# Online supplement to:
# "Shift-Share Designs: Theory and Inference"

Rodrigo Adão* Michal Kolesár† Eduardo Morales‡

August 13, 2018

This online appendix is organized as follows. Appendix C contains details and extensions of the model in Section 3. Appendix D contains additional results for the placebo exercises in Sections 2 and 5. Appendix E contains additional results for the empirical applications in Section 6.

## Appendix C Stylized economic model: details and extensions

In Appendix C.1, we derive the relationship between the sector-level price index and fundamental economic shocks for the model described in Appendix A. In Appendices C.2 and C.3, we provide alternative microfoundations for the equilibrium relationship in eq. (8). Finally, in Appendix C.4, we incorporate migration into the baseline microfoundation described in Appendix A.

### C.1 Sector-specific price index

According to the microfoundation in Appendix A, the price change in every sector $s$, $\hat{P}_s$, depends on the shocks $\hat{A}_{is}$, $\hat{\gamma}_s$ and $\hat{v}_i$ of all sectors and regions of the world economy. Specifically, the change in the sector-specific price index is

$$\hat{P}_s = -\sum_{s'} \theta_{ss'} \sum_{j=1}^{J} x^0_{js'} (\hat{A}_{js'} + \tilde{\lambda}_j \hat{v}_j - \tilde{\lambda}_j \sum_k l^0_{jk} [\hat{\gamma}_k + (\sigma_k - 1)\hat{A}_{jk}]), \tag{C.1}$$

where $\tilde{\lambda}_j \equiv \lambda_j/\phi = \left[\phi + \sum_s l^0_{is}\sigma_s\right]^{-1}$, $\{\theta_{ss'}\}^{S,S}_{s=1,s'=1}$ are positive constants, and $x^0_{js}$ is the share of the world production in sector $s$ that corresponds to region $j$ in the initial equilibrium; i.e. $x^0_{js} \equiv X^0_{js} / \sum_{i=1}^{J} X^0_{is}$.

Imposing that all regions in a country $c$ are small is equivalent to assuming that $x^0_{js} \approx 0$ for all $j \in J_c$ and for $s = 1, \ldots, S$. Therefore, when all regions $j \in J_c$ are small, we can rewrite the change in

*University of Chicago Booth School of Business. Email: rodrigo.adao@chicagobooth.edu
†Princeton University. Email: mkolesar@princeton.edu
‡Princeton University. Email: ecmorales@princeton.edu

the sector-specific price index as

$$\hat{P}_s = -\sum_{s'} \theta_{ss'} \sum_{j \notin J_c} x^0_{js'} (\hat{A}_{js'} + \tilde{\lambda}_j \hat{v}_j - \tilde{\lambda}_j \sum_k l^0_{jk} [\hat{\gamma}_k + (\sigma_k - 1)\hat{A}_{jk}]). \tag{C.2}$$

In this case, $\hat{P}_s$ does not depend on the labor supply shocks and technology shocks in any region $j$ included in country $c$; i.e. $\hat{P}_s$ depends neither on $\{\hat{A}_{js'}\}^S_{j \in J_c, s=1}$ nor on $\{\hat{v}_j\}_{j \in J_c}$

**Proof of eq. (C.1).** Equations (A.7) and (A.16) imply that

$$\hat{P}_s - \sum_k \tilde{\theta}_{sk} \hat{P}_k = \sum_j x^0_{js} (\tilde{\lambda}_j \sum_k l^0_{jk} [\hat{\gamma}_k + (\sigma_k - 1)\hat{A}_{jk}] - \tilde{\lambda}_j \hat{v}_j - \hat{A}_{js}),$$

where $\tilde{\theta}_{sk} \equiv \sum_j x^0_{js} l^0_{jk} \tilde{\lambda}_j (\sigma_k - 1)$. Let us use bold variables to denote vectors, $\boldsymbol{y} \equiv [y_s]_s$, and bar bold variables to denote matrices, $\bar{\boldsymbol{a}} \equiv [a_{sk}]_{s,k}$. Thus,

$$\left(I - \bar{\boldsymbol{\theta}}\right) \hat{\boldsymbol{P}} = \hat{\boldsymbol{\eta}},$$

with $\hat{\eta}_s \equiv \sum_j x^0_{js} \left(\tilde{\lambda}_j \sum_k l^0_{jk} \left[\hat{\gamma}_k + (\sigma_k - 1)\hat{A}_{jk}\right] - \tilde{\lambda}_j \hat{v}_j - \hat{A}_{js}\right)$.

In order to obtain eq. (C.1), it is sufficient to show that $\left(I - \bar{\boldsymbol{\theta}}\right)$ is a nonsingular m-matrix and, therefore, it has a positive inverse matrix. To establish this result, notice that $\tilde{\theta}_{sk} \in (0, 1)$ because $\sigma_k > 1$ and $\phi > 0$ imply that

$$0 < l^0_{jk} \tilde{\lambda}_j (\sigma_k - 1) = \frac{l^0_{jk}(\sigma_k - 1)}{\phi + \sum_k l^0_{jk} \sigma_k} < \frac{l^0_{jk} \sigma_k}{\phi + \sum_k l^0_{jk} \sigma_k} < 1.$$

Finally, to show that $\left(I - \bar{\boldsymbol{\theta}}\right)$ is nonsingular, it is sufficient to establish that it is diagonal dominant:

$$\begin{aligned} |1 - \tilde{\theta}_{sk}| - \sum_{k \neq s} |\tilde{\theta}_{sk}| &= 1 - \sum_j x^0_{js} \frac{l^0_{js}(\sigma_s - 1)}{\phi + \sum_k l^0_{jk} \sigma_k} - \sum_{k \neq s} \sum_j x^0_{js} \frac{l^0_{jk}(\sigma_k - 1)}{\phi + \sum_k l^0_{jk} \sigma_k}, \\ &= \sum_j x^0_{js} \left(1 - \frac{\sum_k l^0_{jk}(\sigma_k - 1)}{\phi + \sum_k l^0_{jk} \sigma_k}\right) \\ &= \sum_j x^0_{js} \left(\frac{\phi + 1}{\phi + \sum_k l^0_{jk} \sigma_k}\right) > 0. \blacksquare \end{aligned}$$

## C.2 Sector-specific factors of production

We extend here the model described in Appendix A to incorporate other factors of production. In particular, we introduce a sector-specific factor, as in Jones (1971) and, more recently, Kovak (2013).

### C.2.1 Environment

The only difference with respect to the setting described in Appendix A.1 is that the production function in eq. (A.1) is substituted for a Cobb-Douglas production function that combines labor and capital inputs:

$$Q_{is} = A_{is} \left(L_{is}\right)^{1-\theta_s} \left(K_{is}\right)^{\theta_s}.$$

We assume that capital is a sector-specific factor of production (sector-$s$ capital has no use in any other sector) and that, for every sector, each region has an endowment of sector-specific capital $\bar{K}_{is}$.

### C.2.2 Equilibrium

**Consumption.** The consumer's problem is identical to that in Appendix A.2.

**Labor supply.** The labor supply decision is identical to that in Appendix A.2.

**Producer's problem.** Conditional on the region-$i$ equilibrium wage $\omega_i$ and rental rate of sector-$s$ capital $R_{is}$, the cost minimization problem of the sector-$s$ region-$i$ representative firm and the market clearing condition for sector-$s$ region-$i$ specific capital imply that

$$\frac{1-\theta_s}{\theta_s}\frac{\bar{K}_{is}}{L_{is}} = \frac{\omega_i}{R_{is}}.$$

Conditional on the sector-$s$ region-$i$ final good price $p_{is}$, the firm's zero profit condition implies that

$$p_{is} A_{is} \tilde{\theta}_s = \left(\omega_i\right)^{1-\theta_{is}} \left(R_{is}\right)^{\theta_{is}},$$

where $\tilde{\theta}_s \equiv \left(\theta_s\right)^{\theta_s} \left(1-\theta_s\right)^{1-\theta_s}$. The combination of these two conditions yields the demand for labor in sector $s$ and region $i$,

$$L_{is} = \frac{1-\theta_s}{\theta_s} \bar{K}_{is} \left(\frac{p_{is} A_{is} \tilde{\theta}_s}{\omega_i}\right)^{\frac{1}{\theta_s}}, \tag{C.3}$$

and the total sales of the sector-$s$ region-$i$ good as a function of the output price $p_{is}$ are

$$X_{is} = \frac{1}{1-\theta_s} \omega_i L_{is} = \frac{\bar{K}_{is}}{\theta_s} \left(p_{is} A_{is} \tilde{\theta}_s\right)^{\frac{1}{\theta_s}} \left(\omega_i\right)^{1-\frac{1}{\theta_s}}. \tag{C.4}$$

**Goods market clearing.** Applying the same normalization as in Appendix A.1, $W = 1$, the total expenditure in the sector-$s$ region-$i$ good is equal to $x_{is}\gamma_s$, with $x_{is}$ defined in eq. (A.7) as a function of the equilibrium prices $p_{is}$. Equating $x_{is}\gamma_s$ and eq. (C.4), we can solve for the equilibrium value of $p_{is}$ as a function of the sector-$s$ price index $P_s$:

$$p_{is} = \left[\frac{\bar{K}_{is}}{\theta_s} \left(A_{is} \tilde{\theta}_s\right)^{\frac{1}{\theta_s}} \left(\omega_i\right)^{1-\frac{1}{\theta_s}} \frac{\left(P_s\right)^{1-\sigma_s}}{\gamma_s}\right]^{-\theta_{is}\eta_{is}}, \tag{C.5}$$

where $\delta_s \equiv \left(1+\theta_s(\sigma_s-1)\right)^{-1} \in (0,1)$. Additionally, combining the expressions in eqs. (C.3) and (C.5), we obtain an expression for labor demand in sector-$s$ region-$i$ as a function of the equilibrium wage

$\omega_i$, the sector-$s$ price $P_s$ and other exogenous determinants:

$$L_{is} = \kappa_{is}\gamma_s^{\delta_s} \left(A_{is}P_s\right)^{(\sigma_s-1)\delta_s} \left(\omega_i\right)^{-\sigma_s\delta_s}, \tag{C.6}$$

where $\kappa_{is} \equiv (1-\theta_s)(\bar{K}_{is}\tilde{\theta}_s^{\frac{1}{\theta_s}}/\theta_s)^{1-\delta_s}$. Note that this labor demand equation is analogous to that in eq. (3), with the region- and sector-specific demand shifter $D_{is}$ defined as

$$D_{is} = \kappa_{is}(\gamma_s)^{\delta_s} \left(A_{is}P_s\right)^{(\sigma_s-1)\delta_s},$$

and with the labor demand elasticity now defined as $\sigma_s\delta_s$. Note that the labor demand elasticity in eq. (3) is identical to that in eq. (C.6) in the specific case in which $\delta_s = 1$, which will hold when $\theta_s = 0$.

