to use three alternative estimators which have better properties with many instruments: the Modified Bias-corrected TSLS (MBTSLS) estimator from Anatolyev (2013) and Kolesar et al. (2015), the Limited Information Maximum Likelihood (LIML) estimator, and the HFUL estimator from Hausman et al. (2012). These estimators may not give the same estimates, as their underlying assumptions are different. Comparing these estimates, along with the Bartik

[14] Some examples of this include Autor, Dorn, and Hanson (2013) and Lucca, Nadauld, and Chen (Forthcoming).

TSLS estimate, provides a useful first pass diagnostic for misspecification concerns. If these estimators agree, then researchers can be more confident in their identifying assumption. Below, we follow Kolesar et al. (2015, pg. 481-2) and interpret differences between HFUL and LIML on the one hand, and MBTSLS and TSLS on the other, as pointing in the direction of potential misspecfication. The reason is that LIML and HFUL are maximum likelihood estimators and so exploit cross-equation restrictions while both MBTSLS and TSLS are two-step estimators and so do not exploit these cross-equation restrictions.

Overidentification tests provide more formal tests for misspecification. These estimators permit test statistics under different assumptions. For the HFUL estimator, we suggest the overidentification test from Chao et al. (2014), for LIML estimator, we use the Anderson, Rubin et al. (1950) chi-squared test and for TSLS we use the Sargan (1958) chi-squared test.[15] Conceptually, the overidentification test asks whether the instruments are correlated with the error term beyond what would be expected by chance, and relies on the validity of at least one of the instruments.

**Heterogeneous effects:** When overidentification tests reject, and when HFUL and LIML differ from MBTSLS and Bartik TSLS, under homogeneous effects these findings point to misspecification. An alternative interpretation of these results is that they point to heterogeneous effects of the form we outlined in Section 4. Under these assumptions, researchers may wish to probe the patterns of heterogeneity and see if there is a reasonable interpretation.

We now outline a visual diagnostic to help researchers assess the pattern of heterogeneity. The fundamental feature of the data that illustrates the heterogeneity is to consider the distribution of the just identified IV estimates (i.e., the $\hat{\beta}_k$). In order to visualize this dispersion, we advocate a particular figure. Here we describe the figure and discuss our reasoning, and below we present examples of it (see Figures 1, 3, and 6).[16] Briefly, the x-axis is the first-stage F-statistic and the y-axis is the $\hat{\beta}_k$ associated with each instrument. So as to not visually overstate dispersion, the figure only includes instruments with reasonable first-stage power (in our applications, we plot instruments with first-stage F-statistics greater than 5). To show how the $\hat{\beta}_k$ compare to the Bartik estimate, the figure include a horizontal line that reflects the overall Bartik estimate. Because first-stage power does not perfectly explain the Rotemberg weights, we weight the individual points of $\beta_k$ by the size of the $\alpha_k$ from the Bartik Rotemberg weights. Finally, to illustrate the role of negative Rotemberg weights, we shade the points differently depending on the sign of the Rotemberg weights.

[15]Code to implement the HFUL overid test is available on request and will be posted on Github.

[16]Code to create this figure is included in the package that computes the Rotemberg weights and will be posted on Github.

Researchers can use this figure to think about three questions. First, why do the overidentification tests reject, and what industries drive the rejection? Intuitively, a researcher might be less concerned by a rejection where the $\beta_k$ are less rather than more dispersed around the Bartik estimate. Similarly, the figure helps isolate which industries are driving the failure of overidentification tests. Researchers should feel comfortable with why the comparisons implied by some instruments are outliers relative to the comparisons implied by other instruments. Second, why does the Bartik estimate end up where it does relative to the underlying $\beta_k$? The relative Rotemberg weights help explain why the Bartik estimate lies where it does relative to the underlying distribution. As we emphasized in Section 3, a researcher should feel comfortable that the largest Rotemberg weight industries make sense with the causal mechanism in the paper. Third, how plausible is it that there are negative weights on some $\beta_l$? Visualizing the industries with the negative Rotemberg weights helps to highlight which industries would potentially generate negative weights on $\beta_l$, as we discussed further in Section 4. Naturally, whether the patterns of heterogeneity make sense will rely on application-specific knowledge, and so we view this figure as providing a useful starting point for an application-specific investigation, rather than an ending point.

**A comment on alternative approaches to overidentifying tests** An alternative approach to overidentifying tests (e.g., by Beaudry, Green, and Sand (2012) and others) is to construct multiple Bartik instruments using different vectors of national growth rates, and then testing whether these different weighted combinations of instruments estimate the same parameter. Often, the correlation between the Bartik instruments constructed with different growth rates is quite low. This fact is interpreted as reassuring because it suggests that exploiting "different sources of variation" gives the same answer.

We recommend instead that researchers use the Rotemberg weights to quantify what variation each Bartik instrument is using, and whether the two Bartik instruments use different sources of variation. Specifically, researchers can report the top-5 Rotemberg weights across the two instruments and also their rank correlation. If these statistics are low, then the two Bartik instruments are likely using different sources of variation and the conclusion discussed above is warranted.[17]

[17] To illustrate the theoretical distinction between looking at correlations between Bartik instruments and comparing Rotemberg weights implied by the two instruments, in Appendix G we produce an example where only one industry has identifying power, but the two instruments are uncorrelated and find the same $\hat{\beta}$. While this example might seem like a theoretical curiosity, in our empirical settings we typically find that a small number of industries provide most of the identifying variation and the variation in the growth rates explains little of the variation in the Rotemberg weights. Hence, there is typically scope for different national growth rates that produce weakly correlated Bartik instruments to rely on the same "identifying variation" (that is, have similar Rotemberg weights).

# 6 Empirical example I: Canonical Setting

We now present three empirical examples to make our theoretical ideas concrete, focusing on our empirical tests from Section 5. Our first example is the canonical setting of estimating the inverse elasticity of labor supply. We begin by reporting the main estimates and then report the industries with the highest Rotemberg weight. We then probe the plausibility of the identifying assumption for these instruments.

## 6.1 Dataset

We use the 5% sample of IPUMS of U.S. Census Data (Ruggles et al. (2015)) for 1980, 1990 and 2000 and we pool the 2009-2011 ACSs for 2010. We look at continental US commuting zones and 3-digit IND1990 industries.[18] 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. We weight all regressions by 1980 population.

We use the leave-one-out means to construct the national growth rates to address the finite sample bias that comes from using own-observation information. Specifically, using own-observation information allows the first-stage to load on the idiosyncratic industry-location component of the growth rate, $\tilde{g}_{lk}$, which is endogeneous. This finite sample bias is generic to overidentified instrumental variable estimators and is the motivation for jackknife instrument variable estimators (e.g. Angrist, Imbens, and Krueger (1999)). In practice, because we have 722 locations, using leave-one-out to estimate the national growth rates matters little in point estimates (compare rows 2 and 3 in Table 3).[19]

## 6.2 Rotemberg weights

We compute the Rotemberg weights of the Bartik estimator with controls, aggregated across time periods. The distribution of sensitivity is skewed, so that a small number of instruments have a large share of the weight. Table 1 shows that the top five instruments account for over fourty percent ($^{0.593}/_{1.375}$) of the positive weight in the estimator. These top five

[18] There are 228 non-missing 3-digit IND1990 industries in 1980. There are 722 continental US commuting zones.

[19] In Appendix H, we show that with a leave-one-out estimator of the $g_k$ component, the Rotemberg weights do not sum to one. In our applications below, when we compute the Rotemberg weights we use simple averages so that the weights sum to one.

instruments are: oil and gas extraction, motor vehicles, other,[20] guided missiles, and blast furnaces.

These weights give a way of describing the research design that reflects the variation in the data that the estimator is using, and hence makes concrete for the reader what types of deviations from the identifying assumption are likely to be important. In this canonical setting, one of the important comparisons is across places with greater and smaller shares of oil and gas extraction. Hence, the estimate is very sensitive to deviations from the identifying assumption related to geographic variation in employment share in oil and gas extraction. Interestingly, a common short-hand to talk about Bartik is to discuss the fate of the automobile industry (e.g. Bound and Holzer (2000, pg. 24)), and this analysis confirms that the motor vehicle industry plays a large role in the Bartik instrument.

Finally, Panel B shows that the national growth rates are weakly correlated with the sensitivity-to-misspecification elasticities. Hence, the growth rates provide a poor guide to understanding what variation in the data drives estimates. In contrast, the elasticities are quite related to the variation in the industry shares across locations (Var($z_{lk}$)). This observation explains why the industries with high weight tend to be tradables: almost by definition, tradables have industry shares that vary across locations, while non-tradables do not.[21]

### 6.3 Discussion of the identifying assumption in terms of the shares

As we discussed in Section 2, a heuristic for figuring out which identifying assumption researchers have in mind is whether they mention particular industries. It is common in the canonical setting to discuss particular industries (e.g., as mentioned above, Bound and Holzer (2000, pg. 24) discuss the automobile industry). Hence, we think that in many settings researchers have in mind this differential exposure design.

### 6.4 Testing the plausibility of the identifying assumption

**Test 1: Correlates of 1980 industry shares** Table 2 shows the relationship between 1980 characteristics of commuting zones and the share of the top 5 industries in Table 1, as well the overall Bartik instrument using 1980 to 1990 growth rates. First, the $R^2$ in these regressions are quite high: for example, we can explain 46% of the variation in share of the "other" industry via our covariates. Second, "other," oil and gas extraction, blast furnaces,

[20] The "Other" industry is the "N/A" code in the IND1990 classification system. Our understanding is that in 1980 the "Other" code includes full-time military personnel. Hence, in 1990 and 2000, we place full-time military personnel in the "Other" category to compute growth rates.

[21] This logic is the basis of Jensen and Kletzer (2005)'s measure of the offshorability of services; as Jensen and Kletzer (2005) recognize, there are other reasons for concentration besides tradability.

and the overall Bartik instrument are statistically significantly correlated with the share of native-born workers. In the immigrant enclave literature, the share of native born (i.e., the complement of the immigrant share) is thought to predict labor supply shocks.

**Test 2: Parallel pre-trends** We note that in this setting there is no pre-period and so it is not possible to test for parallel pre-trends without further assumptions.

**Test 3: Alternative estimators and overidentification tests** Rows 1, 2 and 3 of Table 3 report the OLS and IV estimates (row 2 leaves out the own-CZ growth rate to construct the instrument, while row 3 uses all CZs to construct the growth rates), with and without for the 1980 covariates as controls and 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, though these are not statistically distinguishable.

Rows 4-7 of Table 3 report alternative estimators as well as overidentification tests. We focus on column (2), where we control for covariates. TSLS with the Bartik instrument and LIML are quite similar. This finding is typically viewed as reassuring. In contrast, TSLS and MBTSLS are similar, while HFUL is substantially larger. The different point estimates suggest the presence of misspecification. In column (4), we see that the overidentification tests reject the null that all instruments are exogenous, which also points to misspecification.

**Visualizing the overidentification tests** If one wishes to interpret the failure of the overidentification tests as pointing to heterogeneity of the form outlined in Section 4 rather than as evidence of misspecification, then Figure 1 shows some of the heterogeneity in treatment effects underlying the overall Bartik estimate (Appendix Figure A1 shows the relationship between the Rotemberg weights and the first-stage F-statistic). First, the figure shows that among the "high-powered" (i.e., those with a first stage F-statistic above five) industries, there is substantial dispersion around the Bartik $\hat{\beta}$. Second, the largest weight industries do tend to be closest to the overall Bartik $\hat{\beta}$. Third, if a researcher wishes to adopt a heterogeneous effects interpretation of the rejection of the null in the overidentification tests, then the patterns of heterogeneity suggest that there are likely to be negative weights on some of the underlying location-specific coefficients. In particular, there is substantial dispersion in the $\hat{\beta}_k$ and some of the outlier $\hat{\beta}_k$ have negative weights. Thus, the underlying location-specific effects (the $\beta_l$) that lead to a negative coefficient likely receive negative weights so that the overall Bartik estimate does not reflect convex weights. To see this more generally, the Panel E of Table 1 shows that the mean of the $\beta_k$ among the negative weight industries is very different than the mean of the $\beta_k$ among the industries with positive weights.

# 7 Empirical example II: China shock

We estimate the effect of Chinese imports on manufacturing employment in the United States using the China shock approach of Autor, Dorn, and Hanson (2013) (ADH).

## 7.1 Specification

It is helpful to write the main regression specification of ADH in our notation. The paper is interested in a regression (where we omit covariates for simplicity, but include them in the regressions):

$$y_{lt} = \beta_0 + \beta x_{lt} + \epsilon_{lt}, \tag{7.1}$$

where $y_{lt}$ is the percentage point *change* in manufacturing employment rate, and $x_{lt} = \sum_k z_{lkt} g_{kt}^{US}$ is import exposure, where $z_{lkt}$ is contemporaneous start-of-period industry-location shares, and $g_{kt}^{US}$ is a normalized measure of the *growth* of imports from China to the US in industry $k$. The first stage is:

$$x_{lt} = \gamma_0 + \gamma_1 B_{lt} + \eta_{lt}, \tag{7.2}$$

where $B_{lt} = \sum_k z_{lkt-1} g_{kt}^{\text{high-income}}$, the $z$ are lagged, and $g_{kt}^{\text{high-income}}$ is a normalized measure of the *growth* of imports from China to other high-income countries (mainly in Europe).

We focus on the TSLS estimate in column (6) of Table 3 of ADH, which reports that a \$1,000 increase in import exposure per worker led to a decline in manufacturing employment of 0.60 percentage points. Our replication also produces a coefficient of 0.60 (see Table 6, TSLS (Bartik) row, column (2)) .

## 7.2 Rotemberg weights

As in the canonical setting, despite a very large number of instruments (397 industries) the distribution of sensitivity is skewed so that a small number of instruments get a large share of the weight. Table 4 shows that the top five instruments receive over half of the absolute weight in the estimator ($^{0.532}/_{1.067}$). These instruments are electronic computers, games and toys, household audio and video, telephone apparatus, and computer equipment. Except for games and toys, these industries are different than the ones that ADH emphasize when motivating the empirical strategy.[22] In particular, rather than being low-skill techno-

[22] "The main source of variation in exposure is within-manufacturing specialization in industries subject to different degrees of import competition...there is differentiation according to local labor market reliance on labor-intensive industries...By 2007, China accounted for over 40 percent of US imports in four four-digit SIC industries (luggage, rubber and plastic footwear, games & toys, and die-cut paperboard) and over 30 percent

logically stagnant industries where it is plausible that trade is the main shock hitting the industry, these are higher-skill technologically innovative industries where it is plausible that changes in technology are the main shock hitting the industry.

Relative to the canonical setting, negative weights are less prominent and the variation in the national growth rates (or, imports from China to other high-income countries) explains more of the variation in the sensitivity elasticities. Even so, and consistent with the discussion in the previous paragraph that the growth rates provide an imperfect guide to which industries drive estimates, the $g_k$ component explains less than twenty percent ($0.430^2$, see Table 4, Panel B) of the variance of the Rotemberg weights.

### 7.3 Discussion of the identifying assumption in terms of the shares

Why is it reasonable to interpret this paper as being about the shares? We note first that the paper does not emphasize having a large number of independent shocks (which would be necessary for the shocks interpretation to be plausible). Indeed, it is hard to conceive of a model of an "optimizing China" that would generate random patterns of exports across a wide swathe of the economy. (The random shocks assumption is more plausible in this setting if researchers control for "higher-level" fixed effects and so exploit more idiosyncratic variation. When Borusyak, Hull, and Jaravel (2018, Table 1, column 6) control for 2 digit industries, the estimates are one-sixth the size and no longer statistically significant.) Second, the paper emphasizes particular industries and industries with particular characteristics. That is, our reading of the logic of the paper is that it emphasizes that Chinese exports were concentrated in low-skill, labor-intensive industries. This focus on particular industries is not consistent with identification coming from the shocks.

Is the identification assumption necessarily implausible when viewed in terms of shares? We do not think so. Other papers in the trade literature leverage changes in trade policy and study local labor market effects of these policy changes (e.g., Topalova (2010), Kovak (2013), and Pierce and Schott (2016)). In these papers, the argument is not that the trade policy is literally random, but that the change in trade policy is not correlated with pre-existing trends in outcomes at the local level. The argument does not require that the shares predict nothing in levels, but simply that the shares only predict changes through the causal channel emphasized by the paper.

In the trade policy example, there is some institutional reason to expect that there is a shock in particular industries that only operated through trade policy (because trade policy changed in these industries). By analogy, in the context of ADH this logic would suggest using institutional knowledge to pick industries where there was a large increase in exports

in 28 other industries, including apparel, textiles, furniture, leather goods, electrical appliances, and jewelry" (pg. 2123).

from China because of Chinese comparative advantage (rather than technological change in the industry). Seen in this light, one motivation for ADH to look at imports to other high-income countries might be to isolate the industries where there is strong reason to think that China experienced rapid productivity gains. As we have emphasized, however, the weights that the Bartik estimator places on moments are not solely a function of the growth rates. Indeed, in this example, the growth rates explain less than twenty percent of the variation in the Rotemberg weights. As a result, weighting the shares by growth rates is an imperfect way of isolating the variation that the researcher intends. A research design based on the shares would likely accord more closely with the goals of researchers if there was further pruning of the industries.

