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

Rodrigo Adão* Michal Kolesár† Eduardo Morales‡

August 9, 2019

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

Contents

Appendix A Proofs and additional theoretical results

Appendix A.1 gives proofs and additional details for the results in Sections 4.1 and 4.2. Appendix A.2 gives proofs and additional details for the results in Sections 4.3 and 5.3.

A.1 Proofs and additional details for OLS regression

Since Propositions 1 and 2 are special cases or Propositions 3 and 4, we only prove Propositions 3, 4 and 5. We give the proofs under a slightly more general setup that allows for a linearization error in the potential outcome equation. We introduce this more general setup in Appendix A.1.1, where we also collect the assumptions that we impose on the DGP. We collect some auxiliary Lemmata used in the proofs in Appendix A.1.2, and we prove these propositions in Appendices A.1.3, A.1.3 and A.1.5. Appendix A.1.6 discusses inference when the effects β_{is} are heterogeneous.

Throughout the Appendix, we assume that $\sum_{s=1}^{S} w_{is} \leq 1$ for all i. Thus, $\sum_{s=1}^{S} n_s \leq N$, where $n_s = \sum_{i=1}^{N} w_{is}$ denotes the size of sector s. We use the notation $A_S \preceq B_S$ to denote $A_S = O(B_S)$, i.e. there exists a constant C independent of S such that $A_S \leq CB_S$. Let $\mathcal{F}_0$ denote the σ-field generated by $(\mathcal{Z}, U, Y(0), B, W)$ (for the case with no covariates, $\mathcal{F}_0$ denotes the σ-field generated by $(Y(0), B, W)$). Define $\overline{w}_{st} = \sum_{i=1}^{N} w_{is}w_{it}$, $\tilde{\mathcal{X}}_s = \mathcal{X}_s - \mathcal{Z}_s'\gamma$, and $\sigma_s^2 = \text{var}(\mathcal{X}_s \mid \mathcal{F}_0)$. Finally, let $r_N = (\sum_s n_s^2)^{-1}$, and let E_W denote expectation conditional on W.

A.1.1 General setup and assumptions

We first list and discuss the regularity conditions needed for the results in Section 4.1. We then generalize the setup from Section 4.2 by allowing for a linearization error in the potential outcome equation (11). Unless stated otherwise, all limits are taken as $S \to \infty$. We leave the dependence of the number of regions $N = N_S$ on S implicit.

For the results in Section 4.1, we assume that the observed data (Y, X, W) is generated by the variables $(Y(0), B, W, \mathcal{X})$, which we model as a triangular array, so that the distribution of the data may change with the sample size.[1] The additional regularity conditions we impose on these variables, in addition to Assumptions 1 and 2 as follows:

Assumption A.1. (i) The support of β_{is} is bounded; (ii) $\frac{1}{N}\sum_{i=1}^{N}\sum_{s=1}^{S} \text{var}(\mathcal{X}_s \mid \mathcal{F}_0)w_{is}^2$ converges in probability to a strictly positive non-random limit; (iii) For some $\nu > 0$, $E[|\mathcal{X}_s|^{2+\nu} \mid \mathcal{F}_0]$ exists and is uniformly bounded, and conditional on W, the second moments of $Y_i(0)$ exist, and are bounded uniformly over i; (iv) For some $\nu > 0$, $E[|\mathcal{X}_s|^{4+\nu} \mid \mathcal{F}_0]$ is uniformly bounded, and conditional on W, the fourth moments of $Y_i(0)$ exist, and are bounded uniformly over i.

The bounded support condition on β_{is} in Assumption A.1(i) is made to keep the proofs simple and can be relaxed. Assumption A.1(ii) is a standard regularity condition ensuring that the shocks $\mathcal{X}$ have sufficient variation so that the denominator of $\hat{\beta}$, scaled by N, does not converge to zero.

[1] In other words, to allow the distribution of the data to change with the sample size S, we implicitly index the data by S. Making this index explicit, for each S, the data is thus given by the array $\{(Y_{iS}(0), \beta_{isS}, w_{isS}, \mathcal{X}_{sS}) \colon i = 1, \ldots, N_S, s = 1, \ldots, S\}$.

This requires that there is at least one "non-negligible" sector in most regions in the sense that its share w_{is} is bounded away from zero. This implies that $\sum_{s=1}^{S} n_s/N$ is also bounded away from zero. Assumption A.1(iii) imposes some mild assumptions on the existence of moments of $\mathcal{X}$ and $Y_i(0)$. Assumption A.1(iv), which is only needed for asymptotic normality, strengthens this condition.

For the results in Section 4.2, we generalize the setup in the main text by allowing for a linearization error in the expression for potential outcomes,

$$Y_i(x_1, \dots, x_S) = Y_i(0) + \sum_{s=1}^{S} w_{is} x_s \beta_{is} + L_i(x_1, \dots, x_S), \qquad \sum_{s=1}^{S} w_{is} \leq 1, \tag{A.1}$$

and we weaken Assumption 3(i) by replacing it with the assumption that the observed outcome is given by $Y_i = Y_i(\mathcal{X}_1, \dots, \mathcal{X}_S)$, such that eq. (A.1) holds with $L_i(\mathcal{X}_1, \dots, \mathcal{X}_S) = L_i$.

We assume that the observed data $(Y, \mathcal{X}, Z, W)$ is generated by the triangular array of variables $(Y(0), B, W, U, \mathcal{X}, \mathcal{Z}, L)$. Let $\check{\delta} = (Z'Z)^{-1}Z'(Y - X\beta)$ denote the regression coefficient in a regression of $Y - X\beta$ on Z, that is, the regression coefficient on Z_i in a regression in which $\hat{\beta}$ is restricted to equal to the true value β.

Assumption A.2. (i) $N^{-1} \sum_{i=1}^{N} E[L_i^2]^{1/2} \to 0$, and conditional on W, the second moments of U_i and $\mathcal{Z}_s$ exist and are bounded uniformly over i and s; (ii) $Z'Z/N$ converges in probability to a positive definite non-random limit; (iii) $(\sum_s n_s^2)^{-1/2} \sum_{i=1}^{N} E[L_i^2]^{1/2} \to 0$, $\max_i E[L_i^4 \mid W] \to 0$, and conditional on W, the fourth moments of $\mathcal{Z}_s$, and U_i exist and are bounded uniformly over s and i; (iv) $\check{\delta} - \delta = O_p(q_s)$ for some sequence $q_S \to 0$; (v) $q_S^2 N / \sum_s n_s^2 \cdot \sum_i E[(U_i'\gamma)^2] \to 0$ and $\gamma' U' \epsilon = o_p((\sum_s n_s^2)^{1/2})$.

Assumption A.2(i) imposes some mild moment restrictions on the controls Z_i. It also requires that on average, the variance of the linearization error L_i vanishes with sample size. This ensures that the linearization error does not impact the consistency of $\hat{\beta}$. Assumption A.2(ii) ensures that the controls are not collinear.

Assumptions A.2(iii) to A.2(v) are only needed for asymptotic normality. Assumption A.2(iii) strengthens the moment conditions in Assumption A.2(i). It also imposes a stricter condition on the linearization error: it requires that, on average over N, the standard deviation of L_i is of smaller order than $(\sum_s n_s^2)^{1/2}/N$, the rate of convergence of $\hat{\beta}$. A sufficient condition is that $L_i = o_p(S^{-1/2})$. This ensures that the linearization error is of smaller order than the variance of the estimator, so that the distribution of $\hat{\beta}$ does not suffer from asymptotic bias. This formalizes the assumption that the linearization error is "small". The condition that $\max_i E[L_i^4 \mid W] \to 0$ is only needed for showing consistency of the standard error estimator; it is not needed for asymptotic normality. Assumption A.2(iv) requires that $\check{\delta}$ is consistent, which ensures that the error in estimation of δ does not affect the asymptotic distribution of $\hat{\beta}$. Finally, Assumption A.2(v) imposes conditions on $U_i'\gamma$, the measurement error for controls that matter, which ensure that measurement error in the controls that matter does not impact the asymptotic distribution of $\hat{\beta}$. They are stated as high-level conditions to cover a range of different cases, and depend on the rate of convergence q_S of $\check{\delta}$. In typical cases, the rate will be $q_S = (\sum_s n_s^2)^{1/2}/N$, the same as that of $\hat{\beta}$, and the condition $q_S^2 N / \sum_s n_s^2 \cdot \sum_i E[(U_i'\gamma)^2] \to 0$ is implied by Assumption 3(iii). Let U_{1i} denote the subset of elements of U_i for which $\gamma_k \neq 0$, and let U_{2i} denote the remaining elements. If U_{i1} is mean zero and independent across i conditional on the remaining variables $((Y(0), W, B, \mathcal{Z}, \mathcal{X}, U_2))$, so that these elements are pure measurement error, then the second condition is implied by Assumption 3(iv).

A.1.2 Auxiliary results

Lemma A.1. *$\{\mathcal{A}_{S1}, \dots, \mathcal{A}_{SS}\}_{S=1}^{\infty}$ be a triangular array of random variables. Fix $\eta \geq 1$, and let $A_{Si} = \sum_{s=1}^{S} w_{is}\mathcal{A}_{Ss}$, $i = 1 \dots, N_S$. Suppose $E[|\mathcal{A}_{Ss}|^{\eta} \mid W]$ exists and is uniformly bounded. Then $E[|A_{Si}|^{\eta} \mid W]$ exists and is bounded uniformly over S and i.*

Proof. The result follows by triangle inequality for $\eta = 1$. Suppose therefore that $\eta > 1$. By Hölder's inequality,

$$E[|A_{Si}|^{\eta} \mid W] = E\left[\left|\sum_{s=1}^{S} w_{is}^{\frac{\eta-1}{\eta}} w_{is}^{\frac{1}{\eta}}\mathcal{A}_{Ss}\right|^{\eta} \mid W\right] \leq \left(\sum_{s=1}^{S} w_{is}\right)^{\eta-1} \sum_{s=1}^{S} w_{is} E[|\mathcal{A}_{Ss}|^{\eta} \mid W]$$
$$\leq \max_{s} E[|\mathcal{A}_{Ss}|^{\eta} \mid W] \cdot (\textstyle\sum_{s=1}^{S} w_{is})^{\eta} \leq \max_s E[|\mathcal{A}_{Ss}|^{\eta} \mid W],$$

which yields the result. □

Lemma A.2. *$\{A_{S1}, \dots, A_{SN_S}\}_{S=1}^{\infty}$ be a triangular array of random variables. Suppose $E[A_{Si}^2 \mid W]$ exists and is uniformly bounded. Then $\sum_{s=1}^{S} E\left[(\sum_{i=1}^{N} w_{is}A_{Si})^2 \mid W\right] \preceq \sum_s n_s^2$.*

Proof. By Cauchy-Schwarz inequality,

$$\sum_{s=1}^{S} E\left[\left(\sum_{i=1}^{N} w_{is}A_{Si}\right)^2 \Bigm| W\right] \leq \sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{is}w_{js}E[A_{Si}^2 \mid W]^{1/2}E[A_{Sj}^2 \mid W]^{1/2}$$
$$\preceq \sum_{s=1}^{S}\sum_{i=1}^{N}\sum_{j=1}^{N} w_{is}w_{js} = \sum_{s=1}^{S} n_s^2.$$

□

Lemma A.3. *Let $\{A_{S1}, \dots, A_{SN_S}, B_{S1}, \dots, B_{SN_S}, \mathcal{A}_{S1}, \dots, \mathcal{A}_{SS}\}_{S=1}^{\infty}$ be a triangular array of random variables. Suppose $E[A_{Si}^4 \mid W]$, $E[B_{Si}^4 \mid W]$, and $E[\mathcal{A}_{Ss}^2 \mid W]$ exist and are uniformly bounded. Then $(\sum_s n_s^2)^{-1} \cdot \sum_{i,j,s} w_{is}w_{js}A_{Si}B_{Sj}\mathcal{A}_{Ss} = O_p(1)$.*

Proof. Let $R_S = (\sum_s n_s^2)^{-1}\sum_{i,j,s} w_{is}w_{js}A_{Si}B_{Sj}\mathcal{A}_{Ss}$. By the triangle and Cauchy-Schwarz inequalities,

$$E[|R_S| \mid W] \leq \frac{1}{\sum_s n_s^2}\sum_{i,j,s} w_{is}w_{js}E[|A_{Si}B_{Sj}\mathcal{A}_{Ss}| \mid W]$$
$$\leq \frac{1}{\sum_s n_s^2}\sum_{i,j,s} w_{is}w_{js}E[|B_{Sj}|^4 \mid W]^{1/4}E[|A_{Si}|^4 \mid W]^{1/4}E[\mathcal{A}_{Ss}^2 \mid W]^{1/2} \preceq \frac{1}{\sum_s n_s^2}\sum_{i,j,s} w_{is}w_{js} = 1.$$

The result then follows by Markov inequality. □

A.1.3 Proof of Proposition 3

First we show that

$$Z'W\tilde{\mathcal{X}} = O_p(1/\sqrt{r_N}). \tag{A.2}$$

Conditional on W, the left-hand side has mean zero by Assumption 3(ii), and by Assumption 2(i), the variance of the kth row given by

$$\operatorname{var}\left(\sum_{i,s} w_{is}\tilde{\mathcal{X}}_s Z_{ik} \mid W\right) = \sum_s E_W\sigma_s^2 \left(\sum_i w_{is}Z_{ik}\right)^2 \preceq \sum_s E_W \left(\sum_i w_{is}Z_{ik}\right)^2.$$

By Lemma A.1, Assumption A.2(i), and the C_r-inequality, $E_W[Z_{ik}^2] = E_W[(\sum_s w_{is}\mathcal{Z}_{sk} + U_{ik})^2]$ is uniformly bounded. Therefore, by Lemma A.2, the right-hand side is bounded by $\sum_s n_s^2$, so the result follows by Markov inequality and dominated convergence theorem.

Since $X = W\tilde{\mathcal{X}} + Z\gamma - U\gamma$, it follows from eq. (A.2) and Assumption A.2(ii) that

$$\hat{\gamma} - \gamma = (Z'Z/N)^{-1}Z'W\tilde{\mathcal{X}}/N - (Z'Z/N)^{-1}Z'U\gamma/N = o_p(1), \tag{A.3}$$

where $\hat{\gamma} = (Z'Z)^{-1}Z'X$, and the last equality follows since $\sum_s n_s^2/N^2 \leq \max_s n_s/N \to 0$ by Assumption 2(ii), and since $Z'U\gamma/N = o_p(1)$ by the Cauchy-Schwarz inequality and Assumption 3(iii).

Next, we will show that

$$\ddot{X}'\ddot{X}/N = \frac{1}{N}\sum_{i,s} w_{is}^2\sigma_s^2 + o_p(1). \tag{A.4}$$

To this end, we have

$$\begin{aligned}
\ddot{X}'\ddot{X}/N &= (W\tilde{\mathcal{X}} - U\gamma - Z(\hat{\gamma}-\gamma))'(W\tilde{\mathcal{X}} - U\gamma - Z(\hat{\gamma}-\gamma))/N \\
&= (W\tilde{\mathcal{X}})'(W\tilde{\mathcal{X}})/N + o_p(1) \\
&= \frac{1}{N}\sum_s \overline{w}_{ss}\sigma_s^2 + \frac{2}{N}\sum_{s<t}\overline{w}_{st}\tilde{\mathcal{X}}_s\tilde{\mathcal{X}}_t + \frac{1}{N}\sum_s \overline{w}_{ss}(\tilde{\mathcal{X}}_s^2 - \sigma_s^2) + o_p(1).
\end{aligned}$$

where the first line follows from the decomposition

$$\ddot{X} = X - Z(Z'Z)^{-1}Z'X = X - Z\hat{\gamma} = W\tilde{\mathcal{X}} - U\gamma - Z(\hat{\gamma}-\gamma), \tag{A.5}$$

the second line follows by the Cauchy-Schwarz inequality, Assumption 3(iii), and eq. (A.3), and the third line follows by expanding $(W\tilde{\mathcal{X}})'(W\tilde{\mathcal{X}})/N$. Therefore, to show eq. (A.4), it suffices to show that the second and third term in the above expression are $o_p(1)$. Since the second term has mean zero conditional on W, it suffices to show that its variance converges to zero. To that end,

$$\begin{aligned}
\operatorname{var}\left(\frac{2}{N}\sum_{s<t}\tilde{\mathcal{X}}_s\tilde{\mathcal{X}}_t\overline{w}_{st} \mid W\right) &= \frac{4}{N^2}\sum_{s<t}E_W[\sigma_s^2\sigma_t^2]\overline{w}_{st}^2 \preceq \frac{1}{N^2}\sum_{s,t}\overline{w}_{st}^2 = \frac{1}{N^2}\sum_{i,j,s,t} w_{is}w_{it}w_{js}w_{jt} \\
&\leq \frac{1}{N^2}\sum_{i,j,s,t} w_{is}w_{it}w_{js} \leq \frac{1}{N^2}\sum_{i,j,s} w_{is}w_{js} = \frac{1}{N^2}\sum_s n_s^2 \leq \frac{\max_t n_t \sum_s n_s}{N^2} \to 0.
\end{aligned}$$

where the convergence to 0 follows by Assumption 2(ii). By the inequality of von Bahr and Esseen, Assumption A.1(iii), and the inequality $\overline{w}_{ss} \leq n_s$,

$$E[N^{-1}|\sum_s(\tilde{\mathcal{X}}_s^2 - \sigma_s^2)\overline{w}_{ss}|^{1+\nu/2} \mid \mathcal{F}_0] \leq \frac{2}{N^{1+\nu/2}}\sum_s \overline{w}_{ss}^{1+\nu/2}E[|\tilde{\mathcal{X}}_s^2 - \sigma_s^2|^{1+\nu/2} \mid \mathcal{F}_0]$$
$$\preceq \frac{1}{N^{1+\nu/2}}\sum_s \overline{w}_{ss}^{1+\nu/2} \leq (\max_s n_s/N)^{\nu/2}, \quad \text{(A.6)}$$

which converges to zero by Assumption 2(ii). Equation (A.4) then follows by Markov inequality.

Next, we show that

$$\ddot{X}'Y/N = \frac{1}{N}\sum_{i,s}\sigma_s^2 w_{is}^2\beta_{is} + o_P(1) \quad \text{(A.7)}$$

Using eq. (A.5), we can write the left-hand side as

$$\begin{aligned}
\ddot{X}'Y/N &= \tilde{\mathcal{X}}'W'Y/N - \gamma'U'Y/N - Y'Z/N\cdot(\hat{\gamma}-\gamma)\\
&= \tilde{\mathcal{X}}'W'Y/N + o_p(1)\\
&= \frac{1}{N}\sum_{s,i}w_{is}\tilde{\mathcal{X}}_s L_i + \frac{1}{N}\sum_{s,i}w_{is}^2(\tilde{\mathcal{X}}_s\mathcal{X}_s - \sigma_s^2)\beta_{is} + \frac{1}{N}\sum_{s,i}w_{is}\tilde{\mathcal{X}}_s Y_i(0)\\
&\quad + \frac{1}{N}\sum_{s<t}\sum_i w_{is}w_{it}\tilde{\mathcal{X}}_s\mathcal{X}_t\beta_{it} + \frac{1}{N}\sum_{s<t}\sum_i w_{is}w_{it}\tilde{\mathcal{X}}_t\mathcal{X}_s\beta_{is} + \frac{1}{N}\sum_{s,i}w_{is}^2\sigma_s^2\beta_{is} + o_p(1)
\end{aligned}$$

where the second line follows since by the C_r-inequality, Lemma A.1, Assumptions A.1(i), A.2(i) and A.1(iii), $N^{-1}\sum_i E[Y_i^2]$ is bounded, so that $Y'Z/N = O_p(1)$ and $\gamma'U'Y/N = o_p(1)$ by Cauchy-Schwarz inequality and Assumption 3(iii), and the third line follows by expanding $\tilde{\mathcal{X}}'W'Y$. We therefore need to show that the first five terms in the expression above are $o_p(1)$. By the Cauchy-Schwarz inequality, the expectation of the absolute value of the first term is bounded by

$$N^{-1}\sum_i E[L_i^2]^{1/2}(E\sum_s w_{is}^2\sigma_s^2)^{1/2} \preceq N^{-1}\sum_i E[L_i^2]^{1/2},$$

which converges to zero by Assumption A.2(i). Thus, the first term is $o_p(1)$ by Markov inequality and the dominated convergence theorem. The second term is $o_p(1)$ by an argument analogous to eq. (A.6). The third to fifth terms are mean zero conditional on $\mathcal{F}_0$, so it suffices to show that their variances conditional on W converge to zero. The variance of the third summand is bounded by

$$\operatorname{var}\left(\frac{1}{N}\sum_s\tilde{\mathcal{X}}_s\sum_i w_{is}Y_i(0) \mid W\right) = \frac{1}{N^2}\sum_s E_W\sigma_s^2\left(\sum_i w_{is}Y_i(0)\right)^2 \preceq \frac{1}{N^2}\sum_s E_W\left(\sum_i w_{is}Y_i(0)\right)^2,$$

which converges to zero by Lemma A.2. The variance of the fourth term is bounded by

$$\operatorname{var}\left(\frac{1}{N}\sum_{s<t}\sum_i w_{is}w_{it}\tilde{\mathcal{X}}_s\mathcal{X}_t\beta_{it} \mid W\right) = \frac{1}{N^2}\sum_{s<t,t'}\sum_{i,i'}w_{is}w_{it}\sigma_s^2 E_W[\mathcal{X}_t\mathcal{X}_{t'}]\beta_{it}w_{i's}w_{i't'}\beta_{i't'}$$

$$\preceq \frac{1}{N^2} \sum_{s,t,t',i,i'} w_{is} w_{it} w_{i's} w_{i't'} \leq \frac{1}{N^2} \sum_s n_s^2 \leq \max_s n_s / N \to 0.$$

Variance of the fifth term converges to zero by analogous arguments.