If, without loss of generality, we split the region- and sector-specific productivity $A_{is}$ into a sector component $A_s$ and a residual $\tilde{A}_{is}$, $A_{is} = A_s\tilde{A}_{is}$, and we further consider $P_s$ as our sectoral shock of interest, we can decompose $D_{is}$ as in eq. (4), with

$$\chi_s = P_s, \tag{C.7}$$

$$\rho_s = (\sigma_s - 1)\delta_s, \tag{C.8}$$

$$\mu_s = (A_s)^{(\sigma_s-1)\delta_s}(\gamma_s)^{\delta_s}, \tag{C.9}$$

$$\eta_{is} = \kappa_{is}(\tilde{A}_{is})^{(\sigma_s-1)\delta_s}. \tag{C.10}$$

**Labor market clearing.** Given the sector- and region-specific labor demand in eq. (C.6), total labor demand in region $i$ is

$$L_i = \sum_{s=1}^{S} \kappa_{is}\gamma_s^{\delta_s} \left(A_{is}P_s\right)^{(\sigma_s-1)\delta_s} \left(\omega_i\right)^{-\sigma_s\delta_s}. \tag{C.11}$$

Labor market clearing requires labor supply in eq. (A.8) to equal labor demand in eq. (C.11):

$$v_i(\omega_i)^{\phi} = \sum_{s=1}^{S} \kappa_{is}\gamma_s^{\delta_{is}} \left(A_{is}P_s\right)^{(\sigma_s-1)\delta_{is}} \left(\omega_i\right)^{-\sigma_s\delta_{is}}, \qquad j = 1, \ldots, J. \tag{C.12}$$

**Equilibrium.** Given productivity parameters $\{A_{is}\}_{i=1,s=1}^{J,S}$, sector- and region-specific capital inputs $\{\bar{K}_{is}\}_{i=1,s=1}^{J,S}$, preference and production function parameters $\{\sigma_s, \gamma_s, \theta_s\}_{s=1}^{S}$, labor supply parameters $\{v_i\}_{i=1}^{J}$, and normalizing world income to equal 1, $W = 1$, we can use eq. (A.6), eq. (C.5), and eq. (C.12) to solve for the equilibrium wage in every world region, $\{\omega_i\}_{i=1}^{J}$, the equilibrium price of every sector-region specific good $\{p_{is}\}_{i=1,s=1}^{J,S}$, and the sectoral price indices $\{P_s\}_{s=1}^{S}$. Given these equilibrium wages and sectoral price indices, we can use eq. (C.11) to solve for the equilibrium level of employment in every region, $\{L_i\}_{i=1}^{J}$.

### C.2.3 Labor market impact of sectoral shocks

We assume that the production function parameters $\{\theta_s\}_{s=1}^{S}$ are constant over time and focus on the labor market outcomes in a country $c$ formed by a set of $N$ small open regional economies; see Appendix A.3 for details on the definition of these small regional economies.

Across periods, the microfoundation described in Sections C.2.1 and C.2.2 implies that the changes in labor market outcomes in all $N$ regions of country $c$, $\{\omega_i, L_i\}_{i \in J_c}$, are generated by changes in sectoral prices $\{P_s\}_{s=1}^{S}$, changes in an aggregate of all other sectoral shocks, $\{(A_s)^{(\sigma_s-1)\delta_s}(\gamma_s)^{\delta_s}\}_{s=1}^{S}$, changes in labor supply parameters of all regions in country $c$, $\{v_i\}_{i \in J_c}$, and changes in a composite of sector- and region-specific productivity parameters and capital inputs, $\{\kappa_{is}(\tilde{A}_{is})^{(\sigma_s-1)\delta_s}\}_{i \in J_c, s=1}^{S}$.

**Isomorphism.** Up to a first-order approximation around the initial equilibrium, eqs. (C.11) and (C.12) imply that

$$\hat{L}_i = \sum_{s=1}^{S} l_{is}^0 \left[ \beta_{is}\hat{P}_s + \lambda_i((\sigma_s - 1)\delta_s \hat{A}_s + \delta_s \hat{\gamma}_s) + \lambda_i((\sigma_s - 1)\delta_s \hat{\tilde{A}}_{is} + \hat{\kappa}_{is}) \right] + (1 - \lambda_i)\,\hat{v}_i, \tag{C.13}$$

with $\beta_{is} = (\sigma_s - 1)\delta_s \lambda_i$ and $\lambda_i \equiv \phi(\phi + \sum_s l_{is}^0 \sigma_s \delta_s)^{-1}$. Given the equivalences in eqs. (C.7) to (C.10) and the expression for $\lambda_i$, the expression in eq. (C.13) is identical to that in eq. (8). Consequently, the environment described in Sections C.2.1 and C.2.2 does indeed provide a microfoundation for the equilibrium relationship in eq. (8).

## C.3 Sector-specific preferences

We extend the model described in Appendix A to allow workers to have idiosyncratic preferences for being employed in the different $s = 1, \ldots, S$ sectors and for being non-employed $s = 0$.

### C.3.1 Environment

The only difference with respect to the setting described in Appendix A.1 is that the utility function in eqs. (A.4) and (A.5) is substituted by an alternative utility function that features workers idiosyncratic preferences for being employed in the different $s = 1, \ldots, S$ sectors and for being non-employed $s = 0$. Specifically, we assume here that, conditional on obtaining utility $C_i$ from the consumption of goods, the utility of a worker $\iota$ living in region $i$ is

$$U_{is} = u_s(\iota) C_i, \tag{C.14}$$

and, to simplify the analysis, we assume that $u_s(\iota)$ is i.i.d. across individuals $\iota$ and sectors $s$ with a Fréchet cumulative distribution function; i.e. for every region $i = 1, \ldots, J$ and sector $s = 0, \ldots, S$,

$$F_u(u) = e^{-v_{is} u^{-\phi}}, \qquad \phi > 1. \tag{C.15}$$

A similar modeling of workers' sorting patterns across sectors has been introduced in Galle, Rodríguez-Clare and Yi (2017) and Burstein, Morales and Vogel (2018). See Adão (2016) for a framework that relaxes the distributional assumption in eq. (C.15). Given that individuals have heterogeneous preferences for employment in different sectors, workers are no longer indifferent across sectors and, thus, equilibrium wages $\{\omega_{is}\}_{s=1}^{S}$ may vary across sectors within a region $i$. As in the main text, we assume that workers that choose the non-employment sector $s = 0$ in region $i$ receive non-employment benefits $b_i$.

### C.3.2 Equilibrium

**Consumption.** The consumer's problem is identical to that in Appendix A.2.

**Labor supply.** Conditional on the equilibrium wages $\{\omega_{is}\}_{s=1}^{S}$, the labor supply in sector $s = 1, \ldots, S$ of region $i$ is

$$L_{is} = M_i \frac{v_{is}(\omega_{is})^{\phi}}{\Phi_i} \quad \text{with} \quad \Phi_i \equiv v_{i0} b_i^{\phi} + \sum_{s=1}^{S} v_{is}(\omega_{is})^{\phi}, \tag{C.16}$$

and the labor supply in the non-employment sector $s = 0$ is

$$L_{i0} = M_i \frac{v_{i0}(b_i)^{\phi}}{\Phi_i}. \tag{C.17}$$

**Producer's problem.** In perfect competition, firms must earn zero profits and, therefore,

$$p_{is} = \frac{\omega_{is}}{A_{is}}. \tag{C.18}$$

**Goods market clearing.** The conditions determining the equilibrium in the good's market and, consequently, the region- and sector-specific labor demand equations are identical to those in Appendix A.2.

**Labor market clearing.** Combining the region- and sector-specific labor supply equation (C.16) with the region- and sector-specific labor demand equation in eq. (A.10), and imposing the normalization $W = 1$, the labor market clearing condition in every sector $s = 1, \ldots, S$ and region $i = 1, \ldots, J$ is

$$M_i \frac{v_{is}(\omega_{is})^{\phi}}{\Phi_i} = (\omega_{is})^{-\sigma_s} (A_{is} P_s)^{\sigma_s - 1} \gamma_s. \tag{C.19}$$

**Equilibrium.** Given productivity parameters $\{A_{is}\}_{i=1,s=1}^{J,S}$, preference parameters $\{\sigma_s, \gamma_s\}_{s=1}^{S}$, labor supply parameters $\{v_{is}\}_{i=1,s=0}^{J,S}$, and normalizing world income to equal 1, $W = 1$, we can use eq. (A.6), eq. (C.18), and eq. (C.19) to solve for the equilibrium wage in every sector and region, $\{\omega_{is}\}_{i=1,s=1}^{J,S}$, the equilibrium price of every sector- and region-specific good $\{p_{is}\}_{i=1,s=1}^{J,S}$, and the sectoral price indices $\{P_s\}_{s=1}^{S}$. Given these equilibrium wages and sectoral price indices, we can use eqs. (C.16) and (C.17) to solve for the equilibrium level of employment in every sector and region, $\{L_{is}\}_{i=1,s=0}^{J,S}$.