### 7.4 Testing the plausibility of the identifying assumption

**Test 1: Correlates of 1980 industry shares** Table 5 shows the relationship between the covariates used in ADH and the top industries reported in Table 4. First, relative to the canonical setting, the controls explain less of the variation in shares (lower $R^2$ in the regressions). Second, electronic computers, computer equipment manufacturing as well as the overall measure are concentrated in more college educated areas; in contrast, games and toys is concentrated in places with fewer college educated workers. This pattern emphasizes that researchers should be concerned about other trends potentially affecting manufacturing employment in more educated areas. Interestingly, the identifying assumption related to the computer industry is precisely one that ADH worry about and provide sensitivity analyses related to this industry.[23]

**Test 2: Parallel pre-trends** We construct our pre-trend figures as follows. We use fixed 1980 shares as the instruments, and plot the reduced form effect of each industry on manufacturing employment.[24] We then convert the growth rates to levels and we index the levels in 1970 to 100. Standard errors are constructed using the delta method. For the aggregate Bartik, we use the industry shares fixed in 1980, and combine them using growth rates from 1990 to 2000.

Figure 2 shows the plots and displays several interesting patterns. First, all of the panels diverge from classic pre-trends figures, which show no trends in the pre-periods and then a sharp change at the date of the treatment. Second, as was true in the covariates in Table 5, the patterns in electronic computers (Panel A) and computer equipment (Panel E) are similar to the aggregate, with the decline in manufacturing from 1990 to 2007 undoing

[23] ADH (pg. 2138): "Computers are another sector in which demand shocks may be correlated [across countries], owing to common innovations in the use of information technology."

[24] We use the reduced-form effect because the endogenous variable is not available in the earlier periods. See Appendix Figure A4 for the analogous figures using fixed 1990 shares.

growth from 1970 to 1990. Note that these panels show comparisons of places with more and less of these particular industries in 1980, while the outcome is employment *for all manufacturing industries*.

**Test 3: Alternative estimators and overidentification tests** Rows 1 and 2 of Table 6 report the OLS and IV estimates using Bartik, with and without for the 1980 covariates as controls, though these are not statistically distinguishable for the IV estimates. Rows 3-6 of Table 6 shows alternative estimators as well as overidentification tests. We focus on column (2), where we control for covariates. The estimates range from half the size of the baseline Bartik TSLS estimate (MBTSLS), to several times the size (LIML). The divergence between the two-step estimators (TSLS with Bartik, TSLS and MBTSLS) and the maximum likelihood estimators (LIML and HFUL) is evidence of misspecification. Similarly, the overidentification tests reject. Combined, the movement in the estimates across estimators is not reassuring,[25] and the failure of the overidentification tests points to potential misspecification.

**Visualizing the overidentification tests** If one wishes to interpret the failure of the overidentification tests as pointing to heterogeneity of the form outlined in Section 4 rather than as evidence of misspecification, then Figure 3 shows some of the heterogeneity in treatment effects underlying the overall estimate (Appendix Figure A2 shows the relationship between the Rotemberg weights and the first-stage F-statistic). Relative to the canonical case, the patterns of heterogeneity are less concerning. In particular, visually there is less dispersion in the point estimates among the high-powered industries and the high-weight industries are clustered more closely to the overall point estimate. Finally, while there are negative Rotemberg weights, these industries are a small share of the overall weight, suggesting that there are unlikely to be negative weights on particular location-specific parameters (i.e., $\beta_l$; see also Panel E in Table 4).

# 8 Empirical example III: Immigrant enclave

We estimate the (negative) inverse elasticity of substitution between immigrants and natives following Card (2009). In particular, we focus on the results in Table 6 of that paper (in particular columns (3) and (7)), which provides two sets of results: one for high-school equivalent workers, and one for college-equivalent workers.

[25] Angrist and Pischke (2008, pg. 213) write: "Check overidentified 2SLS estimates with LIML. LIML is less precise than 2SLS but also less biased. If the results come out similar, be happy. If not, worry..."

## 8.1 Specification

It is helpful to convert Card (2009)'s specification into our notation. The paper is interested in a regression:

$$y_{lj} = \beta_0 + \beta \ln x_{lj} + \beta_2 \mathbf{X}_l + \epsilon_{lj}, \tag{8.1}$$

where $l$ is a location (a city) and $j$ is a skill group (either high school- or college-equivalent). Here, $y_{lj}$ is the residual log wage gap between immigrant and native men in skill group $j$, $x_{lj}$ is the ratio of immigrant to native hours in skill group $j$ (of both men and women), and $\mathbf{X}_l$ is a vector of city-level controls. Hence, $\beta$ is the (negative) inverse elasticity of substitution between immigrants and natives in the relevant skill group. Unlike other examples, the controls do not include place and time fixed effects because the paper considers a single cross-section of outcomes in 2000 in 124 cities. The paper does, however, explore robustness to including the lagged dependent variable.

The first stage is:

$$\ln x_{lj} = \gamma_0 + \gamma_1 B_{lj} + \gamma_2 \mathbf{X}_l + \eta_l \tag{8.2}$$

where $B_{lj} = \sum_k z_{lk,1980} g_{kj}$. Here, $z_{lk,1980} = \frac{N_{lk,1980}}{N_{k,1980}} \times \frac{1}{P_{l,2000}}$, where $N_{k,1980}$ is the number of immigrants from one of 38 country (groups) $k$ in the U.S. in 1980, $N_{lk,1980}$ is the number of immigrants from country (group) $k$ in location $l$ in 1980, and $P_{l,2000}$ is the population of location $l$ in 2000. Here, $g_{kj}$ is the number of people arriving in the US from 1990 to 2000 from country (group) $k$ and skill group $j$. Notice that the shares, the immigrant enclave, are not skill-specific, while the shocks, the immigrant inflows, are skill-specific. Relative to our other examples, the shares do not sum to one within a location.

## 8.2 Rotemberg weights

In this setting, there are 38 country groups. For high school-equivalent workers, Panel A of Table 7 shows that the top country is Mexico, which by itself receives almost half the weight, and the the top five countries (in order: Mexico, El Salvador, Philippines, China, and country group of West Europe, Israel, Cyprus, Australia and New Zealand) get almost two-thirds of the overall weight. The large weight on Mexico is perhaps unsurprising. Indeed, Card (2009, pg. 9) emphasizes that one might be concerned that for high-school equivalent workers the instrument is largely just initial Mexican immigrant shares. Unlike in the other examples, all the weights are positive. One reason the weights accord so closely with intuition is that for this instrument the weights are almost perfectly explained by the shocks—the immigrant inflows. Panel II shows that the correlation between the weights and the $g_k$ is 0.991, which is dramatically higher than in the other examples.

For college-equivalent workers, Panel B of Table 7 shows that the top five sending countries receive almost half (45%) of the weight and all the weights are positive. The top five countries are similar to the high-school equivalent workers, with El Salvador replaced by Cuba. The top country is the Philippines, with fifteen percent of the weight. Relative to our other examples, the shocks have much more explanatory power for the weights (the shocks explain about 60% ($= 0.766^2$) of the weights), though this explanatory power is lower than for the high-school equivalent workers.

### 8.3 Discussion of the identifying assumption in terms of the shares

We think that it is typically reasonable to interpret the immigrant enclave setting as having an identifying assumption in terms of the shares. The Card (2009) setting considers a single cross-section but emphasizes the analogy to difference-in-differences by showing robustness to controlling for the lagged dependent variable so that the effect of the instrument is similar to changes. More broadly, a natural way to think of the immigrant enclave instrument is that in any period there are immigrants arriving from different countries and this then naturally affects places differently. For example, even though in Card (1990) the boatlift was not caused by trends in Miami, the shock only hits Miami because of the strong "pull" factor of the immigrant enclave and the discussion of identification is thus about whether Miami would counterfactually have evolved similarly to places without an existing stock of Cuban immigrants. We view it is as reasonable to interpret the immigrant enclave instrument—especially when applied to a particular time period—as pooling this logic. Hence, a researcher should explain and defend why places with different initial stocks of immigrants would have counterfactually evolved in a similar way.

If a researcher does not feel comfortable embracing the shares view, then it is important to understand what the shocks view means in this setting. To embrace the shocks view of identification in the immigrant enclave setting requires not only that there are random "push" factors, but also that there are enough independent push factors that the endogeneity of the shares averages out. Making this case typically requires a large number of independent "push" factors.

### 8.4 Testing the plausibility of the identifying assumption

**Test 1: Correlates of 1980 origin country shares** Table 8 shows the relationship between the 1980 covariates used in Card (2009) and the top origin countries reported in Table 7. First, similar to the canonical setting, the characteristics explain a fair amount of the cross-sectional variation in the shares—especially for the overall instrument. Second, and related to the canonical setting, we tend not to find a significant relationship between manufacturing share and any of the individual country shares or the aggregate instruments (the only exception is West Europe and others).

**Test 2: Parallel pre-trends** We construct our pre-trend figures by replacing the endogenous and outcome variables in equations (8.1) and (8.2) with their 1980, 1990 and 2000 values (that is, we include all the controls in Card (2009) in Table 6, columns (3) and (7)). Hence, the 2000 coefficient corresponds to the $\hat{\beta}_k$ in Table 7 and the TSLS (Bartik) $\hat{\beta}$ in Table 9.

Figure 4 shows that for the high school equivalent native-immigrant wage gap, the variation in 1980 shares of Mexican immigrants did not predict statistically or economically larger wage gaps in 1980 or 1990. That is, conditional on controls, the figures suggest that there was a shock in the 1990s that led to a widening gap in 2000. Given the large weight on Mexico, it is not surprising that the aggregate instrument looks like Mexico. Perhaps more suprisingly, all the other countries look similar to Mexico.

Figure 5 shows less reassuring patterns for the college equivalent regressions. To take the Philippines (the highest weight instrument) as an example, the 1980 variation in the share of people from the Philippines implies as large an effect of the native-immigrant ratio on the native-immigrant wage gap in 1980 and 1990 as in 2000. That is, there is no evidence of change in 2000. Similarly, for other countries and the aggregate there are statistically significant pre-trends. This evidence is consistent with the argument in Jaeger, Ruist, and Stuhler (2018) that the immigrant inflows are typically serially correlated and so the immigrant enclave instrument does not generate a well-defined shock to the supply of immigrants.

**Test 3: Alternative estimators and overidentification tests** Panel A of Table 9 shows the results of alternative estimators and some overidentification tests for high school equivalent workers. Unlike in our other examples, the results are quite stable across estimators, with Bartik, TSLS, LIML, and MBTSLS all giving the same point estimate (HFUL, in contrast, is quite different). Similarly, the overidentification tests on the TSLS estimator fail to reject (though on LIML it does). This result can be approximately anticipated from Table 7 where the $\beta_k$ on each individual instrument are quite similar.

Panel B of Table 9 shows that the results are broadly similar for college equivalent workers. Namely, the results are quite stable across estimators and the overidentification test fails to reject for both TSLS and LIML. Again, this result can be approximately anticipated from Table 7.

**Visualizing the overidentification tests** Given that for several of the estimators the overidentification tests fail to reject, it is not surprising that visually there is not a great deal of dispersion in the point estimates across instruments. Figure 6 shows the heterogeneity in the $\hat{\beta}_k$ and the relationship to the first stage f-statistic. To compare to our other examples, note that the y-axis is dramatically compressed. Moreover, the high-weight industries are all very close to the overall estimate.

## 9 Summary

The central contribution of this paper revolves around understanding identification and the Bartik instrument. Our first set of formal results relate to identification in the sense typically used by econometricians. We show that Bartik is numerically equivalent to a GMM estimator with the industry shares as instruments. We use this equivalence to argue that in many settings the way to interpret the research design implicit in a Bartik instrument is a pooled exposure design. The shares measure the differential exposure to common shocks (the national growth rates), and so the relevant identification assumption—familiar from difference-in-differences—is that there are no other shocks correlated with this differential exposure.

Our second set of formal results relate to identification in the sense often used by practitioners: we show how to compute which of the many instruments "drive" the estimates. Building on Andrews, Gentzkow, and Shapiro (2017) we show that these weights can be interpreted as sensitivity-to-misspecification elasticities and so highlight which identifying assumptions are most worth discussing and probing.

We then elaborated on a number of specification tests that researchers can carry out, and illustrated these tests through a number of applications. Our results clarify the set of reasonable concerns a consumer of the Bartik literature should have. We hope that researchers will use the results and tools in this paper to be clearer about how identification works in their papers: both in the econometric sense of stating the identifying assumption and in the practical sense of showing what variation drives estimates.

## References


Acemoglu, Daron and Joshua Linn. 2004. "Market Size in Innovation: Theory and Evidence from the Pharmaceutical Industry." *Quarterly Journal of Economics* 119 (3):1049–1090.

Adao, Rodrigo, Michal Kolesar, and Eduardo Morales. 2018. "Shift-Share Designs: Theory and Inference." Working paper.

Altonji, Joseph G., Todd E. Elder, and Christopher R. Taber. 2005. "Selection on Observed and Unobserved Variables: Assessing the Effectiveness of Catholic Schools." *Journal of Political Economy* 113 (1):151–184.

Anatolyev, Stanislav. 2013. "Instrumental variables estimation and inference in the presence of many exogenous regressors." *The Econometrics Journal* 16:27–72.

Anderson, Theodore W, Herman Rubin et al. 1950. "The asymptotic properties of estimates of the parameters of a single equation in a complete system of stochastic equations." *The Annals of Mathematical Statistics* 21 (4):570–582.

Andrews, Isaiah. Forthcoming. "On the Structure of IV Estimands." *Journal of Econometrics* .

Andrews, Isaiah, Matthew Gentzkow, and Jesse M. Shapiro. 2017. "Measuring the Sensitivity of Parameter Estimates to Estimation Moments." *Quarterly Journal of Economics* 132 (4):1553–1592.

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

Angrist, Joshua D and Jörn-Steffen Pischke. 2008. *Mostly harmless econometrics: An empiricist's companion*. Princeton university press.

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

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

Baum-Snow, Nathaniel and Fernando Ferreira. 2015. "Causal Inference in Urban and Regional Economics." In *Handbook of Regional and Urban Economics, Volume 5A*, edited by Gilles Duranton, J. Vernon Henderson, and William C. Strange. Elsevier, 3–68.

Beaudry, Paul, David A. Green, and Benjamin Sand. 2012. "Does Industrial Composition Matter for Wages? A Test of Search and Bargaining Theory." *Econometrica* 80 (3):1063–1104.

———. 2018. "In Search of Labor Demand." *American Economic Review* 108 (9):2714–57.

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

Borusyak, Kirill, Peter Hull, and Xavier Jaravel. 2018. "Quasi-experimental Shift-share Research Designs." Working paper.

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

Card, David. 1990. "The Impact of the Mariel Boatlift on the Miami Labor Market." *Industrial and Labor Relations Review* 43 (2):245–257.

———. 2009. "Immigration and Inequality." *American Economic Review: Papers and Proceedings* 99 (2):1–21.

Chao, John C., Jerry A. Hausman, Whitney K. Newey, Norman R. Swanson, and Tiemen Woutersen. 2014. "Testing overidentifying restrictions with many instruments and heteroskedasticity." *Journal of Econometrics* 178:15–21.

Chernozhukhov, Victor and Christian Hansen. 2008. "The reduced form: A simple approach to inference with weak instruments." *Economics Letters* 100:68–71.

Christian, Paul and Christopher B. Barrett. 2017. "Revisiting the Effect of Food Aid on Conflict: A Methodological Caution." Working Paper 8171, World Bank.

Conley, Timothy G., Christian B. Hansen, and Peter E. Rossi. 2012. "Plausibly Exogenous." *Review of Economics and Statistics* 94 (1):260–272.

DiNardo, John and David S Lee. 2011. "Program evaluation and research designs." In *Handbook of labor economics*, vol. 4. Elsevier, 463–536.

Freeman, Richard B. 1980. "An Empirical Analysis of the Fixed Coefficient "Manpower Requirement" Mode, 1960-1970." *Journal of Human Resources* 15 (2):176–199.

Greenstone, Michael, Alexandre Mas, and Hoai-Luu Nguyen. Forthcoming. "Do Credit Market Shocks affect the Real Economy? Quasi-Experimental Evidence from the Great Recession and 'Normal' Economic Times." *American Economic Journal: Economic Policy* .

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

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

Jaeger, David A., Theodore J. Joyce, and Robert Kaestner. Forthcoming. "A Cautionary Tale of Evaluating Identifying Assumptions: Did Reality TV Really Cause a Decline in Teenage Childbearing?" *Journal of Business and Economic Statistics* .

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

Jensen, J. Bradford and Lori G. Kletzer. 2005. "Tradable Services: Understanding the Scope and Impact of Services Offshoring." *Brookings Tade Forum* 2005:75–133.

Kearney, Melissa S. and Phillip B. Levine. 2015. "Media Influences on Social Outcomes: The Impact of MTV's 16 and Pregnant on Teen Childbearing." *American Economic Review* 105 (12):3597–3632.

Kirkeboen, Lars J., Edwin Leuven, and Magne Mogstad. 2016. "Field of Study, Earnings, and Self-Selection." *Quarterly Journal of Economics* 131 (3):1057–1111.

Kolesar, Michal, Raj Chetty, John Friedman, Edward Glaeser, and Guido W. Imbens. 2015. "Identification and inference with many invalid instruments." *Journal of Business and Economic Statistics* 33 (4):474–484.