Combining eq. (A.4) with eq. (A.7) and Assumption A.1(ii) then yields the result.

A.1.4 Proof of Proposition 4

Using eq. (A.5), we have

$$\begin{aligned} r_N^{1/2}(\ddot{X}'\ddot{X})(\hat{\beta} - \beta) &= r_N^{1/2} X'(I - Z(Z'Z)^{-1}Z')(Z\delta + \epsilon) = r_N^{1/2} X'(I - Z(Z'Z)^{-1}Z')\epsilon \\ &= r_N^{1/2} \tilde{\mathcal{X}}' W' \epsilon - r_N^{1/2} \gamma' U' \epsilon - r_N^{1/2} (\hat{\gamma} - \gamma)' Z' \epsilon. \end{aligned}$$

The third term can be written as

$$\begin{aligned} r_N^{1/2}(\hat{\gamma} - \gamma)' Z' \epsilon &= r_N^{1/2} \epsilon' Z (Z'Z)^{-1} (Z'W\tilde{\mathcal{X}} - Z'U\gamma) = r_N^{1/2} (\check{\delta} - \delta)' (Z'W\tilde{\mathcal{X}} - Z'U\gamma) \\ &= (\check{\delta} - \delta)' (O_p(1) - r_N^{1/2} Z'U\gamma) \\ &= o_p(1) - O_p(1) \cdot q_S r_N^{1/2} Z'U\gamma = o_p(1), \end{aligned}$$

where the first line follows from the decomposition in eq. (A.3), the second line follows from eq. (A.2), the third line follows by Assumption A.2(iv), and the last equality follows since by Cauchy-Schwarz inequality and Assumption A.2(v), $q_S r_N^{1/2} E[|Z_k' U \gamma|] \preceq \sqrt{q_S^2 r_N N \sum_i E(U_i'\gamma)^2} \to 0$. Since $r_N^{1/2}\gamma' U' \epsilon = o_p(1)$ by Assumption A.2(v), and since by eq. (A.4) and Assumption A.1(ii), $(\ddot{X}'\ddot{X}/N)^{-1} = (1 + o_p(1)) \cdot (N^{-1} \sum_{i,s} \pi_{is})^{-1}$, it follows that

$$\frac{N}{(\sum_s n_s^2)^{1/2}} (\hat{\beta} - \beta) = (1 + o_p(1)) \frac{1}{N^{-1} \sum_{i,s} \pi_{is}} r_N^{1/2} \sum_{s,i} \tilde{\mathcal{X}}_s w_{is} \epsilon_i + o_p(1).$$

Therefore, it suffices to show

$$r_N^{1/2} \sum_{s,i} \tilde{\mathcal{X}}_s w_{is} \epsilon_i = \mathcal{N}(0, \operatorname{plim} \mathcal{V}_N) + o_p(1). \tag{A.8}$$

Define $V_i = Y_i(0) - Z_i'\delta + \sum_t w_{it} \mathcal{Z}_t' \gamma (\beta_{it} - \beta)$, and

$$a_s = \sum_i w_{is} V_i, \qquad\qquad b_{st} = \sum_i w_{is} w_{it} (\beta_{it} - \beta). \tag{A.9}$$

Then we can write $\epsilon_i = V_i + \sum_t w_{it} \tilde{\mathcal{X}}_t (\beta_{it} - \beta) + L_i$. Since

$$E|r_N^{1/2} \sum_{i,s} \tilde{\mathcal{X}}_s w_{is} L_i| \leq r_N^{1/2} \sum_i (\sum_s E w_{is}^2 \sigma_s^2)^{1/2} E[L_i^2]^{1/2} \preceq r_N^{1/2} \sum_i E[L_i^2]^{1/2} \to 0$$

by Assumption A.2(iii), and since $0 = \sum_{i,s} \pi_{is}(\beta_{is} - \beta) = \sum_s \sigma_s^2 b_{ss}$, we can decompose

$$r_N^{1/2} \sum_{s,i} \tilde{\mathcal{X}}_s w_{is} \epsilon_i = r_N^{1/2} \sum_s \tilde{\mathcal{X}}_s \sum_i w_{is} \left(V_i + \sum_t w_{it} \tilde{\mathcal{X}}_t (\beta_{it} - \beta) + L_i \right) = r_N^{1/2} \sum_s \mathcal{Y}_s + o_P(1),$$

where

$$\mathcal{Y}_s = \tilde{\mathcal{X}}_s a_s + (\tilde{\mathcal{X}}_s^2 - \sigma_s^2) b_{ss} + \sum_{t=1}^{s-1} \tilde{\mathcal{X}}_s \tilde{\mathcal{X}}_t (b_{st} + b_{ts}).$$

Observe that $\mathcal{Y}_s$ is a martingale difference array with respect to the filtration $\mathcal{F}_s = \sigma(\mathcal{X}_1, \dots, \mathcal{X}_s, \mathcal{F}_0)$.

By the dominated convergence theorem and the martingale central limit theorem, it suffices to show that $r_N^{1+\nu/4} \sum_{s=1}^S E_W[\mathcal{Y}_s^{2+\nu/2}] \to 0$ for some $\nu > 0$ so that the Lindeberg condition holds, and that the conditional variance converges,

$$r_N \sum_{s=1}^S E[\mathcal{Y}_s^2 \mid \mathcal{F}_{s-1}] - \mathcal{V}_N = o_p(1).$$

To verify the Lindeberg condition, by the C_r-inequality, it suffices to show that

$$r_N^2 \sum_s E_W[\tilde{\mathcal{X}}_s^4 a_s^4] \to 0, \qquad r_N^{1+\nu/4} \sum_s E_W[(\tilde{\mathcal{X}}_s^2 - \sigma_s^2)^{2+\nu/2} b_{ss}^{2+\nu/2}] \to 0,$$

$$r_N^2 \sum_s E_W \left(\sum_{t=1}^{s-1} \tilde{\mathcal{X}}_s \tilde{\mathcal{X}}_t b_{st} \right)^4 \to 0, \qquad r_N^2 \sum_s E_W \left(\sum_{t=1}^{s-1} \tilde{\mathcal{X}}_s \tilde{\mathcal{X}}_t b_{ts} \right)^4 \to 0.$$

Note that since $E(\sum_t w_{it} \mathcal{Z}_t' \gamma (\beta_{it} - \beta))^4 \preceq (\sum_t w_{it})^4 \preceq 1$, it follows from Assumptions A.2(iii) and A.1(iv), and the C_r inequality that the fourth moment of V_i exists and is bounded. Therefore, by arguments as in the proof of Lemma A.2, $\sum_s E_W[a_s^4] \preceq \sum_s n_s^4$, so that

$$r_N^2 \sum_s E_W[\tilde{\mathcal{X}}_s^4 a_s^4] = r_N^2 \sum_s E_W[E[\tilde{\mathcal{X}}_s^4 \mid \mathcal{F}_0] a_s^4] \preceq r_N^2 \sum_s E_W[a_s^4] \preceq r_N^2 \sum_s n_s^4 \leq \max_s n_s^2 r_N \to 0 \tag{A.10}$$

by Assumption 2(iii). Second, since β_{is} is bounded by Assumption A.1(i), we have $b_{ss} \preceq \sum_i w_{is}^2 \leq n_s$, so that

$$r_N^{1+\nu/4} \sum_s E_W[(\tilde{\mathcal{X}}_s^2 - \sigma_s^2)^{2+\nu/2} b_{ss}^{2+\nu/2}] \preceq r_N^{1+\nu/4} \sum_s n_s^{2+\nu/2} \leq (r_N \max_s n_s^2)^{\nu/4} \to 0.$$

Third, by similar arguments

$$\begin{aligned} r_N^2 \sum_s E_W \left(\sum_{t=1}^{s-1} \tilde{\mathcal{X}}_s \tilde{\mathcal{X}}_t b_{st} \right)^4 &= r_N^2 \sum_s E_W E[\tilde{\mathcal{X}}_s^4 \mid \mathcal{F}_0] E \left[\left(\sum_{t=1}^{s-1} \tilde{\mathcal{X}}_t b_{st} \right)^4 \mid \mathcal{F}_0 \right] \\ &\preceq r_N^2 \sum_s \left(\sum_{t=1}^{s-1} \sum_i w_{is} w_{it} \right)^4 \leq r_N^2 \sum_s n_s^4 \to 0. \end{aligned}$$

The claim that $r_N^2 \sum_s E_W \left(\sum_{t=1}^{s-1} \tilde{\mathcal{X}}_s \tilde{\mathcal{X}}_t b_{ts} \right)^4 \to 0$ follows by similar arguments.

It remains to verify that the conditional variance converges. Since $\mathcal{V}_N$ can be written as

$$\mathcal{V}_N = \frac{1}{\sum_{s=1}^{S} n_s^2} \operatorname{var}\left(\sum_i (X_i - Z_i'\gamma)\epsilon_i \mid \mathcal{F}_0\right) = r_N \sum_s E[\mathcal{Y}_s^2 \mid \mathcal{F}_0] + o_P(1)$$
$$= r_N \sum_s \left[E\left[(\tilde{\mathcal{X}}_s a_s + (\tilde{\mathcal{X}}_s^2 - \sigma_s^2)b_{ss})^2 \mid \mathcal{F}_0\right] + \sum_{t=1}^{s-1} \sigma_s^2\sigma_t^2(b_{st}+b_{ts})^2\right] + o_p(1),$$

we can decompose

$$r_N \sum_s E[\mathcal{Y}_s^2 \mid \mathcal{F}_{s-1}] - \mathcal{V}_N = 2D_1 + D_2 + 2D_3 + o_p(1),$$

where

$$D_1 = r_N \sum_s (\sigma_s^2 a_s + E[\tilde{\mathcal{X}}_s^3 \mid \mathcal{F}_0]b_{ss}) \sum_{t=1}^{s-1} \tilde{\mathcal{X}}_t (b_{st}+b_{ts}),$$
$$D_2 = r_N \sum_s \sigma_s^2 \sum_{t=1}^{s-1} (\tilde{\mathcal{X}}_t^2 - \sigma_t^2)(b_{st}+b_{ts})^2,$$
$$D_3 = r_N \sum_s \sigma_s^2 \sum_{t=1}^{s-1}\sum_{u=1}^{t-1} \tilde{\mathcal{X}}_t\tilde{\mathcal{X}}_u (b_{st}+b_{ts})(b_{su}+b_{us}).$$

It therefore suffices to show that $D_j = o_p(1)$ for $j = 1, 2, 3$. Since $E[D_j \mid \mathcal{F}_0] = 0$, it suffices to show that $\operatorname{var}(D_j \mid W) = E_W[\operatorname{var}(D_j \mid \mathcal{F}_0)]$ converges to zero. Since $b_{st} + b_{ts} \preceq \overline{w}_{st}$, and since $E_W[|a_s a_t|] \preceq n_s n_t$, and $|b_{ss}| \preceq \overline{w}_{ss} \leq n_s$, it follows that

$$\operatorname{var}(D_1 \mid W) = r_N^2 \sum_t E_W\left[\sigma_t^2 \left(\sum_{s=t+1}^{S} (b_{st}+b_{ts})(\sigma_s^2 a_s + E[\tilde{\mathcal{X}}_s^3 \mid \mathcal{F}_0]b_{ss})\right)^2\right]$$
$$\preceq r_N^2 \sum_t \left(\sum_{s=t+1}^{S} \overline{w}_{st} n_s\right)^2 \leq r_N^2 \max_s n_s^2 \sum_t \left(\sum_s \overline{w}_{st}\right)^2 = r_N \max_s n_s^2 \to 0,$$

where the convergence to zero follows by Assumption 2(iii). By similar arguments, since $\overline{w}_{st} \leq n_s$

$$\operatorname{var}(D_2 \mid W) = r_N^2 \sum_t E_W(\tilde{\mathcal{X}}_t^2 - \sigma_t^2)^2 \left(\sum_{s=t+1}^{S} \sigma_s^2 (b_{st}+b_{ts})^2\right)^2 \preceq r_N^2 \sum_t \left(\sum_{s=t+1}^{S} \overline{w}_{st}^2\right)^2$$
$$\leq r_N^2 \sum_t \left(\sum_{s=1}^{S} n_s \overline{w}_{st}\right)^2 \leq r_N \max_s n_s^2 \to 0.$$

Finally,

$$\operatorname{var}(D_3 \mid W) = r_N^2 \sum_t \sum_{u=t+1}^{S} E_W \sigma_t^2\sigma_u^2 \left(\sum_{s=u+1}^{S} \sigma_s^2 (b_{st}+b_{ts})(b_{su}+b_{us})\right)^2$$
$$\preceq r_N^2 \sum_t \sum_{u=t+1}^{S} \left(\sum_{s=u+1}^{S} \overline{w}_{st}\overline{w}_{su}\right)^2 \leq r_N^2 \sum_{s,t,u,v} \overline{w}_{st}\overline{w}_{su}\overline{w}_{vt}\overline{w}_{vu} \leq r_N \max_s n_s^2 \to 0,$$

where the last line follows from the fact that since $\sum_s \overline{w}_{st} = n_t$ and $\overline{w}_{st} \leq n_s$,

$$\sum_{s,t,u,v} \overline{w}_{st}\overline{w}_{su}\overline{w}_{vt}\overline{w}_{vu} \leq \max_s n_s \sum_{s,t,u,v} \overline{w}_{su}\overline{w}_{vt}\overline{w}_{vu} = \max_s n_s \sum_{u,v} n_u n_v \overline{w}_{vu}$$
$$\leq \max_s n_s^2 \sum_{u,v} n_v \overline{w}_{vu} = \max_s n_s^2 / r_N. \quad \text{(A.11)}$$

Consequently, $D_j = o_p(1)$ for $j = 1, 2, 3$, the conditional variance converges, and the theorem follows.

A.1.5 Proof of Proposition 5

We'll prove a more general result that doesn't assume constant treatment effects. In particular, we will show that under the conditions of the proposition when the condition $\beta_{is} = \beta$ is dropped, the variance estimator $\hat{\mathcal{V}}_N = r_N \sum_s \widehat{\mathcal{X}}_s \hat{R}_s^2$, where $r_N = 1/\sum_{s=1}^{S} n_s^2$ satisfies

$$\hat{\mathcal{V}}_N = r_N \sum_{s=1}^{S} E[\tilde{\mathcal{X}}_s^2 R_s^2 \mid \mathcal{F}_0] + o_p(1), \quad \text{(A.12)}$$

where, using the definitions of a_s and b_{st} in eq. (A.9),

$$R_s = \sum_{i=1}^{N} w_{is}\epsilon_i = a_s + \sum_{i=1}^{N} w_{is} L_i + \sum_{t=1}^{S} \tilde{\mathcal{X}}_t b_{st}.$$

Since under constant treatment effects, $\mathcal{V}_N = r_N \sum_{s=1}^{S} E[\tilde{\mathcal{X}}_s^2 R_s^2 \mid \mathcal{F}_0]$, the assertion of the proposition follows from eq. (A.12).

Throughout the proof, we write $E_{\mathcal{F}_0}[\cdot]$ and $E_W[\cdot]$ to denote expectations conditional on $\mathcal{F}_0$, and W, respectively. Let $\tilde{\theta} = (\tilde{\beta}, \tilde{\delta}')'$, $\theta = (\beta, \delta)$, $M_i = (X_i, Z_i')'$. We can decompose the variance estimator as

$$\hat{\mathcal{V}}_N = r_N \sum_s (\widehat{\mathcal{X}}_s^2 - \tilde{\mathcal{X}}_s^2)\hat{R}_s^2 + r_N \sum_s \tilde{\mathcal{X}}_s^2(\hat{R}_s^2 - R_s^2) + r_N \sum_s (\tilde{\mathcal{X}}_s^2 R_s^2 - E_{\mathcal{F}_0}[\tilde{\mathcal{X}}_s^2 R_s^2]) + r_N \sum_s E_{\mathcal{F}_0}[\tilde{\mathcal{X}}_s^2 R_s^2]. \quad \text{(A.13)}$$

We need to show that the first three terms are $o_p(1)$. Since $\tilde{\epsilon}_i = \epsilon_i + M_i'(\theta - \tilde{\theta})$, with $\epsilon_i = V_i + L_i + \sum_t w_{it}\tilde{\mathcal{X}}_t(\beta_{it} - \beta)$, we can decompose

$$\hat{R}_s^2 = \sum_{i,j} w_{is}w_{js}\tilde{\epsilon}_i\tilde{\epsilon}_j = R_s^2 + 2\sum_{i,j} w_{js}w_{is}M_i'(\theta - \tilde{\theta})\epsilon_j + \sum_{i,j} w_{is}w_{js}M_i'(\theta - \tilde{\theta})M_j'(\theta - \tilde{\theta}). \quad \text{(A.14)}$$

Therefore, the second term in eq. (A.13) satisfies

$$r_N \sum_s \tilde{\mathcal{X}}_s^2(\hat{R}_s^2 - R_s^2) = 2(\theta - \tilde{\theta})' \left[r_N \sum_{s,i,j} w_{is}w_{js}\tilde{\mathcal{X}}_s^2 M_i\epsilon_j \right] + (\theta - \tilde{\theta})' \left[r_N \sum_{s,i,j} \tilde{\mathcal{X}}_s^2 w_{is}w_{js}M_iM_j' \right] (\theta - \tilde{\theta})$$
$$= (\theta - \tilde{\theta})' O_p(1) + (\theta - \tilde{\theta})' O_p(1)(\theta - \tilde{\theta}) = o_p(1),$$

where the second line follows by applying Lemma A.3 to the terms in square brackets. Next, the third term in (A.13) can be decomposed as

$$r_N \sum_s (\tilde{\mathcal{X}}_s^2 R_s^2 - E_{\mathcal{F}_0}[\tilde{\mathcal{X}}_s^2 R_s^2]) =$$

$$+ r_N \sum_s b_{ss}^2 (\tilde{\mathcal{X}}_s^4 - E_{\mathcal{F}_0}[\mathcal{X}_s^4]) + r_N \sum_{s<t} (b_{st}^2 + b_{ts}^2)(\tilde{\mathcal{X}}_s^2 \tilde{\mathcal{X}}_t^2 - \sigma_s^2 \sigma_t^2) + 2r_N \sum_s \sum_{t<u} b_{st} b_{su} \tilde{\mathcal{X}}_s^2 \tilde{\mathcal{X}}_t \tilde{\mathcal{X}}_u$$

$$+ r_N \sum_s (\tilde{\mathcal{X}}_s^2 - \sigma_s^2) a_s^2 + r_N \sum_{i,j,s} w_{js} w_{is} (\tilde{\mathcal{X}}_s^2 L_i L_j - E_{\mathcal{F}_0}[\tilde{\mathcal{X}}_s^2 L_i L_j]) + 2r_N \sum_{i,s} w_{is} a_s (\tilde{\mathcal{X}}_s^2 L_i - E_{\mathcal{F}_0}[\tilde{\mathcal{X}}_s^2 L_i])$$

$$+ 2r_N \sum_{s<t} a_s b_{st} \tilde{\mathcal{X}}_s^2 \tilde{\mathcal{X}}_t + 2r_N \sum_{s<t} a_t b_{ts} \tilde{\mathcal{X}}_t^2 \tilde{\mathcal{X}}_s + 2r_N \sum_s a_s b_{ss} (\tilde{\mathcal{X}}_s^3 - E_{\mathcal{F}_0}[\tilde{\mathcal{X}}_s^3])$$

$$+ r_N \sum_{i,s,t} w_{is} b_{st} (\tilde{\mathcal{X}}_s^2 \tilde{\mathcal{X}}_t L_i - E_{\mathcal{F}_0}[\tilde{\mathcal{X}}_s^2 \tilde{\mathcal{X}}_t L_i]). \quad \text{(A.15)}$$

We will show that all terms are of the order $o_p(1)$. By the inequality of von Bahr and Esseen, since b_{ss} is bounded by a constant times $\overline{w}_{ss} \leq n_s$,

$$E_{\mathcal{F}_0} | r_N \sum_s b_{ss}^2 (\tilde{\mathcal{X}}_s^4 - E_{\mathcal{F}_0}[\mathcal{X}_s^4]) |^{1+\nu/4} \preceq r_N^{1+\nu/4} \sum_s n_s^{2+\nu/2} E_{\mathcal{F}_0} |(\tilde{\mathcal{X}}_s^4 - E_{\mathcal{F}_0}[\mathcal{X}_s^4])|^{1+\nu/4} \leq (\max_s n_s^2 r_N)^{\nu/4} \to 0$$

by Assumption 2(iii), so that the first term is $o_p(1)$. The second term can be written as

$$r_N \sum_{s<t} (b_{st}^2 + b_{ts}^2)(\tilde{\mathcal{X}}_s^2 - \sigma_s^2)(\tilde{\mathcal{X}}_t^2 - \sigma_t^2) + r_N \sum_{s \neq t} (b_{st}^2 + b_{ts}^2)(\tilde{\mathcal{X}}_s^2 - \sigma_s^2) \sigma_t^2$$