### C.3.3 Labor market impact of sectoral shocks

We focus on the labor market outcomes in a country $c$ formed by a set of $N$ small open regional economies; see Appendix A.3 for details on the definition of these small regional economies.

Across periods, the microfoundation described in Sections C.3.1 and C.3.2 implies that the changes in labor market outcomes in all $N$ regions of country $c$, $\{\omega_i, L_i\}_{i\in J_c}$, are generated by changes in sectoral prices $\{P_s\}_{s=1}^S$, changes in an aggregate of all other sectoral shocks, $\{(A_s)^{(\sigma_s-1)}\gamma_s\}_{s=1}^S$, changes in labor supply parameters and productivities of all regions in country $c$ and all sectors, $\{v_{is}, (\tilde{A}_{is})^{(\sigma_s-1)}\}_{i\in J_c, s=1}^S$, and changes in benefits in all regions in country $c$, $\{b_i\}_{i\in J_c}$.

**Isomorphism.** Given that the total population of a region, $M_i$, is fixed across time periods, it holds that, to a first-order approximation, $l_{i0}^0\hat{L}_{i0} + (1-l_{i0}^0)\hat{L}_i = 0$, where $\hat{L}_i$ denotes the log-change in total population in region $i$. Therefore, the change in total employment in region $i$ may be written as

$$\begin{aligned}
\hat{L}_i &= -\frac{l_{i0}^0}{1-l_{i0}^0}\hat{L}_{i0} \\
&= \frac{l_{i0}^0}{1-l_{i0}^0}(\hat{\Phi}_i - \phi\hat{b}_i - \hat{v}_{i0}) \\
&= \frac{l_{i0}^0}{1-l_{i0}^0}(\sum_{s=0}^{S} l_{is}^0\hat{v}_{is} + \phi l_{i0}^0\hat{b}_i + \phi\sum_{s=1}^{S} l_{is}^0\hat{\omega}_{is} - \phi\hat{b}_i - \hat{v}_{i0}).
\end{aligned} \tag{C.20}$$

From eq. (C.19), we can express the changes in wages in every sector and every region of country $c$ as

$$\hat{\omega}_{is} = (\phi+\sigma_s)^{-1}\left(\hat{\Phi}_i + \hat{\gamma}_s + (\sigma_s-1)(\hat{A}_{is}+\hat{P}_s) - \hat{v}_{is}\right). \tag{C.21}$$

Combining eqs. (C.20) and (C.21), we can re-express the change in total employment in region $i$ as

$$\hat{L}_i = \sum_{s=1}^{S} l_{is}^0[\beta_{is}\hat{P}_s + \lambda_i(\phi+\sigma_s)^{-1}((\sigma_s-1)\hat{A}_s + \hat{\gamma}_s) + \lambda_i(\phi+\sigma_s)^{-1}(\sigma_s-1)\hat{\tilde{A}}_{is}] + \hat{v}_i, \tag{C.22}$$

where $\hat{v}_i = l_{i0}^0(1-l_{i0}^0)^{-1}(\hat{v}_i - \phi\hat{b}_i - \hat{v}_{i0})$, $\hat{v}_i = (1-\phi\sum_{s=1}^S l_{is}^0(\phi+\sigma_s)^{-1})^{-1}(\phi l_{i0}^0\hat{b}_i + l_{i0}^0\hat{v}_{i0} + \sum_{s=1}^S l_{is}^0\sigma_s(\phi+\sigma_s)^{-1}\hat{v}_{is})$, $\beta_{is} = (\sigma_s-1)(\phi+\sigma_s)^{-1}\lambda_i$, and $\lambda_i = \phi l_{i0}^0(1-l_{i0}^0)^{-1}(1-\phi\sum_{s=1}^S l_{is}^0(\phi+\sigma_s)^{-1})^{-1}$.

The expression in eq. (C.22) is identical to that in eq. (8) under the following equivalences

$$\begin{aligned}
\chi_s &= P_s, \\
\rho_s &= (\sigma_s-1)(\phi+\sigma_s)^{-1}, \\
\mu_s &= (A_s)^{(\sigma_s-1)(\phi+\sigma_s)^{-1}}(\gamma_s)^{(\phi+\sigma_s)^{-1}}, \\
\eta_{is} &= (\tilde{A}_{is})^{(\sigma_s-1)(\phi+\sigma_s)^{-1}},
\end{aligned}$$

and the adjustment of the expression for $\lambda_i$ and $\hat{v}_i$. Consequently, the environment described in Sections C.3.1 and C.3.2 does indeed provide a microfoundation for the equilibrium relationship in eq. (8).

## C.4 Allowing for regional migration

We extend here the baseline environment described in Appendix A.1 to allow for mobility of individuals across regions within a single country $c$. We still assume that the number of individuals living in each country $c$ is fixed and equal to $M_c$.

### C.4.1 Environment

The only difference with respect to the setting described in Appendix A.1 is that the mass of individuals living in a region $i$, $M_i$, is no longer fixed. Instead, we assume that, before the realization of the shock $u(\iota)$ in eq. (A.4), individuals must decide their preferred region of residence taking into account their idiosyncratic preferences for local amenities in each region. Specifically, we assume that the utility to individual $\iota$ of residing in region $i$ is

$$U(\iota) = \tilde{u}_i(\iota)\,(\bar{U}_i(\omega_i/P, b_i/P) - 1) \tag{C.23}$$

where $\bar{U}_i(\omega_i/P, b_i/P)$ is the expected utility of residing in region $i$, as determined by eqs. (A.4) and (A.5), and $\tilde{u}_i(\iota)$ is the idiosyncratic amenity level of region $i$ for individual $\iota$. For simplicity, we assume that individuals draw their idiosyncratic amenity level independently (across individuals and regions) from a Type I extreme value distribution:

$$\tilde{u}_i(\iota) \sim F_{\tilde{u}}(\tilde{u}) = e^{-\tilde{u}^{-\tilde{\phi}}}, \qquad \tilde{\phi} > 0. \tag{C.24}$$

A similar modeling of labor mobility has been previously imposed, among others, in Allen and Arkolakis (2016), Redding (2016), Allen, Arkolakis and Takahashi (2018), Monte, Redding and Rossi-Hansberg (2018) and Fajgelbaum et al. (2018), among others. See Redding and Rossi-Hansberg (2017) for additional references.

### C.4.2 Equilibrium

**Consumption.** The consumer's problem is identical to that in Appendix A.2.

**Labor supply.** To characterize the labor supply in region $i$, we first compute $\bar{U}_i(w_i/P, b_i/P)$:

$$\begin{aligned}
\bar{U}_i(\omega_i/P, b_i/P) &= \frac{\omega_i}{P}\int_{b_i/\omega_i}^{\infty} u dF_u(u) + \frac{b_i}{P}\int_{\nu_i}^{b_i/\omega_i} dF_u(u),\\
&= \phi\frac{\omega_i}{P}\int_{b_i/\omega_i}^{\infty}\left(\frac{u}{\nu_i}\right)^{-\phi} du + \frac{b_i}{P}\int_{\nu_i}^{b_i/\omega_i}\frac{\phi}{\nu_i}\left(\frac{u}{\nu_i}\right)^{-\phi-1} du,\\
&= \frac{\phi}{\phi-1}\frac{\omega_i}{P}\nu_i^{\phi}\left(\frac{\omega_i}{b_i}\right)^{\phi-1} + \frac{b_i}{P}\left(1-\nu_i^{\phi}\left(\frac{\omega_i}{b_i}\right)^{\phi}\right),\\
&= \frac{b_i}{P}\left(1+\frac{1}{\phi-1}\nu_i^{\phi}\left(\frac{\omega_i}{b_i}\right)^{\phi}\right).
\end{aligned}$$

To simplify the analysis, we assume that the unemployment benefit is identical in all regions and equal to the price index $P$; i.e. $b_i = P$ for all $i \in J$. Defining $v_i \equiv (\nu_i/b_i)^{\phi}$ as in eq. (A.8), the assumption that $b_i = P$ for all $i \in J$ implies that $v_i \equiv \nu_i / P$ and, thus,

$$\bar{U}_i(\omega_i/P, b_i/P) = 1 + \frac{1}{\phi - 1} v_i \left(\frac{\omega_i}{P}\right)^{\phi},$$

and the share of national population in region $i$ is

$$\begin{aligned} M_i &= \Pr\left[\tilde{u}_i(\iota)\left(\bar{U}_i(\omega_i/P, b_i/P) - 1\right) > \tilde{u}_j(\iota)\left(\bar{U}_j(\omega_j/P, b_j/P) - 1\right), \ \ \forall j \in J_c\right] \\ &= \Pr\left[\tilde{u}_i(\iota) v_i (\omega_i)^{\phi} > \tilde{u}_j(\iota) v_j (\omega_j)^{\phi}, \ \ \forall j \in J_c\right]. \end{aligned}$$

Given the distributional assumption in eq. (C.24), it holds that

$$M_i = \frac{v_i(\omega_i)^{\phi_m}}{\Phi_c} M_c \ \text{ such that } \ \Phi_c = \sum_{j \in J_c} v_j(\omega_j)^{\phi_m} \ \text{ and } \ \phi_m \equiv \tilde{\phi}\phi. \tag{C.25}$$

Given the value of $M_i$, total employment in region $i$ is determined as in eq. (A.8). Therefore, the total labor supply in region $i$ is

$$L_i = \frac{v_i(\omega_i)^{\phi_m}}{\sum_{j \in J_c} v_j(\omega_j)^{\phi_m}} M_c v_i (\omega_i)^{\phi}. \tag{C.26}$$

**Producer's problem.** The producer's problem is identical to that in Appendix A.2.

**Goods market clearing.** The conditions determining the equilibrium in the good's market and, consequently, the region- and sector-specific labor demand equations are identical to those in Appendix A.2.