Kovak, Brian K. 2013. "Regional Effects of Trade Reform: What is the Correct Measure of Liberalization?" *American Economic Review* 103 (5):1960–1976.

Lucca, David O, Taylor Nadauld, and Karen Chen. Forthcoming. "Credit supply and the rise in college tuition: Evidence from the expansion in federal student aid programs." *Review of Financial Studies* .

Monte, Ferdinando, Stephen J. Redding, and Esteban Rossi-Hansberg. 2018. "Commuting, Migration and Local Employment Elasticities." *American Economic Review* 108 (12).

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

Oster, Emily. Forthcoming. "Unobservable Selection and Coefficient Stability: Theory and Evidence." *Journal of Business and Economic Statistics* :1–18.

Perloff, Harvey S. 1957. "Interrelations of State Income and Industrial Structure." *Review of Economics and Statistics* 39 (2):162–171.

Pierce, Justin R. and Peter K. Schott. 2016. "The Surprisingly Swift Decline of US Manufacturing Employment." *American Economic Review* 106 (7):1632–1662.

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

Ruggles, Steven, Katie Genadek, Ronald Goeken, Josiah Grover, and Matthew Sobek. 2015. *Integrated Public Use Microdata Series: Version 6.0 [Machine-readable database]*. Minneapolis: University of Minnesota.

Sargan, John D. 1958. "The estimation of economic relationships using instrumental variables." *Econometrica: Journal of the Econometric Society* :393–415.

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

Table 1: Summary of Rotemberg weights: canonical setting

| **Panel A: Negative and positive weights** | | | | | |
|---|---|---|---|---|---|
| | Sum | Mean | Share | | |
| Negative | -0.368 | -0.004 | 0.212 | | |
| Positive | 1.368 | 0.010 | 0.788 | | |
| **Panel B: Correlations** | | | | | |
| | $\hat{\alpha}_k$ | $g_k$ | $\hat{\beta}_k$ | $\hat{F}_k$ | Var($z_k$) |
| $\hat{\alpha}_k$ | 1 | | | | |
| $g_k$ | -0.015 | 1 | | | |
| $\hat{\beta}_k$ | 0.017 | -0.495 | 1 | | |
| $\hat{F}_k$ | 0.476 | -0.032 | 0.016 | 1 | |
| Var($z_k$) | 0.549 | -0.036 | -0.003 | 0.316 | 1 |
| **Panel C: Variation across years in $\hat{\alpha}_k$** | | | | | |
| | Sum | Mean | | | |
| 1980 | 0.458 | 0.002 | | | |
| 1990 | 0.182 | 0.001 | | | |
| 2000 | 0.360 | 0.002 | | | |
| **Panel D: Top 5 Rotemberg weight industries** | | | | | |
| | $\hat{\alpha}_k$ | $g_k$ | $\hat{\beta}_k$ | 95 % CI | Ind Share |
| Oil+Gas Extraction | 0.229 | 0.034 | 1.170 | (0.80,1.90) | 0.568 |
| Motor Vehicles | 0.140 | -0.017 | 1.525 | (1.30,1.90) | 1.404 |
| Other | 0.091 | -0.062 | 0.759 | (0.10,1.70) | 1.697 |
| Guided Missiles | 0.069 | 0.047 | 0.115 | (-2.20,0.70) | 0.236 |
| Blast furnaces | 0.058 | -0.078 | 1.084 | (0.60,5.10) | 0.800 |
| **Panel E: Estimates of $\beta_k$ for positive and negative weights** | | | | | |
| | $\alpha$-weighted Sum | Share of overall $\beta$ | Mean | | |
| Negative | -0.074 | -0.061 | 1.622 | | |
| Positive | 1.290 | 1.061 | -0.584 | | |

*Notes:* This table reports statistics about the Rotemberg weights. In all cases, we report statistics about the aggregated weights with normalized growth rates, where we aggregate a given industry across years as discussed in Section 3.3 and normalize growth rates to the per-period average as discussed in Section 3.2. Panel A reports the share and sum of negative weights. Panel B reports correlations between the weights ($\hat{\alpha}_k$), the national component of growth ($g_k$), the just-identified coefficient estimates ($\hat{\beta}_k$), the first-stage F-statistic of the industry share ($\hat{F}_k$), and the variation in the industry shares across locations (Var($z_k$)). Panel C reports variation in the weights across years. Panel D reports the top five industries according to the Rotemberg weights. The $g_k$ is the national industry growth rate, $\hat{\beta}_k$ is the coefficient from the just-identified regression, the 95% confidence interval is the weak instrument robust confidence interval using the method from Chernozhukhov and Hansen (2008) over a range from -10 to 10, and Ind Share is the industry share (multiplied by 100 for legibility). Panel E reports statistics about how the values of $\hat{\beta}_k$ vary with the positive and negative Rotemberg weights. The "Other" industry is the "N/A" code in the IND1990 classification system and includes full-time military personnel.

Table 2: Relationship between industry shares and characteristics: canonical setting

| | Oil and Gas Extraction | Motor Vehicles | Other | Guided Missiles | Blast furnaces | Bartik (1980 shares) |
|---|---|---|---|---|---|---|
| Male | 1.319 | -0.501 | 4.076 | 0.126 | 0.344 | -0.178 |
| | (0.242) | (0.160) | (0.600) | (0.063) | (0.159) | (0.035) |
| White | 0.043 | -0.714 | -1.310 | 0.057 | -0.681 | -0.088 |
| | (0.102) | (0.653) | (0.281) | (0.043) | (0.256) | (0.029) |
| Native Born | 0.364 | -0.129 | 0.824 | -0.157 | -0.312 | -0.172 |
| | (0.092) | (0.110) | (0.281) | (0.133) | (0.129) | (0.019) |
| 12th Grade Only | -1.096 | 1.283 | 1.040 | -0.193 | 0.202 | 0.036 |
| | (0.218) | (0.392) | (0.356) | (0.091) | (0.150) | (0.030) |
| Some College | -0.311 | 0.687 | 1.060 | 0.033 | -0.808 | 0.376 |
| | (0.143) | (0.520) | (0.288) | (0.072) | (0.254) | (0.042) |
| Veteran | -0.295 | 0.895 | -5.793 | 0.202 | 2.526 | 0.000 |
| | (0.227) | (0.917) | (0.879) | (0.126) | (0.714) | (0.072) |
| # of Children | -0.043 | 0.954 | -2.409 | -0.006 | 0.003 | -0.070 |
| | (0.142) | (0.538) | (0.558) | (0.047) | (0.223) | (0.034) |
| $R^2$ | 0.24 | 0.08 | 0.46 | 0.27 | 0.23 | 0.77 |
| N | 722 | 722 | 722 | 722 | 722 | 722 |

*Notes:* Each column reports results of a single regression of a 1980 industry share on 1980 characteristics. The final column is the Bartik instrument constructed using the growth rates from 1980 to 1990. Results are weighted by 1980 population. Standard errors in parentheses. The "Other" industry is the "N/A" code in the IND1990 classification system and includes full-time military personnel.

Table 3: OLS and IV estimates: canonical setting

| | Δ Emp | | Coefficient Equal | Over ID test |
|---|---|---|---|---|
| | (1) | (2) | (3) | (4) |
| OLS | 0.71 | 0.63 | [0.04] | |
| | (0.06) | (0.07) | | |
| TSLS (Leave-Out Bartik) | 1.76 | 1.28 | [0.23] | |
| | (0.33) | (0.42) | | |
| TSLS (Bartik) | 1.65 | 1.22 | [0.19] | |
| | (0.34) | (0.15) | | |
| TSLS | 0.74 | 0.67 | [0.10] | 1014.05 |
| | (0.05) | (0.07) | | [0.00] |
| MBTSLS | 0.76 | 0.69 | [0.13] | |
| | (0.06) | (0.07) | | |
| LIML | 1.60 | 1.42 | [0.76] | 2820.96 |
| | (0.00) | (0.57) | | [0.00] |
| HFUL | 2.85 | 2.69 | [0.00] | 804.19 |
| | (0.14) | (0.13) | | [0.00] |
| Year and CZone FE | Yes | Yes | | |
| Controls | No | Yes | | |
| Observations | 2,166 | 2,166 | | |

*Notes:* This table reports a variety of estimates of the inverse elasticity of labor supply. The regressions are at the commuting zone level and the instruments are 3-digit industry-time periods (1980-1990, 1990-2000, and 2000-2010). Column (1) does not contain controls, while column (2) does. The TSLS (Bartik) row uses the Bartik instrument. The TSLS row uses each industry share (times time period) separately as instruments. The MBTSLS row uses the estimator of Anatolyev (2013) and Kolesar et al. (2015) with the same set of instruments. The LIML row shows estimates using the limited information maximum likelihood estimator with the same set of instruments. Finally, the HFUL row uses the HFUL estimator of Hausman et al. (2012) with the same set of instruments. The J-statistic for HFUL comes from Chao et al. (2014). 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. Results are weighted by 1980 population. Standard errors are in parentheses and are constructed by bootstrap over commuting zones. p-values are in brackets.

Table 4: Summary of Rotemberg weights: China shock

| | | | | | |
|---|---|---|---|---|---|
| **Panel A: Negative and positive weights** | | | | | |
| | Sum | Mean | Share | | |
| Negative | -0.067 | -0.000 | 0.059 | | |
| Positive | 1.067 | 0.004 | 0.941 | | |
| **Panel B: Correlations** | | | | | |
| | $\hat{\alpha}_k$ | $g_k$ | $\hat{\beta}_k$ | $\hat{F}_k$ | Var($z_k$) |
| $\hat{\alpha}_k$ | 1 | | | | |
| $g_k$ | 0.430 | 1 | | | |
| $\hat{\beta}_k$ | 0.003 | -0.320 | 1 | | |
| $\hat{F}_k$ | 0.192 | 0.027 | 0.017 | 1 | |
| Var($z_k$) | 0.102 | -0.141 | 0.157 | 0.229 | 1 |
| **Panel C: Variation across years in $\hat{\alpha}_k$** | | | | | |
| | Sum | Mean | | | |
| 1990 | 0.017 | 0.000 | | | |
| 2000 | 0.983 | 0.002 | | | |
| **Panel D: Top 5 Rotemberg weight industries** | | | | | |
| | $\hat{\alpha}_k$ | $g_k$ | $\hat{\beta}_k$ | 95 % CI | Ind Share |
| Electronic Computers | 0.183 | 186.231 | -0.619 | (-1.50,-0.20) | 0.137 |
| Games, Toys, and Children's Vehicles | 0.138 | 243.794 | -0.126 | (-0.60,0.30) | 0.044 |
| Household Audio and Video Equipment | 0.085 | 187.718 | 0.174 | (-0.20,1.80) | 0.046 |
| Telephone and Telegraph Apparatus | 0.066 | 92.922 | -0.315 | N/A | 0.100 |
| Computer Peripheral Equipment, NEC | 0.060 | 34.982 | -0.303 | (-1.20,-0.20) | 0.100 |
| **Panel E: Estimates of $\beta_k$ for positive and negative weights** | | | | | |
| | $\alpha$-weighted Sum | Share of overall $\beta$ | Mean | | |
| Negative | -0.014 | 0.024 | -0.036 | | |
| Positive | -0.582 | 0.976 | -1.170 | | |

*Notes:* This table reports statistics about the Rotemberg weights. In all cases, we report statistics about the aggregated weights, where we aggregate a given industry across years as discussed in Section 3.3. Panel A reports the share and sum of negative Rotemberg weights. Panel B reports correlations between the weights ($\hat{\alpha}_k$), the national component of growth ($g_k$), the just-identified coefficient estimates ($\hat{\beta}_k$), the first-stage F-statistic of the industry share ($\hat{F}_k$), and the variation in the industry shares across locations (Var($z_k$)). Panel C reports variation in the weights across years. Panel D reports the top five industries according to the Rotemberg weights. The $g_k$ is the national industry growth rate, $\hat{\beta}_k$ is the coefficient from the just-identified regression, the 95% confidence interval is the weak instrument robust confidence interval using the method from Chernozhukhov and Hansen (2008) over a range from -10 to 10 (N/A indicates that these standard errors are not well-defined), and Ind Share is the industry share (multiplied by 100 for legibility). Panel E reports statistics about how the values of $\hat{\beta}_k$ vary with the positive and negative Rotemberg weights.

Table 5: Relationship between industry shares and characteristics: China shock

| | Electronic Computers | Games, Toys, and Children's Vehicles | Household Audio and Video Equipment | Telephone and Telegraph Apparatus | Computer Peripheral Equipment, NEC | China to other |
|---|---|---|---|---|---|---|
| Share Empl in Manufacturing | 0.016 | 0.002 | 0.006 | 0.002 | 0.009 | 0.099 |
| | (0.008) | (0.001) | (0.003) | (0.003) | (0.004) | (0.011) |
| Share College Educated | 0.016 | -0.001 | 0.002 | 0.001 | 0.012 | 0.068 |
| | (0.006) | (0.001) | (0.002) | (0.003) | (0.003) | (0.014) |
| Share Foreign Born | 0.004 | 0.001 | -0.001 | -0.005 | 0.002 | 0.052 |
| | (0.003) | (0.001) | (0.001) | (0.003) | (0.002) | (0.009) |
| Share Empl of Women | -0.002 | 0.003 | -0.006 | -0.003 | 0.000 | 0.031 |
| | (0.006) | (0.002) | (0.003) | (0.006) | (0.004) | (0.017) |
| Share Empl in Routine | -0.083 | 0.006 | 0.010 | -0.010 | -0.046 | -0.051 |
| | (0.041) | (0.003) | (0.007) | (0.015) | (0.018) | (0.084) |
| Avg Offshorability | 0.410 | -0.027 | -0.022 | 0.248 | 0.182 | -1.173 |
| | (0.214) | (0.022) | (0.039) | (0.076) | (0.091) | (0.460) |
| $R^2$ | 0.18 | 0.02 | 0.01 | 0.04 | 0.12 | 0.22 |
| N | 1444 | 1444 | 1444 | 1444 | 1444 | 1444 |

*Notes:* Each column reports a separate regression. The regressions are two pooled cross-sections, where one cross section is 1980 shares on 1990 characteristics, and one is 1990 shares on 2000 characteristics. The final column is constructed using 1990 to 2000 growth rates. Results are weighted by the population in the period the characteristics are measured. Standard errors in parentheses.

Table 6: OLS and IV estimates: China shock

| | Δ Emp (1) | (2) | Coefficients Equal (3) | Over ID Test (4) |
|---|---|---|---|---|
| OLS | -0.38 | -0.17 | [0.00] | |
| | (0.07) | (0.04) | | |
| TSLS (Bartik) | -0.73 | -0.60 | [0.05] | |
| | (0.06) | (0.09) | | |
| TSLS | -0.45 | -0.21 | [0.00] | 917.36 |
| | (0.06) | (0.04) | | [0.00] |
| MBTSLS | -0.56 | -0.29 | [0.00] | |
| | (0.06) | (0.04) | | |
| LIML | -1.47 | -1.94 | [0.83] | 1868.95 |
| | (0.71) | (3.33) | | [0.00] |
| HFUL | -1.15 | -1.13 | [0.47] | 968.37 |
| | (0.05) | (0.04) | | [0.00] |
| Year and Census Division FE | Yes | Yes | | |
| Controls | No | Yes | | |
| Observations | 1,444 | 1,444 | | |

*Notes:* This table reports a variety of estimates of the effect of rising imports from China on US manufacturing employment. The regressions are at the CZ level and include two time periods (1990 to 2000, and 2000 to 2007). The TSLS row is our replication of Column (1) and Column (6) of Table 3 in ADH. Column (1) does not contain controls, while column (2) does. The TSLS (Bartik) row uses the Bartik instrument. The TSLS row uses each industry share (times time period) separately as instruments. The MBTSLS row uses the estimator of Anatolyev (2013) and Kolesar et al. (2015) with the same set of instruments. The LIML row shows estimates using the limited information maximum likelihood estimator with the same set of instruments. Finally, the HFUL row uses the HFUL estimator of Hausman et al. (2012) with the same set of instruments. The J-statistic for HFUL comes from Chao et al. (2014). The p-value for the equality of coefficients compares the adjacent columns with and without controls. The controls are the contemporaneous characteristics displayed in Table 5. Results are weighted by start of period population. Standard errors are in parentheses and are constructed by bootstrap over commuting zones. p-values are in brackets.