The conditional variance of both summands is bounded by a constant times $r_N^2 \sum_s (\sum_t \overline{w}_{st}^2)^2 \leq r_N^2 \cdot \sum_s n_s^4 \to 0$, so that the second term is also $o_p(1)$. The third term admits the decomposition

$$2r_N \sum_s \sum_{t<u} b_{st} b_{su} \tilde{\mathcal{X}}_s^2 \tilde{\mathcal{X}}_t \tilde{\mathcal{X}}_u = 2r_N \sum_{s,t} \sum_{s \notin \{t,u\}} b_{st} b_{su} \tilde{\mathcal{X}}_s^2 \tilde{\mathcal{X}}_t \tilde{\mathcal{X}}_u + 2r_N \sum_{t \neq u} b_{tt} b_{tu} E_{\mathcal{F}_0}[\tilde{\mathcal{X}}_t^3] \tilde{\mathcal{X}}_u$$
$$2r_N \sum_{u<t} b_{tt} b_{tu} (\tilde{\mathcal{X}}_t^3 - E_{\mathcal{F}_0}[\tilde{\mathcal{X}}_t^3]) \tilde{\mathcal{X}}_u + 2r_N \sum_{t<u} b_{tt} b_{tu} (\tilde{\mathcal{X}}_t^3 - E_{\mathcal{F}_0}[\tilde{\mathcal{X}}_t^3]) \tilde{\mathcal{X}}_u.$$

The conditional variance of the first summand is bounded by a constant times $r_N^2 \sum_{t,u,s,v} \overline{w}_{st} \overline{w}_{su} \overline{w}_{vt} \overline{w}_{vu}$, which converges to zero by the inequality in eq. (A.11). The conditional variance of the second summand is bounded by a constant times $r_N^2 \sum_{s,t,u} \overline{w}_{tt} \overline{w}_{tu} \overline{w}_{ss} \overline{w}_{su} \leq r_N^2 \max_s n_s^2 \sum_s n_s^2 \to 0$. Since $(\tilde{\mathcal{X}}_t^3 - E_{\mathcal{F}_0}[\tilde{\mathcal{X}}_t^3]) \sum_{u=1}^{t-1} b_{tt} b_{tu} \tilde{\mathcal{X}}_u$ and $\tilde{\mathcal{X}}_u \sum_{t=1}^{u-1} b_{tt} b_{tu} (\tilde{\mathcal{X}}_t^3 - E_{\mathcal{F}_0}[\tilde{\mathcal{X}}_t^3])$ are both martingale differences, by the inequality of von Bahr and Esseen, the 4/3-th absolute moment of the last two terms is bounded by a constant times $r_N^{4/3} \sum_{s,t} \overline{w}_{tt}^{4/3} \overline{w}_{ts}^{4/3} \leq (\max_s n_s^2 r_N)^{1/3} r_N \sum_t n_t^2 \to 0$. Thus, all summands in the above display are of the order $o_p(1)$, and the third term in eq. (A.15) is therefore also $o_p(1)$. The fourth term is $o_p(1)$ by arguments in eq. (A.10). By the triangle and Cauchy-Schwarz inequalities, the conditional expectation of the absolute value of the fifth term is bounded by

$$2r_N \sum_{i,j,s} w_{js} w_{is} E_W[\tilde{\mathcal{X}}_s^4]^{1/2} E_W[L_i^4]^{1/4} E_W[L_j^4]^{1/4} \preceq \max_i E_W[L_j^4]^{1/2} \to 0.$$

Similarly, conditional expectation of the absolute value of the sixth term is bounded by

$$4r_N \sum_{i,j,s} w_{is} w_{js} E_W[V_j^4]^{1/4} E[\tilde{\mathcal{X}}_s^4]^{1/2} E_W[L_i^4]^{1/4} \preceq \max_i E_W[L_j^4]^{1/4} \to 0.$$

Thus, by the Markov inequality, the fifth and sixth terms are both of the order $o_p(1)$. The conditional variance of the seventh and eighth terms is bounded by a constant times $r_N^2 \sum_{s,t,u} n_s n_u \overline{w}_{st} \overline{w}_{ut} \leq r_N \max_s n_s^2 \to 0$, so that they are both $o_p(1)$ by Markov inequality. By the inequality of von Bahr and Esseen, the 4/3-th absolute moment of the last ninth term is bounded by a constant times $r_N^{4/3} \sum_s E_W[|a_s|^{4/3}] n_s^{4/3} \preceq (\max_s n_s^2 r_N)^{1/3} \to 0$, since by Jensen's inequality, $E|a_s|^{4/3} \leq (Ea_s^2)^{2/3}$, which is bounded by a constant times $n_s^{4/3}$. Finally, the expectation of the absolute value of the last term in eq. (A.15) is bounded by a constant times

$$r_N \sum_{i,s,t} w_{is} \overline{w}_{st} E_W[\tilde{\mathcal{X}}_s^4]^{1/2} E_W[\tilde{\mathcal{X}}_t^4]^{1/4} E_W[L_i^4]^{1/4} \preceq \max_i E_W[L_i^4]^{1/4} \to 0.$$

It remains to show that the first term in eq. (A.13) is $o_p(1)$. It follows from eq. (A.5) and eq. (24) that

$$\widehat{\mathcal{X}} = (W'W)^{-1} W' \ddot{X} = \tilde{\mathcal{X}} - (W'W)^{-1} W' U (\hat{\gamma} - \gamma) - \mathcal{Z}(\hat{\gamma} - \gamma) - (W'W)^{-1} W' U \gamma,$$

where $\hat{\gamma} = (Z'Z)^{-1} Z' X$. Let $\mathcal{U} = (W'W)^{-1} W' U$, and denote the sth row by $\mathcal{U}_s'$. Since $\mathcal{U}_{sk}^4 = (\sum_i ((W'W)^{-1} W')_{si} U_{ik})^4$, it follows by the Cauchy-Schwarz inequality that

$$E[\mathcal{U}_{sk}^4 \mid W] \leq \max_s E[(\sum_i ((W'W)^{-1} W')_{si} U_{ik})^4 \mid W] \preceq \max_s (\sum_i |((W'W)^{-1} W')_{si}|)^4,$$

which is bounded assumption of the proposition. Therefore, the fourth moments of $\mathcal{U}_s$ are bounded uniformly over s. Observe also that $E_W[\epsilon_i^4]$ is bounded uniformly over s by assumptions of the proposition. Therefore, by applying Lemma A.3 after using the expansion in eq. (A.14), we get

$$\begin{aligned}
r_N \sum_s (\widehat{\mathcal{X}}_s^2 - \tilde{\mathcal{X}}_s^2) \hat{R}_s^2 &= r_N \sum_s \hat{R}_s^2 (\mathcal{U}_s' \gamma)^2 - 2 r_N \sum_s \hat{R}_s^2 \tilde{\mathcal{X}}_s \mathcal{U}_s' \gamma \\
&\quad + r_N \sum_s \hat{R}_s^2 \left[2\mathcal{U}_s' \gamma - 2\tilde{\mathcal{X}}_s + (\mathcal{Z}_s + \mathcal{U}_s)'(\hat{\gamma} - \gamma) \right] (\mathcal{Z}_s + \mathcal{U}_s)'(\hat{\gamma} - \gamma) \\
&= r_N \sum_s R_s^2 (\mathcal{U}_s' \gamma)^2 - 2 r_N \sum_s R_s^2 \tilde{\mathcal{X}}_s \mathcal{U}_s' \gamma + O_p(1)(\hat{\gamma} - \gamma) + o_p(1).
\end{aligned}$$

By Cauchy-Schwarz inequality,

$$r_N \sum_s E_W |R_s^2 (\mathcal{U}_s' \gamma)^2| \leq r_N \sum_s (E_W[R_s^4])^{1/2} (E_W (\mathcal{U}_s' \gamma)^4)^{1/2} \preceq \max_s (E_W (\mathcal{U}_s' \gamma)^4)^{1/2} r_N \sum_s n_s^2 \to 0,$$

since $\max_s E_W[(\mathcal{U}_s' \gamma)^4] \preceq \max_i E_W (U_i' \gamma)^4 \max_s (\sum_i |((W'W)^{-1} W')_{si}|)^4$, which converges to zero by assumption of the proposition. By similar arguments, $2 r_N \sum_s E_W |R_s^2 \tilde{\mathcal{X}}_s \mathcal{U}_s' \gamma| \to 0$ also, so that

$$r_N \sum_s (\widehat{\mathcal{X}}_s^2 - \tilde{\mathcal{X}}_s^2) \hat{R}_s^2 = o_p(1) + O_p(1)(\hat{\gamma} - \gamma) = o_p(1),$$

where the second equality follows from eq. (A.3).

A.1.6 Inference under heterogeneous effects

For valid (but perhaps conservative) inference under heterogeneous effects, we need to ensure that when $\beta_{is} \neq \beta$, eq. (32) holds with inequality, that is,

$$\frac{\sum_{s=1}^{S} \widehat{\mathcal{X}}_s^2 \hat{R}_s^2}{\sum_{s=1}^{S} n_s^2} \geq \mathcal{V}_N + o_p(1). \tag{A.16}$$

To discuss conditions under which this is the case, suppose, for simplicity, that $L_i = 0$ so that eq. (11) holds, and $R_s = \sum_s w_{is}\epsilon_i$, where $\epsilon_i = Y_i(0) - Z_i'\delta + \sum_s \mathcal{X}_s w_{is}(\beta_{is} - \beta)$ is the regression residual. Then the "middle sandwich" in the asymptotic variance sandwich formula, $\mathcal{V}_N$, as defined in Proposition 4, can be decomposed into three terms:

$$\begin{aligned}\mathcal{V}_N = \frac{\operatorname{var}\left(\sum_s \tilde{\mathcal{X}}_s R_s \mid \mathcal{F}_0\right)}{\sum_{s=1}^{S} n_s^2} &= \frac{\sum_s E[\tilde{\mathcal{X}}_s^2 R_s^2 \mid \mathcal{F}_0]}{\sum_{s=1}^{S} n_s^2} - \frac{\sum_s E[\tilde{\mathcal{X}}_s R_s \mid \mathcal{F}_0]^2}{\sum_{s=1}^{S} n_s^2} + \frac{\sum_{s \neq t} \operatorname{cov}(\tilde{\mathcal{X}}_s R_s, \tilde{\mathcal{X}}_t R_t \mid \mathcal{F}_0)}{\sum_{s=1}^{S} n_s^2} \\ &= D_1 + D_2 + D_3, \end{aligned} \tag{A.17}$$

where

$$D_1 = \frac{\sum_s E[\tilde{\mathcal{X}}_s^2 R_s^2 \mid \mathcal{F}_0]}{\sum_{s=1}^{S} n_s^2}, \qquad D_2 = -\frac{\sum_s \left(\sum_i \sigma_s^2 w_{is}^2 (\beta_{is} - \beta)\right)^2}{\sum_{s=1}^{S} n_s^2},$$

$$D_3 = \frac{\sum_{s \neq t} \sigma_s^2 \sigma_t^2 \sum_{i,j} w_{is} w_{it} (\beta_{it} - \beta) w_{jt} w_{js} (\beta_{js} - \beta)}{\sum_{s=1}^{S} n_s^2}.$$

As shown in the proof of Proposition 5 (see eq. (A.12)), the standard error estimator consistently estimates D_1. Under homogeneous effects, $D_2 = D_3 = 0$, and it follows that the standard error estimator is consistent. To ensure valid inference under heterogeneous effects, one needs to ensure that $D_2 + D_3 \leq o_p(1)$. This is the case under several sufficient conditions, and we give two such conditions below.

The term D_2 reflects the variability of the treatment effect and it is always negative. It therefore makes the variance estimate that we propose conservative if $D_3 = o_p(1)$. An analogous term, also reflecting the variability of the treatment effect, is present in randomized, and cluster-randomized trials, which is why the robust and cluster-robust standard error estimators yield conservative inference in these settings (see, for example Imbens and Rubin, 2015, Chapter 6). The term D_3 reflects correlation between the treatment effects. It arises due to aggregating the sectoral shocks $\mathcal{X}_s$ to a regional level to form the shifter X_i, and it has no analog in cluster-randomized trials. Indeed, in the example with "concentrated sectors", which is analogous to cluster-randomized trials if there are no covariates, the term equals zero, since in that case $w_{is}w_{it} = 0$ for $s \neq t$. Our standard errors are thus valid, although conservative, in this case.

More generally, a sufficient condition for validity of our standard error estimator under treatment effect heterogeneity is that $T_N = \sum_{s \neq t} (\sum_i w_{is} w_{it})^2 / \sum_s n_s^2 \to 0$, since $D_3 = O_p(T_N)$. The condition $T_N \to 0$ requires that the shares are sufficiently concentrated so that not too many regions "specialize" in more than one sector (in the sense that the sectoral share w_{is} is bounded away from zero as $S \to \infty$

for more than one sector). For example, $T_N \to 0$ if the share of the second-largest sector goes to zero as $S \to \infty$, that is $\max_{i,s\neq s_i} w_{is} \to 0$, where s_i denotes the largest sector in region i. This follows from the inequalities

$$\begin{aligned}\sum_{i,j}\sum_{s\neq t} w_{is}w_{it}w_{js}w_{jt} &= \sum_{i,j,s,t} I(s=s_i, t\neq s_i)w_{is}w_{it}w_{js}w_{jt} + \sum_{i,j}\sum_{s\neq t} I(s\neq s_i)w_{is}w_{it}w_{js}w_{jt} \\ &\leq \sum_{i,j,s,t} I(t\neq s_i)w_{is}w_{it}w_{js}w_{jt} + \sum_{i,j,s,t} I(s\neq s_i)w_{is}w_{it}w_{js}w_{jt} \\ &\leq 2\max_{i,s\neq s_i} w_{is}\sum_{i,j,s,t} w_{it}w_{js}w_{jt} \leq 2\max_{i,s\neq s_i} w_{is}\sum_t n_t^2 = o(r_N).\end{aligned}$$

For illustration, in the empirical application in Section 7.1, $T_N = 0.0014$.

A second sufficient condition for the asymptotic negligibility of D_3 is that the conditional variance of the shifters $\mathcal{X}_s$, $\sigma_s^2 = E[(\mathcal{X}_s - \mathcal{Z}_s'\gamma)^2 \mid \mathcal{F}_0]$ and the weighted treatment effects $\sigma_s^2\beta_{is}$ are mean-independent of the shares W, provided some additional mild regularity conditions are satisfied, as shown in the lemma below. Importantly, this condition still allows the treatment effects to depend on the controls Z, or other aspects of the model, such as $Y_i(0)$: the covariance assumptions in the lemma allow the treatment effects β_{is} to be correlated within a region and/or within a sector. The assumption that $\sum_i\sum_{s\neq t} w_{is}^2 w_{it}^2/\sum_{s'} n_{s'}^2 \to 0$ holds if either a vanishing fraction of regions "specialize" in more than one sector (in the sense that the sectoral share w_{is} is bounded away from zero as $S\to\infty$ for more than one sector). It also holds if $S/\sum_s n_s \to 0$, that is, the number of regions grows faster than the number of sectors.[2] For illustration, the quantity equals 0.00022 in the empirical example in Section 7.1. The lemma uses the notation defined at the beginning of Appendix A.1.5.

Lemma A.4. *Suppose that the assumptions of Proposition 4 hold. Suppose, in addition, that the conditional expectations* $E[\sigma_s^2\beta_{is} \mid W] = E[(\mathcal{X}_s - \mathcal{Z}_s'\gamma)^2\beta_{is} \mid W]$ *and* $E[\sigma_s^2 \mid W] = E[(\mathcal{X}_s - \mathcal{Z}_s'\gamma)^2 \mid W]$ *do not depend on* W, *i, or s. Suppose also that* $\operatorname{cov}(\sigma_s^2\beta_{is}, \sigma_t^2\beta_{jt} \mid W) = 0$ *unless* $i=j$ *or* $s=t$, *that* $\operatorname{cov}((\sigma_s^2\beta_{is}, \sigma_s^2), \sigma_t^2 \mid W) = 0$ *unless* $s=t$, *and that* $\sum_{s\neq t}\sum_i w_{is}^2 w_{it}^2/\sum_s n_s^2 \to 0$. *Then* $D_3 = o_p(1)$.

Proof. By Assumptions A.1(i) and A.1(iii),

$$r_N\sum_{s\neq t}\sum_i E_W|\sigma_s^2\sigma_t^2 w_{is}^2 w_{it}^2(\beta_{it}-\beta)(\beta_{js}-\beta)| \preceq r_N\sum_{s\neq t}\sum_i w_{is}^2 w_{it}^2,$$

and the right-hand side converges to zero by assumption of the lemma. Therefore, by Markov inequality, $D_3 = r_N\sum_{s\neq t}\sum_{i\neq j} w_{is}w_{it}\sigma_t^2(\beta_{it}-\beta)w_{jt}w_{js}\sigma_s^2(\beta_{js}-\beta) + o_p(1)$. By Assumptions A.1(i), A.2(iii) and A.1(iv), and assumptions of the lemma, the variance of $\sum_{i,s} w_{is}^2\sigma_s^2\beta_{is}/N$ and of $\sum_{i,s} w_{is}^2\sigma_s^2/N$ conditional on W is bounded by a constant times $\sum_{i,j,s} w_{is}^2 w_{js}^2/N^2 + \sum_{i,s,t} w_{is}^2 w_{it}^2/N^2 \leq 2\max_s n_s/N \to 0$. Therefore, by Assumption A.1(ii), $\beta = \mu/\sigma + o_p(1)$, where $\mu = E_W[(\mathcal{X}_s - \mathcal{Z}_s'\gamma)^2\beta_{is}]$ and $\sigma = E_W[\sigma_s^2]$. It then follows that

$$D_3 = r_N\sum_{s\neq t}\sum_{i\neq j} w_{is}w_{it}w_{jt}w_{js}(\sigma_s^2\beta_{js}-\mu)(\sigma_t^2\beta_{it}-\mu) - 2r_N\sum_{s\neq t}\sum_{i\neq j} w_{is}w_{it}w_{jt}w_{js}(\mu - \sigma_t^2\mu/\sigma)(\sigma_s^2\beta_{js}-\mu)$$

[2] This follows from the inequalities $\sum_{i,s,t} w_{is}^2 w_{it}^2 \leq \sum_s n_s$, and $\sum_s n_s^2 \geq (\sum_s n_s)^2/S$.

$$+ r_N \sum_{s\neq t}\sum_{i\neq j} w_{is}w_{it}w_{jt}w_{js}(\mu - \sigma_s^2\mu/\sigma)(\mu - \sigma_t^2\mu/\sigma) + o_p(1).$$

Each term in the above display has mean zero, and variance bounded by a constant times

$$r_N^2 \sum_{s\neq t}(\sum_{i\neq j} w_{is}w_{it}w_{jt}w_{js})^2 + r_N^2 \sum_{i\neq j}(\sum_{s\neq t} w_{is}w_{it}w_{jt}w_{js})^2 \\ \leq r_N^2 \max_s n_s^2 \sum_{i,j,s,t} w_{is}w_{it}w_{js}w_{jt} + r_N^2 \sum_{i,j,s,t} w_{it}w_{jt}w_{is}w_{js} \leq 2r_N \max_s n_s^2 \to 0.$$

Therefore, $D_3 = o_p(1)$ by Markov inequality and dominated convergence theorem. □

Although both the condition $T_N \to 0$ and the conditions in Lemma A.4 may be restrictive in some applications, note that both of these conditions are merely sufficient, but not necessary for $D_3 + D_2 \leq o_p(1)$.

A.2 Proofs and additional details for IV regression

We prove eqs. (38) and (45), and show that the bias of the estimator $\tilde{\alpha}$ is of the order $\frac{1}{N}\sum_{i,s} w_{is}\check{w}_{is}/\check{n}_s$. We also discuss how the case with estimated shifters relates to the literature on many instruments.

A.2.1 Assumptions

To compactly state the assumptions, let $\mathcal{F}_0 = (\mathcal{Z}, U, Y_1(0), Y_2(0), B, W, \check{W})$, and put $\check{W} = W$, and $\psi_{is} = 0$ if the shifters $\mathcal{X}$ are observed.

We impose an instrumental variables version of the regularity conditions Assumptions A.1 and A.2:

Assumption A.3. (i) For some $\nu > 0$, $E[\mathcal{X}_s^{2+\nu} \mid \mathcal{F}_0]$ exists and is uniformly bounded. The support of β_{is} is bounded. Conditional on $(W, \check{W})$, the second moments of $Y_{1i}(0), Y_{2i}(0), U_i$ and $\mathcal{Z}_s$ exist, and are bounded uniformly over i and s. $Z'Z/N$ converges in probability to a positive definite non-random limits; (ii) For some $\nu > 0$, $E[|\mathcal{X}_s|^{4+\nu} \mid \mathcal{F}_0, \Psi]$ is uniformly bounded, and $\mathcal{X}_s$ are independent across s conditional on $(\mathcal{F}_0, \Psi)$, with $E[\mathcal{X}_s \mid \mathcal{F}_0, \Psi] = E[\mathcal{X}_s \mid \mathcal{Z}]$. Conditional on $(W, \check{W})$, the fourth moments of $Y_{1i}(0)$, U_i and $\mathcal{Z}_s$ exist, and are bounded uniformly over i and s. Assumption A.2(iv) and Assumption A.2(v) hold $\delta = E[Z'Z]^{-1}E[Z'Y_1(0)]$, $\check{\delta} = (Z'Z)^{-1}Z'Y_1(0)$, and $\epsilon_i = Y_{1i} - Y_{2i}\alpha - Z_i'\delta$.

Assumption A.3(i) is needed for consistency, and Assumption A.3(ii) is needed for asymptotic normality. When the shifters are observed, these assumptions are natural analogs of the regularity conditions in the OLS case that are needed for consistency (Assumptions A.1(i) and A.1(iii) and Assumptions A.2(i) and A.2(ii)) and asymptotic normality (Assumption A.1(iv) and Assumptions A.2(iii) to A.2(v)). When the shifters are not directly observed, Assumption A.3(ii) strengthens Assumption 4(ii) so that it holds conditionally on Ψ also.