**Labor market clearing.** Combining the region- and sector-specific labor supply equation in eq. (C.26) with the aggregate labor demand equation in eq. (A.15), and imposing the normalization $W = 1$, the labor market clearing condition in every region $i \in J_c$ is

$$\frac{v_i(\omega_i)^{\phi_m}}{\sum_{j \in J_c} v_j(\omega_j)^{\phi_m}} M_c v_i(\omega_i)^{\phi} = \sum_s (\omega_i)^{-\sigma_s} \left(A_{is} P_s\right)^{\sigma_s - 1} \gamma_s, \tag{C.27}$$

or, equivalently,

$$(\Phi_c)^{-1} M_c v_i (\omega_i)^{\phi + \phi_m} = \sum_s (\omega_i)^{-\sigma_s} \left(A_{is} P_s\right)^{\sigma_s - 1} \gamma_s, \tag{C.28}$$

for every region $i$ in every country $c$.

**Equilibrium.** Given productivity parameters $\{A_{is}\}_{i=1,s=1}^{J,S}$, preference parameters $\{\sigma_s, \gamma_s\}_{s=1}^{S}$, labor supply parameters $\{v_i\}_{i=1}^{J}$, and normalizing world income to equal 1, $W = 1$, we can use eq. (A.6), eq. (A.9), and eq. (C.28) to solve for the equilibrium wage in every region, $\{\omega_i\}_{j=1}^{J}$, the equilibrium price of every sector- and region-specific good $\{p_{is}\}_{i=1,s=1}^{J,S}$, and the sectoral price indices $\{P_s\}_{s=1}^{S}$.

Given these equilibrium wages and sectoral price indices, we can use eq. (C.26) to solve for the equilibrium level of employment in every region, $\{L_i\}_{i=1}^{J}$.

### C.4.3 Labor market impact of sectoral shocks

We focus on the labor market outcomes in a country $c$ formed by a set of $N$ small open regional economies; see Appendix A.3 for details on the definition of these small regional economies.

Across periods, the microfoundation described in Sections C.4.1 and C.4.2 implies that the changes in labor market outcomes in all $N$ regions of country $c$, $\{\omega_i, L_i\}_{i\in J_c}$, are generated by changes in sectoral prices $\{P_s\}_{s=1}^{S}$, changes in an aggregate of all other sectoral shocks, $\{(A_s)^{(\sigma_s-1)}\gamma_s\}_{s=1}^{S}$, changes in productivities of all regions in country $c$ and all sectors, $\{(\tilde{A}_{is})^{(\sigma_s-1)}\}_{i\in J_c, s=1}^{S}$, and changes in labor supply parameters in all regions in country $c$, $\{v_i\}_{i\in J_c}$.

**Isomorphism.** According to eq. (C.26), the change in employment in any region $i$ in country $c$ is

$$\hat{L}_i = 2\hat{v}_i + (\phi+\phi_m)\hat{\omega}_i - \hat{\Phi}_c. \tag{C.29}$$

Assuming that $\{M_c\}_{c=1}^{C}$, $\{\sigma_s\}_{s=1}^{S}$, and $(\phi, \phi_m)$ are fixed and totally differentiating eq. (C.27) with respect to the remaining determinants of $\hat{\omega}_i$, we can express the changes in wages in every region $i$ of country $c$ as

$$\hat{\omega}_i = \lambda_i\hat{\Phi}_c + \lambda_i\sum_{s=1}^{S} l_{is}^0\left[\hat{\gamma}_s + (\sigma_s-1)(\hat{A}_{is}+\hat{P}_s)\right] - \lambda_i\hat{v}_i, \tag{C.30}$$

where $\lambda_i \equiv (\phi+\phi_m+\sum_s l_{is}^0\sigma_s)^{-1}$. Using the expression in eq. (C.25), we can also express

$$\hat{\Phi}_c = \sum_{i\in J_c} m_i^0\left(\phi_m\hat{\omega}_i + \hat{v}_i\right), \tag{C.31}$$

where $m_i^0$ is the share of individuals living in country $c$ that had residence in region $i$ at the initial period 0; i.e. $m_i^0 \equiv M_i^0/M_c^0$, with $M_c^0 \equiv \sum_{i\in J_c} M_i^0$.

Combining eqs. (C.29) and (C.30), we can express the change in total employment in region $i$ as

$$\begin{aligned}\hat{L}_i = {} & [(\phi+\phi_m)\lambda_i - 1]\hat{\Phi}_c + \sum_{s=1}^{S} l_{is}^0[\beta_{is}\hat{P}_s + \lambda_i(\phi+\phi_m)((\sigma_s-1)\hat{A}_s+\hat{\gamma}_s) + \lambda_i(\phi+\phi_m)(\sigma_s-1)\hat{\tilde{A}}_{is}] \\ & + [2-(\phi+\phi_m)\lambda_i]\hat{v}_i\end{aligned} \tag{C.32}$$

where $\beta_{is} = (\sigma_s-1)(\phi+\phi_m)\lambda_i$. If it were to be the case that $\hat{\Phi}_c = 0$, the expression in eq. (C.32) would be identical to that in eq. (8) under the following equivalences

$$\begin{aligned}\chi_s &= P_s, \\ \rho_s &= (\sigma_s-1)(\phi+\phi_m), \\ \mu_s &= (A_s)^{(\sigma_s-1)(\phi+\phi_m)}(\gamma_s)^{(\phi+\phi_m)},\end{aligned}$$

$$\eta_{is} = (\tilde{A}_{is})^{(\sigma_s - 1)(\phi + \phi_m)},$$

and the necessary adjustment of the expression for $\lambda_i$ and $\hat{v}_i$. However, the term $\hat{\Phi}_c$ will generally not be zero and, as indicated in eq. (C.31), it will generally capture the effect of shocks to all regions in the same country $c$ as the region of interest $i$. In the specific case in which $\sigma_s = \sigma$ for all sectors $s$, it will be the case that $\lambda_i = \lambda$ for all regions $i$, and, consequently, the term $[(\phi + \phi_m)\lambda_i - 1]\hat{\Phi}_c$ will be common to all regions $i$ belonging to the same country $c$. In this special case, the parameter $\beta_{is}$ will no longer capture the total effect of the price shifters $\hat{P}_s$ but the differential effect of this price shifter on region $i$ relative to all other regions in the same country $c$.

# Appendix D Placebo Exercise

In Appendix D.1, we present results illustrating the impact of confounding sector-level shocks on different estimators of the coefficient on the shift-share covariate of interest. In Appendix D.2, we provide additional results on the placebo exercises described in Sections 2 and 5.

## D.1 Confounding sector-level shocks: omitted variable bias and solutions

In this appendix, we investigate the consequences of violations of the assumption that observed sectoral shocks of interest $(\mathcal{X}_1, \ldots, \mathcal{X}_S)$ are independent from other sectoral shocks affecting the outcome variable of interest. We study in this section the impact that violations of this assumption have on the properties of the OLS estimator of the coefficient on the shift-share regressor of interest. We also consider the properties of two solutions to this problem: (i) the inclusion of regional controls as a proxy for sector-level unobserved shocks (discussed in Section 4.2), and (ii) the use of a shift-share instrumental variable constructed as a weighted average of exogenous sector-level shocks (discussed in Section 4.3.2).

To generate both confounding sectoral shocks and an instrument for the sectoral shock of interest, we extend the baseline placebo exercise and, for each sector $s$ and simulation $m$, we take a draw of a three-dimensional vector

$$(\mathcal{X}_s^{a,m}, \mathcal{X}_s^{b,m}, \mathcal{X}_s^{c,m}) \sim N(0; \tilde{\Sigma}),$$

where $\mathcal{X}_s^a$ is the variable of interest, $\mathcal{X}_s^b$ is the unobserved confounding effect, $\mathcal{X}_s^c$ is an observed instrumental variable. Specifically, the matrix $\tilde{\Sigma}$ is such that $var(\mathcal{X}_s^a) = var(\mathcal{X}_s^b) = var(\mathcal{X}_s^c) = \tilde{\sigma}$, $cov(\mathcal{X}_s^a, \mathcal{X}_s^b) = cov(\mathcal{X}_s^a, \mathcal{X}_s^c) = \tilde{\rho}\tilde{\sigma}$, and $cov(\mathcal{X}_s^b, \mathcal{X}_s^c) = 0$. Thus, we impose that $\mathcal{X}_s^a$ has a correlation of $\tilde{\rho}$ with both $\mathcal{X}_s^b$ and $\mathcal{X}_s^c$, but $\mathcal{X}_s^b$ and $\mathcal{X}_s^c$ are independent. In our simulations, we set $\tilde{\rho} = 0.7$ and $\tilde{\delta} = 12$.