Table 7: Summary of Rotemberg weights: immigrant enclave

**Panel A: High school equivalent**

**Panel I: Correlations**

| | $\hat{\alpha}_k$ | $g_k$ | $\hat{\beta}_k$ | $\hat{F}_k$ | Var($z_k$) |
|---|---|---|---|---|---|
| $\hat{\alpha}_k$ | 1 | | | | |
| $g_k$ | 0.991 | 1 | | | |
| $\hat{\beta}_k$ | 0.169 | 0.164 | 1 | | |
| $\hat{F}_k$ | 0.203 | 0.173 | 0.181 | 1 | |
| Var($z_k$) | 0.043 | -0.032 | -0.106 | -0.260 | 1 |

**Panel II: Top 5 Rotemberg weight origin countries**

| | $\hat{\alpha}_k$ | $g_k$ | $\hat{\beta}_k$ | 95 % CI |
|---|---|---|---|---|
| Mexico | 0.482 | 4.95e+06 | -0.026 | (-0.040,0.000) |
| El Salvador | 0.054 | 4.65e+05 | -0.046 | (-0.070,-0.030) |
| Phillipines | 0.050 | 5.31e+05 | -0.023 | (-0.040,0.130) |
| China | 0.038 | 4.28e+05 | -0.041 | (-0.070,-0.010) |
| West Europe and Others | 0.031 | 6.41e+05 | -0.067 | (-0.110,-0.050) |

**Panel B: College equivalent**

**Panel I: Correlations**

| | $\hat{\alpha}_k$ | $g_k$ | $\hat{\beta}_k$ | $\hat{F}_k$ | Var($z_k$) |
|---|---|---|---|---|---|
| $\hat{\alpha}_k$ | 1 | | | | |
| $g_k$ | 0.766 | 1 | | | |
| $\hat{\beta}_k$ | 0.293 | 0.255 | 1 | | |
| $\hat{F}_k$ | -0.028 | -0.055 | 0.230 | 1 | |
| Var($z_k$) | 0.033 | -0.381 | -0.075 | -0.225 | 1 |

**Panel II: Top 5 Rotemberg weight origin countries**

| | $\hat{\alpha}_k$ | $g_k$ | $\hat{\beta}_k$ | 95 % CI |
|---|---|---|---|---|
| Phillipines | 0.151 | 6.32e+05 | -0.065 | (-0.125,-0.040) |
| Mexico | 0.102 | 5.44e+05 | -0.062 | (-0.095,0.000) |
| China | 0.082 | 3.74e+05 | -0.084 | (-0.125,-0.060) |
| West Europe and Others | 0.066 | 5.31e+05 | -0.090 | (-0.145,-0.065) |
| Cuba | 0.049 | 1.86e+05 | -0.008 | (-0.045,0.500) |

*Notes:* This table reports statistics about the Rotemberg weights, which are all positive in this application. Panels A.I and B.I reports correlations between the weights ($\hat{\alpha}_k$), the national component of growth ($g_k$), the just-identified coefficient estimates ($\hat{\beta}_k$), the first-stage F-statistics ($\hat{F}_k$), and the variation in the origin country shares across locations (Var($z_k$)). Panels A.II and B.II report the top five origin countries according to the Rotemberg weights. The "Others" are Australia, Cyprus, Israel, and New Zealand. The $g_k$ is the number of immigrants from 1990 to 2000, $\hat{\beta}_k$ is the coefficient from the just-identified regression, the 95% confidence interval is the weak instrument robust confidence interval using the method from Chernozhukhov and Hansen (2008) over a range from -10 to 10.

Table 8: Relationship between origin country shares and characteristics: immigrant enclave

| | Mexico | Philippines | El Salvador | China | Cuba | West Europe & Others | Bartik High School | Bartik College |
|---|---|---|---|---|---|---|---|---|
| City size | 0.054 | 0.026 | 0.106 | 0.057 | 0.049 | 0.039 | 0.059 | 0.023 |
| | (0.018) | (0.021) | (0.027) | (0.019) | (0.060) | (0.009) | (0.009) | (0.004) |
| College share | -0.545 | 0.559 | 0.692 | 1.318 | -0.828 | 0.530 | -0.021 | 0.157 |
| | (0.370) | (0.416) | (0.554) | (0.389) | (1.206) | (0.175) | (0.189) | (0.072) |
| Mean wage residuals | 0.601 | -0.428 | 0.595 | -0.212 | -0.199 | 0.052 | 0.267 | 0.041 |
| for all natives | (0.388) | (0.437) | (0.582) | (0.408) | (1.266) | (0.184) | (0.199) | (0.076) |
| Mean wage residuals | -0.652 | 0.596 | -0.856 | 0.152 | -0.079 | -0.209 | -0.385 | -0.061 |
| for all immigrants | (0.361) | (0.406) | (0.540) | (0.379) | (1.175) | (0.170) | (0.185) | (0.070) |
| Mfg share | 0.059 | -0.379 | 0.268 | -0.192 | -0.653 | 0.230 | -0.006 | -0.010 |
| | (0.202) | (0.228) | (0.303) | (0.213) | (0.660) | (0.096) | (0.104) | (0.039) |
| N | 124 | 124 | 124 | 124 | 124 | 124 | 124 | 124 |
| $R^2$ | 0.150 | 0.095 | 0.216 | 0.246 | 0.020 | 0.294 | 0.371 | 0.430 |

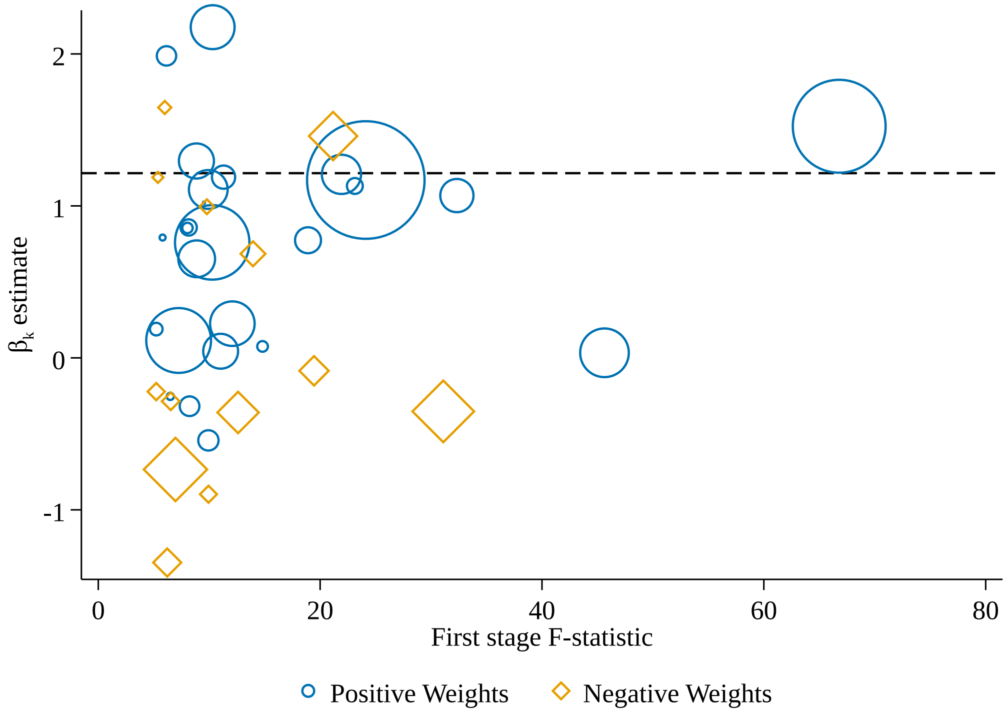
*Notes:* Each column reports results of a single regression of a 1980 origin country share on 1980 characteristics. Results are weighted by 1990 population. Standard errors in parentheses. For legibility, coefficients and standard errors of the first six columns are multiplied by 10,000,000. Coefficients and standard errors of the last two columns are not scaled. The "Others" are Australia, Cyprus, Israel, and New Zealand.

Table 9: OLS and IV estimates: immigrant enclave

| **Panel A. High school equivalent** | | | | |
|---|---|---|---|---|
| | Δ Emp | | Coefficients Equal | Over ID Test |
| | (1) | (2) | (3) | (4) |
| OLS | -0.02 | -0.03 | [0.00] | |
| | (0.01) | (0.01) | | |
| TSLS (Bartik) | -0.02 | -0.04 | [0.07] | |
| | (0.01) | (0.01) | | |
| TSLS | -0.02 | -0.04 | [0.02] | 43.30 |
| | (0.01) | (0.01) | | [0.22] |
| MBTSLS | -0.03 | -0.04 | [0.08] | |
| | (0.01) | (0.01) | | |
| LIML | -0.03 | -0.04 | [0.06] | 73.16 |
| | (0.01) | (0.01) | | [0.00] |
| HFUL | 0.03 | 0.02 | [0.26] | 82.45 |
| | (0.01) | (0.00) | | [0.00] |
| **Panel B. College equivalent** | | | | |
| OLS | -0.06 | -0.06 | [0.65] | |
| | (0.01) | (0.01) | | |
| TSLS (Bartik) | -0.08 | -0.08 | [0.93] | |
| | (0.01) | (0.01) | | |
| TSLS | -0.06 | -0.06 | [0.71] | 35.54 |
| | (0.01) | (0.01) | | [0.54] |
| MBTSLS | -0.06 | -0.07 | [0.71] | |
| | (0.01) | (0.01) | | |
| LIML | -0.06 | -0.06 | [0.72] | 33.67 |
| | (0.01) | (0.01) | | [0.63] |
| HFUL | 0.04 | 0.04 | [0.23] | 67.95 |
| | (0.01) | (0.00) | | [0.00] |
| Controls | No | Yes | | |
| Observations | 124 | 124 | | |

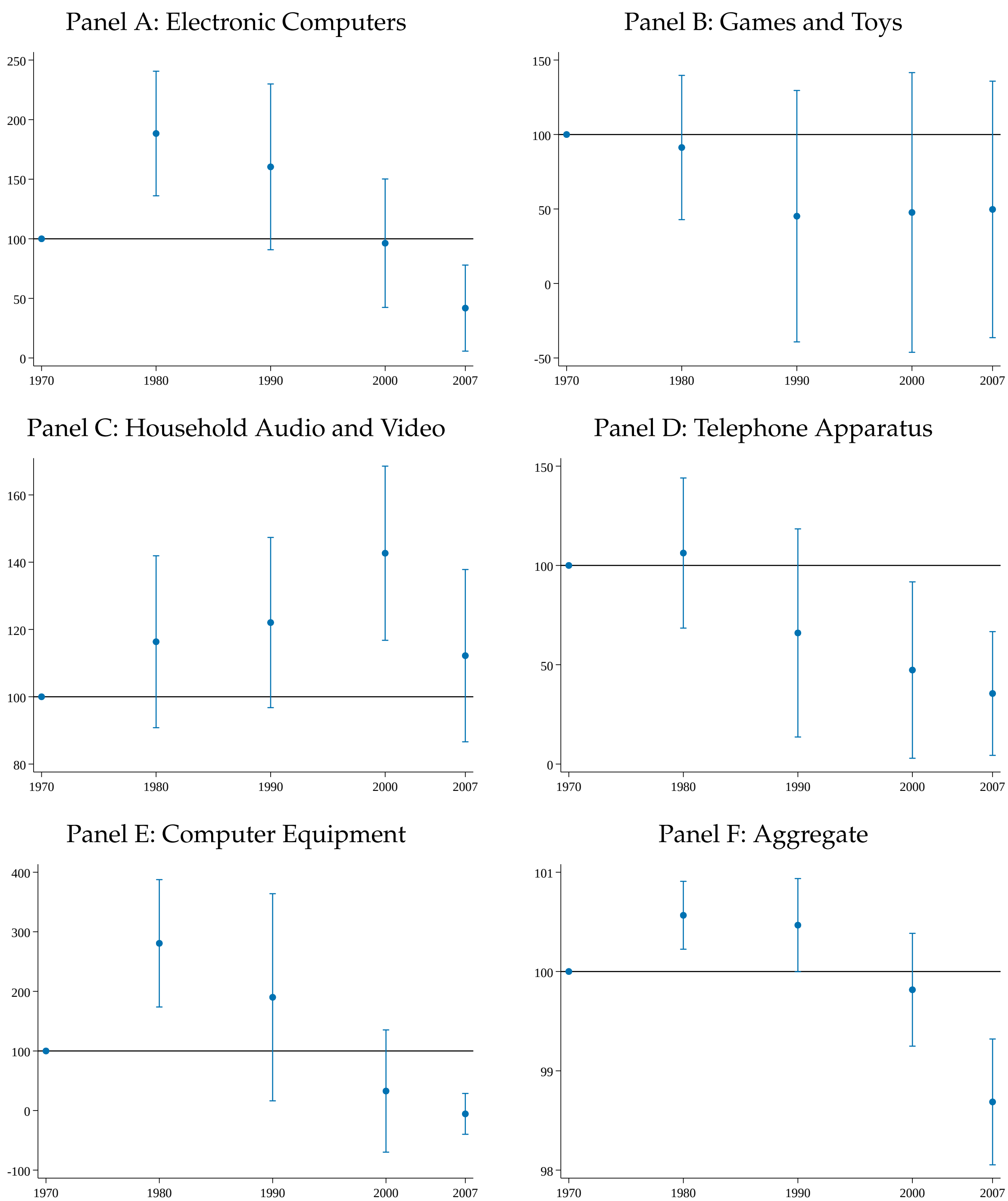
*Notes:* This table reports a variety of estimates of the negative of the inverse elasticity of substitution between immigrants and natives. The regressions are at the city level and include a single time period (2000). The TSLS row is our replication of Column (3) and Column (7) of Table 6 in Card (2009). Column (1) does not contain controls, while column (2) does. The TSLS (Bartik) row uses the Bartik instrument. The TSLS row uses each origin country share separately as instruments. The MBTSLS row uses the estimator of Anatolyev (2013) and Kolesar et al. (2015) with the same set of instruments. The LIML row shows estimates using the limited information maximum likelihood estimator with the same set of instruments. Finally, the HFUL row uses the HFUL estimator of Hausman et al. (2012) with the same set of instruments. The J-statistic for HFUL comes from Chao et al. (2014). The p-value for the equality of coefficients compares the adjacent columns with and without controls. The controls are the contemporaneous characteristics displayed in Table 8. Results are weighted by 1990 population. Standard errors are in parentheses and are constructed by bootstrap over commuting zones. p-values are in brackets.

Figure 1: Heterogeneity of $\beta_k$: canonical setting



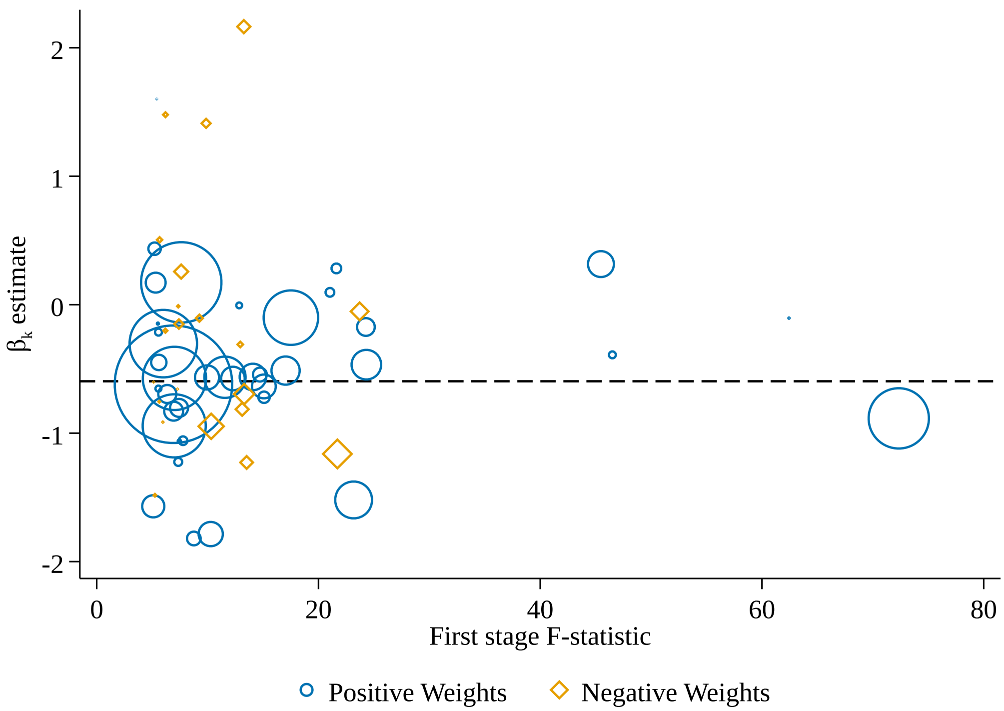

*Notes:* This figure plots the relationship between each instruments' $\hat{\beta}_k$, first stage F-statistics and the Rotemberg weights. Each point is a separate instrument's estimates (industry share). The figure plots the estimated $\hat{\beta}_k$ for each instrument on the y-axis and the estimated first-stage F-statistic on the x-axis. The size of the points are scaled by the magnitude of the Rotemberg weights, with the circles denoting positive Rotemberg weights and the diamonds denoting negative weights. The horizontal dashed line is plotted at the value of the overall $\hat{\beta}$ reported in the second column in the TSLS (Bartik) row in Table 3. The figure excludes instruments with first-stage F-statistics below 5.