If X_i is not observed, we need to impose additional conditions on ψ_{is} and the weights $\check{w}_{is}$:

Assumption A.4. Let A_{-i} denote the vector A with the ith element removed. Let $\mathcal{F}_{-i} = \sigma(Y_{1,-i}(0)$, $Y_{2,-i}(0), U_{-i}, W, \check{W}, \mathcal{Z})$. (i) For all s and i, $E[\check{w}_{is}\psi_{is} \mid \mathcal{F}_{-i}] = 0$, and $E[\check{w}_{is}^2\psi_{is}^2 \mid \mathcal{F}_0]$ is bounded by a universal constant times $\check{w}_{is}^2$; (ii) For all s, t, and all $i \neq j$, $E[\check{w}_{is}\check{w}_{jt}\psi_{is}\psi_{jt} \mid \mathcal{F}_{-i}] = 0$; (iii) $\max_{i,s} \check{w}_{is} / \sum_{j=1}^{N} \check{w}_{js}$ is bounded away from 1; (iv) $\max_i \sum_s \frac{n_s}{\check{n}_s}\check{w}_{is}$ is bounded; (v) There exist variables $\{C_i, \eta_i\}_{i=1}^{N}$ such that $(Y_{i1}(0), U_i) = C_i + \eta_i$, and conditional on $(C, W, \mathcal{Z})$, $\{\check{w}_{i1}\psi_{i1}, \ldots, \check{w}_{iS}\psi_{iS}, \eta_i\}$ are independent across i, with uniformly bounded second moments, and $E[(\check{w}_{is}\psi_{is}, \eta_i) \mid C, W, \check{W}, \mathcal{Z}] = 0$. Conditional on $(W, \check{W})$, the fourth moments of η_i and C_i are uniformly bounded; (vi) $E_{W,\check{W}}[\check{w}_{is}\psi_{js}]^4$ is bounded by a constant times $\check{w}_{is}^4$; (vii) $N/(\sum_s n_s^2)^2 \to 0$.

Assumption A.4(i) requires that the local shock ψ_{is} in region i is mean zero, and unrelated to the regional variables $(Y_{1j}(0), Y_{2j}(0), U_j)$ in other regions. Importantly, it allows these local shocks to be correlated with the regional variables in region i. In particular, in some applications, it may be the case that $Y_{2i} = \sum_s w_{is} X_{is} + \eta_i$, with the additional term η_i potentially zero. In this case ψ_{is} is always mechanically correlated with Y_{2i} (and hence also Y_{1i} if there is endogeneity). As we will show below, this correlation causes bias in the estimator $\tilde{\alpha}$ that ignores the estimation error in the shifters.

Assumption A.4(ii) requires that these local shocks are uncorrelated across regions: this ensures consistency of the leave-one-out estimator. One could relax this assumption and instead only require no correlation across clusters of regions, in which case one would have to leave out region i's cluster when constructing an estimate of X_i. The local shocks are allowed to be correlated across industries in the same region. The scaling by $\check{w}_{is}$ in the statement of the assumption allows for the possibility that X_{is} gives an uninformative signal about $\mathcal{X}_s$ if $\check{w}_{is} = 0$. Assumption A.4(iii) imposes two mild regularity conditions on the weights; it ensures that no single weight $\check{w}_{is}$ is so large that it dominates a particular sector, which is necessary for the leave-one-out estimator to be well-defined.

Assumption A.4(iv) ensures that the weights $\check{w}_{is}$ are balanced in the sense that no single region i is asymptotically non-negligible. The condition holds under equal weighting, $\check{w}_{is} = 1$, since in this case $\sum_s n_s\check{w}_{is}/\check{n}_s = \sum_s n_s/N \leq 1$. Oftentimes, the weights $\check{w}_{is}$ take the form $\check{w}_{is} = L_i w_{is}$, where L_i is a measure of the size or region i. In this case, $\sum_s n_s\check{w}_{is}/\check{n}_s = \sum_s \frac{L_i w_{is}}{\bar{L}_s}$, where $\bar{L}_s = \check{n}_s/n_s = \sum_i L_i w_{is} / \sum_j w_{js}$ is the sector-weighted average size of a region. Thus, the condition requires that the sector-weighted size of region i, $w_{is}L_i$, is non-negligible relative to the national average for at most a fixed number of sectors. Since $\sum_s n_s\check{w}_{is}/\check{n}_s \leq \frac{\max_i L_i}{\min_j L_j}$, a sufficient condition is that the ratio of the largest to the smallest region is bounded.

Assumptions A.4(v) to A.4(vii) are only needed for asymptotic normality. Assumption A.4(v) effectively imposes that only the part of $(Y_{i1}(0), U_i)$ that's independent of ψ_i is allowed to be correlated across i; the part that's related to ψ_i must be independent across i. Assumption A.4(vii) imposes a very mild condition on the sector sizes, and holds, for example, if $n_s \geq 1$.

A.2.2 Asymptotic results

When the shifters are observed, we obtain the following result, which implies eq. (38) in the main text:

Proposition A.1. *Suppose that Assumptions 2(i) and 2(ii) and Assumption 4 hold with $\mathcal{F}_0 = (\mathcal{Z}, U, Y_1(0), Y_2(0), B, W)$, and that Assumption A.3(i) holds. Then the estimator $\hat{\alpha}$ in eq. (36) is consistent. If, in addition, Assumption 2(iii) and Assumption A.3(ii) hold, then $\hat{\alpha}$ satisfies eq. (38), provided $\mathcal{V}_N$ converges to a non-random limit.*

The consistency result follows since by arguments analogous to those in the proof of Proposition 3 (see, in particular, eq. (A.7)), $N^{-1}\sum_i \ddot{X}_i Y_{1i}(0) = o_p(1)$, and $N^{-1}\sum_i \ddot{X}_i Y_{2i}(0) = N^{-1}\sum_{i,s}\sigma_s^2 w_{is}^2 \beta_{is} + o_p(1)$. Furthermore, since $N^{-1}\sum_{i,s}\sigma_s^2 w_{is}^2\beta_{is} \neq 0$ by Assumption 4(iv), it follows by Slutsky's lemma that

$$\hat{\alpha} - \alpha = \frac{N^{-1}\sum_i \ddot{X}_i Y_{1i}(0)}{N^{-1}\sum_i \ddot{X}_i Y_{2i}(0)} = o_p(1).$$

The asymptotic normality result follows since $r_N^{1/2}\sum_i \ddot{X}_i Y_{1i}(0) = \mathcal{N}(0, \mathcal{V}_N) + o_p(1)$ by arguments analogous to those in proof of Proposition 4 (see, in particular, eq. (A.8)).

Proposition A.2. *Suppose that Assumptions 2(i) and 2(ii) and Assumption 4 hold with $\mathcal{F}_0 = (\mathcal{Z}, U, Y_1(0), Y_2(0), B, W, \check{W})$, and that Assumption A.3(i) and Assumptions A.4(i) to A.4(iv) hold. Then the estimator $\hat{\alpha}_-$ is consistent for α. Furthermore, the estimator $\tilde{\alpha}$ satisfies $\tilde{\alpha} = \alpha + O_p\left(\frac{1}{N}\sum_{i,s}\frac{w_{is}\check{w}_{is}}{\check{n}_s}\right)$, provided that $(\ddot{\hat{X}}'Y_2/N)^2$ converges to a strictly positive probability limit.*

The asymptotic bias $\tilde{\alpha}$ is analogous to the own observation bias of the two-stage least squares (2SLS) estimator in settings with many instruments. To see the connection, consider the special case in which $Y_{2i} = \sum_s w_{is}X_{is} = \sum_s w_{is}\mathcal{X}_s + \sum_s w_{is}\psi_{is}$, and each region specializes in a single sector, $w_{is} = \mathbb{I}\{s(i) = s\}$, with $\check{w}_{is} = w_{is}$. Then we can write $Y_{2i} = \mathcal{X}_{s(i)} + \psi_{is(i)}$, and $\hat{X}_i = \frac{1}{n_s}\sum_i \mathbb{I}\{s(i) = s\}Y_{2i}$. This setting is isomorphic to a many instrument setting, where the instruments are group indicators $\mathbb{I}\{s(i) = s\}$, individuals are assigned to groups, and the average treatment intensity depends on group membership (for example, the endogenous variable may be the length of a sentence, the groups are groups of individuals assigned to the same judge, and judges differ in their average sentencing severity $\mathcal{X}_s$). Then the first-stage predictor used by the 2SLS estimator is $\hat{X}_i$. Since $\hat{X}_i$ puts weight $1/n_s$ on the first-stage regression error $\psi_{is(i)}$, this generates a bias in the 2SLS estimate, which persists in large samples unless the weight $1/n_s$ is negligible. In our setting, Proposition A.2 shows that the bias is of the order $\frac{1}{N}\sum_{i,s}\frac{w_{is}\check{w}_{is}}{\check{n}_s} \leq \frac{1}{N}\sum_{i,s}\frac{\check{w}_{is}}{\check{n}_s} = S/N$. Thus, a sufficient condition for consistency is that the number of sectors grows more slowly than the number of regions. This is analogous to the requirement for 2SLS consistency in the many instruments literature that the number of instruments grows more slowly than the number of observations.

Proposition A.3. *Suppose that Assumptions 2 and 4 hold with $\mathcal{F}_0 = (\mathcal{Z}, U, Y_1(0), Y_2(0), B, W, \check{W})$, and that Assumptions A.3 and A.4 hold. Suppose that $\mathcal{V}_N$ and $\mathcal{W}_N$, defined in eq.* (45), *converge in probability to non-random limits. Then*

$$\frac{N}{\sqrt{\sum_{s=1}^{S} n_s^2}}(\hat{\alpha}_- - \alpha) = \mathcal{N}\left(0, \frac{\mathcal{V}_N + \mathcal{W}_N}{\left(\frac{1}{N}\sum_i \ddot{X}_i Y_{2i}\right)^2}\right) + o_p(1).$$

The additional term $\mathcal{W}_N$ in the expression for the asymptotic variance of $\hat{\alpha}_-$, which is absent if $\mathcal{X}$ is observed, is of the order

$$\frac{1}{\sum_s n_s^2}\sum_j\left(\sum_s \frac{n_s \check{w}_{js}}{\check{n}_s}\right)^2 + \frac{1}{\sum_s n_s^2}\sum_{i,j,s,t}\frac{w_{is}\check{w}_{js}}{\check{n}_s}\frac{w_{jt}\check{w}_{it}}{\check{n}_t} \preceq \frac{N+S}{\sum_s n_s^2} \preceq S/N + (S/N)^2,$$

where the second inequality follows Assumption A.4(iv), and the last inequality follows by ℓ_1-ℓ_2 norm inequality $\sqrt{S\sum_s n_s^2} \geq \sum_s n_s$, and we assume that $\sum_s w_{is}$ is bounded away from zero so that $\sum_s n_s$ is of the same order as N. Therefore, if the number of regions grows faster than the number of sectors, the term will be asymptotically negligible. This is similar to the result in the many IV literature that the usual standard error formula for the jackknife IV estimator is valid if the number of instruments grows more slowly than the sample size. The term $\mathscr{W}_N$ also has a similar structure to the many-instrument term in the standard error for jackknife IV (see Chao et al. (2012)).

A.2.3 Proof of Proposition A.2

By the arguments in the proof of Proposition 3, for the first part of the proposition, it suffices to show that $(\hat{\ddot{X}}_{-} - \ddot{X})'Y_1/N = o_p(1)$ and $(\hat{\ddot{X}}_{-} - \ddot{X})'Y_2/N = o_p(1)$, which in turn follows if we can show that for $A_i \in \{Y_{1i}, Y_{2i}, Z_i\}$,

$$\frac{1}{N}\sum_i (\hat{X}_{i,-} - X_i)A_i = \frac{1}{N}\sum_{j,i,s} \mathbb{I}\{j \neq i\} \frac{w_{is}\breve{w}_{js}}{\breve{n}_{s,-i}} \psi_{js} A_i = o_p(1), \tag{A.18}$$

where $\breve{n}_{s,-i} = \sum_{j=1}^N \breve{w}_{js} - \breve{w}_{is}$. By Assumption A.4(i), conditional on W, this term has mean zero. Since by Assumption A.4(ii), $\mathbb{I}\{j \neq j'\}\,\mathbb{I}\{j \neq i\}\,\mathbb{I}\{j' \neq i'\} E_{W,\breve{W}}[w_{js}\psi_{js}A_i \cdot w_{j't}\psi_{j't}A_{i'}] = 0$ unless $j = i'$ and $j' = i$, the variance of this term is given by

$$\frac{1}{N^2}\sum_{j,i,i',s,t} \mathbb{I}\{j \neq i, i'\} w_{is} w_{i't} \frac{E_{W,\breve{W}}[\breve{w}_{js}\psi_{js}A_i \breve{w}_{jt}\psi_{jt}A_{i'}]}{\breve{n}_{s,-i}\breve{n}_{t,-i'}} + \frac{1}{N^2}\sum_{j,i,s,t} \mathbb{I}\{j \neq i\} w_{is} w_{jt} \frac{E_{W,\breve{W}}[\breve{w}_{js}\psi_{js}\breve{w}_{it}\psi_{it}A_i A_j]}{\breve{n}_{s,-i}\breve{n}_{t,-j}}.$$

Now, by Assumption A.3(i), $E_{W,\breve{W}}[\breve{w}_{js}\psi_{js}A_i\breve{w}_{jt}\psi_{jt}A_{i'}] \preceq \breve{w}_{js}\breve{w}_{jt}E_{W,\breve{W}}[A_iA_{i'}]$, which is bounded by a constant times $\breve{w}_{js}\breve{w}_{jt}$ since the second moment of A_i is uniformly bounded by Assumption A.4(i). Similarly, $E_{W,\breve{W}}[\breve{w}_{js}\psi_{js}\breve{w}_{it}\psi_{it}A_iA_j]$ is bounded by a constant times $\breve{w}_{js}\breve{w}_{it}$. Therefore, the expression in the preceding display is bounded by a constant times

$$\frac{1}{N^2}\sum_{j,i,i',s,t} w_{is}w_{i't}\frac{\breve{w}_{js}\breve{w}_{jt}}{\breve{n}_{s,-i}\breve{n}_{t,-i'}} + \frac{1}{N^2}\sum_{j,i,s,t} w_{is}w_{jt}\frac{\breve{w}_{js}\breve{w}_{it}}{\breve{n}_{s,-i}\breve{n}_{t,-j}} \leq \frac{1}{N^2}\max_{is}\frac{\breve{n}_s^2}{\breve{n}_{s,-i}^2}\left[\sum_j\left(\sum_s n_s\frac{\breve{w}_{js}}{\breve{n}_s}\right)^2 + N\right] \preceq \frac{1}{N},$$

where the first inequality follows since $\sum_{j,i,s,t} w_{is}w_{jt}\frac{\breve{w}_{js}\breve{w}_{it}}{\breve{n}_s\breve{n}_t} \leq \sum_{j,i,s,t} w_{is}w_{jt}\frac{\breve{w}_{js}}{\breve{n}_s} \leq \sum_{j,s} n_s\frac{\breve{w}_{js}}{\breve{n}_s} = N$, and the second inequality follows since Assumption A.4(iii) implies $\max_{is}\breve{n}_s/\breve{n}_{s,-i} = 1/(1-\max_{is}\breve{w}_{is}/\breve{n}_{is})$ is bounded, and since Assumption A.4(iv) implies that $\sum_j\left(\sum_s n_s\frac{\breve{w}_{js}}{\breve{n}_s}\right)^2 \preceq \sum_j 1 = N$. Therefore, eq. (A.18) holds by Markov inequality and the dominated convergence theorem.

To show the second part of the proposition, decompose

$$\frac{1}{N}\sum_i A_i(\hat{X}_i - \hat{X}_{i,-}) = \frac{1}{N}\sum_{i,s}\frac{w_{is}\check{w}_{is}}{\check{n}_s}\psi_{is}A_i - \frac{1}{N}\sum_{i,j,s}\mathbb{I}\{j\neq i\}\frac{\check{w}_{is}}{\check{n}_s}\frac{w_{is}\check{w}_{js}}{\check{n}_{s,-i}}\psi_{js}A_i.$$

By arguments similar to those above, conditional on $(W,\check{W})$, the second term has mean zero and variance that converges to zero. By Assumption A.4(i) and Jensen's inequality, the mean of the first term is of the order $\frac{1}{N}\sum_{i,s}\frac{w_{is}\check{w}_{is}}{\check{n}_s}$. Consequently, provided that $(\ddot{\hat{X}}'Y_2/N)^2$ converges to a strictly positive limit, we have

$$\tilde{\alpha}-\alpha = \frac{O_p(\frac{1}{N}\sum_{i,s}\frac{w_{is}\check{w}_{is}}{\check{n}_s})}{\ddot{\hat{X}}'Y_2/N} = O_p\left(\frac{1}{N}\sum_{i,s}\frac{w_{is}\check{w}_{is}}{\check{n}_s}\right),$$

as required.

A.2.4 Proof of Proposition A.3

Since $Nr_N^{1/2}(\hat{\alpha}_- - \alpha) = r_N^{1/2}\ddot{\hat{X}}_-'Y_1(0)/\ddot{\hat{X}}_-'Y_2/N = r_N^{1/2}\ddot{\hat{X}}_-'Y_1(0)\cdot(\beta_{FS}N^{-1}\sum_{i,s}w_{is}^2\sigma_s^2)^{-1}(1+o_P(1))$, it suffices to show that

$$r_N^{1/2}\ddot{\hat{X}}_-'Y_1(0) = \mathcal{N}(0,\mathcal{V}_N+\mathcal{W}_N)+o_p(1).$$

By arguments as in the proof of Proposition 4,

$$\begin{aligned} r_N^{1/2}\ddot{\hat{X}}_-'Y_1(0) &= r_N^{1/2}(W\tilde{\mathcal{X}} - U\gamma + (\hat{X}_- - X))'(Z(\delta-\check{\delta})+\epsilon_\Delta)\\ &= r_N^{1/2}(W\tilde{\mathcal{X}})'\epsilon_\Delta + r_N^{1/2}(\hat{X}_- - X)'(Z(\delta-\check{\delta})+\epsilon_\Delta)+o_p(1)\\ &= r_N^{1/2}(W\tilde{\mathcal{X}}+(\hat{X}_- - X))'\epsilon_\Delta + o_p(1), \end{aligned}$$

where the last line follows since $(\hat{X}_- - X)'Z/N = o_p(1)$ by eq. (A.18). Let $C_{\Delta,i} = C_{iY(0)} - C_{iU}'\delta - \sum_s w_{is}\mathcal{Z}_s'\delta$ and $\eta_{\Delta,i} = \eta_{iY(0)} - \eta_{iU}'\delta$, so that $\epsilon_{\Delta,i} = Y_{i1}(0) - Z_i'\delta = \eta_{\Delta,i} + C_{\Delta,i}$. Then we can decompose

$$r_N^{1/2}(W\tilde{\mathcal{X}}+(\hat{X}_- - X))'\epsilon_\Delta = r_N^{1/2}\sum_{j=1}^{N+S}\mathcal{Y}_j,$$

where

$$\mathcal{Y}_j = \begin{cases} \sum_{i=1}^N\sum_{s=1}^S w_{is}\check{w}_{js}\frac{\mathbb{I}\{j\neq i\}\psi_{js}C_{\Delta,i}}{\check{n}_{s,-i}} + \sum_{i=1}^{j-1}\sum_{s=1}^S\left[\frac{w_{is}\check{w}_{js}\psi_{js}\eta_{\Delta,i}}{\check{n}_{s,-i}} + \frac{\check{w}_{is}w_{js}\eta_{\Delta,j}\psi_{is}}{\check{n}_{s,-j}}\right], & j=1,\dots,N,\\ \tilde{\mathcal{X}}_{j-N}\sum_i w_{i,j-N}\epsilon_{\Delta,i}, & j=N+1,\dots,N+S. \end{cases}$$

Let H denote the matrix with rows η_i', and define the σ-fields $\mathcal{G}_i = \sigma(W,\check{W},\mathcal{Z},C,\eta_1,\dots,,\eta_i,\psi_1,\dots,\psi_i)$, $i=1,\dots,N$, $\mathcal{G}_i = \sigma(W,\check{W},\mathcal{Z},C,H,\Psi,\mathcal{X}_1,\dots,\mathcal{X}_{j-N})$, $j=N+1,\dots,N+S$. Then, under Assumption A.4(v), $\mathcal{Y}_j$ is a martingale difference array with respect to the filtration $\mathcal{G}_j$. Since by the arguments in the proof of Proposition 4, $r_N^{1+\nu/4}\sum_{j=N+1}^{N+S}E_{W,\check{W}}[\mathcal{Y}_j^{2+\nu/2}]\to 0$, and $r_N\sum_{j=N+1}^{N+S}E[\mathcal{Y}_j^2\mid\mathcal{G}_{j-1}] - \mathcal{V}_N = o_p(1)$, it suffices to show that $r_N^2\sum_{j=1}^N E_{W,\check{W}}[\mathcal{Y}_j^4]\to 0$, and $r_N\sum_{j=1}^N E[\mathcal{Y}_j^2\mid\mathcal{G}_{j-1}] - \mathcal{W}_N = o_p(1)$. The result then follows by a martingale central limit theorem.