To assign the role of a confounding effect to $\mathcal{X}_s^b$, we generate an outcome variable as

$$Y_i^m = Y_i^{obs} + \delta \sum_{s=1}^{S} w_{is} \mathcal{X}_s^{b,m},$$

where $Y_i^{obs}$ is the observed 2000–2007 change in the employment rate in CZ $i$, and $\delta$ is a parameter controlling the impact of the unobserved sectoral shocks $(\mathcal{X}_1^b, \dots, \mathcal{X}_S^b)$ on the simulated outcome $Y_i^m$. Thus, the parameter $\delta$ captures the magnitude of the confounding effect of the unobserved shock $\mathcal{X}_s^b$ on the dependent variable. We explore the impact of confounding effects by simulating data both with $\delta = 0$ and with $\delta = 6$.

In addition, we assume that we observe a regional variable that is a noisy measure of CZ $i$'s exposure to the unobserved sectoral shocks $(\mathcal{X}_1^b, \dots, \mathcal{X}_S^b)$,

$$X_i^{b,m} = u_i^m + \sum_s w_{is} \mathcal{X}_s^{b,m} \quad \text{where} \quad u_i^m \sim N(0, \sigma_u).$$

The parameter $\sigma_u$ thus modulates the measurement error in $X_i^b$ as a proxy for the impact of the unobserved sectoral shocks $(\mathcal{X}_1^b, \dots, \mathcal{X}_S^b)$ in CZ $i$. We explore the impact of $\sigma_u$ by simulating data both with $\sigma_u = 0$ and with $\sigma_u = 6$.[1]

For each set of parameters $(\delta, \sigma_u)$ and for each simulation draw, we compute three estimators of the impact of $X_i^a \equiv \sum_{s=1}^S w_{is} \mathcal{X}_s^a$ on $Y_i$. First, we ignore the possible endogeneity problem and compute the OLS estimator without controls; i.e. the estimator in eq. (11). Second, we consider the OLS estimator in a regression that includes $X_i^b$ as a proxy for the vector of unobserved confounding sectoral shocks; i.e. the estimator in eq. (19). Third, we consider the IV estimator that uses $X_i^c \equiv \sum_i w_{is} \mathcal{X}_s^c$ as the instrumental variable; i.e. the estimator in eq. (32). For each of these three estimators, we compute four estimates of its standard error: *Robust*, *St-cluster*, *AKM* and *AKM0*. All results are reported in Table D.1.

When there is no confounding sectoral shock ($\delta = 0$), Panel A shows that all three estimators yield an average coefficient close to zero. Panels B and C report results in the presence of confounding sectoral shocks ($\delta > 0$), in which case the OLS estimator in a simple regression of $Y_i$ on $X_i^a$ without additional covariates is positively biased ($\hat{\beta} = 4.2$). The introduction of the regional control only yields unbiased estimates when it is a good proxy for the underlying confounding sectoral shock (i.e. if $\sigma_u = 0$ as in Panel B). In contrast, the IV estimate always yields an average estimated coefficient of zero due to the orthogonality between the sector-level instrumental variable and the sector-level unobserved confounding shock.

Similarly to our baseline placebo, traditional inference methods under-predict the dispersion of estimated coefficients both in the case of the OLS and the IV estimators. As discussed above, this is driven by the correlation between the unobservable residuals of regions with similar sector employment compositions. By allowing for such a correlation, our proposed methods yield, on average, estimates of the median length of the 95% confidence interval equal or higher to the standard deviation of the empirical distribution of estimates. As a result, Table D.2 in Appendix D reports that, while traditional methods overreject the null $H_0 : \beta = 0$ in the context of both OLS and IV estimation, our methods yield the correct test size for both estimators.

[1]Using the notation in Section 4.2, the simulated variable $\mathcal{X}_s^a$ corresponds to $\mathcal{X}_s$, the simulated variable $\mathcal{X}_s^b$ is a column in the matrix $\mathcal{Z}$ (which also includes a column of ones), $u_i$ corresponds to $U_i$, and $X_i^b$ to $Z_i$. The value of the parameter $\gamma$ in eq. (22) is thus equal to $\tilde{\rho}$.

Table D.1: Magnitude of standard errors—Confounding effects

| | Estimates | | Median effective std. error | | | |
|---|---|---|---|---|---|---|
| | Average | Std. dev | Robust | St-cluster | AKM | AKM0 |
| **Panel A: No confounding effect ($\delta = 0$)** | | | | | | |
| OLS without controls | $-0.01$ | 1.29 | 0.47 | 0.59 | 1.23 | 1.43 |
| OLS with regional control | 0.00 | 1.80 | 0.67 | 0.83 | 1.72 | 1.97 |
| 2SLS | 0.00 | 1.85 | 0.69 | 0.85 | 1.75 | 2.02 |
| **Panel B: Confounding effect ($\delta = 6$) and perfect regional control ($\sigma_u = 0$)** | | | | | | |
| OLS without controls | 4.22 | 1.48 | 0.58 | 0.70 | 1.37 | 1.60 |
| OLS with regional control | 0.02 | 1.8 | 0.67 | 0.83 | 1.71 | 1.97 |
| 2SLS | 0.02 | 1.84 | 0.69 | 0.85 | 1.75 | 2.02 |
| **Panel C: Confounding effect ($\delta = 6$) and imperfect regional control ($\sigma_u = 2$)** | | | | | | |
| OLS without controls | 4.21 | 1.47 | 0.58 | 0.69 | 1.37 | 1.60 |
| OLS with regional control | 4.11 | 1.46 | 0.58 | 0.69 | 1.39 | 1.60 |
| 2SLS | $-0.21$ | 2.45 | 0.93 | 1.10 | 2.11 | 2.66 |

Notes: All estimates in this table use the total employment share in each CZ as the outcome variable $Y_i$. For the inference procedure indicated in the first row, "median effective std. error" refers to the median length of the 95% confidence interval across the simulated datasets divided by $2 \times 1.96$. *Robust* is the Eicker-Huber-White standard error; *St-cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in eq. (25); *AKM0* is the confidence interval Remark 6. 30,000 simulation draws.

Table D.2: Rejection rate of $H_0\colon \beta = 0$ with significance level of 5%—Confounding effects

| | Estimates | | Rejection rate of $H_0\colon \beta = 0$ at 5% | | | |
|---|---|---|---|---|---|---|
| | Average | Std. Dev | Robust | St-cluster | AKM | AKM0 |
| **Panel A: No confounding effect ($\delta = 0$)** | | | | | | |
| OLS without controls | $-0.01$ | 1.29 | 48.1% | 37.6% | 7.9% | 4.8% |
| OLS with regional control | 0.00 | 1.80 | 48.1% | 38.1% | 8.1% | 4.8% |
| 2SLS | 0.00 | 1.85 | 48.0% | 38.1% | 8.0% | 4.9% |
| **Panel B: Confounding effect ($\delta = 6$) and perfect regional control ($\sigma_u = 0$)** | | | | | | |
| OLS without controls | 4.22 | 1.48 | 97.9% | 96.8% | 81.7% | 72.9% |
| OLS with regional control | 0.02 | 1.80 | 48.1% | 37.9% | 8.1% | 5.0% |
| 2SLS | 0.02 | 1.84 | 47.6% | 37.7% | 8.0% | 4.8% |
| **Panel C: Confounding effect ($\delta = 6$) and imperfect regional control ($\sigma_u = 2$)** | | | | | | |
| OLS without controls | 4.21 | 1.47 | 98.0% | 96.8% | 81.6% | 72.7% |
| OLS with regional control | 4.11 | 1.46 | 97.7% | 96.4% | 79.5% | 71.3% |
| 2SLS | $-0.21$ | 2.45 | 41.1% | 33.8% | 8.3% | 4.9% |

Notes: All estimates in this table use the total employment share in each CZ as the outcome variable $Y_i$. For the inference procedure indicated in the first row, this table indicates the percentage of the simulated datasets for which we reject the null hypothesis $H_0 : \beta = 0$ using a 5% significance level test. *Robust* is the Eicker-Huber-White standard error; *St-cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in eq. (25); *AKM0* is the test in Remark 6. 30,000 simulation draws.

### D.2 Additional results

Figure D.1 reports the empirical distribution of the estimated coefficients when: (a) the dependent variable is the 2000–2007 change in each CZ's employment rate; in each simulation draw $m$, we draw a random vector $(\mathcal{X}_1^m, \ldots, \mathcal{X}_{S-1}^m)$ of i.i.d. normal random variables with zero mean and variance $\mathrm{var}(\mathcal{X}_s^m) = 5$, and set $\mathcal{X}_S^m = 0$; and (c) the vector of controls $Z_i$ only includes a constant. The empirical distribution of the estimated coefficients resembles a normal distribution centered around $\beta = 0$. For more details in the placebo exercise that generates this distribution of estimated coefficients, see Section 2.