Figure 2: Pre-trends for high Rotemberg weight industries: China shock



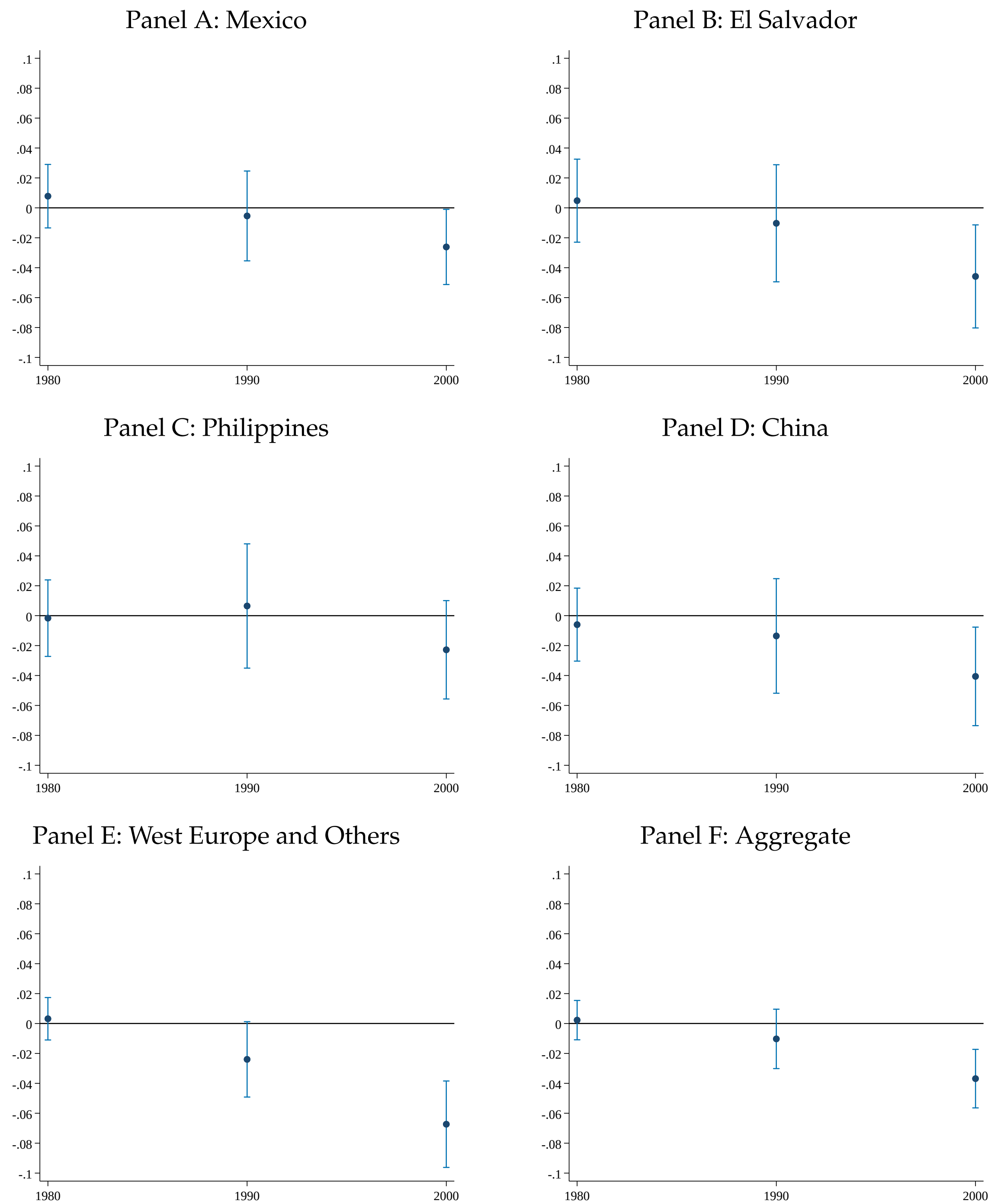

*Notes:* These figures report pre-trends for the overall instrument and the top-5 Rotemberg weight industries as reported in Table 4. The figures fix industry shares at the 1980 values and report the effect of these industry shares on manufacturing employment. We run regressions in growth rates and then convert to levels. We normalize 1970 to 100, and compute the standard errors using the delta method. For the aggregate panel, we use the Bartik estimate for 1980.

Figure 3: Heterogeneity of $\beta_k$: China shock



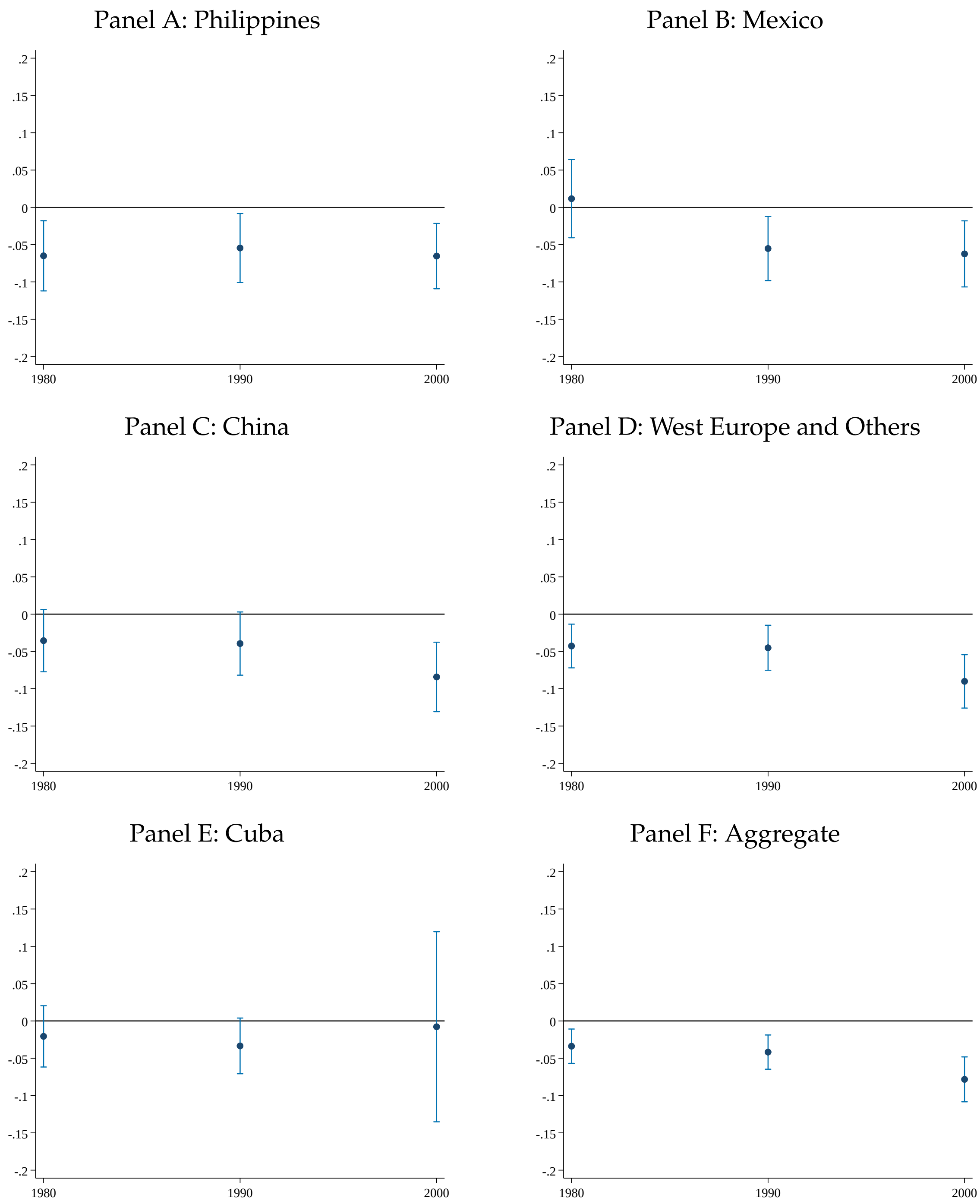

*Notes:* This figure plots the relationship between each instruments' $\hat{\beta}_k$, first stage F-statistics and the Rotemberg weights. Each point is a separate instrument's estimates (industry share). The figure plots the estimated $\hat{\beta}_k$ for each instrument on the y-axis and the estimated first-stage F-statistic on the x-axis. The size of the points are scaled by the magnitude of the Rotemberg weights, with the circles denoting positive Rotemberg weights and the diamonds denoting negative weights. The horizontal dashed line is plotted at the value of the overall $\hat{\beta}$ reported in the second column in the TSLS (Bartik) row in Table 6. The figure excludes instruments with first-stage F-statistics below 5.

Figure 4: Pre-trends for high Rotemberg weight origin countries: immigrant enclave, high school equivalent




*Notes:* These figures report pre-trends for the overall instrument and the top-5 Rotemberg weight origin countries as reported in Panel B of Table 7. We replace the endogenous and outcome variables in equations (8.1) and (8.2) with their 1980, 1990 and 2000 values (that is, we include all the controls in Card (2009) in Table 6, columns (3) and (7)). The 2000 coefficient corresponds to the $\hat{\beta}_k$ in Table 7 and the TSLS(Bartik) $\hat{\beta}$ in Table 9. The "Others" are Cyprus, New Zealand, Israel and Australia.

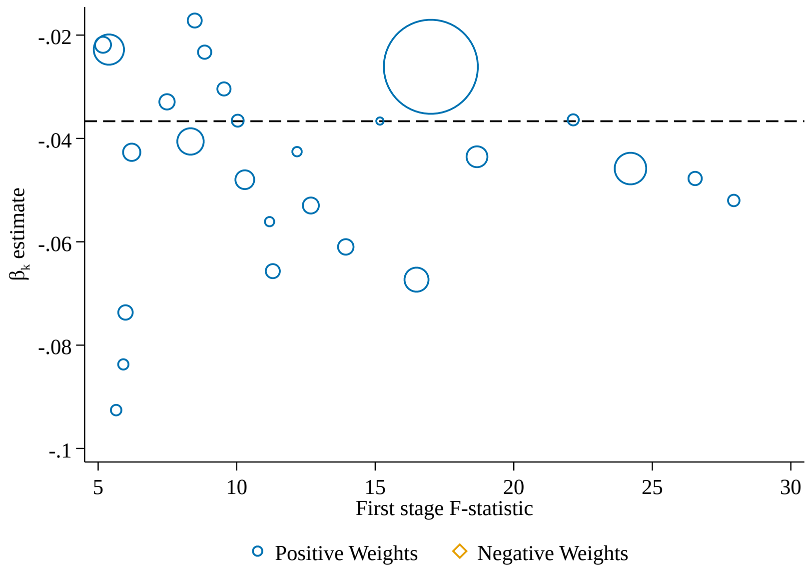
Figure 5: Pre-trends for high Rotemberg weight origin countries: immigrant enclave, college equivalent



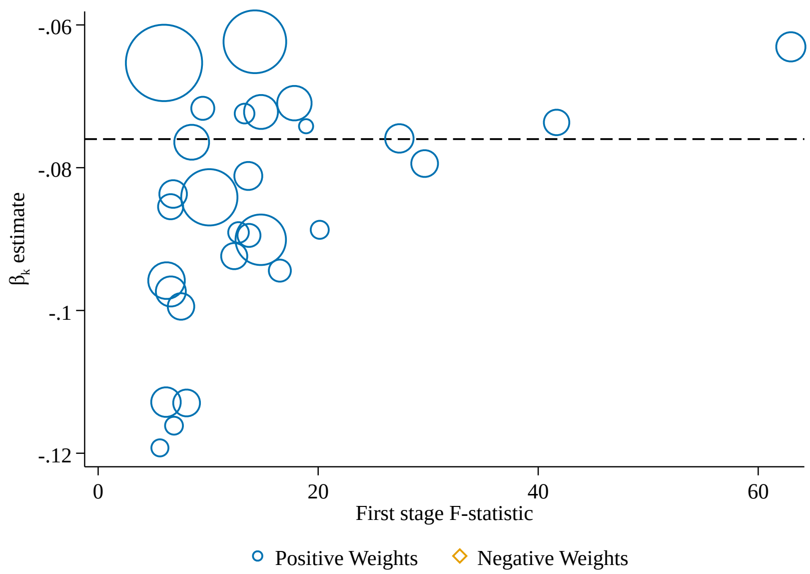

*Notes:* These figures report pre-trends for the overall instrument and the top-5 Rotemberg weight origin countries as reported in Panel D of Table 7. We replace the endogenous and outcome variables in equations (8.1) and (8.2) with their 1980, 1990 and 2000 values (that is, we include all the controls in Card (2009) in Table 6, columns (3) and (7)). The 2000 coefficient corresponds to the $\hat{\beta}_k$ in Table 7 and the TSLS(Bartik) $\hat{\beta}$ in Table 9. The "Others" are Cyprus, New Zealand, Israel and Australia.

Figure 6: Heterogeneity of $\beta_k$: immigrant enclave

Panel A: High school equivalent




Panel B: College equivalent




*Notes:* This figure plots the relationship between each instruments' $\hat{\beta}_k$, first stage F-statistics and the Rotemberg weights. Each point is a separate instrument's (country of origin) estimates. The figure plots the estimated $\hat{\beta}_k$ for each instrument on the y-axis and the estimated first-stage F-statistic on the x-axis. The size of the points are scaled by the magnitude of the Rotemberg weights, with the circles denoting positive Rotemberg weights and the diamonds denoting negative weights. The horizontal dashed line is plotted at the value of the overall $\hat{\beta}$ reported in the second column in the TSLS (Bartik) row in Table 9. The figure excludes instruments with first-stage F-statistics below 5.

# A Instruments encompassed by our structure

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

## A.1 Bank lending relationships

Greenstone, Mas, and Nguyen (Forthcoming) 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 $\hat{g}_{-l,k} = \frac{1}{L-1} \sum_{l' \neq l} g_{l'k}$. However, Greenstone, Mas, and Nguyen (Forthcoming) 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 $\hat{g}_l$ captures the change in bank lending that is common to a county, while $\hat{g}_k$ 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).

## A.2 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 instrument 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.

# B Omitted proofs

## Proposition 1.1

**Proof.**

$$\begin{aligned}\hat{\beta}_{GMM} &= \frac{X^{\perp\prime} Z G G' Z' Y^{\perp}}{X^{\perp\prime} Z G G' Z' X^{\perp}} \\ &= \frac{X^{\perp\prime} B B' Y^{\perp}}{X^{\perp\prime} B B' X^{\perp}} \\ &= \hat{\beta}_{Bartik},\end{aligned}$$

where $X^{\perp\prime} B$ is a scalar and so cancels. □

## Proposition 3.1

We use slightly more general notation than in the body of the paper. Let $\hat{W}$ be an arbitrary weight matrix and let

$$\hat{C}(\hat{W}) = \hat{W} Z' X^{\perp} \text{ and } \hat{c}_k(\hat{W}) = \hat{W}_k Z' X^{\perp},$$

where $\hat{W}_k$ is the $k^{th}$ row of $\hat{W}$. We index a solution for $\hat{\beta}$ by $\hat{W}$: $\hat{\beta}(\hat{W})$. The more general version of the proposition stated in the text is:

PROPOSITION B.1. *Let*

$$\hat{\beta}(\hat{W}) = \frac{\hat{C}(\hat{W})' Z' Y^{\perp}}{\hat{C}(\hat{W})' Z' X^{\perp}},\ \hat{\alpha}_k(\hat{W}) = \frac{\hat{c}_k(\hat{W}) Z_k' X^{\perp}}{\sum_{k'} \hat{c}_{k'}(\hat{W}) Z_{k'}' X^{\perp}},\ \textit{and } \hat{\beta}_k = (Z_k' X^{\perp})^{-1} Z_k' Y^{\perp}.$$

*Then:*

$$\hat{\beta}(\hat{W}) = \sum_{k=1}^{K} \hat{\alpha}_k(\hat{W}) \hat{\beta}_k,$$

*where* $\sum_{k=1}^{K} \hat{\alpha}_k(\hat{W}) = 1$.

**Proof.** The proof is just algebra:

$$\hat{\alpha}_k(\hat{W})\hat{\beta}_k = \frac{\hat{c}_k(\hat{W}) Z_k' X^{\perp}}{\sum_{k=1}^{K} \hat{c}_k(\hat{W}) Z_k' X^{\perp}} (Z_k' X^{\perp})^{-1} Z_k' Y^{\perp} = \frac{\hat{c}_k(\hat{W}) Z_k' Y^{\perp}}{\sum_{k=1}^{K} \hat{c}_k(\hat{W}) Z_k' X^{\perp}} \tag{A1}$$

$$\sum_{k=1}^{K} \hat{\alpha}_k(\hat{W})\hat{\beta}_k = \frac{\sum_{k=1}^{K} \hat{c}_k(\hat{W}) Z_k' Y^{\perp}}{\sum_{k=1}^{K} \hat{c}_k(\hat{W}) Z_k' X^{\perp}} \tag{A2}$$

$$= \frac{\hat{C}(\hat{W})' Z' Y^{\perp}}{\hat{C}(\hat{W})' Z' X^{\perp}}. \tag{A3}$$

□

The proposition stated in the text comes from substituting in for the Bartik definition of $\hat{W}$.