Since $\check{n}_s/\check{n}_{s,-i}$ is bounded, and $\sum_s w_{js} \leq 1$, and since $\sum_{s=1}^S \frac{n_s \check{w}_{js}}{\check{n}_s}$ is bounded by Assumption A.4(iv), we have the bound

$$r_N^2 \sum_{j=1}^N E_{W,\check{W}} \left(\sum_{i=1}^{j-1} \sum_{s=1}^S w_{is} \check{w}_{js} \frac{\psi_{js}\eta_{\Delta,i}}{\check{n}_{s,-i}} \right)^4 \preceq r_N^2 \sum_j \left(\sum_{i=1}^{j-1} \sum_{s=1}^S \frac{w_{is}\check{w}_{js}}{\check{n}_s} \right)^4 \leq r_N^2 \sum_{j=1}^N \left(\sum_{s=1}^S \frac{n_s \check{w}_{js}}{\check{n}_s} \right)^4 \leq r_N^2 N,$$

which converges to zero by Assumption A.4(vii). By an analogous argument, the conditional expectation of $r_N^2 \sum_{j=1}^N \left(\sum_{i=1}^N \sum_{s=1}^S w_{is} \check{w}_{js} \frac{\mathbb{I}\{j \neq i\} \psi_{js} C_{\Delta,i}}{\check{n}_{s,-i}} \right)^4$ and of $r_N^2 \sum_{j=1}^N \left(\sum_{i=1}^{j-1} \sum_{s=1}^S \check{w}_{is} w_{js} \frac{\eta_{\Delta,j}\psi_{is}}{\check{n}_{s,-j}} \right)^4$ is also bounded by $r_N^2 N$, so that $r_N^2 \sum_{j=1}^N E_{W,\check{W}}[\mathcal{Y}_j^4] \to 0$ by C_r-inequality.

It remains to show that the conditional variance $r_N \sum_{j=1}^N E[\mathcal{Y}_j^2 \mid \mathcal{G}_{j-1}]$ converges. Expanding the expectation yields

$$\begin{aligned}
r_N \sum_{j=1}^N E[\mathcal{Y}_j^2 \mid \mathcal{G}_{j-1}] = {} & 2r_N \sum_{i,j,s,t} \sum_{i'}^{j-1} \frac{\mathbb{I}\{j \neq i\} E_{\mathcal{G}_0}[\check{w}_{js}\check{w}_{jt}\psi_{js}\psi_{jt}]}{\check{n}_{s,-i}} \frac{w_{is}w_{i't}C_{\Delta,i}\eta_{\Delta,i'}}{\check{n}_{t,-i'}} \\
& + 2r_N \sum_{i,j,s,t} \sum_{i'=1}^{j-1} \frac{\mathbb{I}\{j \neq i\} E_{\mathcal{G}_0}[\check{w}_{js} w_{jt}\psi_{js}\eta_{\Delta,j}]}{\check{n}_{s,-i}} \frac{w_{is}\check{w}_{i't}C_{\Delta,i}\psi_{i't}}{\check{n}_{t,-j}} \\
& + r_N \sum_{j,s,t} \sum_{i=1}^{j-1} \sum_{i'=1}^{j-1} \mathbb{I}\{i \neq i'\} \frac{E_{\mathcal{G}_0}[\check{w}_{js}\check{w}_{jt}\psi_{js}\psi_{jt}]}{\check{n}_{s,-i}} \frac{w_{i't}w_{is}\eta_{\Delta,i}\eta_{\Delta,i'}}{\check{n}_{t,-i'}} \\
& + 2r_N \sum_{j,s,t} \sum_{i=1}^{j-1} \sum_{i'=1}^{j-1} \mathbb{I}\{i \neq i'\} \frac{E_{\mathcal{G}_0}[\check{w}_{js} w_{jt}\psi_{js}\eta_{\Delta,j}]}{\check{n}_{s,-i}} \frac{w_{is}\check{w}_{i't}\eta_{\Delta,i}\psi_{i't}}{\check{n}_{t,-j}} \\
& r_N \sum_{j,s,t} \sum_{i=1}^{j-1} \sum_{i'=1}^{j-1} \mathbb{I}\{i \neq i'\} \frac{w_{jt}w_{js}E_{\mathcal{G}_0}[\eta_{\Delta,j}^2]}{\check{n}_{s,-j}} \frac{\check{w}_{i't}\check{w}_{is}\psi_{i't}\psi_{is}}{\check{n}_{t,-j}} \\
& + r_N \sum_{j,s,t} \sum_{i=1}^{j-1} \frac{w_{jt}w_{js}E_{\mathcal{G}_0}[\eta_{\Delta,j}\eta_{\Delta,j}]}{\check{n}_{s,-j}} \frac{\check{w}_{is}\check{w}_{it}\psi_{is}\psi_{it}}{\check{n}_{t,-j}} + 2r_N \sum_{j,s,t} \sum_{i=1}^{j-1} \frac{E_{\mathcal{G}_0}[\check{w}_{js}w_{jt}\psi_{js}\eta_{\Delta,j}]}{\check{n}_{s,-i}} \frac{w_{is}\check{w}_{it}\eta_{\Delta,i}\psi_{it}}{\check{n}_{t,-j}} \\
& + r_N \sum_{j,s,t} \sum_{i=1}^{j-1} \frac{E_{\mathcal{G}_0}[\check{w}_{js}\check{w}_{jt}\psi_{js}\psi_{jt}]}{\check{n}_{s,-i}} \frac{w_{is}w_{it}\eta_{\Delta,i}^2}{\check{n}_{t,-i}} + r_N \sum_{j=1}^N E_{\mathcal{G}_0} \left(\sum_{i=1}^N \sum_{s=1}^S \frac{\mathbb{I}\{j \neq i\} w_{is}\check{w}_{js}\psi_{js}C_{\Delta,i}}{\check{n}_{s,-i}} \right)^2.
\end{aligned}$$

Conditional on $(W, \check{W})$, the first five terms are mean zero. The variance of the first term is bounded by a constant times

$$r_N^2 \sum_{i'} \left(\sum_{i,j,s,t} \frac{w_{is}w_{i't}\check{w}_{js}\check{w}_{jt}}{\check{n}_s \check{n}_t} \right)^2 = r_N^2 \sum_{i'} \left(\sum_{j,t} \frac{w_{i't}\check{w}_{jt}}{\check{n}_t} \sum_s \frac{n_s \check{w}_{js}}{\check{n}_s} \right)^2 \preceq r_N^2 N.$$

Similarly, the variance of the second, third, fourth, and fifth term can be shown to be bounded by a constant times $r_N^2 N$. Next, the expectation conditional on $(W, \check{W})$ of the absolute value of the sixth term is bounded by a constant times

$$r_N \sum_{i,j} \left(\sum_s \frac{\check{w}_{is}w_{js}}{\check{n}_s} \right) \left(\sum_t \frac{w_{jt}\check{w}_{it}}{\check{n}_t} \right) \leq r_N \sum_i \max_{i'} \sum_j \left(\sum_s \frac{\check{w}_{is}w_{js}}{\check{n}_s} \right) \left(\sum_t \frac{w_{jt}\check{w}_{i't}}{\check{n}_t} \right)$$

$$= r_N \sum_i \max_{i'} \sum_j \left(\sum_s \frac{\check{w}_{is} w_{js}}{\check{n}_s} \right) \left(\sum_t \frac{w_{jt} \check{w}_{i't}}{\check{n}_t} \right)$$

Consequently, by Markov inequality,

$$r_N \sum_{j=1}^N E[\mathcal{Y}_j^2 \mid \mathcal{G}_{j-1}] =$$
$$r_N \sum_{j,s,t} \sum_{i=1}^{j-1} \frac{w_{jt} w_{js} E_{\mathcal{G}_0}[\eta_{\Delta,j} \eta_{\Delta,j}]}{\check{n}_{s,-j}} \frac{\check{w}_{is} \check{w}_{it} \psi_{is} \psi_{it}}{\check{n}_{t,-j}} + 2 r_N \sum_{j,s,t} \sum_{i=1}^{j-1} \frac{E_{\mathcal{G}_0}[\check{w}_{js} w_{jt} \psi_{js} \eta_{\Delta,j}]}{\check{n}_{s,-i}} \frac{w_{is} \check{w}_{it} \eta_{\Delta,i} \psi_{it}}{\check{n}_{t,-j}}$$
$$+ r_N \sum_{j,s,t} \sum_{i=1}^{j-1} \frac{E_{\mathcal{G}_0}[\check{w}_{js} \check{w}_{jt} \psi_{js} \psi_{jt}]}{\check{n}_{s,-i}} \frac{w_{is} w_{it} \eta_{\Delta,i}^2}{\check{n}_{t,-i}} + r_N \sum_{j=1}^N E_{\mathcal{G}_0} \left(\sum_{i=1}^N \sum_{s=1}^S \frac{\mathbb{I}\{j \neq i\} w_{is} \check{w}_{js} \psi_{js} C_{\Delta,i}}{\check{n}_{s,-i}} \right)^2 + o_p(1). \quad \text{(A.19)}$$

Similarly, expanding the expression for $\mathcal{W}_N$ yields

$$\mathcal{W}_N = \frac{1}{r_N} \sum_{i,i',j,s,t} \mathbb{I}\{j \neq i, i'\} \mathbb{I}\{i \neq i'\} \frac{\check{w}_{js} \check{w}_{jt} \psi_{jt} \psi_{js}}{\check{n}_{s,-i}} \frac{w_{is} w_{i't} \eta_{\Delta,i} \eta_{\Delta,i'}}{\check{n}_{t,-i'}}$$
$$+ \frac{2}{r_N} \sum_{i,i',j,s,t} \mathbb{I}\{j \neq i, i'\} \frac{\check{w}_{js} \check{w}_{jt} \psi_{jt} \psi_{js}}{\check{n}_{s,-i}} \frac{w_{is} w_{i't} C_{\Delta,i} \eta_{\Delta,i'}}{\check{n}_{t,-i'}}$$
$$+ \frac{1}{r_N} \sum_{i,j,s,t} \mathbb{I}\{i \neq j\} \frac{w_{is} \check{w}_{js} \psi_{it} C_{\Delta,i}}{\check{n}_{s,-i}} \frac{w_{jt} \check{w}_{it} \psi_{js} \eta_{\Delta,j}}{\check{n}_{t,-j}} + \frac{1}{r_N} \sum_{i,j,s,t} \mathbb{I}\{i \neq j\} \frac{w_{is} \check{w}_{js} \psi_{it} \eta_{\Delta,i}}{\check{n}_{s,-i}} \frac{w_{jt} \check{w}_{it} \psi_{js} C_{\Delta,j}}{\check{n}_{t,-j}}$$
$$+ \frac{1}{r_N} \sum_{i,j,s,t} \mathbb{I}\{i \neq j\} \frac{w_{is} \check{w}_{js} \psi_{it} C_{\Delta,i}}{\check{n}_{s,-i}} \frac{w_{jt} \check{w}_{it} \psi_{js} C_{\Delta,j}}{\check{n}_{t,-j}}$$
$$+ \frac{1}{r_N} \sum_{i,j,s,t} \mathbb{I}\{j \neq i\} \frac{\check{w}_{js} \psi_{js} \check{w}_{jt} \psi_{jt}}{\check{n}_{s,-i}} \frac{w_{is} w_{it} \eta_{\Delta,i}^2}{\check{n}_{t,-i}} + \frac{2}{r_N} \sum_{i,j,s,t} \mathbb{I}\{i < j\} \frac{w_{is} \check{w}_{js} \psi_{it} \eta_{\Delta,i}}{\check{n}_{s,-i}} \frac{w_{jt} \check{w}_{it} \psi_{js} \eta_{\Delta,j}}{\check{n}_{t,-j}}$$
$$+ \frac{1}{r_N} \sum_j \left(\sum_{i,s} \mathbb{I}\{i \neq j\} \frac{w_{is} \check{w}_{js} \psi_{js} C_{\Delta,i}}{\check{n}_{s,-i}} \right)^2.$$

Conditional on $(W, \check{W})$, the first five terms are mean zero. The variance of the first term is bounded by a constant times

$$\frac{1}{r_N^2} \sum_{i'} \left(\sum_{i,j,s,t} \frac{\check{w}_{js} \check{w}_{jt}}{\check{n}_s} \frac{w_{is} w_{i't}}{\check{n}_t} \right)^2 + \frac{1}{r_N^2} \sum_{i'} \left(\sum_{i,j,s,t} \frac{\check{w}_{js} \check{w}_{jt}}{\check{n}_s} \frac{w_{is} w_{i't}}{\check{n}_t} \right) \left(\sum_{i_2,j_2,s_2,t_2} \frac{\check{w}_{j_2 s_2} \check{w}_{j_2 t_2}}{\check{n}_{s_2}} \frac{w_{i' s_2} w_{i_2 t_2}}{\check{n}_{t_2}} \right)$$
$$+ \frac{1}{r_N^2} \sum_{i'} \left(\sum_{i,j,s,t} \frac{\check{w}_{js} \check{w}_{jt}}{\check{n}_s} \frac{w_{is} w_{i't}}{\check{n}_t} \right) \left(\sum_{i_2,j_2,s_2,t_2} \frac{\check{w}_{i' s_2} \check{w}_{i' t_2}}{\check{n}_{s_2}} \frac{w_{i s_2} w_{j_2 t_2}}{\check{n}_{t_2}} \right) \preceq \frac{N}{r_N^2}.$$

Similarly, the variance of the second, third, fourth and fifth term can also be shown to be bounded by a constant times $N r_N^2$. Therefore by Markov inequality, in view of eq. (A.19),

$$r_N \sum_{j=1}^N E[\mathcal{Y}_j^2 \mid \mathcal{G}_{j-1}] - \mathcal{W}_N = r_N \sum_{j,s,t} \sum_{i=1}^{j-1} \frac{w_{jt} w_{js} E_{\mathcal{G}_0}([\eta_{\Delta,j}^2] - \eta_{\Delta,j}^2)}{\check{n}_{s,-j}} \frac{\check{w}_{is} \check{w}_{it} \psi_{is} \psi_{it}}{\check{n}_{t,-j}}$$

$$+ 2r_N \sum_{j,s,t} \sum_{i=1}^{j-1} \frac{\check{w}_{js} w_{jt} (E_{\mathcal{G}_0}[\psi_{js}\eta_{\Delta,j}] - \psi_{js}\eta_{\Delta,j})}{\check{n}_{s,-i}} \frac{w_{is}\check{w}_{it}\eta_{\Delta,i}\psi_{it}}{\check{n}_{t,-j}}$$

$$+ r_N \sum_{j,s,t} \sum_{i=1}^{j-1} \frac{\check{w}_{js}\check{w}_{jt} (E_{\mathcal{G}_0}[\psi_{js}\psi_{jt}] - \psi_{js}\psi_{jt})}{\check{n}_{s,-i}} \frac{w_{is}w_{it}\eta_{\Delta,i}^2}{\check{n}_{t,-i}}$$

$$+ r_N \sum_{j=1}^{N} \sum_{i,i',s,t} \mathbb{I}\{j \neq i, i'\} \frac{w_{is}C_{\Delta,i}w_{i't}C_{\Delta,i'}}{\check{n}_{s,-i}} \frac{\check{w}_{jt}\check{w}_{js} (E_{\mathcal{G}_0}[\psi_{js}\psi_{jt}] - \psi_{js}\psi_{jt})}{\check{n}_{t,-i'}} + o_p(1).$$

All terms in this expression have mean zero conditional on W, and the variance of each term can be shown to be bounded by a constant times $r_N N$, so that $r_N \sum_{j=1}^{N} E[\mathcal{Y}_j^2 \mid \mathcal{G}_{j-1}] - \mathcal{W}_N = o_p(1)$ as required.

Appendix B Stylized economic model: baseline microfoundation

Appendices B.1 and B.2 provide a microfoundation for the stylized economic model presented in Section 3.1. In Appendix B.3, we use this microfoundation to derive expressions analogous to those in eqs. (8) and (9) in Section 3.2. In Appendix B.4, we exploit again our microfoundation and outline a set of restrictions on the model fundamentals such our main identification restriction, Assumption 1(ii) in Section 4.1, holds.

B.1 Environment

We consider a model with multiple sectors $s = 1, \ldots, S$ and multiple regions $i, j = 1, \ldots, N$. Regions are partitioned into countries indexed by $c = 1, \ldots, C$, and we denote the set of regions located in a country c by N_c. Region i has a population of M_i individuals who cannot move across regions. Each individual belongs to a different group, $g = 1, \ldots, G$. The share of group g in the population of region i is n_{ig}.

Production. Each sector s in region i has a representative firm that produces a differentiated good using only local labor. For simplicity, we assume that workers of different groups are perfect substitutes in production. The quantity Q_{is} produced by sector s in region i is produced using labor with productivity A_{is}; i.e.

$$Q_{is} = A_{is} L_{is}, \tag{B.1}$$

where L_{is} denotes the number of workers (irrespective of their group) employed by the representative firm in this sector-region pair. Regions thus differ in terms of their sector-specific productivity A_{is}.

Preferences for consumption goods. Every individual has identical nested preferences over the sector- and region-specific differentiated goods. Specifically, we assume that individuals have Cobb-Douglas preferences over sectoral composite goods,

$$C_j = \prod_{s=1}^{S} (C_{js})^{\gamma_s}, \tag{B.2}$$

where C_j is the utility level of a worker located in region j that obtains utility C_{js} from consuming goods in sector s, and C_{js} is a CES aggregator of the sector s goods produced in different regions:

$$C_{js} = \left[\sum_{i=1}^{N} \left(c_{ijs} \right)^{\frac{\sigma_s - 1}{\sigma_s}} \right]^{\frac{\sigma_s}{\sigma_s - 1}}, \qquad \sigma_s \in (1, \infty), \tag{B.3}$$

where c_{ijs} denotes the consumption in region j of the sector s good produced in region i. This preference structure has been previously used in Armington (1969), Anderson (1979) and multiple papers since (e.g. Anderson and van Wincoop, 2003; Arkolakis, Costinot and Rodríguez-Clare, 2012).

Preferences for sectors and non-employment. Individuals of every group g have the choice of being employed in one of the sectors $s = 1, \ldots, S$ of the economy or opting for non-employment, which we index as $s = 0$. Conditional on being employed, all workers of group g have identical homogeneous preferences over their sector of employment, but workers differ in their preferences for non-employment. Specifically, conditional on obtaining utility C_j from the consumption of goods, the utility of a worker ι of group g living in region j is

$$U(\iota \mid C_j) = \begin{cases} u(\iota) C_j & \text{if employed in any sector } s = 1, \ldots, S, \\ C_j & \text{if not employed } (s = 0). \end{cases} \tag{B.4}$$

We assume that each individual ι belonging to group g and living in a region located in country c independently draws $u(\iota)$ from a Pareto distribution with scale parameter v_{cg} and shape parameter ϕ, so that the cumulative distribution function of $u(\iota)$ is given by

$$F_{ig}^{u}(u) = 1 - \left(\frac{u}{v_{cg}} \right)^{-\phi}, \qquad u \geq v_{cg}, \qquad \phi > 1. \tag{B.5}$$

If a worker living in region j chooses to be employed, she will earn wage ω_j. In equilibrium, wages are equalized across sectors and groups because (i) firms are indifferent between workers of different groups, (ii) workers are indifferent about the sector of employment, and (iii) workers are freely mobile across sectors. If a worker chooses to not be employed, she receives a benefit b_j. We denote the total number of employed workers of group g in region j by L_{jg}, the total employment in region j as $L_j = \sum_{g=1}^{G} L_{jg}$, and the employment rate in j as $E_j \equiv L_j / M_j$. [3]

Market structure. Goods and labor markets are perfectly competitive.

Trade costs. We assume that there are no trade costs, which implies that the equilibrium price of the good produced in a region is the same in every other region; i.e. $p_{ijs} = p_{is}$ for $j = 1, \ldots, N$. Thus,

[3] We assume that benefits are paid by a national government that imposes a flat tax χ_c on all income earned in country c. The budget constraint of the government is thus $\sum_{j \in N_c} \{ \chi_c (\omega_j E_j + b_j (1 - E_j)) M_j \} = \sum_{j \in N_c} \{ b_j (1 - E_j) M_j \}$. Alternatively, we could think of the option $s = 0$ as home production and assume that workers that opt for home production in region j obtain b_j units of the final good, which they consume. This alternative model is isomorphic to that in the main text.

for every sector s there is a composite sectoral good that has identical price P_s in all regions; i.e.

$$(P_s)^{1-\sigma_s} = \sum_{s=1}^{S} (p_{is})^{1-\sigma_s}, \tag{B.6}$$

and the final good's price is $P = \prod_{s=1}^{S} (P_s)^{\gamma_s}$.

B.2 Equilibrium

We now characterize the equilibrium wage ω_j and total employment L_j of all regions $j = 1, \ldots, N$.

Consumption. We first solve the expenditure minimization problem of an individual residing in region j. Given the sector-level utility in eq. (B.3) and the condition that $p_{ijs} = p_{is}$ for $j = 1, \ldots, N$, all regions j have identical spending shares x_{is} on goods from region i, given by

$$x_{is} = \left(\frac{p_{is}}{P_s}\right)^{1-\sigma_s}. \tag{B.7}$$

Labor supply. Every worker maximizes the utility function in eq. (B.4) in order to decide whether to be employed. Consequently, conditional on the wage ω_i and the non-employment benefit b_i, the total employment of individuals of group g in region i is $L_{ig} = n_{ig} M_i \Pr\left[u_i(\iota)\omega_i > b_i\right]$. It therefore follows from eq. (B.5) that $L_i = \sum_{g=1}^{G} L_{jg}$ is

$$L_i = \omega_i^{\phi} v_i \tag{B.8}$$

such that

$$v_i = \nu_i \sum_{g=1}^{G} n_{ig} \nu_{cg} \tag{B.9}$$

with $\nu_i \equiv M_i b_i^{-\phi}$, and $\nu_{cg} \equiv v_{cg}^{\phi}$.