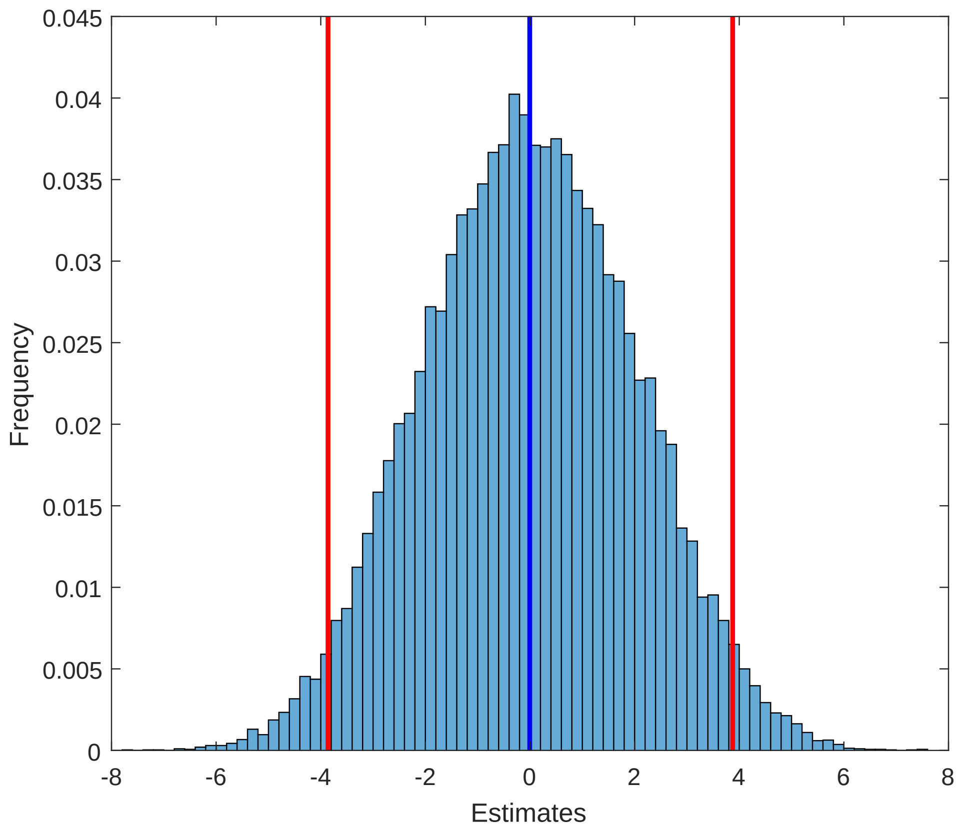


Figure D.1: Empirical distribution of estimated coefficients in the placebo exercise.

Notes: The blue line indicates the average estimated coefficient; the red lines indicate the 2.5% and 97.5% percentiles of distribution of $\hat{\beta}^m$ across the $m = 1, \ldots, 30,000$ simulations. The dependent variable is the 2000–2007 change in the employment rate.

For the placebo exercise based on the change in the CZ's employment rate, Figure D.2 reports the empirical distribution of the effective standard errors, its median and, for comparison, the standard deviation of the OLS estimates $\{\hat{\beta}^m\}_{m=1}^{30,000}$. The top bracket in each histogram combines all draws whose effective standard error is above 5. Our new standard error estimators have a larger dispersion than the traditional ones, but are centered close to the standard deviation of estimated parameters.

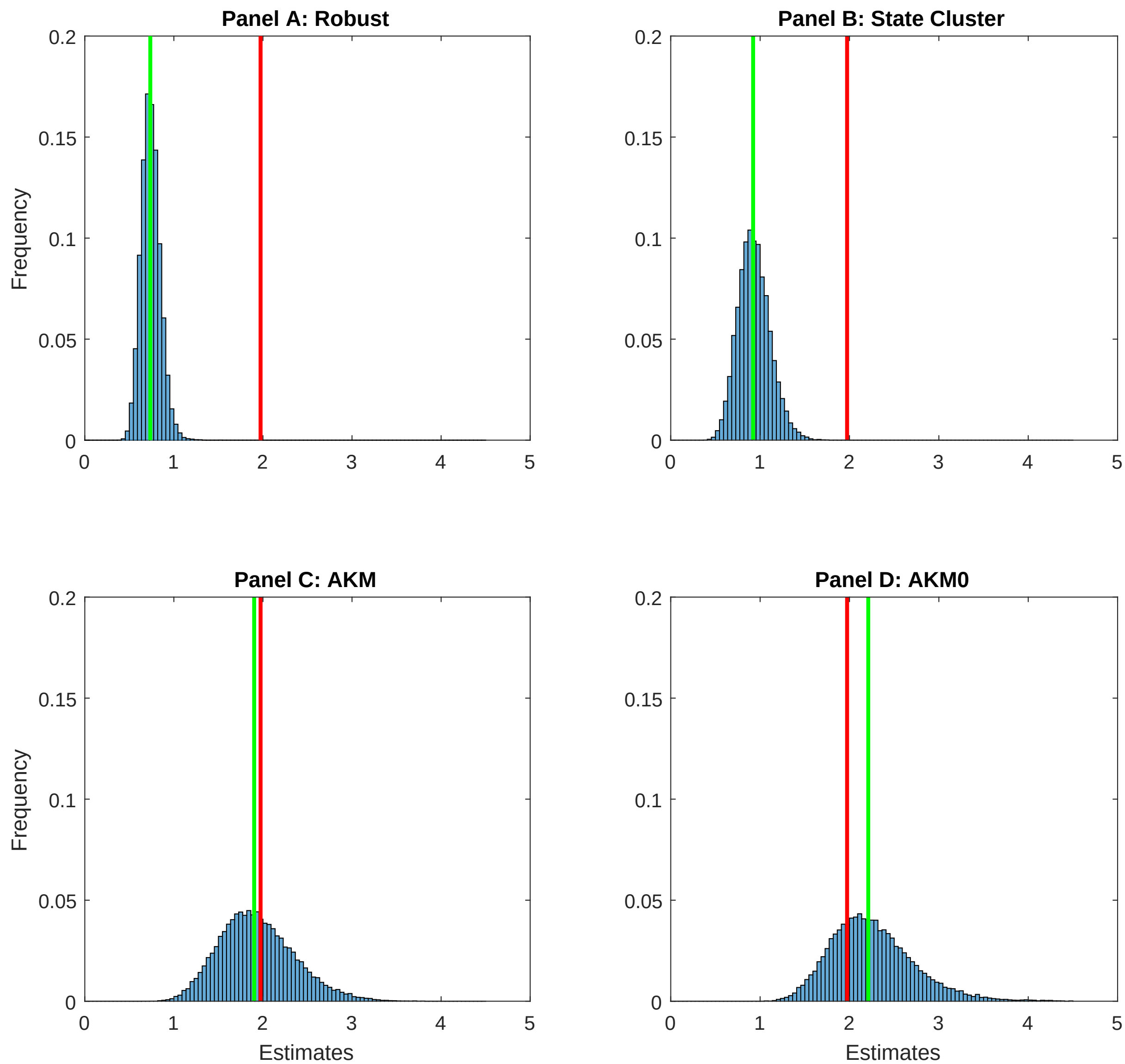


Figure D.2: Empirical distribution of effective standard errors for different standard error estimators in the placebo exercise.

Notes: In each of the four panels, the red line indicates the standard deviation of the empirical distribution of estimated coefficients represented in Figure D.1; the green line indicates median of the estimated effective standard error (95% confidence interval divided by $2 \times 1.96$) and the light blue bars represent the distribution of the effective standard errors across the 30,000 simulations. The dependent variable is the 2000–2007 change in the employment rate in all four panels.

Table D.3 investigates the sensitivity of our results to the definitions of geographic units. We report results of our main placebo exercise using counties as the geographic unit $i$ rather than CZs. As in our baseline exercise, we use the CBP data to construct employment by county and sector using the imputation procedure in Autor, Dorn and Hanson (2013). Since this procedure does not yield wage bill information at the county-level, suppressed payroll values prevent the construction of wage variables for a large fraction of counties. For this reason, we only implement the placebo exercise for the employment variables in Panel A. Results indicate that rejection rates for all inference procedures are very similar to those obtained in the baseline placebo exercise reported in Panel A of Tables 1 and 2.

Table D.3: Rejection rate of $H_0\colon \beta = 0$ at 5% significance level: County-level analysis.

| | Robust | St-cluster | AKM | AKM0 |
|---|---|---|---|---|
| **Panel A: Change in the share of working-age population** | | | | |
| Employed | 47.7% | 36.7% | 7.9% | 4.7% |
| Employed in manufacturing | 66.5% | 51.9% | 8.2% | 4.6% |
| Employed in non-manufacturing | 28.0% | 25.1% | 8.6% | 4.6% |

Notes: This table reports rejection rates of $H_0\colon \beta = 0$ at 5% significance level. Content analogous to that in columns (5)–(6) of Tables 1 and 2. The only difference is that the placebo exercise is based on actual labor market outcomes in 3107 counties instead of 722 CZs. 30,000 simulations draws.

Finally, we investigate the performance of different inference procedures when we consider simulated random shocks to 331 occupations. In this exercise, we construct occupation-region shares using actual occupation employment shares in 1990 for the 722 U.S. CZs. Table D.4 shows that the overrejection of traditional inference methods is more severe with shift-share regressors based on occupations. In this case, our inference procedure under the null yields the correct test size.

Table D.4: Rejection rate of $H_0\colon \beta = 0$ at 5% significance level: Occupation employment shares.

| | Robust | St-cluster | AKM | AKM0 |
|---|---|---|---|---|
| **Panel A: Change in the share of working-age population** | | | | |
| Employed | 83.6% | 62.5% | 25.4% | 4.2% |
| Employed in manufacturing | 90.0% | 75.8% | 33.7% | 3.3% |
| Employed in non-manufacturing | 65.7% | 39.2% | 18.1% | 4.0% |
| **Panel B: Change in average log weekly wage** | | | | |
| Employed | 85.4% | 63.5% | 31.0% | 3.5% |
| Employed in manufacturing | 56.2% | 26.8% | 11.6% | 4.8% |
| Employed in non-manufacturing | 85.4% | 65.1% | 32.3% | 3.3% |

Notes: This table reports rejection rates of $H_0\colon \beta = 0$ at 5% significance level. Content analogous to that in columns (5)–(6) of Tables 1 and 2. The only difference is that the placebo exercise is based on actual labor market outcomes of 311 occupations in 722 CZs. 30,000 simulations draws.