**Proposition 4.1**

**Proof.** For a given $k$,

$$\hat{\beta}_k = \frac{\sum_l z_{lk} x_l^{\perp} \beta_l}{\sum_l z_{lk} x_l^{\perp}} + \frac{\sum_l z_{lk} \epsilon_l^{\perp}}{\sum_l z_{lk} x_l^{\perp}} \tag{A4}$$

$$= \frac{\sum_l z_{lk} x_l^{\perp} \beta_l}{\sum_l z_{lk} x_l^{\perp}} + o_p(1) \tag{A5}$$

$$= \frac{\sum_l z_{lk}^{\perp,2} \pi_{lk} \beta_l + z_{lk} u_{lk}^{\perp}}{\sum_l z_{lk}^{\perp,2} \pi_{lk} + z_{lk} u_{lk}^{\perp}} + o_p(1) \tag{A6}$$

$$= \sum_l \omega_{lk} \beta_l + o_p(1), \tag{A7}$$

where $\omega_{lk} = z_{lk}^{\perp,2} \pi_{lk} / \sum_l z_{lk}^{\perp,2} \pi_{lk}$. □

# C Equivalence with $K$ industries, $L$ locations, and controls

The two stage least squares system of equations is:

$$y_{lt} = D_{lt}\rho + x_{lt}\beta + \epsilon_{lt} \tag{A1}$$

$$x_{lt} = D_{lt}\tau + B_{lt}\gamma + \eta_{lt}, \tag{A2}$$

where $D_{lt}$ is a $1 \times S$ vector of controls. Typically in a panel context, $D_{lt}$ 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 $y_{lt}$, $\mathbf{D}$ denote the $n \times L$ stacked vector of $D_{lt}$ 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 $\mathbf{P_D} = \mathbf{D}(\mathbf{D}'\mathbf{D})^{-1}\mathbf{D}'$ as the $n \times n$ projection matrix of $\mathbf{D}$, and $\mathbf{M_D} = \mathbf{I}_n - \mathbf{P_D}$ as the annhilator matrix. Then, because this is an exactly identified instrumental variable our estimator is

$$\hat{\beta}_{Bartik} = \frac{B'\mathbf{M_D}Y}{B'\mathbf{M_D}X}. \tag{A3}$$

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

$$y_{lt} = D_{lt}\rho + x_{lt}\beta + \epsilon_{lt} \tag{A4}$$

$$x_{it} = D_{lt}\tau + Z_{lt}\gamma_t + \eta_{lt}, \tag{A5}$$

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 Section 1.2, the $t$ subscript allows the effect of a given industry share to be time-varying. In matrix notation, we write

$$Y = \mathbf{D}\rho + X\beta + \epsilon \tag{A6}$$
$$X = \mathbf{D}\tau + \tilde{Z}\mathbf{\Gamma} + \eta, \tag{A7}$$

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

$$\mathbf{\Gamma} = [\gamma_1 \cdots \gamma_T], \tag{A8}$$

and $\tilde{\mathbf{Z}}$ is a stacked $n \times (T \times K)$ matrix such that

$$\tilde{\mathbf{Z}} = \left[ \begin{array}{ccc} \mathbf{Z} \odot 1_{t=1} & \cdots & \mathbf{Z} \odot 1_{t=T} \end{array} \right], \tag{A9}$$

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. Then, using the $\tilde{\mathbf{Z}}$ as instruments, the GMM estimator is:

$$\hat{\beta}_{GMM} = \frac{X'\mathbf{M_D}\tilde{\mathbf{Z}}\Omega\tilde{\mathbf{Z}}'\mathbf{M_D}Y}{X'\mathbf{M_D}\tilde{\mathbf{Z}}\Omega\tilde{\mathbf{Z}}'\mathbf{M_D}X}, \tag{A10}$$

where $\Omega$ is a $KT \times KT$ weight matrix.

PROPOSITION C.1. *If $\Omega = GG'$, then $\hat{\beta}_{GMM} = \hat{\beta}_{Bartik}$.*

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

$$\hat{\beta}_{Bartik} = \frac{B'\mathbf{M_D}Y}{B'\mathbf{M_D}X} \tag{A11}$$
$$= \frac{G'\tilde{\mathbf{Z}}'\mathbf{M_D}Y}{G'\tilde{\mathbf{Z}}'\mathbf{M_D}X} \tag{A12}$$
$$= \frac{X'\mathbf{M_D}\tilde{\mathbf{Z}}GG'\tilde{\mathbf{Z}}'\mathbf{M_D}Y}{X'\mathbf{M_D}\tilde{\mathbf{Z}}GG'\tilde{\mathbf{Z}}'\mathbf{M_D}X}, \tag{A13}$$

where the second equality follows from the definition of $B$, and the third equality follows because $X'\mathbf{M_D}\tilde{\mathbf{Z}}G$ is a scalar. By inspection, if $\Omega = GG'$, then $\hat{\beta}_{GMM} = \hat{\beta}_{Bartik}$. □

# D Interpreting the Rotemberg weights

To interpret the Rotemberg weights, we move from finite samples to population limits. We first state the standard assumptions such that GMM estimators are consistent for all sequences of $\hat{W}$ matrices. We then consider local-to-zero asymptotics (e.g., Conley, Hansen, and Rossi (2012)) to interpret the Rotemberg weights in terms of sensitivity-to-misspecification as discussed in AGS. As such, the results in this section are largely special cases of AGS.

The Rotemberg weights depend on the choice of weight matrix, $\hat{W}$. Given standard assumptions, the choice of weight matrix does not affect consistency or bias of the estimates, and only affects the asymptotic variance of the estimator (there is a rich literature studying how to optimize this choice).

When some of the instruments are not exogeneous, however, the population version of the Rotemberg weights measures how much the overidentified estimate of $\beta_0$ is affected by this misspecification. To allow for this interpretation, we modify our estimating equation:

$$y_{lt} = D_{lt}\rho + x_{lt}\beta_0 + V_{lt}\kappa + \epsilon_{lt},$$

where we assume that for some $k$, $\mathbb{E}[Z_{lkt}V_{lt}|D_{lt}] \neq 0$. We follow Conley, Hansen, and Rossi (2012, Section III.C) and AGS (pg. 1569) and allow $\kappa$ to be proportional to $L^{-1/2}$ such that we have local misspecification. We make the following standard regularity assumptions:

ASSUMPTION 4 (Identification and Regularity). *(i) the data* $\{\{x_{lt}, Z_{lt}, D_{lt}, V_{lt}, \epsilon_{lt}\}_{t=1}^{T}\}_{l=1}^{L}$ *are independent and identically distributed with K and T fixed, and L going to infinity;*

*(ii)* $\mathbb{E}[\epsilon_{lt}] = 0$, $\mathbb{E}[V_{lt}] = 0$ *and* $Var(\tilde{\epsilon}) < \infty$;

*(iii)* $\mathbb{E}[z_{lkt}\epsilon_{lt}|D_{lt}] = 0$ *for all values of k;* $\mathbb{E}[z_{lt}V_{lt}] = \Sigma_{ZV}$, *where* $\Sigma_{ZV}$ *is a* $1 \times K$ *covariance vector with at least one non-zero entry; and* $\mathbb{E}[Z_{lt}x_{lt}^{\perp}] = \Sigma_{ZX^{\perp}}$ *is a* $1 \times K$ *covariance vector with all non-zero entries (*$x_{lt}$ *is a scalar), and* $\Sigma_{ZX^{\perp},k}$ *is the* $k^{th}$ *entry; and*

*(iv)* $Var(z_{lkt}\epsilon_{lt}) < \infty$, $Var(z_{lkt}V_{lt}) < \infty$ *and* $Var(z_{lkt}x_{lt}^{\perp}) < \infty$ *for all values of k.*

We first establish the population version of $\hat{\alpha}_k(\hat{W})$:

LEMMA D.1. *If Assumption 4 holds and* $\operatorname{plim}_{L\to\infty} \hat{W}_L = W$ *where W is a positive semi-definite matrix, then*

$$\operatorname*{plim}_{L\to\infty} \hat{\alpha}_k(\hat{W}) = \alpha_k(W) = \frac{\Sigma_{ZX^{\perp}} W_k \Sigma_{ZX^{\perp},k}}{\Sigma_{ZX^{\perp}} W \Sigma'_{ZX^{\perp}}}.$$

**Proof.** Note that

$$\hat{\alpha}_k(\hat{W}) = \frac{X^{\perp\prime} Z \hat{W}_k Z_k' X^{\perp}}{X^{\perp\prime} Z \hat{W} Z' X^{\perp}} \tag{A1}$$

$$= \frac{\left(\sum_{l,t} x_{lt}^{\perp} Z_{lt}\right) \hat{W}_k \left(\sum_{l,t} z_{lkt} x_{lt}^{\perp}\right)}{\left(\sum_{l,t} x_{lt}^{\perp} Z_{lt}\right) \hat{W} \left(\sum_{l,t} Z_{lt} x_{lt}^{\perp}\right)}. \tag{A2}$$

Since our data is i.i.d. and the variance of $x_{lt}^{\perp} Z_{lt}$ is bounded, the law of large numbers holds as $L \to \infty$. □

We now present results about the asymptotic behavior of our estimators with misspecification.

PROPOSITION D.1. *We assume that Assumption 4 holds and* $\operatorname{plim}_{L\to\infty} \hat{W}_L = W$ *where W is a positive semi-definite matrix.*

*If* $\kappa = L^{-1/2}$, *then*

*(a)* $\sqrt{L}(\hat{\beta}_k - \beta_0)$ *converges in distribution to a random variable* $\tilde{\beta}_k$, *with* $\mathbb{E}[\tilde{\beta}_k] = \frac{\Sigma_{ZV,k}}{\Sigma_{ZX^{\perp},k}}$ *and*

*(b)* $\sqrt{L}(\hat{\beta} - \beta_0)$ *converges in distribution to a random variable* $\tilde{\beta}$, *with* $\mathbb{E}[\tilde{\beta}] = \sum_{k=1}^{K} \alpha_k(W)\mathbb{E}[\tilde{\beta}_k] = \sum_{k=1}^{K} \alpha_k(W) \frac{\Sigma_{ZV,k}}{\Sigma_{ZX^{\perp},k}}$.

**Proof.** First, note that

$$\hat{\beta}_k = \frac{\sum_{l,t} z_{lkt} y_{lt}^{\perp}}{\sum_{l,t} z_{lkt} x_{lt}^{\perp}} = \beta_0 + \frac{\sum_{l,t} z_{lkt}(L^{-1/2} V_{lt} + \epsilon_{lt})}{\sum_{l,t} z_{lkt} x_{lt}}$$
$$\hat{\beta}_k - \beta_0 = L^{-1/2} \frac{\sum_{l,t} z_{lkt} V_{lt}}{\sum_{l,t} z_{lkt} x_{lt}} + \frac{\sum_{l,t} z_{lkt} \epsilon_{lt}}{\sum_{l,t} z_{lkt} x_{lt}}.$$

The second term goes to zero because $\mathbb{E}[z_{lkt}\epsilon_{lt}] = 0$. The first term goes to zero as $L \to \infty$. Finally, since our summand terms have bounded variance, the law of large numbers holds. A similar argument holds for the broader summand.

The asymptotic bias of $\tilde{\beta}_k$ follows from Proposition 3 of AGS. A sketch of the proof for this case follows:

$$\sqrt{L}(\hat{\beta}_k - \beta_0) = \frac{\sum_{l,t} z_{lkt} V_{lt}}{\sum_{l,t} z_{lkt} x_{lt}} + \sqrt{L} \frac{\sum_{l,t} z_{lkt} \epsilon_{lt}}{\sum_{l,t} z_{lkt} x_{lt}}$$
$$\sqrt{L}(\hat{\beta}_k - \beta_0) - \frac{\sum_{l,t} z_{lkt} V_{lt}}{\sum_{l,t} z_{lkt} x_{lt}} = \sqrt{L} \frac{\sum_{l,t} z_{lkt} \epsilon_{lt}}{\sum_{l,t} z_{lkt} x_{lt}}.$$

Since $\frac{\sum_{l,t} z_{lkt} V_{lt}}{\sum_{l,t} z_{lkt} x_{lt}}$ converges to $\frac{\Sigma_{ZV,k}}{\Sigma_{ZX^{\perp},k}}$, this implies that $\sqrt{L}(\hat{\beta}_k - \beta_0)$ converges in distribution to a normally distributed random variable $\tilde{\beta}_k$ with $\mathbb{E}[\tilde{\beta}_k] = \frac{\Sigma_{ZV,k}}{\Sigma_{ZX^{\perp},k}}$. Finally, since $\hat{\alpha}_k(\hat{W})$ converges in probability to $\alpha_k(W)$, by a similar argument this implies that $\sqrt{L}(\hat{\beta} - \beta_0)$ converges in distribution to a normally distributed random variable $\tilde{\beta}$ with $\mathbb{E}[\tilde{\beta}] = \sum_k \alpha_k(W) \frac{\Sigma_{ZV,k}}{\Sigma_{ZX^{\perp},k}} = \sum_k \alpha_k(W)\mathbb{E}[\tilde{\beta}_k]$. □

This proposition shows that in the presence of misspecification, the estimator is asymptotically biased. Two useful corollaries follow:

COROLLARY D.1. *Suppose that $\beta_0 \neq 0$. Then the percentage bias can be written in terms of the Rotemberg weights:*

$$\frac{\mathbb{E}[\tilde{\beta}]}{\beta_0} = \sum_k \alpha_k(W) \frac{\mathbb{E}[\tilde{\beta}_k]}{\beta_0}. \tag{A3}$$

COROLLARY D.2. *Under the Bartik weight matrix ($W = GG'$),*

$$\frac{\mathbb{E}[\tilde{\beta}]}{\beta_0} = \sum_k \frac{g_k \Sigma_{ZX^{\perp},k}}{G'\Sigma'_{ZX^{\perp}}} \frac{\mathbb{E}[\tilde{\beta}_k]}{\beta_0}. \tag{A4}$$

The first corollary interprets the $\alpha_k(W)$ as a sensitivity-to-misspecification *elasticity*. Because of the linear nature of the estimator, it rescales the AGS sensitivity parameter to be unit-invariant, and hence is comparable across instruments.[26] Specifically, $\alpha_k(W)$ is the

[26] AGS (pg. 1558) write: "The second limitation is that the units of [our sensitivity vector] are contingent on the units of [the moment condition]. Changing the measurement of an element [$j$ of the moment condition] from, say, dollars to euros, changes the corresponding elements of [the sensitivity vector]. This does not affect the bias a reader would estimate for specific alternative assumptions, but it does matter for qualitative conclusions about the relative importance of different moments."

percentage point shift in the bias of the over-identified estimator given a percentage point change in the bias from a single industry. The second corollary gives the population version of Bartik's Rotemberg weights.

An alternative approach to measuring sensitivity is to drop an instrument and then re-estimate the model. Let $\hat{\beta}(\hat{W}_{-k})$ be the same estimator as $\hat{\beta}(\hat{W})$, except excluding the $k^{th}$ instrument and define the bias term for $\hat{\beta}(\hat{W}_{-k})$ as $\tilde{\beta}(\hat{W}_{-k}) = \hat{\beta}(\hat{W}_{-k}) - \beta$.

PROPOSITION D.2. *The difference in the bias from the full estimator and the estimator that leaves out the $k^{th}$ industry is:*

$$\frac{\mathbb{E}\left[\tilde{\beta}(\hat{W}) - \tilde{\beta}(\hat{W}_{-k})\right]}{\beta} = \alpha_k(W)\frac{\mathbb{E}[\tilde{\beta}_k]}{\beta} - \frac{\alpha_k(W)}{1-\alpha_k(W)}\sum_{k'\neq k}\alpha_{k'}(W)\frac{\mathbb{E}[\tilde{\beta}_{k'}]}{\beta}.$$

*If $\mathbb{E}[\tilde{\beta}_{k'}] = 0$ for $k' \neq k$, then we get a simpler expression:*

$$\frac{\mathbb{E}\left[\tilde{\beta}(\hat{W}) - \tilde{\beta}(\hat{W}_{-k})\right]}{\beta} = \alpha_k(W)\frac{\mathbb{E}[\tilde{\beta}_k]}{\beta}.$$

**Proof.** Consider the difference in the bias for the two estimators:

$$\mathbb{E}\left[\tilde{\beta}(\hat{W}) - \tilde{\beta}(\hat{W}_{-k})\right] = \sum_{k'}\alpha_{k'}(W)\mathbb{E}[\tilde{\beta}_{k'}] - \sum_{k'\neq k}\alpha_{k'}(W_{-k})\mathbb{E}[\tilde{\beta}_{k'}] \tag{A5}$$

$$= \alpha_k(W)\mathbb{E}[\tilde{\beta}_k] + \sum_{k'\neq k}(\alpha_{k'}(W) - \alpha_{k'}(W_{-k}))\mathbb{E}[\tilde{\beta}_{k'}]. \tag{A6}$$

Now, consider $\alpha_{k'}(W) - \alpha_{k'}(W_{-k})$. If $W = GG'$, then $C(W) = GB'X^{\perp}$ and $\alpha_{k'}(W) = \frac{g_{k'}Z_{k'}X^{\perp}}{\sum_{k'}g_{k'}Z_{k'}X^{\perp}}$. If $W_{-k} = G_{-k}G'_{-k}$, then $\alpha_{k'}(W_{-k}) = \frac{g_{k'}Z_{k'}X^{\perp}}{\sum_{k'\neq k}g_{k'}Z_{k'}X^{\perp}}$, or $\alpha_{k'}(W_{-k}) = \alpha_{k'}(W)/(1 - \alpha_k(W))$.[27] This gives:

$$\mathbb{E}\left[\tilde{\beta}(\hat{W}) - \tilde{\beta}(\hat{W}_{-k})\right] = \alpha_k(W)\mathbb{E}[\tilde{\beta}_k] + \sum_{k'\neq k}\left(\alpha_{k'}(W) - \frac{\alpha_{k'}(W)}{1-\alpha_k(W)}\right)\mathbb{E}[\tilde{\beta}_{k'}] \tag{A7}$$

$$= \alpha_k(W)\mathbb{E}[\tilde{\beta}_k] - \frac{\alpha_k(W)}{1-\alpha_k(W)}\sum_{k'\neq k}\left(\alpha_{k'}(W)\right)\mathbb{E}[\tilde{\beta}_{k'}]. \tag{A8}$$

□

As emphasized by AGS (Appendix A.1), dropping an instrument and seeing how estimates change does not directly measure sensitivity. Instead, this measure combines two forces: the sensitivity of the instrument to misspecification, and how misspecificed the instrument is relative to the remaining instruments.