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

$$p_{is} = \frac{\omega_i}{A_{is}}. \tag{B.10}$$

Goods market clearing. Given that labor is the only factor of production and firms earn no profits, the income of all individuals living in region i is $W_i \equiv \sum_s \omega_i L_{is}$, and world income is $W \equiv \sum_i W_i$. We normalize world income to one, $W = 1$. Given preferences in eq. (B.2), all individuals spend a share γ_s of their income on sector s, so that world demand for the differentiated good s produced in region i is $x_{is}\gamma_s$. Goods market clearing requires world demand for good s produced in region i to equal total revenue of the representative firm operating in sector s in region i, $\omega_i L_{is}$. Thus, using the expression in eq. (B.7), we obtain

$$L_{is} = (\omega_i)^{-\sigma_s} (A_{is}P_s)^{\sigma_s - 1} \gamma_s. \tag{B.11}$$

Note that this labor demand equation is analogous to that in eq. (2) of Section 3, with the region- and sector-specific demand shifter D_{is} defined as $D_{is} = (A_{is}P_s)^{\sigma_s-1} \gamma_s$.

If, without loss of generality, we split the region- and sector-specific productivity A_{is} into a country and sector-specific component A_{cs} and a residual $\tilde{A}_{is}$,

$$A_{is} = A_{cs}\tilde{A}_{is}. \tag{B.12}$$

Labor market clearing. Given the sector- and region-specific labor demand in eq. (B.11), total labor demand in region i is

$$L_i = \sum_{s=1}^{S} (\omega_i)^{-\sigma_s} (A_{is}P_s)^{\sigma_s - 1} \gamma_s. \tag{B.13}$$

Labor market clearing requires labor supply in eq. (B.8) to equal labor demand in eq. (B.13):

$$v_i(\omega_i)^{\phi} = \sum_{s=1}^{S} (\omega_i)^{-\sigma_s} (A_{is}P_s)^{\sigma_s - 1} \gamma_s. \tag{B.14}$$

Equilibrium. Given technology parameters $\{A_{cs}\}_{c=1,s=1}^{C,S}$ and $\{\tilde{A}_{is}\}_{i=1,s=1}^{N,S}$, preference parameters $\{(\sigma_s, \gamma_s)\}_{s=1}^{S}$, labor supply parameters ϕ, $\{\nu_i\}_{i=1}^{N}$, $\{n_{ig}\}_{i=1,g=1}^{N,G}$ and $\{\nu_{cg}\}_{c=1,g=1}^{C,G}$, and normalizing world income to equal 1, $W = 1$, we can use eqs. (B.6), (B.9), (B.10), (B.12) and (B.14) to solve for the equilibrium wage in every world region, $\{\omega_i\}_{i=1}^{N}$, the equilibrium price of every sector-region specific good $\{p_{is}\}_{i=1,s=1}^{N,S}$, and the sectoral price indices $\{P_s\}_{s=1}^{S}$. Given these equilibrium wages and sectoral price indices, we can use eq. (B.13) to solve for the equilibrium level of employment in every region, $\{L_i\}_{i=1}^{N}$.

B.3 Labor market impact of sectoral shocks: equilibrium relationships

We assume that, in every period, the model described in Appendices B.1 and B.2 characterizes the labor market equilibrium in every region $i = 1, \ldots, N$. Across periods, we assume that the parameters $\{\sigma_s\}_{s=1}^{S}$, $\{n_{ig}\}_{i=1,g=1}^{N,G}$, and ϕ are fixed, and that all changes in the labor market outcomes $\{\omega_i, L_i\}_{i=1}^{N}$ are generated by changes in technology $\{A_{cs}\}_{c=1,s=1}^{C,S}$ and $\{\tilde{A}_{is}\}_{i=1,s=1}^{N,S}$, sectoral preferences $\{\gamma_s\}_{s=1}^{S}$, and labor supply parameters $\{\nu_i\}_{i=1}^{N}$ and $\{\nu_{cg}\}_{c=1,g=1}^{C,G}$.

We focus here on understanding how changes in these exogenous parameters affect the labor market equilibrium in all regions located in a given country c; i.e. all regions belonging to the set N_c.

In our model, the sectoral prices mediate the impact of all foreign technology and labor supply shocks on the labor market equilibrium of every region in country c; i.e. the changes in $\{(\omega_i, L_i)\}_{i \in N_c}$ depend on the changes in $\{\tilde{A}_{is}\}_{s=1, i \notin N_c}^{S}$, $\{\nu_i\}_{i \notin N_c}$, and $\{\nu_{c'g}\}_{g=1,c' \neq c}^{G}$ only through changes in $\{P_s\}_{s=1}^{S}$. Therefore, we can write the changes in wages and employment in every region i of the population of interest N_c as a function of the changes in the sectoral prices, and the changes in the productivity and labor supply shocks in region i.

Isomorphism. As in Section 3.2, we use $\hat{z} = \log(z^t/z^0)$ to denote log-changes in any given variable z between some initial period $t = 0$ and any other period t. Up to a first-order approximation around the initial equilibrium, eqs. (B.13) and (B.14) imply that

$$\hat{L}_i = \sum_{s=1}^{S} l_{is}^0 \left[\theta_{is}\hat{P}_s + \lambda_i((\sigma_s - 1)\hat{A}_{cs} + \hat{\gamma}_s) + \lambda_i((\sigma_s - 1)\hat{\tilde{A}}_{is})\right] + (1-\lambda_i)(\sum_{g=1}^{G} \tilde{w}_{ig}\hat{v}_{cg} + \hat{v}_i), \tag{B.15}$$

with $l_{is}^0 \equiv L_{is}^0/L_i^0$, $\tilde{w}_{ig} \equiv L_{ig}^0/L_i^0$, $\theta_{is} = (\sigma_s - 1)\lambda_i$ and $\lambda_i \equiv \phi\left[\phi + \sum_s l_{is}^0\sigma_s\right]^{-1}$. Combining eqs. (B.8), (B.9) and (B.15), we can similarly obtain

$$\hat{\omega}_i = \sum_{s=1}^{S} l_{is}^0 \left[\theta_{is}\hat{P}_s + \lambda_i((\sigma_s - 1)\hat{A}_{cs} + \hat{\gamma}_s) + \lambda_i((\sigma_s - 1)\hat{\tilde{A}}_{is})\right] - \phi^{-1}\lambda_i(\sum_{g=1}^{G} \tilde{w}_{ig}\hat{v}_{cg} + \hat{v}_i). \tag{B.16}$$

Given our emphasis on understanding the changes in labor market outcomes for regions located in the same country, all regions in the population of interest will share the same value of A_{cs} for every sector s, and the same value of $\hat{v}_{cg}$ for every labor group g; thus, we can simplify the notation by writing $\hat{A}_{cs} = \hat{A}_s$ and $\hat{v}_{cg} = \hat{v}_g$ for all s and g, respectively. Given this notational simplification and the following equivalences

$$\chi_s = P_s, \tag{B.17}$$
$$\mu_s = (A_s)^{\sigma_s - 1}\gamma_s, \tag{B.18}$$
$$\eta_{is} = (\tilde{A}_{is})^{\sigma_s - 1}, \tag{B.19}$$

we can easily see that the expressions in eqs. (B.15) and (B.16) are identical to those in eqs. (8) and (9) in Section 3.2, respectively. Consequently, the environment described in Appendices B.1 and B.2 does indeed provide a microfoundation for the equilibrium relationships in eqs. (8) and (9).

B.4 Identification of labor market impact of sectoral prices

As the mapping in eq. (B.17) illustrates, we may think of the changes in sectoral prices $\{\hat{P}_s\}_{s=1}^S$ as our sectoral shocks of interest. Given data on changes in a labor market outcome (e.g. changes in the employment rate $\hat{L}_i$) for all units of a population of interest formed by all regions of a particular country c, and data on the changes in sectoral prices $\{\hat{P}_s\}_{s=1}^S$, Assumption 1(ii) in Section 4.1 indicates that identifying the coefficient in front of a shift-share term that aggregates these sectoral price changes requires that these are as good as randomly allocated.

In the context of the equilibrium relationship in eq. (B.15), the sectoral price changes $\{\hat{P}_s\}_{s=1}^S$ will satisfy Assumption 1(ii) if they are mean independent of: country c-specific sectoral productivity changes $\{\hat{A}_{cs}\}_{s=1}^S$; country c-specific labor-group supply shocks $\{\hat{v}_{cg}\}_{g=1}^G$; region and sector-specific productivity shocks, for all sectors and all regions in country c, $\{\hat{\tilde{A}}_{is}\}_{s=1,i\in N_c}^S$; region-specific labor supply shocks, for all regions in country c, $\{\hat{v}_i\}_{i\in N_c}$. This mean independence restriction will hold if the following two conditions are satisfied.

First, country c is "small"; i.e. all labor demand and labor supply shocks in country c have no impact on the changes in sectoral prices $\{\hat{P}_s\}_{s=1}^S$.

Second, labor demand and labor supply shocks affecting any region i in the country or population of interest c are mean independent of any labor demand and labor supply shock affecting any other region of the world economy that is "large" (i.e. any other region whose labor demand and supply shocks have an impact on the changes in sectoral prices).

In summary, if the vector of shifters of interest $\{\mathcal{X}_s\}_{s=1}^S$ corresponds to the sectoral price changes $\{\hat{P}_s\}_{s=1}^S$, the researcher is interested on the impact of these shifters on a collection of "small" regions, and labor market shocks in these "small" regions are independent of the corresponding shocks in any "large" region, then the identification condition in Assumption 1(ii) is satisfied.

B.4.1 Impact of labor demand and supply shocks on sector-specific price indices

In general equilibrium, the price change in every sector s, $\hat{P}_s$, depends on the shocks A_{cs}, $\hat{\hat{A}}_{is}$, $\hat{\gamma}_s$, $\hat{v}_i$, and ν_{cg} of all sectors, labor groups, and regions in the world economy. Specifically, the change in the sector-specific price index is

$$\hat{P}_s = -\sum_{s'} \alpha_{ss'} \sum_{j=1}^{N} x_{js'}^0 (\hat{A}_{js'} + \tilde{\lambda}_j \hat{v}_j - \tilde{\lambda}_j \sum_k l_{jk}^0 [\hat{\gamma}_k + (\sigma_k - 1)\hat{A}_{jk}]), \tag{B.20}$$

where $\tilde{\lambda}_j \equiv \left[\phi + \sum_s l_{is}^0 \sigma_s\right]^{-1}$, $\{\alpha_{ss'}\}_{s=1,s'=1}^{S,S}$ are positive constants, and x_{js}^0 is the share of the world production in sector s that corresponds to region j in the initial equilibrium; i.e. $x_{js}^0 \equiv X_{js}^0 / \sum_{i=1}^N X_{is}^0$. Imposing that all regions in a country c verify that $x_{js}^0 \approx 0$ for all $j \in N_c$ and for $s = 1, \ldots, S$, we can rewrite the change in the sector-specific price index as

$$\hat{P}_s = -\sum_{s'} \alpha_{ss'} \sum_{j \notin N_c} x_{js'}^0 (\hat{A}_{js'} + \tilde{\lambda}_j \hat{v}_j - \tilde{\lambda}_j \sum_k l_{jk}^0 [\hat{\gamma}_k + (\sigma_k - 1)\hat{A}_{jk}]). \tag{B.21}$$

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 $\{\hat{A}_{cs}\}_{s=1}^S$, nor $\{\hat{v}_{cg}\}_{g=1}^G$, nor $\{\hat{\hat{A}}_{is}\}_{s=1,i\in N_c}^S$, nor $\{\hat{v}_i\}_{i\in N_c}$.

Proof of eq. (B.20). Equations (B.7) and (B.14) imply that

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

where $\tilde{\alpha}_{sk} \equiv \sum_j x_{js}^0 l_{jk}^0 \tilde{\lambda}_j (\sigma_k - 1)$. 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, we can rewrite the equation above in matrix form as

$$(I - \bar{\boldsymbol{\alpha}})\, \hat{\boldsymbol{P}} = \hat{\boldsymbol{\eta}},$$

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. (B.20), it is sufficient to show that $(I - \tilde{\boldsymbol{\alpha}})$ is a nonsingular m-matrix and, therefore, it has a positive inverse matrix. To establish this result, notice first that $\tilde{\alpha}_{sk} \in (0,1)$ for every s and k; to show this, it is sufficient to show that, for every j, k, and s, it holds that $0 < x^0_{js} < 1$ and

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

where the last two inequalities arise from $\sigma_k > 1$ and $\phi > 0$.

Finally, to show that $(I - \tilde{\boldsymbol{\alpha}})$ is nonsingular, it is sufficient to establish that it is diagonal dominant:

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

Appendix C Stylized economic model: Extensions

In Appendices C.1 and C.2, we provide alternative microfoundations for the equilibrium relationship in eq. (8). Finally, in Appendix C.3, we incorporate migration into the baseline microfoundation described in Appendix B.

C.1 Sector-specific factors of production

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

C.1.1 Environment

The only difference with respect to the setting described in Appendix B.1 is that the production function in eq. (B.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.1.2 Equilibrium

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

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

Producer's problem. Conditional on the region-i equilibrium wage ω_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

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

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

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

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

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

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

$$X_{is} = \frac{1}{1-\alpha_s}\omega_i L_{is} = \frac{\bar{K}_{is}}{\alpha_s}(p_{is}A_{is}\tilde{\alpha}_s)^{\frac{1}{\alpha_s}}(\omega_i)^{1-\frac{1}{\alpha_s}}. \tag{C.2}$$

Goods market clearing. Applying the same normalization as in Appendix B.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. (B.7) as a function of the equilibrium prices p_{is}. Equating $x_{is}\gamma_s$ and eq. (C.2), we can solve for the equilibrium value of p_{is} as a function of the sector-s price index P_s:

$$p_{is} = \left[\frac{\bar{K}_{is}}{\alpha_s}(A_{is}\tilde{\alpha}_s)^{\frac{1}{\alpha_s}}(\omega_i)^{1-\frac{1}{\alpha_s}}\frac{(P_s)^{1-\sigma_s}}{\gamma_s}\right]^{-\theta_{is}\eta_{is}}, \tag{C.3}$$

where $\delta_s \equiv (1+\alpha_s(\sigma_s-1))^{-1} \in (0,1)$. Additionally, combining eqs. (C.1) and (C.3), we obtain an expression for labor demand in sector-s region-i as a function of the equilibrium wage ω_i, the sector-s price P_s and other exogenous determinants:

$$L_{is} = \kappa_{is}\gamma_s^{\delta_s}(A_{is}P_s)^{(\sigma_s-1)\delta_s}(\omega_i)^{-\sigma_s\delta_s}, \tag{C.4}$$

where $\kappa_{is} \equiv (1-\alpha_s)(\bar{K}_{is}\tilde{\alpha}_s^{\frac{1}{\alpha_s}}/\alpha_s)^{1-\delta_s}$. Note that this labor demand equation is analogous to that in eq. (2), with the region- and sector-specific demand shifter D_{is} defined as

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

and with the labor demand elasticity now defined as $\sigma_s\delta_s$. Note that the labor demand elasticity in eq. (2) is identical to that in eq. (C.4) in the specific case in which $\delta_s = 1$, which will hold when

$\alpha_s = 0$. Without loss of generality, we split the region- and sector-specific productivity A_{is} according to eq. (B.12).

Labor market clearing. Given the sector- and region-specific labor demand in eq. (C.4), total labor demand in region i is

$$L_i = \sum_{s=1}^{S} \kappa_{is} \gamma_s^{\delta_s} \left(A_{is} P_s\right)^{(\sigma_s - 1)\delta_s} \left(\omega_i\right)^{-\sigma_s \delta_s}. \tag{C.5}$$

Labor market clearing requires labor supply in eq. (B.8) to equal labor demand in eq. (C.5):

$$v_i (\omega_i)^{\phi} = \sum_{s=1}^{S} \kappa_{is} \gamma_s^{\delta_{is}} \left(A_{is} P_s\right)^{(\sigma_s - 1)\delta_{is}} \left(\omega_i\right)^{-\sigma_s \delta_{is}}, \qquad j = 1, \ldots, N. \tag{C.6}$$

Equilibrium. Given the technology parameters $\{\alpha_s\}_{s=1}^{S}$, $\{A_{cs}\}_{c=1,s=1}^{C,S}$ and $\{\tilde{A}_{is}\}_{i=1,s=1}^{N,S}$, sector- and region-specific capital inputs $\{\bar{K}_{is}\}_{i=1,s=1}^{N,S}$, preference parameters $\{(\sigma_s, \gamma_s)\}_{s=1}^{S}$, labor supply parameters ϕ, $\{\nu_i\}_{i=1}^{N}$, $\{n_{ig}\}_{i=1,g=1}^{N,G}$ and $\{\nu_{cg}\}_{c=1,g=1}^{C,G}$, and normalizing world income to equal 1, $W = 1$, we can use eqs. (B.6), (B.9), (B.12), (C.3) and (C.6) to solve for the equilibrium wage in every world region, $\{\omega_i\}_{i=1}^{N}$, the equilibrium price of every sector-region specific good $\{p_{is}\}_{i=1,s=1}^{N,S}$, and the sectoral price indices $\{P_s\}_{s=1}^{S}$. Given these equilibrium wages and sectoral price indices, we can use eq. (C.5) to solve for the equilibrium level of employment in every region, $\{L_i\}_{i=1}^{N}$.

C.1.3 Labor market impact of sectoral shocks

We assume that, in every period, the model described in Appendices C.1.1 and C.1.2 characterizes the labor market equilibrium in every region $i = 1, \ldots, N$. Across periods, we assume that the parameters $\{(\sigma_s, \alpha_s)\}_{s=1}^{S}$, $\{n_{ig}\}_{i=1,g=1}^{N,G}$, and ϕ are fixed, and that all changes in the labor market outcomes $\{\omega_i, L_i\}_{i=1}^{N}$ are generated by changes in technology $\{A_{cs}\}_{c=1,s=1}^{C,S}$ and $\{\tilde{A}_{is}\}_{i=1,s=1}^{N,S}$, sectoral preferences $\{\gamma_s\}_{s=1}^{S}$, and labor supply parameters $\{\nu_i\}_{i=1}^{N}$ and $\{\nu_{cg}\}_{c=1,g=1}^{C,G}$. We focus here on understanding how changes in these exogenous parameters affect the labor market equilibrium in all regions located in a given country c; i.e. all regions belonging to the set N_c.

Isomorphism. Following steps analogous to those in Appendix B.3, we can show that eqs. (C.5) and (C.6) imply that

$$\begin{aligned}\hat{L}_i = &\sum_{s=1}^{S} l_{is}^{0} \left[\theta_{is} \hat{P}_s + \lambda_i ((\sigma_s - 1)\delta_s \hat{A}_{cs} + \delta_s \hat{\gamma}_s) + \lambda_i ((\sigma_s - 1)\delta_s \hat{\tilde{A}}_{is} + \hat{\kappa}_{is})\right] \\ &+ (1 - \lambda_i) \left(\sum_{g=1}^{G} \tilde{\omega}_{ig} \hat{\nu}_{cg} + \hat{\nu}_i\right),\end{aligned} \tag{C.7}$$

with $\theta_{is} = (\sigma_s - 1)\delta_s \lambda_i$ and $\lambda_i \equiv \phi(\phi + \sum_s l_{is}^{0} \sigma_s \delta_s)^{-1}$. As in Appendix B.3, given our emphasis on understanding the changes in labor market outcomes for regions located in the same country, all regions in the population of interest will share the same value of A_{cs} for every sector s, and the same value of $\hat{\nu}_{cg}$ for every labor group g; thus, we can simplify the notation by writing $\hat{A}_{cs} = \hat{A}_s$

and $\hat{v}_{cg} = \hat{v}_g$ for all s and g, respectively. Given this notational simplification and the following equivalences

$$\chi_s = P_s, \tag{C.8}$$

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

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

we can easily see that the expression in eq. (C.7) is identical to that in eq. (8) in Section 3.2. Consequently, the environment described in Appendices C.1.1 and C.1.2 does indeed provide a microfoundation for the equilibrium relationship in eq. (8).

C.2 Sector-specific preferences

We extend the model described in Appendix B to allow workers to have idiosyncratic preferences for being employed in the different $s = 1, \ldots, S$ sectors and for being non-employed $s = 0$. In order to maintain the analysis simple, we assume here that there is a single worker group $G = 1$.

C.2.1 Environment

The only difference with respect to the setting described in Appendix B.1 is that the utility function in eqs. (B.4) and (B.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 ι living in region i is

$$U_{is} = u_s(\iota)C_i, \tag{C.11}$$

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

$$F_u(u) = e^{-v_{is}u^{-\phi}}, \qquad \phi > 1. \tag{C.12}$$

This modeling of workers' sorting patterns across sectors is similar to that in Galle, Rodríguez-Clare and Yi (2018) and Burstein, Morales and Vogel (2019). See Adão (2016) for a framework that relaxes the distributional assumption in eq. (C.12). 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.2.2 Equilibrium

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

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

$$L_{is} = M_i \frac{v_{is}(\omega_{is})^{\phi}}{\Phi_i} \quad \text{with} \quad \Phi_i \equiv v_{i0} b_i^{\phi} + \sum_{s=1}^{S} v_{is}(\omega_{is})^{\phi}, \tag{C.13}$$

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

$$L_{i0} = M_i \frac{v_{i0}(b_i)^{\phi}}{\Phi_i}. \tag{C.14}$$

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

$$p_{is} = \frac{\omega_{is}}{A_{is}}. \tag{C.15}$$

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 B.2.