# Appendix E Empirical application: additional results

## E.1 Effect of Chinese exports on U.S. labor market outcomes

Table E.1: Effect of Chinese on U.S. Commuting Zones in Autor, Dorn and Hanson (2013): Reduced-Form Regression

| | Change in the employment share | | | Change in avg. log weekly wage | | |
|---|---|---|---|---|---|---|
| | All (1) | Manuf. (2) | Non-Manuf. (3) | All (4) | Manuf. (5) | Non-Manuf. (6) |
| **Panel A: All Workers** | | | | | | |
| $\hat{\beta}$ | -0.49 | -0.38 | -0.11 | -0.48 | 0.10 | -0.48 |
| Robust | [-0.71,-0.27] | [-0.48,-0.28] | [-0.31,0.08] | [-0.80,-0.16] | [-0.5,0.69] | [-0.83,-0.13] |
| St-cluster | [-0.64,-0.34] | [-0.45,-0.30] | [-0.27,0.05] | [-0.78,-0.18] | [-0.51,0.70] | [-0.81,-0.15] |
| AKM (indep.) | [-0.80,-0.18] | [-0.52,-0.23] | [-0.33,0.11] | [-0.83,-0.13] | [-0.46,0.65] | [-0.87,-0.09] |
| $\text{AKM}_0$ (indep.) | [-1.10,-0.25] | [-0.64,-0.26] | [-0.52,0.07] | [-1.07,-0.16] | [-0.91,0.57] | [-1.22,-0.17] |
| AKM (4d cluster) | [-0.79,-0.18] | [-0.53,-0.23] | [-0.33,0.11] | [-0.86,-0.10] | [-0.49,0.68] | [-0.89,-0.08] |
| $\text{AKM}_0$ (4d cluster) | [-1.12,-0.25] | [-0.66,-0.25] | [-0.53,0.07] | [-1.16,-0.15] | [-1.00,0.59] | [-1.29,-0.16] |
| AKM (3d cluster) | [-0.81,-0.16] | [-0.52,-0.23] | [-0.35,0.13] | [-0.87,-0.09] | [-0.50,0.69] | [-0.92,-0.04] |
| $\text{AKM}_0$ (3d cluster) | [-1.25,-0.24] | [-0.68,-0.25] | [-0.64,0.08] | [-1.26,-0.12] | [-1.15,0.60] | [-1.46,-0.13] |
| **Panel B: College Graduates** | | | | | | |
| $\hat{\beta}$ | -0.27 | -0.37 | 0.11 | -0.48 | 0.29 | -0.47 |
| Robust | [-0.42,-0.12] | [-0.48,-0.26] | [-0.04,0.25] | [-0.82,-0.13] | [-0.10,0.68] | [-0.83,-0.11] |
| St-cluster | [-0.39,-0.14] | [-0.48,-0.27] | [-0.04,0.26] | [-0.83,-0.13] | [-0.14,0.72] | [-0.81,-0.12] |
| AKM (indep.) | [-0.45,-0.08] | [-0.50,-0.25] | [-0.03,0.24] | [-0.81,-0.14] | [-0.11,0.69] | [-0.82,-0.12] |
| $\text{AKM}_0$ (indep.) | [-0.58,-0.10] | [-0.56,-0.24] | [-0.11,0.24] | [-0.99,-0.14] | [-0.35,0.67] | [-1.06,-0.15] |
| AKM (4d cluster) | [-0.45,-0.08] | [-0.51,-0.23] | [-0.04,0.25] | [-0.84,-0.11] | [-0.14,0.71] | [-0.84,-0.10] |
| $\text{AKM}_0$ (4d cluster) | [-0.59,-0.10] | [-0.59,-0.23] | [-0.12,0.25] | [-1.07,-0.13] | [-0.42,0.70] | [-1.13,-0.14] |
| AKM (3d cluster) | [-0.46,-0.08] | [-0.52,-0.23] | [-0.04,0.25] | [-0.86,-0.09] | [-0.14,0.72] | [-0.88,-0.06] |
| $\text{AKM}_0$ (3d cluster) | [-0.63,-0.09] | [-0.59,-0.20] | [-0.17,0.25] | [-1.19,-0.10] | [-0.46,0.73] | [-1.31,-0.11] |
| **Panel C: Non-College Graduates** | | | | | | |
| $\hat{\beta}$ | -0.70 | -0.37 | -0.34 | -0.51 | -0.06 | -0.52 |
| Robust | [-1.02,-0.38] | [-0.48,-0.25] | [-0.60,-0.07] | [-0.90,-0.13] | [-0.69,0.56] | [-0.94,-0.10] |
| St-cluster | [-0.92,-0.48] | [-0.47,-0.26] | [-0.55,-0.12] | [-0.84,-0.19] | [-0.53,0.40] | [-0.87,-0.17] |
| AKM (indep.) | [-1.19,-0.21] | [-0.55,-0.18] | [-0.69,0.02] | [-1.07,0.04] | [-0.69,0.56] | [-1.14,0.10] |
| $\text{AKM}_0$ (indep.) | [-1.71,-0.33] | [-0.73,-0.22] | [-1.03,-0.06] | [-1.59,-0.07] | [-1.27,0.44] | [-1.8,-0.05] |
| AKM (4d cluster) | [-1.18,-0.22] | [-0.55,-0.18] | [-0.68,0.01] | [-1.08,0.05] | [-0.69,0.57] | [-1.14,0.10] |
| $\text{AKM}_0$ (4d cluster) | [-1.72,-0.34] | [-0.75,-0.22] | [-1.03,-0.06] | [-1.64,-0.07] | [-1.31,0.44] | [-1.82,-0.05] |
| AKM (3d cluster) | [-1.22,-0.18] | [-0.55,-0.18] | [-0.71,0.04] | [-1.08,0.05] | [-0.71,0.58] | [-1.15,0.12] |
| $\text{AKM}_0$ (3d cluster) | [-1.97,-0.31] | [-0.80,-0.22] | [-1.22,-0.04] | [-1.79,-0.06] | [-1.54,0.44] | [-2.03,-0.04] |

Notes: $N = 1,444$ (722 CZs $\times$ two time periods). Models are weighted by start of period CZ share of national population. All regressions include the full vector of baseline controls in ADH. 95% confidence intervals in square brackets. *Robust* is the Eicker-Huber-White standard error; *St-cluster* is the standard error that clusters of CZs in the same state; *AKM* is the standard error in eq. (29) with 3-digit SIC clusters; *AKM0* is the confidence interval with 3-digit SIC clusters described in the last sentence of Section 4.3.1.

Table E.2: Effect of Chinese on U.S. Commuting Zones in Autor, Dorn and Hanson (2013): 2SLS Regression

| | Change in the employment share | | | Change in avg. log weekly wage | | |
|---|---|---|---|---|---|---|
| | All (1) | Manuf. (2) | Non-Manuf. (3) | All (4) | Manuf. (5) | Non-Manuf. (6) |
| **Panel A: All Workers** | | | | | | |
| $\hat{\beta}$ | -0.77 | -0.60 | -0.18 | -0.76 | 0.15 | -0.76 |
| Robust | [-1.10,-0.45] | [-0.78,-0.41] | [-0.47,0.12] | [-1.23,-0.29] | [-0.81,1.11] | [-1.27,-0.25] |
| St-cluster | [-1.12,-0.42] | [-0.79,-0.40] | [-0.45,0.10] | [-1.26,-0.26] | [-0.81,1.11] | [-1.28,-0.24] |
| AKM (indep.) | [-1.20,-0.35] | [-0.81,-0.38] | [-0.50,0.15] | [-1.28,-0.24] | [-0.75,1.05] | [-1.31,-0.22] |
| $\text{AKM}_0$ (indep.) | [-1.42,-0.41] | [-0.90,-0.39] | [-0.66,0.11] | [-1.46,-0.26] | [-1.14,0.98] | [-1.57,-0.28] |
| AKM (4d cluster) | [-1.20,-0.35] | [-0.84,-0.35] | [-0.51,0.15] | [-1.32,-0.19] | [-0.79,1.09] | [-1.34,-0.18] |
| $\text{AKM}_0$ (4d cluster) | [-1.48,-0.42] | [-0.97,-0.38] | [-0.67,0.12] | [-1.59,-0.23] | [-1.25,1.03] | [-1.68,-0.26] |
| AKM (3d cluster) | [-1.25,-0.29] | [-0.85,-0.35] | [-0.54,0.18] | [-1.36,-0.16] | [-0.80,1.10] | [-1.41,-0.12] |
| $\text{AKM}_0$ (3d cluster) | [-1.72,-0.39] | [-1.02,-0.36] | [-0.85,0.13] | [-1.76,-0.19] | [-1.49,1.03] | [-1.97,-0.21] |
| **Panel B: College Graduates** | | | | | | |
| $\hat{\beta}$ | -0.42 | -0.59 | 0.17 | -0.76 | 0.46 | -0.74 |
| Robust | [-0.64,-0.20] | [-0.81,-0.37] | [-0.08,0.41] | [-1.29,-0.22] | [-0.19,1.11] | [-1.29,-0.20] |
| St-cluster | [-0.67,-0.18] | [-0.84,-0.34] | [-0.07,0.41] | [-1.37,-0.14] | [-0.22,1.14] | [-1.34,-0.15] |
| AKM (indep.) | [-0.70,-0.15] | [-0.83,-0.36] | [-0.07,0.40] | [-1.28,-0.23] | [-0.22,1.13] | [-1.26,-0.22] |
| $\text{AKM}_0$ (indep.) | [-0.79,-0.16] | [-0.87,-0.33] | [-0.15,0.40] | [-1.42,-0.21] | [-0.45,1.12] | [-1.45,-0.24] |
| AKM (4d cluster) | [-0.70,-0.15] | [-0.85,-0.33] | [-0.08,0.41] | [-1.32,-0.19] | [-0.26,1.18] | [-1.29,-0.20] |
| $\text{AKM}_0$ (4d cluster) | [-0.83,-0.16] | [-0.93,-0.32] | [-0.15,0.42] | [-1.54,-0.20] | [-0.54,1.18] | [-1.55,-0.23] |
| AKM (3d cluster) | [-0.72,-0.13] | [-0.86,-0.32] | [-0.08,0.42] | [-1.36,-0.16] | [-0.25,1.17] | [-1.36,-0.13] |
| $\text{AKM}_0$ (3d cluster) | [-0.92,-0.14] | [-0.96,-0.28] | [-0.23,0.41] | [-1.70,-0.15] | [-0.61,1.21] | [-1.81,-0.18] |
| **Panel C: Non-College Graduates** | | | | | | |
| $\hat{\beta}$ | -1.11 | -0.58 | -0.53 | -0.81 | -0.10 | -0.82 |
| Robust | [-1.58,-0.64] | [-0.76,-0.40] | [-0.93,-0.13] | [-1.35,-0.28] | [-1.07,0.87] | [-1.41,-0.23] |
| St-cluster | [-1.61,-0.61] | [-0.77,-0.39] | [-0.94,-0.13] | [-1.28,-0.34] | [-0.84,0.63] | [-1.31,-0.33] |
| AKM (indep.) | [-1.77,-0.45] | [-0.83,-0.33] | [-1.03,-0.04] | [-1.60,-0.03] | [-1.08,0.88] | [-1.70,0.06] |
| $\text{AKM}_0$ (indep.) | [-2.15,-0.57] | [-0.96,-0.36] | [-1.29,-0.11] | [-2.00,-0.12] | [-1.55,0.77] | [-2.20,-0.09] |
| AKM (4d cluster) | [-1.77,-0.46] | [-0.85,-0.31] | [-1.02,-0.04] | [-1.62,-0.01] | [-1.08,0.88] | [-1.70,0.06] |
| $\text{AKM}_0$ (4d cluster) | [-2.22,-0.58] | [-1.03,-0.36] | [-1.31,-0.11] | [-2.11,-0.12] | [-1.63,0.77] | [-2.28,-0.09] |
| AKM (3d cluster) | [-1.87,-0.35] | [-0.87,-0.30] | [-1.09,0.03] | [-1.66,0.04] | [-1.11,0.91] | [-1.77,0.12] |
| $\text{AKM}_0$ (3d cluster) | [-2.66,-0.52] | [-1.14,-0.35] | [-1.62,-0.07] | [-2.43,-0.10] | [-1.99,0.76] | [-2.70,-0.06] |