[27] Note that with TSLS, these results would not hold, as the estimates for the first stage parameters after dropping an industry would be different.

## E Normalization

This appendix presents results to understand the role of normalizations. Following Remark 1.1 we always "drop" industry $k$ by subtracting off $g_k$ from all the growth rates. Proposition E.1 shows that the bias coming each instrument can be written as a weighted average of the bias coming from the remaining $K-1$ instruments. Corollary E.1 shows how the Rotemberg weight gets shifted across instruments depending on which instrument is dropped. Finally, corollary E.2 shows that the average of the $K$ normalizations is to set the unweighted mean of the growth rates to zero.

PROPOSITION E.1. *If the $\sum_{k=1}^{K} z_{lk} = 1 \forall l$, then we can write*

$$\mathbb{E}[\tilde{\beta}_k] = \sum_{j \neq k} \omega_{j,k} \mathbb{E}[\tilde{\beta}_j]$$

*where $\omega_{j,k} = \frac{\Sigma_{ZX_j^{\perp}}}{\sum_{j' \neq k} \Sigma_{ZX_{j'}^{\perp}}}$ and $\mathbb{E}[\tilde{\beta}_j] = \frac{\Sigma_{ZV_j}}{\Sigma_Z X_j^{\perp}}$.*

**Proof.** Recall from Proposition D.1 that

$$\mathbb{E}[\tilde{\beta}_k] = \frac{\Sigma_{ZV_k}}{\Sigma_Z X_k^{\perp}}.$$

When $\sum_{k=1}^{K} z_{lk} = 1$, then $\sum_{k=1}^{K} \Sigma_{ZX_k^{\perp}} = 0$ and $\sum_{k=1}^{K} \Sigma_{ZV_k} = 0$. Then we can write

$$\Sigma_{ZV_k} = -\sum_{j \neq k} \Sigma_{ZV_j}$$

and

$$\Sigma_{ZX_k^{\perp}} = -\sum_{j \neq k} \Sigma_{ZX_j^{\perp}}.$$

Then:

$$\begin{aligned}
\mathbb{E}[\tilde{\beta}_k] &= \frac{\Sigma_{ZV_k}}{\Sigma_{ZX_k^{\perp}}} && \text{(A1)}\\
&= \sum_{j \neq k} \frac{\Sigma_{ZV_j}}{\sum_{j' \neq k} \Sigma_{ZX_{j'}^{\perp}}} && \text{(A2)}\\
&= \sum_{j \neq k} \frac{\Sigma_{ZX_j^{\perp}}}{\sum_{j' \neq k} \Sigma_{ZX_{j'}^{\perp}}} \frac{\Sigma_{ZV_j}}{\Sigma_{ZX_j^{\perp}}} && \text{(A3)}\\
&= \sum_{j \neq k} \omega_{j,k} \mathbb{E}[\tilde{\beta}_j], && \text{(A4)}
\end{aligned}$$

where $\omega_{j,k} = \frac{\Sigma_{ZX_j^{\perp}}}{\sum_{j' \neq k} \Sigma_{ZX_{j'}^{\perp}}}$ and $\mathbb{E}[\tilde{\beta}_j] = \frac{\Sigma_{ZV_j}}{\Sigma_Z X_j^{\perp}}$. $\square$

COROLLARY E.1. *Let* $\sum_{k=1}^{K} z_{lk} = 1 \forall l$. *Let* $\{\alpha_k(GG')\}_{k=1}^{K}$ *be the set of sensitivity-to-misspecification elasticities given a weight matrix formed by a set of growth rates G. Now renormalize the growth rates by subtracting off* $g_k$. *Define* $\alpha_{j,k}(GG') = \alpha_j((G - g_k)(G - g_k)')$ *to be the resulting sensitivity-to-misspecification elasticities (which imply that we have "zeroed out" the* $k^{th}$ *instrument). Then:*

$$\alpha_{j,k}(GG') = \alpha_j(GG') + \omega_{j,k}\alpha_k(GG'),$$

*where* $\omega_{j,k} = \frac{\Sigma_{ZX_j^{\perp}}}{\sum_{j' \neq k} \Sigma_{ZX_{j'}^{\perp}}}$.

**Proof.** Write:

$$\alpha_{j,k}(GG') = \frac{(g_j - g_k)\Sigma_{ZX_j^{\perp}}}{\sum_{j'}(g_{j'} - g_k)\Sigma_{ZX_{j'}^{\perp}}} \tag{A5}$$

$$= \frac{g_j\Sigma_{ZX_j^{\perp}}}{\sum_{j'}(g_{j'} - g_k)\Sigma_{ZX_{j'}^{\perp}}} - \frac{g_k\Sigma_{ZX_j^{\perp}}}{\sum_{j'}(g_{j'} - g_k)\Sigma_{ZX_{j'}^{\perp}}} \tag{A6}$$

$$= \frac{g_j\Sigma_{ZX_j^{\perp}}}{\sum_{j'} g_{j'}\Sigma_{ZX_{j'}^{\perp}}} - \frac{g_k\Sigma_{ZX_j^{\perp}}}{\sum_{j'} g_{j'}\Sigma_{ZX_{j'}^{\perp}}}, \tag{A7}$$

because $g_k \sum_{j'} \Sigma_{ZX_{j'}^{\perp}} = 0$. Then:

$$\alpha_{j,k}(GG') = \alpha_j(GG') - \frac{g_k\Sigma_{ZX_j^{\perp}}}{\sum_{j'} g_{j'}\Sigma_{ZX_{j'}^{\perp}}}\frac{\Sigma_{ZX_k^{\perp}}}{\Sigma_{ZX_k^{\perp}}} \tag{A8}$$

$$= \alpha_j(GG') - \alpha_k(GG')\frac{\Sigma_{ZX_j^{\perp}}}{\Sigma_{ZX_k^{\perp}}}. \tag{A9}$$

Recall that $\Sigma_{ZX_k^{\perp}} = -\sum_{j \neq k} \Sigma_{ZX_j^{\perp}}$. So that: $-\frac{\Sigma_{ZX_j^{\perp}}}{\Sigma_{ZX_k^{\perp}}} = \frac{\Sigma_{ZX_j^{\perp}}}{\sum_{j \neq k} \Sigma_{ZX_j^{\perp}}} = \omega_{j,k}$. Hence:

$$\alpha_{j,k}(GG') = \alpha_j(GG') + \omega_{j,k}\alpha_k(GG').$$

□

COROLLARY E.2. *The average of the K normalizations is:*

$$\alpha_j(GG')^{avg} = \alpha_j(GG') - \frac{\Sigma_{ZX_j^{\perp}}}{K}\left[\frac{\sum_{k=1}^{K} g_k}{\sum_{k=1}^{K} g_k\Sigma_{ZX_k^{\perp}}}\right].$$

*If* $\sum_{k=1}^{K} g_k = 0$, *then* $\alpha_j(GG')^{avg} = \alpha_j(GG')$.

**Proof.** Note that we have two expressions for $\omega_{j,k} = -\frac{\Sigma_{ZX_j^\perp}}{\Sigma_{ZX_k^\perp}} = \frac{\Sigma_{ZX_j^\perp}}{\sum_{j \neq k} \Sigma_{ZX_j^\perp}}$

$$\alpha_j(GG')^{avg} = \frac{1}{K}\sum_{k=1}^{K} \alpha_{j,k}(GG') \tag{A10}$$

$$= \frac{1}{K}\sum_{k=1}^{K} \left[\alpha_j(GG') + \omega_{j,k}\alpha_k(G'G)\right] \tag{A11}$$

$$= \frac{1}{K}\sum_{k=1}^{K} \left[\alpha_j(GG') - \frac{\Sigma_{ZX_j^\perp}}{\Sigma_{ZX_k^\perp}}\alpha_k(G'G)\right] \tag{A12}$$

$$= \alpha_j(GG') - \frac{1}{K}\sum_{k=1}^{K} \left[\frac{\Sigma_{ZX_j^\perp}}{\Sigma_{ZX_k^\perp}} \frac{g_k \Sigma_{ZX_k^\perp}}{\sum_{j'=1}^{K} g_{j'} \Sigma_{ZX_{j'}^\perp}}\right] \tag{A13}$$

$$= \alpha_j(GG') - \frac{\Sigma_{ZX_j^\perp}}{K}\left[\frac{\sum_{k=1}^{K} g_k}{\sum_{k=1}^{K} g_k \Sigma_{ZX_k^\perp}}\right]. \tag{A14}$$

□

# F 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:

$$\ln N_{lt}^S = \sigma_{lt} + \theta \ln w_{lt}. \tag{A1}$$

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

$$\ln N_{lkt}^D = T_{lk}\alpha_{lkt} - \phi \ln w_{lt}. \tag{A2}$$

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:

$$\ln N_{lt}^D = \alpha_{lt} - \phi \ln w_{lt}. \tag{A3}$$

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:[28]

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

Define $z_{lk0} \equiv \frac{exp(T_{lk}\alpha_{lk0})}{\sum_{k'} exp(T_{lk'}\alpha_{lk'0})}$ to be the industry shares in period 0.[29] 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 estimating is:

$$(d\tilde{w}_{lt+1} - d\tilde{w}_{lt}) = (\tau_{t+1} - \tau_t) + \beta(d\tilde{N}_{lt+1} - d\tilde{N}_{lt}) + (\epsilon_{lt+1} - \epsilon_{lt}) \quad \text{(A4)}$$

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 $\beta = \frac{1}{\theta}$. Traditional OLS estimation of equation (A4) is subject to concerns of endogeneity and hence the Bartik instrument may provide a way to estimate $\beta$ consistently.

## F.1 The model's empirical analogue

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

$$\hat{\beta}_{OLS} = \frac{1}{\theta} \frac{\frac{\theta}{(\theta+\phi)^2} Var(d\alpha_{lt+1} - d\alpha_{lt}) - \frac{\phi}{(\theta+\phi)^2} Var(d\sigma_{lt+1} - d\sigma_{lt}) + \frac{\phi-\theta}{\phi+\theta} Cov(d\alpha_{lt+1} - d\alpha_{lt}, d\sigma_{lt+1} - d\sigma_{lt})}{\frac{\theta}{(\theta+\phi)^2} \underbrace{Var(d\alpha_{lt+1} - d\alpha_{lt})}_{\text{demand}} + \frac{\phi}{\theta}\frac{\phi}{(\theta+\phi)^2} \underbrace{Var(d\sigma_{lt+1} - d\sigma_{lt})}_{\text{supply}} + \frac{\phi}{(\theta+\phi)^2} \underbrace{Cov(d\alpha_{lt+1} - d\alpha_{lt}, d\sigma_{lt+1} - d\sigma_{lt})}_{\text{covariance}}}$$

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

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 $Cov[d\sigma_{lt+1} - d\sigma_{lt}, \sum_k z_{lk0}(d\alpha_{kt+1} - d\alpha_{kt})] = 0$. Bartik is invalid if the innovations in the supply shocks are predicted by industry composition. For example, Bartik would not be valid if $d\sigma_{lt+1} - d\sigma_{lt} = d\tilde{\sigma}_{lt+1} - d\tilde{\sigma}_{lt} + \sum_k z_{lk0}(d\sigma_{kt+1} - d\sigma_{kt})$. The relevance condition is that $Cov[d\alpha_{lt+1} - d\alpha_{lt}, \sum_k z_{lk0}(d\alpha_{kt+1} - d\alpha_{kt})] \neq 0$. 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

[28] Combine equation (A1) and (A3) to have the following equilibrium wage equation: $\ln w_{lt} = \frac{1}{\theta+\phi}\alpha_{lt} - \frac{1}{\theta+\phi}\sigma_{lt}$. Then substitute in to equation (A2) for the equilibrium wage, take differences, and add and subtract a $d\alpha_{kt}$.

[29] Note that $\frac{N^D_{lkt}}{N^D_{lt}} = \frac{exp(T_{lk}\alpha_{lkt} - \phi \ln w_{lt})}{exp(\alpha_{lt} - \phi \ln w_{lt})} = \frac{exp(T_{lk}\alpha_{lkt})}{exp(\alpha_{lt})} = \frac{exp(T_{lk}\alpha_{lkt})}{exp(\ln(\sum_k \exp\{T_{lk}\alpha_{lkt}\}))} = \frac{exp(T_{lk}\alpha_{lkt})}{\sum_k \exp\{T_{lk}\alpha_{lkt}\}}$.

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 $Cov[d\alpha_{lt+1} - d\alpha_{lt}, z_{lk0}] \neq 0$, which says that the industry share is correlated with the innovations in the demand shocks.

# G Using growth rates to test overidentification restrictions

We consider a setting where only one instrument has first stage power. We consider a researcher choosing two sets of weights. We show that given one set of weights, denoted by $G_1$, and all but one entry in a second vector $G_2$, it is possible to generate two instruments that have a covariance of 0 and lead to identical parameter estimates. In this case, however, both Bartik instruments use the same identifying variation and so finding that they are uncorrelated does not imply that they leverage different sources of variation.

PROPOSITION G.1. *Suppose that $Z'Z$ is full rank. Suppose that only the first entry in $Z'X$ (a $K \times 1$ vector) is non-zero. Since we assume that the $Z$ constitute a valid instrument, then only the first entry in $Z'Y$ is non-zero. Suppose that we are given two sets of weights, $G_1$ and $G_2$, with $G_{1,1} \neq 0$ and $G_{2,1} \neq 0$. Suppose we leave the last entry of the second vector unknown ($G_{2,K}$). Use these two sets of weights to construct two Bartik instruments: $B_1 = ZG_1$ and $B_2 = ZG_2$. Assume further that all the entries in $G_1' Var(Z)$ are non-zero. Then it is always possible to find $G_{2,K}$ such that:*

1. *The two Bartik instruments lead to identical parameter estimates.*
2. *The two Bartik instruments are uncorrelated.*

The proof shows that the first constraint is always satisfied, and derives an expression for the second constraint.

**Proof.** The first constraint is that:

$$\hat{\beta}_1 = \hat{\beta}_2 \tag{A1}$$

where for $j \in \{1, 2\}$ $\hat{\beta}_j = G_j' Z'Y(G_j' Z'X)^{-1}$. Since only the first entries in $Z'X$ and $Z'Y$ are nonzero, we have:

$$G_j' Z'Y(G_j' Z'X)^{-1} = \frac{\sum_k G_{j,k} Z_k' Y}{\sum_k G_{j,k} Z_k' X} \tag{A2}$$

$$= \frac{G_{j,1} Z_1' Y + \sum_{k=2}^K G_{j,k} Z_k' Y}{G_{j,1} Z_1' X + \sum_{k=2}^K G_{j,k} Z_k' X} \tag{A3}$$

$$= \frac{G_{j,1} Z_1' Y + \sum_{k=2}^K G_{j,k} 0}{G_{j,1} Z_1' X + \sum_{k=2}^K G_{j,k} 0} \tag{A4}$$

$$= \frac{Z_1' Y}{Z_1' X}, \tag{A5}$$

where this derivation uses the fact that only the first entry in $Z'X$ (and $Z'Y$) is nonzero. Hence, if $G_{1,1} \neq 0$ and $G_{2,1} \neq 0$, $\hat{\beta}_1 = \hat{\beta}_2$, which is true by assumption. Hence, the first constraint always holds.

The second constraint is that the covariance between the two Bartik instruments is zero:

$$\begin{aligned}
Cov(B_1, B_2) &= \mathbb{E}[B_1 B_2] - \mathbb{E}[B_1]\mathbb{E}[B_2] && \text{(A6)}\\
&= \mathbb{E}[(ZG_1)(ZG_2)] - \mathbb{E}[ZG_1]\mathbb{E}[ZG_2] && \text{(A7)}\\
&= \mathbb{E}[(ZG_1)'(ZG_2)] - \mathbb{E}[ZG_1]\mathbb{E}[ZG_2] && \text{(A8)}\\
&= G_1'\mathbb{E}[Z'Z]G_2 - G_1'\mathbb{E}[Z']\mathbb{E}[Z]G_2 && \text{(A9)}\\
&= G_1'[\mathbb{E}[Z'Z] - \mathbb{E}[Z']\mathbb{E}[Z]]G_2 && \text{(A10)}\\
&= G_1' Var(Z) G_2, && \text{(A11)}
\end{aligned}$$

where this exploits the fact that $B_{1,l}$ is a scalar so we can take the transpose, and $G_1$ and $G_2$ are non-stochastic so that we can pull them out of the expectation. Let $T = G_1'\Sigma_Z$, where $\Sigma_Z = Var(Z)$. So we can write this first constraint as:

$$TG_2 = 0. \tag{A12}$$

Note that $T$ is $1 \times K$. By assumption, the last entry in $T$ are nonzero. We now construct an expression for this entry. To make $TG_2 = 0$, we need $\sum_{k=1}^{K} T_k G_{2,k} = 0 \Rightarrow G_{2,K} = -\frac{\sum_{k=1}^{K-1} T_k G_{2,k}}{T_K}$.