Labor market clearing. Combining the region- and sector-specific labor supply in eq. (C.13) with the region- and sector-specific labor demand in eq. (B.11), and imposing the normalization $W = 1$, the labor market clearing condition in every sector $s = 1, \dots, S$ and region $i = 1, \dots, N$ is

$$M_i \frac{v_{is}(\omega_{is})^{\phi}}{\Phi_i} = (\omega_{is})^{-\sigma_s} (A_{is} P_s)^{\sigma_s - 1} \gamma_s. \tag{C.16}$$

Equilibrium. Given productivity parameters $\{A_{cs}\}_{c=1,s=1}^{C,S}$ and $\{\tilde{A}_{is}\}_{i=1,s=1}^{N,S}$, preference parameters $\{\sigma_s, \gamma_s\}_{s=1}^{S}$, labor supply parameters ϕ and $\{v_{is}\}_{i=1,s=0}^{N,S}$, and normalizing world income to equal 1, $W = 1$, we can use eqs. (B.6), (B.12), (C.15) and (C.16) to solve for the equilibrium wage in every sector and region, $\{\omega_{is}\}_{i=1,s=1}^{N,S}$, the equilibrium price of every sector- and region-specific good $\{p_{is}\}_{i=1,s=1}^{N,S}$, and the sectoral price indices $\{P_s\}_{s=1}^{S}$. Given these equilibrium wages and sectoral price indices, we can use eqs. (C.13) and (C.14) to solve for the equilibrium level of employment in every sector and region, $\{L_{is}\}_{i=1,s=0}^{N,S}$.

C.2.3 Labor market impact of sectoral shocks

We assume that, in every period, the model described in Appendices C.2.1 and C.2.2 characterizes the labor market equilibrium in every region $i = 1, \dots, N$. Across periods, we assume that the parameters $\{\sigma_s\}_{s=1}^{S}$, and ϕ are fixed, and that all changes in the labor market outcomes $\{\omega_i, L_i\}_{i=1}^{N}$ are generated by changes in technology $\{A_{cs}\}_{c=1,s=1}^{C,S}$ and $\{\tilde{A}_{is}\}_{i=1,s=1}^{N,S}$, sectoral preferences $\{\gamma_s\}_{s=1}^{S}$, and labor supply parameters $\{v_{is}\}_{i=1,s=1}^{N,S}$. We focus here on understanding how changes in these exogenous parameters affect the labor market equilibrium in all regions located in a given country c; i.e. all regions belonging to the set N_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

$$\begin{aligned}\hat{L}_i &= -\frac{l_{i0}^0}{1 - l_{i0}^0}\hat{L}_{i0} \\ &= \frac{l_{i0}^0}{1 - l_{i0}^0}(\hat{\Phi}_i - \phi\hat{b}_i - \hat{v}_{i0}) \\ &= \frac{l_{i0}^0}{1 - l_{i0}^0}(\sum_{s=0}^{S} l_{is}^0 \hat{v}_{is} + \phi l_{i0}^0 \hat{b}_i + \phi \sum_{s=1}^{S} l_{is}^0 \hat{\omega}_{is} - \phi\hat{b}_i - \hat{v}_{i0}). \end{aligned} \tag{C.17}$$

From eq. (C.16), we can express the changes in wages in every sector and every region of country c as

$$\hat{\omega}_{is} = (\phi + \sigma_s)^{-1}\left(\hat{\Phi}_i + \hat{\gamma}_s + (\sigma_s - 1)(\hat{A}_{is} + \hat{P}_s) - \hat{v}_{is}\right). \tag{C.18}$$

Combining eqs. (C.17) and (C.18), we can re-express the change in total employment in region i as

$$\hat{L}_i = \sum_{s=1}^{S} l_{is}^0[\theta_{is}\hat{P}_s + \lambda_i(\phi + \sigma_s)^{-1}((\sigma_s - 1)\hat{A}_{cs} + \hat{\gamma}_s) + \lambda_i(\phi + \sigma_s)^{-1}(\sigma_s - 1)\hat{\tilde{A}}_{is}] + \hat{\nu}_i, \tag{C.19}$$

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

As in Appendix B.3, given our emphasis on understanding the changes in labor market outcomes for regions located in the same country, all regions in the population of interest will share the same value of A_{cs} for every sector s; thus, we can simplify the notation by writing $\hat{A}_{cs} = \hat{A}_s$ for all s. Given this notational simplification, the following equivalences

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

and the adjustment of the expression for λ_i and $\hat{\nu}_i$, the expression in eq. (C.19) is identical to that in eq. (8) in Section 3.2. Consequently, the environment described in Appendices C.2.1 and C.2.2 does indeed provide a microfoundation for the equilibrium relationship in eq. (8).

C.3 Allowing for regional migration

We extend here the baseline environment described in Appendix B.1 to allow for mobility of individuals across regions within a single country c. As in Appendix C.2, to maintain the analysis simple, we focus on the special case with a single worker group, $G = 1$.

C.3.1 Environment

We still assume that the number of individuals living in each country c is fixed and equal to M_c. The only difference with respect to the setting described in Appendix B.1 is that the mass of individuals living in a region i, M_i, is no longer fixed. We assume that, before the realization of the shock $u(\iota)$ in eq. (B.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 ι of residing in region i is

$$U(\iota) = \tilde{u}_i(\iota)\left(\bar{U}_i(\omega_i/P, b_i/P) - 1\right) \tag{C.20}$$

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

$$\tilde{u}_i(\iota) \sim F_{\tilde{u}}(\tilde{u}) = e^{-\tilde{u}^{-\tilde{\phi}}}, \qquad \tilde{\phi} > 0. \tag{C.21}$$

A similar modeling of labor mobility has been previously imposed, among others, in Allen and Arkolakis (2016), Redding (2016), Allen, Arkolakis and Takahashi (2018), and Fajgelbaum et al. (2019), among others. See Redding and Rossi-Hansberg (2017) for additional references.

C.3.2 Equilibrium

Consumption. The consumer's problem is identical to that in Appendix B.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 $i \in N$. Defining $v_i \equiv (\nu_i/b_i)^{\phi}$ as in eq. (B.8), the assumption that $b_i = P$ for all $i \in N$ 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

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

Given the distributional assumption in eq. (C.21), it holds that

$$M_i = \frac{v_i(\omega_i)^{\phi_m}}{\Phi_c} M_c \ \text{ such that } \ \Phi_c = \sum_{j \in N_c} v_j(\omega_j)^{\phi_m} \ \text{ and } \ \phi_m \equiv \tilde{\phi}\phi. \tag{C.22}$$

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

$$L_i = \frac{v_i(\omega_i)^{\phi_m}}{\sum_{j \in N_c} v_j(\omega_j)^{\phi_m}} M_c v_i(\omega_i)^{\phi}. \tag{C.23}$$

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

Labor market clearing. Combining the region- and sector-specific labor supply in eq. (C.23) with the aggregate labor demand in eq. (B.13), and imposing the normalization $W = 1$, the labor market clearing condition in every region $i \in N_c$ is

$$\frac{v_i(\omega_i)^{\phi_m}}{\sum_{j \in N_c} v_j(\omega_j)^{\phi_m}} M_c v_i(\omega_i)^{\phi} = \sum_s (\omega_i)^{-\sigma_s} \left(A_{is} P_s\right)^{\sigma_s - 1} \gamma_s, \tag{C.24}$$

or, equivalently,

$$(\Phi_c)^{-1} M_c v_i(\omega_i)^{\phi + \phi_m} = \sum_s (\omega_i)^{-\sigma_s} \left(A_{is} P_s\right)^{\sigma_s - 1} \gamma_s, \tag{C.25}$$

for every region i in every country c.

Equilibrium. Given productivity parameters $\{A_{cs}\}_{c=1,s=1}^{C,S}$ and $\{\tilde{A}_{is}\}_{i=1,s=1}^{N,S}$, preference parameters $\{\sigma_s, \gamma_s\}_{s=1}^{S}$, labor supply parameters, ϕ, ϕ_m, and $\{v_i\}_{i=1}^{N}$, and normalizing world income to equal 1, $W = 1$, we can use eqs. (B.6), (B.10), (B.12) and (C.25) to solve for the equilibrium wage in every region, $\{\omega_i\}_{j=1}^{N}$, the equilibrium price of every sector- and region-specific good $\{p_{is}\}_{i=1,s=1}^{N,S}$, and the sectoral price indices $\{P_s\}_{s=1}^{S}$. Given these equilibrium wages and sectoral price indices, we can use eq. (C.23) to solve for the equilibrium level of employment in every region, $\{L_i\}_{i=1}^{N}$.

C.3.3 Labor market impact of sectoral shocks

We assume that, in every period, the model described in Appendices C.3.1 and C.3.2 characterizes the labor market equilibrium in every region $i = 1, \ldots, N$. Across periods, we assume that the parameters $\{\sigma_s\}_{s=1}^S$, ϕ and ϕ_m are fixed, and that all changes in the labor market outcomes $\{\omega_i, L_i\}_{i=1}^N$ are generated by changes in technology $\{A_{cs}\}_{c=1,s=1}^{C,S}$ and $\{\tilde{A}_{is}\}_{i=1,s=1}^{N,S}$, sectoral preferences $\{\gamma_s\}_{s=1}^S$, and labor supply parameters $\{v_i\}_{i=1}^N$. We focus here on understanding how changes in these exogenous parameters affect the labor market equilibrium in all regions located in a given country c; i.e. all regions belonging to the set N_c.

Isomorphism. According to eq. (C.23), the change in employment in any region i in country c is

$$\hat{L}_i = 2\hat{v}_i + (\phi + \phi_m)\hat{\omega}_i - \hat{\Phi}_c. \tag{C.26}$$

Assuming that $\{M_c\}_{c=1}^C$, $\{\sigma_s\}_{s=1}^S$, and (ϕ, ϕ_m) are fixed and totally differentiating eq. (C.24) 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

$$\hat{\omega}_i = \lambda_i \hat{\Phi}_c + \lambda_i \sum_{s=1}^S l_{is}^0 \left[\hat{\gamma}_s + (\sigma_s - 1)(\hat{A}_{is} + \hat{P}_s)\right] - \lambda_i \hat{v}_i, \tag{C.27}$$

where $\lambda_i \equiv (\phi + \phi_m + \sum_s l_{is}^0 \sigma_s)^{-1}$. Using the expression in eq. (C.22), we can also express

$$\hat{\Phi}_c = \sum_{i \in N_c} m_i^0 \left(\phi_m \hat{\omega}_i + \hat{v}_i\right), \tag{C.28}$$

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 $M_c^0 \equiv \sum_{i \in N_c} M_i^0$.

Combining eqs. (C.26) and (C.27), we can express the change in total employment in region i as

$$\begin{aligned}\hat{L}_i = {} & [(\phi + \phi_m)\lambda_i - 1]\hat{\Phi}_c + \sum_{s=1}^S l_{is}^0 [\theta_{is}\hat{P}_s + \lambda_i(\phi + \phi_m)((\sigma_s - 1)\hat{A}_{cs} + \hat{\gamma}_s) + \lambda_i(\phi + \phi_m)(\sigma_s - 1)\hat{\tilde{A}}_{is}] \\ & + [2 - (\phi + \phi_m)\lambda_i]\hat{v}_i\end{aligned} \tag{C.29}$$

where $\theta_{is} = (\sigma_s - 1)(\phi + \phi_m)\lambda_i$. As in Appendix B.3, given our emphasis on understanding the changes in labor market outcomes for regions located in the same country, all regions in the population of interest will share the same value of A_{cs} for every sector s; thus, we can simplify the notation used in eq. (C.29) by writing $\hat{A}_{cs} = \hat{A}_s$ for all s. Given this notational simplification, if it were to be the case that $\hat{\Phi}_c = 0$, the expression in eq. (C.29) would be analogous to that in eq. (8) under the following equivalences

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

and the necessary adjustment of the expression for λ_i and $\hat{v}_i$. However, the term $\hat{\Phi}_c$ will generally not be zero and, as indicated in eq. (C.28), 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 θ_{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 Additional placebo exercises

This section presents additional placebo exercises that complement the results in Sections 2 and 6. Appendix D.1 reports the empirical distribution of the estimated coefficients and standard errors of the baseline placebo exercise in Sections 2 and 6. Appendix D.2 investigates the importance of controlling for the size of the residual sector in shift-share specifications. In Appendix D.3, we present results illustrating the impact of confounding sector-level shocks on different estimators of the coefficient on the shift-share covariate of interest. Appendix D.4 investigates the consequences of serial correlation in panel data applications of shift-share specifications. Appendix D.5 analyzes the consequences of misspecification of our baseline linearly additive potential outcome framework. Appendix D.6 reports results investigating the performance of inference procedures in the presence of unobserved shift-share components whose shares differ from those of the shift-share variable of interest. Appendix D.7 studies the consequences of treatment heterogeneity. In Appendix D.8, we provide additional results for the placebo exercises described in Sections 2 and 6. <

D.1 Placebo exercise: empirical distributions

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 $\text{var}(\mathcal{X}_s^m) = 5$, 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.

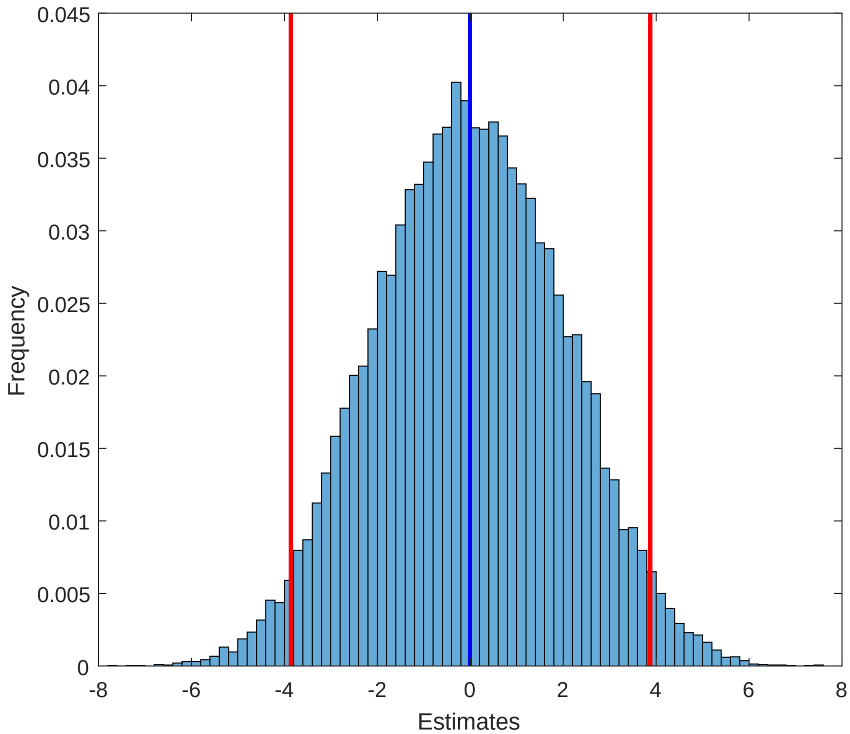


Figure D.1: Empirical distribution of estimated coefficients in the placebo exercise.

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.

D.2 Controlling for size of the residual sector

In the placebo simulations described in Tables 1 to 3, we have drawn the shifters from a mean-zero distribution. In Table D.1, we depart from the mean-zero assumption.

As discussed in Section 4.2, controlling for the region-specific sum of shares, $\sum_{s=1}^{S} w_{is}$, is important if the shifters have non-zero mean. In our placebo setting, this is equivalent to controlling for the CZ-specific share of employment in the non-manufacturing sector in 1990, $1 - \sum_{s=1}^{S} w_{is}$; we refer to this control here as the "residual sector control". Panel A in Table D.1 shows that, when the shifters are mean zero, the mean of $\hat{\beta}$ is not affected by whether we include the residual sector control. However, including the residual sector control attenuates the overrejection problem of traditional inference methods. Intuitively, this control soaks part of the correlation in residuals that traditional inference methods do not take into account. Panel B in Table D.1 shows that, if the shifter mean is non-zero, the OLS estimate of β in eq. (1) suffers from substantial bias when the residual sector control is not included in the regression; this bias disappears once it is included. Specifically, in Panel B, $\mathcal{X}_s^m \sim \mathcal{N}(1, 5)$, and the estimator in the first row of this panel suffers from negative bias because the positive mean of the shifters creates a positive correlation between the shift-share regressor of interest and the control $\sum_{s=1}^{S} w_{is}$, which captures the larger secular decline in the employment rate in regions initially specialized in manufacturing production.

Table D.1: Controlling for the size of the residual sector in each CZ

	Estimate		Median eff. s.e.				Rejection rate			
	Mean (1)	Std. dev (2)	Robust (3)	Cluster (4)	AKM (5)	AKM0 (6)	Robust (7)	Cluster (8)	AKM (9)	AKM0 (10)
Panel A: Shifters with mean equal to zero										
No controls	0.01	1.99	0.74	0.92	1.91	2.23	48.0%	37.7%	7.6%	4.5%
Control: $1-\sum_{s=1}^{S} w_{is}$	-0.02	1.43	0.74	0.84	1.31	1.52	33.6%	28.4%	11.2%	4.7%
Panel B: Shifters with mean different from zero										
No controls	-4.67	1.28	0.71	0.94	1.48	1.66	99.1%	97.8%	85.4%	87.6%
Control: $1-\sum_{s=1}^{S} w_{is}$	0.00	1.43	0.74	0.84	1.31	1.52	33.3%	27.8%	11.1%	4.6%

Notes: All estimates in this table use the change in the share of the working-age population employed in each CZ as the outcome variable Y_i in eq. (1). This table indicates the median and standard deviation of the OLS estimates of β in eq. (1) across the placebo samples (columns (1) and (2)), the median effective standard error estimates (columns (3) to (6)), and the percentage of placebo samples for which we reject the null hypothesis $H_0\colon \beta = 0$ using a 5% significance level test (columns (7) to (10)). *Robust* is the Eicker-Huber-White standard error; *Cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5; *AKM0* is the confidence interval in Remark 6. For each inference procedure, the median effective standard error is equal to the median length of the corresponding 95% confidence interval divided by 2×1.96. Results are based on 30,000 placebo samples. In Panel A, $(\mathcal{X}_1^m, \dots, \mathcal{X}_S^m)$ is drawn i.i.d. from a normal distribution with zero mean and variance equal to five in each placebo sample. In Panel B, $(\mathcal{X}_1^m, \dots, \mathcal{X}_S^m)$ is drawn i.i.d. from a normal distribution with mean equal to one and variance equal to five in each placebo sample. For each of the two panels, the first row presents results in which no control is accounted for in the estimating equation; the second row presents results in which we control for the size of the residual sector.

D.3 Confounding sector-level shocks: omitted variable bias and solutions

In this appendix, we illustrate the consequences of violations of the assumption that the shifters $(\mathcal{X}_1, \dots, \mathcal{X}_S)$ are independent of other shocks affecting the outcome variable of interest. Specifically, we show the impact that the presence of latent sector-specific shocks correlated with the shifters $(\mathcal{X}_1, \dots, \mathcal{X}_S)$ has on the properties of the OLS estimator of the coefficient on the shift-share regressor of interest $X_i \equiv \sum_{s=1}^{S} \mathcal{X}_s$. We also illustrate the properties of two solutions to this problem: (i) the inclusion of regional controls as a proxy for sector-level unobserved shocks (see Section 4.2), and (ii) the use of a shift-share instrumental variable constructed as a weighted average of exogenous sector-level shocks (see Section 4.3).

To generate the shifters of interest, the confounding sectoral shocks, and the exogenous sector-specific shocks that will enter the instrumental variable, 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 shifter of interest, $\mathcal{X}_s^b$ is the unobserved confounding shock, $\mathcal{X}_s^c$ is an exogenous shifter. 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 $\tilde{\sigma} = 12$.

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 δ is a parameter controlling the impact of the unobserved sectoral shocks $(\mathcal{X}_1^b, \ldots, \mathcal{X}_S^b)$ on the simulated outcome Y_i^m. Thus, the parameter δ captures the magnitude of the impact that the unobserved shocks $(\mathcal{X}_1^b, \ldots, \mathcal{X}_S^b)$ have on the outcome variable. We simulate 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 $(\mathcal{X}_1^b, \ldots, \mathcal{X}_S^b)$,

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

The parameter σ_u thus modulates the measurement error in X_i^b as a proxy for the impact of the unobserved shocks $(\mathcal{X}_1^b, \ldots, \mathcal{X}_S^b)$ on CZ i. We simulate data both with $\sigma_u = 0$ and with $\sigma_u = 6$.[4]

For each set of parameters (δ, σ_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. (13). Second, we consider the OLS estimator of the coefficient on X_i^a 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. (23). 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. (36). For each of these three estimators, we implement four inference procedures: *Robust*, *Cluster*, *AKM* and *AKM0*. All results are reported in Table D.2.

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 this case, the OLS estimator of the coefficient on X_i^a 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 latent confounding sectoral shock (i.e. if $\sigma_u = 0$ as in Panel B). In contrast, the IV estimate always yields an average estimated coefficient close to zero.

As illustrated in Table D.2, traditional inference methods always under-predict the dispersion in the estimated coefficient. As discussed in eq. (21), this is driven by the correlation between the unobservable residuals of regions with similar sector employment compositions. The *AKM* and *AKM0* inference procedures impose no assumption on the cross-regional pattern of correlation in the regression residuals and yield, on average, estimates of the median length of the 95% confidence interval that are equal or higher to the standard deviation of the empirical distribution of estimates. As a result, as Table D.2 reports, while traditional methods overreject the null $H_0 : \beta = 0$ in the context of both OLS and IV estimation procedures, our methods yield the correct test size for both estimators.

[4] 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 an element of $\mathcal{Z}_s$, u_i corresponds to U_i, and X_i^b to Z_i. The value of the parameter γ in eq. (26) is thus equal to $\tilde{\rho}$.