Notes: $N = 1,444$ (722 CZs $\times$ two time periods). Models are weighted by start of period CZ share of national population. All regressions include the full vector of baseline controls in ADH. *Robust* is the Eicker-Huber-White standard error; *St-cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5 (for independent shocks), or that in eq. (29) (for correlated shocks). *AKM0* is the test in Remark 6 (for independent shocks) or that described in the last sentence of Section 4.3.1 (for correlated shocks). The table reports confidence intervals computed under the assumption of independent shocks, 4-digit SIC correlated shocks, and 3-digit SIC correlated shocks.

Table E.3: Rejection rate of $H_0\colon \beta = 0$ with significance level of 5%. Placebo exercise based on the first-stage regression in Autor, Dorn and Hanson (2013):

| Estimates | | Rejection rate of $H_0\colon \beta = 0$ at 5% | | | |
|---|---|---|---|---|---|
| Average | St Dev | Robust | St-cluster | AKM | AKM0 |
| **Panel A: Baseline simulation without controls** | | | | | |
| 0.01 | 1.74 | 42.1% | 37.3% | 6.9% | 4.5% |
| **Panel B: Controlling for ADH IV** | | | | | |
| 0.00 | 0.79 | 18.3% | 17.0% | 6.9% | 4.5% |
| **Panel C: Controlling for ADH IV and baseline controls** | | | | | |
| 0.00 | 0.67 | 13.8% | 13.6% | 5.4% | 3.5% |

Notes: Dependent variable is the "shift-share" regressor in ADH contructed from the interaction of CZ's employment share in 4-digit SIC manufacturing industries and the normalized U.S. imports from China in the same industries. The first row indicates the inference procedure employed to compute the share of the 30,000 simulated datasets for which we reject the null hypothesis $H_0 : \beta = 0$ using a 5% significance level. *Robust* is the Eicker-Huber-White standard error; *St-cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5; *AKM0* is the test in Remark 6.

## E.2 Effect of Immigration on U.S. local labor markets

Table E.4: Origin Countries

| | | | |
|---|---|---|---|
| Afghanistan | France | Liechtenstein and Lux. | Scandinavia |
| Africa | Greece | Malaysia | Scotland |
| Albania | Gulf States | Maldives | Singapore |
| Andorra and Gibraltar | India | Malta | South America |
| Austria | Indonesia | Mexico | Spain |
| Belgium | Iran | Nepal | Switzerland |
| Brunei | Iraq | Netherlands | Syria |
| Cambodia | Ireland | Oceania | Thailand |
| Canada | Israel/Palestine | Other | Turkey |
| Central America | Italy | Other Europe | Vietnam |
| China | Japan | Other USSR and Russia | Wales |
| Cuba and West Indies | Jordan | Philippines | Yemen |
| Cyprus | Korea | Portugal | |
| Eastern Europe | Laos | Rest of Asia | |
| England | Lebanon | Saudi Arabia | |

Table E.5: Effect of Immigration on U.S. Commuting Zones

| | Change in log native employment | Change in avg. log weekly wage | | |
|---|---|---|---|---|
| | | All workers | High-Skill. | Low-Skill |
| **Panel A: 2SLS Regression** | | | | |
| $\hat{\beta}$ | $-0.49$ | 0.13 | 0.27 | $-0.2$ |
| Robust | $[-1.12, 0.14]$ | $[-0.37, 0.63]$ | $[-0.09, 0.64]$ | $[-0.85, 0.44]$ |
| St-cluster | $[-0.98, 0.01]$ | $[-0.15, 0.41]$ | $[0.08, 0.47]$ | $[-0.49, 0.08]$ |
| AKM | $[-1.05, 0.08]$ | $[-0.39, 0.65]$ | $[-0.13, 0.67]$ | $[-0.80, 0.39]$ |
| AKM0 | $[-1.09, 0.83]$ | $[-0.33, 1.56]$ | $[-0.08, 1.41]$ | $[-0.76, 1.37]$ |
| **Panel B: OLS Reduced-Form Regression** | | | | |
| $\hat{\beta}$ | $-0.19$ | 0.05 | 0.11 | $-0.08$ |
| Robust | $[-0.39, 0.02]$ | $[-0.16, 0.26]$ | $[-0.07, 0.28]$ | $[-0.30, 0.14]$ |
| St-cluster | $[-0.37, 0.00]$ | $[-0.07, 0.17]$ | $[0.01, 0.20]$ | $[-0.17, 0.01]$ |
| AKM | $[-0.41, 0.04]$ | $[-0.16, 0.26]$ | $[-0.08, 0.29]$ | $[-0.29, 0.13]$ |
| AKM0 | $[-0.64, 0.50]$ | $[-0.10, 1.91]$ | $[-0.02, 1.72]$ | $[-0.24, 1.54]$ |
| **Panel C: 2SLS First-Stage** | | | | |
| $\hat{\beta}$ | | 0.38 | | |
| Robust | | $[0.20, 0.57]$ | | |
| St-cluster | | $[0.27, 0.49]$ | | |
| AKM | | $[0.22, 0.55]$ | | |
| AKM0 | | $[0.24, 1.49]$ | | |

Notes: $N = 2,166$ (722 CZs $\times$ three time periods). Models are weighted by start of period CZ share of national population. All regressions include period dummies. 95% confidence intervals in square brackets. *Robust* is the Eicker-Huber-White standard error; *St-cluster* is the standard error that clusters of CZs in the same state; *AKM* is the standard error in Remark 5; and *AKM0* is the confidence interval in Remark 6

## References


**Adão, Rodrigo**, "Worker Heterogeneity, Wage Inequality, and International Trade: Theory and Evidence from Brazil," *University of Chicago mimeo*, 2016.

**Allen, Treb and Costas Arkolakis**, "Trade and the Topography of the Spatial Economy," *Quarterly Journal of Economics*, 2016, *129* (3), 1085–1140.

**_ , _ , and Yuta Takahashi**, "Universal Gravity," *Dartmouth University mimeo*, 2018.

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

**Burstein, Ariel, Eduardo Morales, and Jonathan Vogel**, "Changes in Between-Group Inequality: Computers, Occupations, and International Trade," *University of California Los Angeles mimeo*, 2018.

**Fajgelbaum, Pablo, Eduardo Morales, Juan Carlos Suárez Serrato, and Owen Zidar**, "State Taxes and Spatial Misallocation," *University of California Los Angeles mimeo*, 2018.

**Galle, Simon, Andrés Rodríguez-Clare, and Moises Yi**, "Slicing the Pie: Quantifying the Aggregate and Distributional Effects of Trade," *University of California Berkeley mimeo*, 2017.

**Jones, Ronald W.**, "A Three-Factor Model in Theory, Trade and History," in Jagdish Bhagwati, Ronald Jones, Robert Mundell, and Jaroslav Vanek, eds., *Trade, Balance of Payments and Growth*, North-Holland, 1971.

**Kovak, Brian K**, "Regional effects of trade reform: What is the correct measure of liberalization?," *American Economic Review*, 2013, *103* (5), 1960–76.

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

**Redding, Stephen J.**, "Goods Trade, Factor Mobility and Welfare," *Journal of International Economics*, 2016, *101*, 148–167.

**_ and Esteban Rossi-Hansberg**, "Quantitative Spatial Economics," *Annual Review of Economics*, 2017, *9*, 21–58.