□

# H The Rotemberg weights with leave-one-out

The formulas we present in Section 3 apply to the case where the weights are common to all locations (i.e., we compute the national industry growth rates using a weighted average that included all locations). Here we present the formulas for the $\alpha_k$ that obtain when we use leave-one-out growth rates to construct the Bartik estimator. We note a few things. First, the numerical equivalence between GMM and Bartik obtains in the limit as the number of locations goes to infinity when we use a leave-one-out estimator. Second, when we use a leave-one-out estimator, the weights sum to one in the limit as the number of locations goes to infinity. (For notational simplicity we suppress notation that residualizes for controls.)

First, we derive how the leave-location-$l$-out estimator of $G$, which we denote by $G_{-l}$, relates to the overall average, $G$ and the location-specific $G_l$ ($L$ is the number of locations):

$$G = \frac{L-1}{L} G_{-l} + \frac{1}{L} G_l \Rightarrow G_{-l} = \frac{L}{L-1} G - \frac{1}{L-1} G_l.$$

Second, we derive a version of Proposition 3.1 with the leave-one-out estimator of $G$. Note that the instrument constructed using leave-$l$-out growth rates in location $l$ is: $B_{l,-l} = Z_l \left(\frac{L}{L-1} G - \frac{1}{L-1} G_l\right)$ where $G$ and $G_l$ are $K \times 1$ vectors and $Z_l$ is a $1 \times K$ vector (and $Z$ will be the $L \times K$ stacked matrix). Then:

$$B_{l,-l} = Z_l \left( \frac{L}{L-1} G_L - \frac{1}{L-1} G_l \right) \tag{A1}$$

$$B_{l,-l} = \frac{L}{L-1} Z_l G - \frac{1}{L-1} Z_l G_l \tag{A2}$$

$$B_{l,-l} = \frac{L}{L-1} B_l - \frac{1}{L-1} X_l, \tag{A3}$$

where the observation is that $Z_l G_l = X_l$. Then the stacked version is:

$$B_{-l} = \frac{L}{L-1} B - \frac{1}{L-1} X,$$

where $B$ is the vector of $B_l$ and $B_{-l}$ is the vector of $B_{l,-l}$.

Then:

$$\hat{\beta} = \frac{B'_{-l} Y}{B'_{-l} X} \tag{A4}$$

$$= \frac{\left( \frac{L}{L-1} B - \frac{1}{L-1} X \right)' Y}{\left( \frac{L}{L-1} B - \frac{1}{L-1} X \right)' X} \tag{A5}$$

$$= \frac{\left( \frac{L}{L-1} (ZG) - \frac{1}{L-1} X \right)' Y}{\left( \frac{L}{L-1} (ZG) - \frac{1}{L-1} X \right)' X}. \tag{A6}$$

As before:

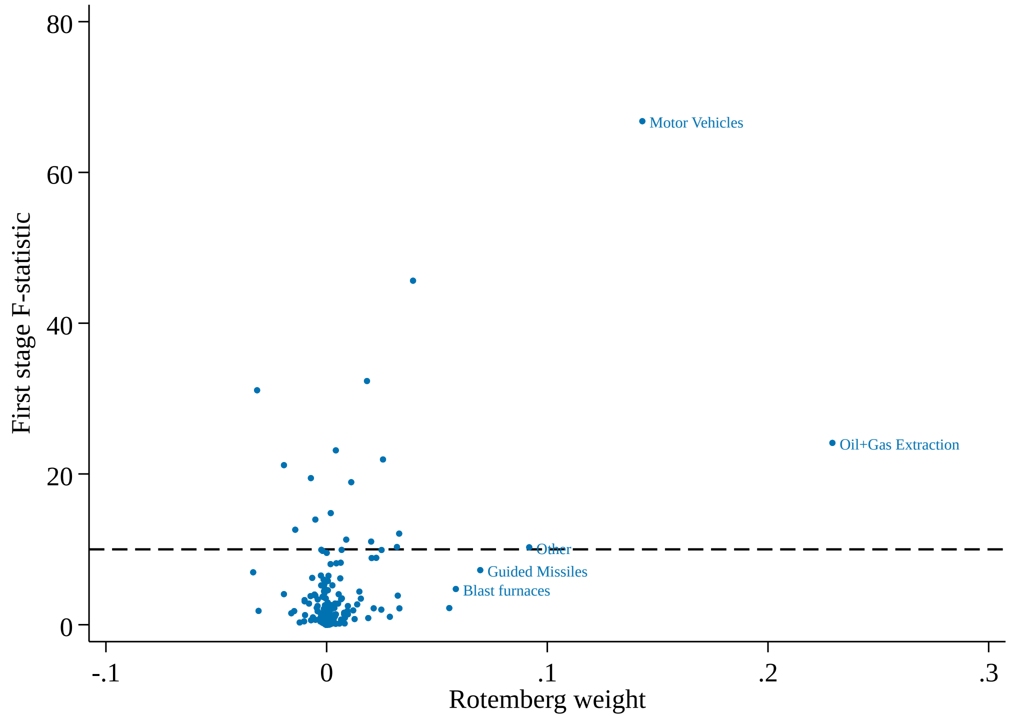
$$\beta_k = \frac{Z'_k Y}{Z'_k X}. \tag{A7}$$

Then one can show:

$$\alpha_k = \frac{\frac{L}{L-1} g_k Z'_k X - \frac{1}{L-1} X' Y \beta_k^{-1}}{\sum_k \frac{L}{L-1} g_k Z'_k X - \frac{1}{L-1} X' X}. \tag{A8}$$

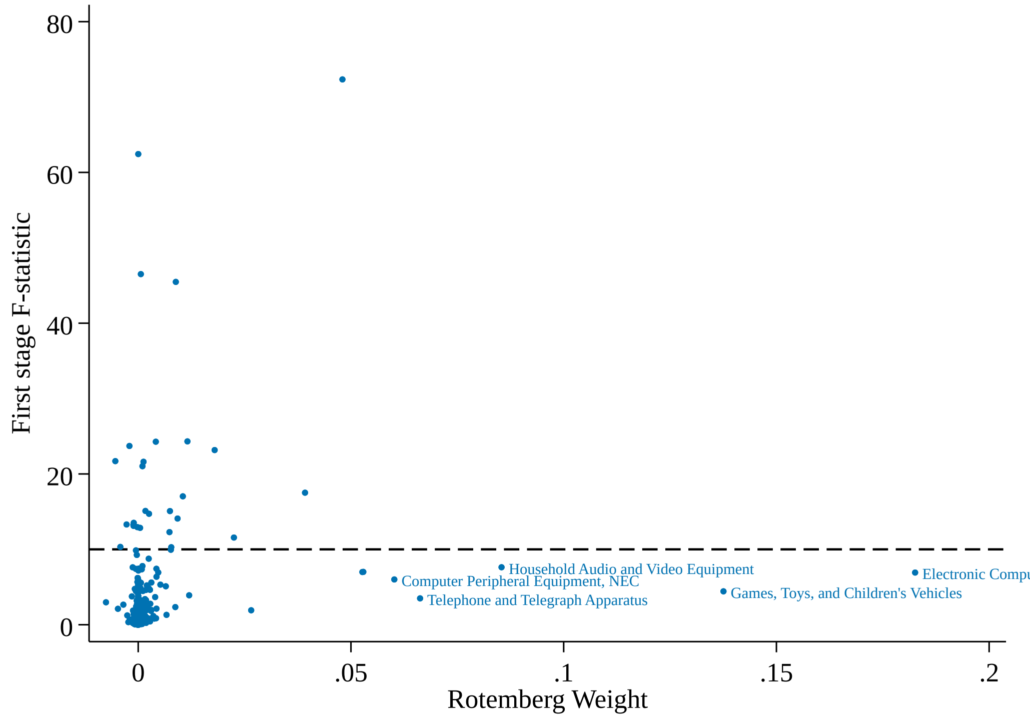
By inspection, $\sum_k \alpha_k \neq 1$. However, as $L \to \infty$ the sum converges to 1 as the leave-one-out terms drop out.

Figure A1: First stage versus Rotemberg weights: canonical setting




*Notes:* This figure plots each instrument's Rotemberg weight against the first stage F-statistic. Each point represents the estimates for an instrument, where instruments are aggregated across time periods following Section 3.3. The labelled industries correspond to the five highest Rotemberg weight industries from Table 1. The dashed horizontal line is equal to 10.

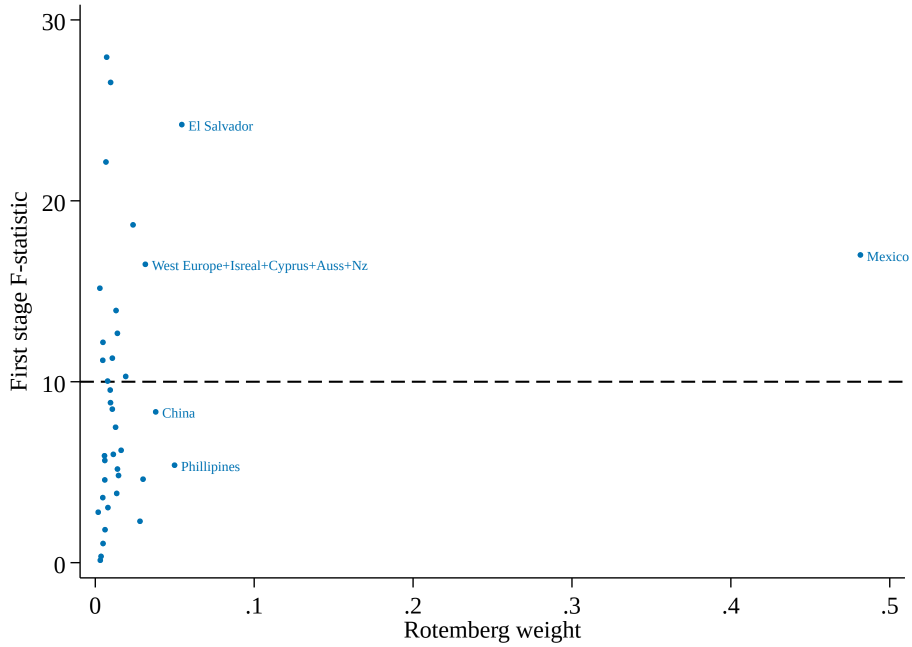
Figure A2: First stage versus Rotemberg weights: China shock



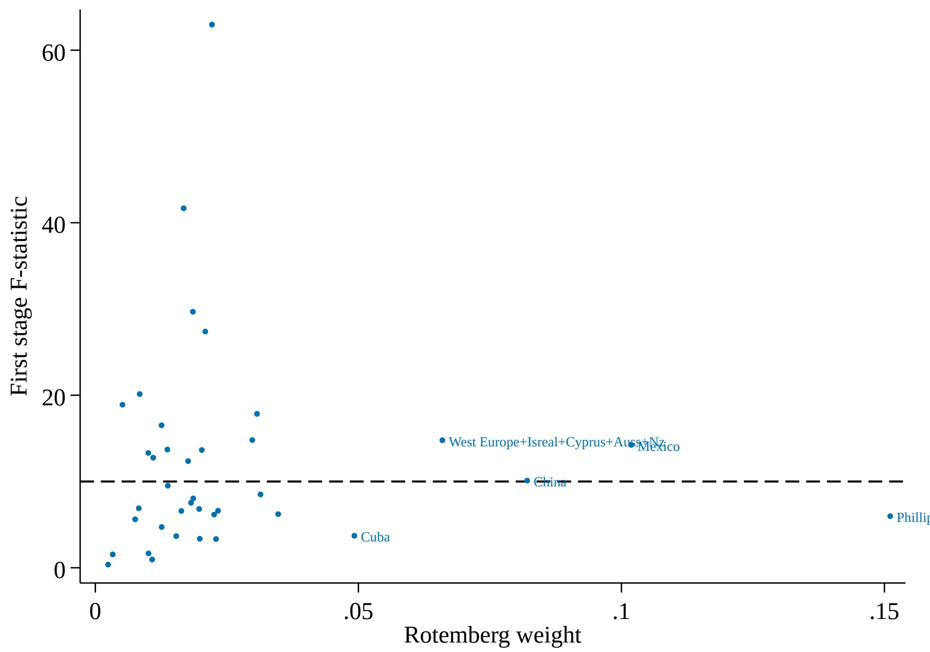

*Notes:* This figure plots each instrument's Rotemberg weight against the first stage F-statistic. Each point represents the estimates for an instrument, where instruments are aggregated across time periods following Section 3.3. The labelled industries correspond to the five highest Rotemberg weight industries from Table 4. The dashed horizontal line is equal to 10.

Figure A3: First stage versus Rotemberg weights: immigrant enclave

Panel A: High school equivalent

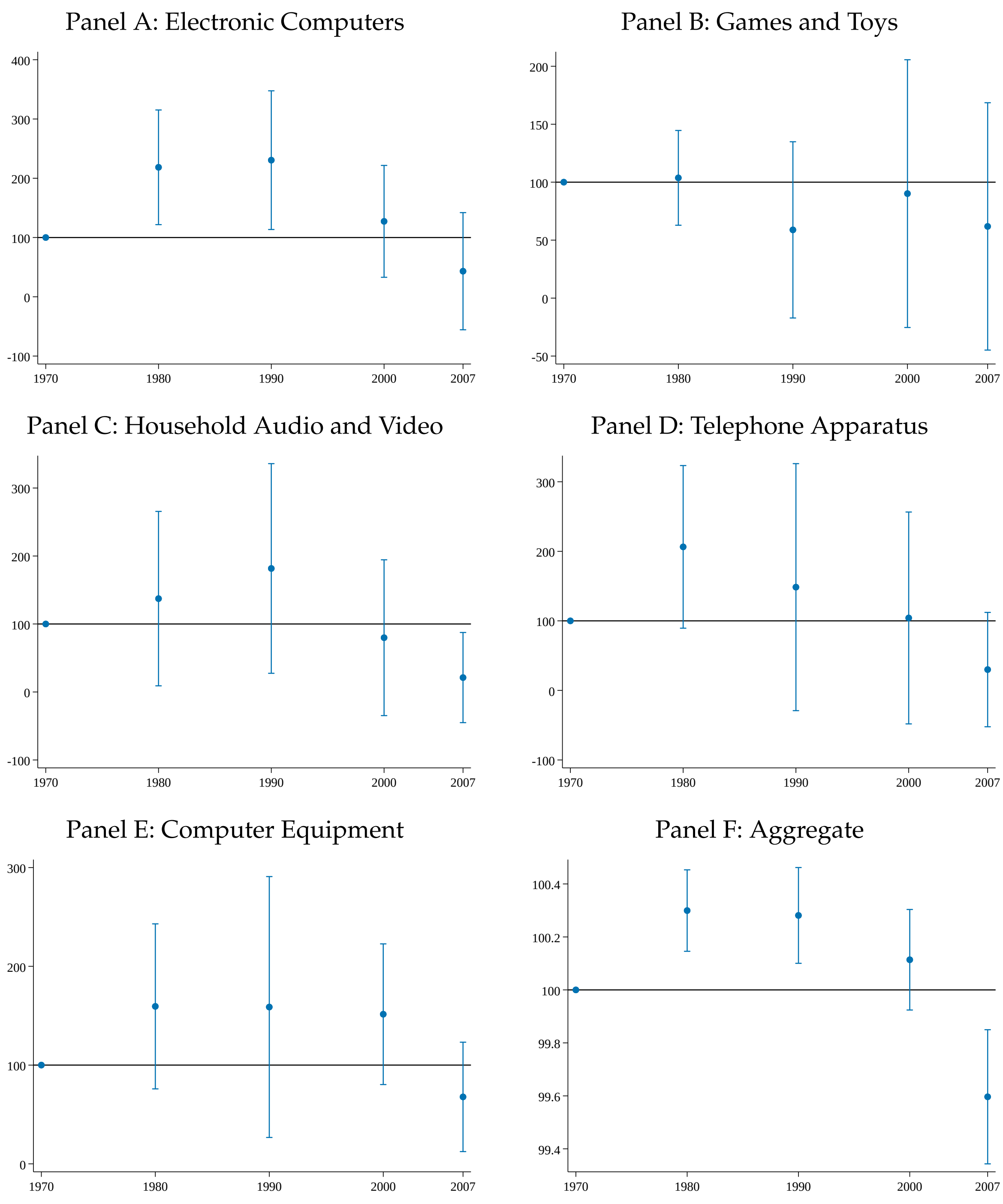



Panel B: College equivalent




*Notes:* This figure plots each instrument's Rotemberg weight against the first stage F-statistic. Each point represents the estimates for an instrument, where instruments are aggregated across time periods following Section 3.3. The labelled industries correspond to the five highest Rotemberg weight industries from Table 7. The dashed horizontal line is equal to 10.

Figure A4: Pre-trends for high Rotemberg weight industries (1990 shares): China shock




*Notes:* These figures report pre-trends for the overall instrument and the top-5 Rotemberg weight industries as reported in Table 4. The Figures fix industry shares at the 1990 values and report the effect of these industry shares on manufacturing employment. We run regressions in growth rates and then convert to levels. We normalize 1970 to 100, and compute the standard errors using the delta method. For the aggregate panel, we use the Bartik estimate for 1990.