Table D.2: Magnitude of standard errors and rejection rates—Confounding effects

	Estimate		Median eff. s.e.				Reject. $H_0\colon \beta = 0$ at 5%			
	Mean (1)	Std. dev (2)	Robust (3)	Cluster (4)	AKM (5)	AKM0 (6)	Robust (7)	Cluster (8)	AKM (9)	AKM0 (10)
Panel A: No confounding effect ($\delta = 0$)										
OLS no controls	0.00	1.28	0.47	0.59	1.23	1.43	48.2%	37.6%	7.7%	4.5%
OLS with controls	0.00	1.80	0.67	0.83	1.72	1.97	47.6%	37.9%	7.9%	4.7%
2SLS	0.00	1.84	0.69	0.85	1.76	2.02	47.7%	37.7%	7.7%	4.6%
Panel B: Confounding effect ($\delta = 6$) and perfect regional control ($\sigma_u = 0$)										
OLS no controls	4.19	1.47	0.58	0.70	1.38	1.60	97.9%	96.8%	80.9%	72.2%
OLS with controls	−0.01	1.81	0.67	0.83	1.72	1.97	48.2%	38.3%	8.1%	4.6%
2SLS	−0.01	1.85	0.69	0.85	1.75	2.02	48.1%	38.3%	8.0%	4.7%
Panel C: Confounding effect ($\delta = 6$) and imperfect regional control ($\sigma_u = 2$)										
OLS no controls	4.20	1.47	0.58	0.70	1.37	1.60	97.9%	96.8%	81.4%	72.6%
OLS with controls	4.10	1.46	0.58	0.70	1.39	1.61	97.7%	96.3%	79.4%	71.3%
2SLS	−0.22	2.46	0.93	1.10	2.12	2.66	41.7%	34.0%	8.1%	4.6%

Notes: All estimates in this table use the change in the share of the working-age population employed in each CZ as the outcome variable Y_i in eq. (1). This table indicates the median and standard deviation of the OLS estimates of β in eq. (1) across the placebo samples (columns (1) and (2)), the median effective standard error estimates (columns (3) to (6)), and the percentage of placebo samples for which we reject the null hypothesis $H_0\colon \beta = 0$ using a 5% significance level test (columns (7) to (10)). The median effective standard error refers to the median length of the 95% confidence interval across the simulated datasets divided by 2×1.96 *Robust* is the Eicker-Huber-White standard error; *Cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5; *AKM0* is the confidence interval in Remark 6. All results are based on 30,000 simulation draws.

D.4 Panel data: serial correlation in residuals and shifters

In this appendix, we focus on panel data applications and perform several placebo exercises that illustrate the consequences of serial correlation in either the shifters $(\mathcal{X}_1, \ldots, \mathcal{X}_S)$ or the regression residuals on the properties of several standard error estimates. For each of our placebo exercises, we generate 30,000 placebo samples indexed by m. Each of them contains 722 regions, 397 sectors, and 2 periods: the first period corresponds to 1990–2000 changes, and the second period corresponds to 2000–2007 changes. As in the baseline placebo, each region corresponds to a U.S. Commuting Zone (CZ), and each sector corresponds to a 4-digit SIC manufacturing industry. We index each region by j and each sector by k. When implementing the *AKM* and *AKM0* in this context, we follow the approach in Section 5.2 by defining "generalized regions" as $i = (j, t)$, "generalized sectors" as $s = (k, t)$, and shares w_{is} as in eq. (41).

As in our baseline placebo, each simulated sample m has identical values of the shares $\{w_{is}\}_{i=1,s=1}^{N,S}$. Specifically, the shares in periods 1 and 2 correspond to employment shares in 1990 and 2000, respectively. Depending on the placebo exercise, the placebo samples may differ across simulated samples in terms of the outcomes $\{Y_i\}_{i=1}^{N}$. Finally, all placebo samples always differ in the shifters $(\mathcal{X}_1, \ldots, \mathcal{X}_S)$.

For each simulated sample m, we draw the random vector of shifters $(\mathcal{X}_1^m, \ldots, \mathcal{X}_S^m)$ from the joint distribution

$$(\mathcal{X}_1^m, \ldots, \mathcal{X}_S^m) \sim \mathcal{N}\left(0, \Sigma^2\right), \tag{D.1}$$

where Σ^2 is a $S \times S$ covariance matrix with $\Sigma^2_{sk} = (1-\rho^2)\sigma^2\,\mathbb{I}\{s = k\} + \rho^2\sigma^2\,\mathbb{I}\{c(s) = c(k)\}$ and, for every s, $c(s)$ indicates the "cluster" that the generalized sector s belongs to. We incorporate serial correlation in the sector-level shocks by defining clusters of generalized sectors associated with the same underlying sector in different periods. We follow the baseline placebo by setting $\sigma^2 = 5$. The value of ρ^2 controls the degree of correlation across shifters of different generalized sectors that correspond to the same underlying industrial sector at different points in time.

For each simulated sample m, we generate the outcome of region i in the placebo sample m as

$$Y_i^m = Y_i + \eta_i^m, \tag{D.2}$$

where Y_i denotes the change in the employment rate in the generalized region i. By changing the distribution from which the term η_i^m is drawn, we change the distribution of the regression residuals. We implement different placebo exercises in which $\{\eta_i^m\}_{i=1}^{N}$ is drawn from different distributions.

In some placebo exercises, we allow for serial correlation in η_i for every region i but impose that η_i is independent of η_j for any two different regions i and j; specifically,

$$(\eta_1^m, \ldots, \eta_{1444}^m) \sim \mathcal{N}(0, \Sigma^1), \tag{D.3}$$

where Σ^1 is a 1444×1444 covariance matrix with $\Sigma^1_{ii'} = (1-\rho^1)\sigma^1\,\mathbb{I}\{i = i'\} + \rho^1\sigma^1\,\mathbb{I}\{j(i) = j(i')\}$ and $j(i)$ is the region associated with the generalized observation i. We set $\sigma^1 = Var(Y_i)/2$ and generate different placebo samples for different values of ρ^1. The value of ρ^1 controls the degree of correlation across regression residuals of different generalized regions that correspond to the same geographic region at different points in time.

In some other placebo exercises, we assume that η_i^m has a shift-share structure with shares identical to those entering the shift-share component of interest. Specifically, we assume that

$$\eta_i^m = \sum_{s=1}^{S} w_{is} \mu_s^m \quad \text{such that} \quad (\mu_1^m, \dots, \mu_S^m) \sim \mathcal{N}\left(0, \Sigma^2\right), \tag{D.4}$$

where Σ^2 is identical to the variance matrix of the shifters $(\mathcal{X}_1^m, \dots, \mathcal{X}_S^m)$ introduced in eq. (D.1).

We start by evaluating the robustness of our results to the existence of serial correlation in regional outcomes or regression residuals. In Panel A of Table D.3, we implement a placebo exercise in which the shifters are drawn according to eq. (D.1) with $\rho^2 = 0$ (i.e. no serial correlation in sectoral shifters) and the outcome variables are drawn according to eqs. (D.2) and (D.3) with three different values of ρ^1 (i.e. different degree of serial correlation in the regression residuals). The rejection rates of all six inference procedures we consider (*Robust*, *Cluster*, *AKM* and *AKM0*, the last two both in a version that assumes that the shifters are independent, and in a version that allows them to be serially correlated) are robust to different degrees of serial correlation in the regression residuals. The reason is that, as illustrated in column (4) of Table D.3, the standard deviation of the estimator $\hat{\beta}$ is invariant to these patterns of serial correlation in the regression residuals.

In Panel B, we implement a placebo exercise in which the shifters are drawn according to eq. (D.1) with ρ^2 equal to either 0, 0.5 or 1 (i.e. different degrees of serial correlation in sectoral shifters) and the distribution of the simulated outcome variables is identical to their empirical distribution (i.e. $\eta_i^m = 0$ for every region i and placebo sample m). The results indicate that the larger the serial correlation in the sector-level shifters, the larger the rejection rates implied by the *Robust* and *Cluster* standard errors, as well as those implied by an implementation of the *AKM* and *AKM0* inference procedures that wrongly assumes that the shifters are independent across generalized sectors. Conversely, as illustrated in columns (15) and (16) in Panel B of Table D.3, the *AKM* and *AKM0* become very close to the nominal rejection rate of 5% once we cluster across generalized sectors that correspond to the same underlying sector at different points in time.

In Panel C, we depart from the setting described in Panel B in that we draw values of η_i^m according to the distribution described in eq. (D.4). The sector-level shifters entering the shift-share covariate of interest X_i^m and the term η_i^m are thus drawn from the same distribution. The results are very similar to those in Panel B.

Finally, in Panel D, we draw shifters $(\mathcal{X}_1^m, \dots, \mathcal{X}_S^m)$ that are not only serially correlated but also correlated across 4-digit industries belonging to the same 3-digit sector. Columns (11) to (14) show that, when ignored by the corresponding inference procedure, such correlation patterns in the shifters of interest lead to an overrejection problem, the severity of which depends on the correlation in the shifters. Columns (15) and (16) show that this overrejection problem disappears when we implement the *AKM* and *AKM0* inference procedures clustering across all generalized shifters that correspond to pairs of a 4-digit sectors and time period such that the 4-digit sector is associated to the same 3-digit industry.

Table D.3: Magnitude of standard errors and rejection rates: panel data with serially correlated shifters

Serial Correl.		Estimate		Median eff. s.e.						Rejection rate of $H_0: \beta = 0$ at 5%					
ρ^1	ρ^2	Mean	Std. dev	Robust	Cluster	AKM	AKM0	AKM (cluster)	AKM0 (cluster)	Robust	Cluster	AKM	AKM0	AKM (cluster)	AKM0 (cluster)
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)	(16)
Panel A: Correlation over time in a residual region-level component															
0	0	−0.01	1.12	0.55	0.67	1.06	1.15	1.05	1.17	34.3%	24.0%	6.7%	5.0%	7.0%	4.8%
0.5	0	−0.01	1.10	0.55	0.66	1.07	0.00	1.06	1.18	32.7%	24.6%	6.6%	5.0%	6.7%	4.8%
1	0	−0.01	1.09	0.55	0.66	1.05	0.00	1.05	1.16	32.7%	24.2%	5.9%	4.7%	6.5%	4.3%
Panel B: Correlation over time in shifter of interest															
0	0	−0.01	1.05	0.46	0.60	1.02	1.1	1.01	1.12	39.4%	27.0%	6.4%	4.6%	6.7%	4.5%
0	0.5	0.00	1.13	0.47	0.59	1.04	1.12	1.09	1.22	42.9%	31.6%	8.1%	6.1%	6.9%	4.7%
0	1	0.00	1.22	0.47	0.57	1.07	1.14	1.18	1.37	46.2%	36.5%	10.1%	7.9%	7.6%	4.5%
Panel C: Correlation over time in shifter of interest and in a residual shift-share component															
0	0	0.00	1.06	0.47	0.60	1.03	1.11	1.02	1.14	39.7%	27.5%	6.7%	4.8%	6.9%	4.7%
0	0.5	0.00	1.14	0.47	0.59	1.05	1.12	1.1	1.23	42.6%	31.5%	8.0%	6.2%	6.8%	4.5%
0	1	0.01	1.22	0.48	0.58	1.07	1.14	1.19	1.38	45.8%	36.1%	9.6%	7.5%	7.2%	4.2%
Panel D: Correlation over time and within 3-digit sectors in shifter of interest and in a residual shift-share component															
0	0	0.00	1.06	0.46	0.60	1.03	1.11	1.01	1.16	39.6%	27.1%	6.6%	4.9%	7.2%	4.7%
0	0.5	0.02	1.24	0.47	0.61	1.03	1.1	1.19	1.39	47.3%	34.6%	11.8%	9.8%	7.4%	4.5%
0	1	0.00	1.40	0.47	0.61	1.02	1.09	1.35	1.68	53.5%	41.7%	16.9%	14.5%	8.1%	4.2%

Notes: All estimates in this table use the change in the share of the working-age population employed in each CZ as the outcome variable Y_i in eq. (1). This table indicates the median and standard deviation of the OLS estimates of β in eq. (1) across the placebo samples (columns (1) and (2)), the median effective standard error estimates (columns (3) to (8)), and the percentage of placebo samples for which we reject the null hypothesis $H_0: \beta = 0$ using a 5% significance level test (columns (8) to (15)). The median effective standard error refers to the median length of the 95% confidence interval across the simulated datasets divided by 2×1.96 *Robust* is the Eicker-Huber-White standard error; *Cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5; *AKM0* is the confidence interval in Remark 6. In Panels A, B, and C, *AKM (cluster)* and *AKM0 (cluster)* assume that the shifter corresponding to each 4-digit SIC shifter is distributed independently of those corresponding to other 4-digit shifter, but allow for correlation over time in these 4-digit SIC shifters. In Panel D, *AKM (cluster)* and *AKM0 (cluster)* additionally allow for correlation across 4-digit SIC shifters that belong to the same 3-digit SIC sector. All results are based on 30,000 simulation draws.

D.5 Misspecification in linearly additive potential outcome framework

In this appendix section, we study the consequences of potential misspecification in the linearly additive potential outcome framework introduced in eq. (11) in Section 3.3. The extent to which this linearly additive framework is misspecified obviously depends on what the true potential outcome framework is. Inspired by the economic model described in Section 3, we outline a nonlinear potential outcome framework in Appendix D.5.1. In Appendix D.5.2, we determine theoretically the asymptotic properties of the OLS estimator of the coefficient on the shift-share component in the linearly additive potential outcome framework; specifically, we compare the treatment effects implied by the linear framework to those implied by the nonlinear one. In Appendix D.5.3, we present simulation results that quantify the bias in the estimation of treatment effects that arise from assuming a linearly additive potential outcome framework when the true one corresponds to the nonlinear framework described in Appendix D.5.1.

D.5.1 Nonlinear potential outcome framework

Consider the special case of the model of Section 3 in which the labor demand elasticity is identical in all sectors, i.e. $\sigma_s = \sigma$ for all s. We also set $\rho_s \equiv 1$ for all s. In this case, region i's labor demand in sector s is

$$L_{is} = (\omega_i)^{-\sigma} \left(\chi_s \mu_s \eta_{is}\right),$$

which implies that the total labor demand in region i is

$$L_i = (\omega_i)^{-\sigma} \sum_{s=1}^{S} \left(\chi_s \mu_s \eta_{is}\right).$$

By equalizing this expression with the expression for region i's labor supply in eq. (4) in Section 3, we obtain the following relationship between equilibrium wages in region i and both labor supply and labor demand shocks in i:

$$\log \omega_i = \check{\beta} \log \left(\sum_{s=1}^{S} \left(\chi_s \mu_s \eta_{is}\right) \right) - \check{\beta} \log \nu_i \tag{D.5}$$

where $\check{\beta} \equiv (\phi + \sigma)^{-1}$.

We focus here on determining the impact on log-changes in regional wages ω_i of log-changes in the sectoral demand shifters $\{\chi_s\}_{s=1}^{S}$; i.e. using the notation introduced in Section 3.2, we focus on characterizing the impact of $\{\hat{\chi}_s\}_{s=1}^{S}$ on $\hat{\omega}_i$. Because of the nonlinear nature of the relationship between labor demand shocks and wages in eq. (D.5), the impact of $\{\hat{\chi}_s\}_{s=1}^{S}$ on $\hat{\omega}_i$ depends on the changes in all other labor demand and supply shocks. For simplicity, we focus on the case in which all these other labor demand and supply shocks remain constant at their initial level. From eq. (D.5), the wages in the new and old equilibria are given by

$$\log \omega_i = \check{\beta} \log \left(\sum_{s=1}^{S} \chi_s^0 \mu_s^0 \eta_{is}^0 e^{\hat{\chi}_s} \right) - \check{\beta} \log \nu_i^0,$$

$$\log \omega_i^0 = \check{\beta} \log \left(\sum_{s=1}^{S} \chi_s^0 \mu_s^0 \eta_{is}^0 \right) - \check{\beta} \log \nu_i^0,$$

where we use a superscript zero to denote the value of the variables in the initial equilibrium and the absence of superscript denotes the value of the corresponding variable in the new equilibrium. By taking the difference between these two expressions,

$$\hat{\omega}_i = \check{\beta} \log \left(\sum_{s=1}^{S} \frac{\chi_s^0 \mu_s^0 \eta_{is}^0}{\sum_{k=1}^{S} \chi_k^0 \mu_k^0 \eta_{ik}^0} e^{\hat{\chi}_s} \right) = \check{\beta} \log \left(\sum_{s=1}^{S} \frac{L_{is}^0 \left(\omega_i^0\right)^{\sigma}}{\sum_{k=1}^{S} L_{ik}^0 \left(\omega_i^0\right)^{\sigma}} e^{\hat{\chi}_s} \right) = \check{\beta} \log \left(\sum_{s=1}^{S} \frac{L_{is}^0}{L_i^0} e^{\hat{\chi}_s} \right) \tag{D.6}$$

where the second equality follows from rearranging the terms in the labor demand expression in eq. (2) in Section 3 to obtain the equality $\chi_s^0 \mu_s^0 \eta_{is}^0 = L_{is}^0 \left(\omega_i^0\right)^{\sigma}$ for every region and sector, and the third equality follows from the fact that labor market clearing yields $L_i^0 = \sum_{s=1}^{S} L_{is}^0$.

Note that, by using data on the labor allocation across sectors for every region in some initial equilibrium (i.e. L_{is}^0 / L_i^0, for every i and s), the expression in eq. (D.6) allows to compute the effect of changes in the sector-specific labor demand shifters $\{\chi_s\}_{s=1}^{S}$ while calibrating the value of the overall labor demand shifter $(\chi_s^0)^{\rho_s} \mu_s^0 \eta_{is}^0$ at the initial equilibrium. Furthermore, the last expression in eq. (D.6) has the advantage that, conditional on values of $\{\hat{\chi}_s\}_{s=1}^{S}$ that are of interest, it depends exclusively on the parameter $\check{\beta}$; specifically, it does not depend on the labor demand parameter σ.

We can map the expression in eq. (D.6) to a nonlinear potential outcome framework by setting $\mathcal{X}_s = \hat{\chi}_s$, $Y_i = \hat{\omega}_i$, and $w_{is} = L_{is}^0 / L_i^0$ for every region and sector; i.e.

$$Y_i(\mathcal{X}_1, \ldots, \mathcal{X}_S) = \check{\beta} \log \left(\sum_{s=1}^{S} w_{is} e^{\mathcal{X}_s} \right). \tag{D.7}$$

According to the model in Section 3, this nonlinear potential outcome function yields the exact expression for the change in wages implied by a change in the labor demand shifters $\{\hat{\chi}_s\}_{s=1}^{S}$. Using eq. (D.7) we can also compute the treatment effect on region i of changing the shifters from $\{\mathcal{X}_s\}_{s=1}^{S}$ to $\{\mathcal{X}_s'\}_{s=1}^{S}$,

$$Y_i(\mathcal{X}_1, \ldots, \mathcal{X}_S) - Y_i(\mathcal{X}_1', \ldots, \mathcal{X}_S') = \check{\beta} \Big[\log \Big(\sum_{s=1}^{S} w_{is} e^{\mathcal{X}_s} \Big) - \log \Big(\sum_{s=1}^{S} w_{is} e^{\mathcal{X}_s} \Big) \Big]. \tag{D.8}$$

and the average treatment effect

$$\bar{Y}(\mathcal{X}_1, \ldots, \mathcal{X}_S) - \bar{Y}(\mathcal{X}_1', \ldots, \mathcal{X}_S') = \check{\beta} \frac{1}{N} \sum_{i=1}^{N} \Big[\log \Big(\sum_{s=1}^{S} w_{is} e^{\mathcal{X}_s} \Big) - \log \Big(\sum_{s=1}^{S} w_{is} e^{\mathcal{X}_s} \Big) \Big]. \tag{D.9}$$

The linearly additive function in eq. (11) in Section 3.3 provides a first-order approximation to the nonlinear function in eq. (D.8). In the next two subsections, we study the extent to which the linear expression in eq. (11) provides an accurate approximation to the nonlinear one in eq. (D.7). Specifically, we explore the extent to which the treatment effects in eqs. (D.8) and (D.9) are well approximated by those computed on the basis of the linear potential outcome framework introduced in Section 3.3.

The extent to which the linear approximation is accurate will depend on the distribution of $\{\mathcal{X}_s\}_{s=1}^S$. Throughout this section, we assume that $\{\mathcal{X}_s\}_{s=1}^S$ are independently drawn from a normal distribution,

$$\mathcal{X}_s \sim \mathcal{N}(0, \gamma^2), \tag{D.10}$$

so that $e^{\mathcal{X}_s}$ is log-normally distributed with $E[(e^{\mathcal{X}_s})^k] = e^{k^2\gamma^2/2}$.

D.5.2 Asymptotic properties of the shift-share linear specification

We consider here the asymptotic properties of the OLS estimator of β in the linear shift-share regression,

$$Y_i = \alpha + \beta \sum_{s=1}^{S} w_{is}\mathcal{X}_s + \epsilon_i, \tag{D.11}$$

when the distribution of $\mathcal{X}_s$ for every sector s is given by eq. (D.10), and the distribution of Y_i for every region i is given by the potential outcome framework in eq. (D.7). Since $\mathcal{X}_s$ has mean zero, the constant does not affect the regression estimand, which is given by

$$\beta = \frac{\sum_{i=1}^N E[X_i Y_i]}{\sum_{i=1}^N E[X_i^2]}, \tag{D.12}$$

where, under eqs. (D.7) and (D.10),

$$\sum_{i=1}^{N} E[X_i^2] = \gamma^2 \sum_{i=1}^{N}\sum_{s=1}^{S} w_{is}^2, \tag{D.13}$$

and

$$\sum_{i=1}^{N} E[X_i Y_i] = \check{\beta} E \sum_{i=1}^{N}\sum_{s=1}^{S} w_{is}\mathcal{X}_s \log\left(\sum_{k=1}^{S} w_{ik} e^{\mathcal{X}_k}\right). \tag{D.14}$$

Using eqs. (D.12) to (D.14), we can obtain an expression for β, the OLS estimand in a regression of Y_i on $\sum_{s=1}^S w_{is}\mathcal{X}_s$,

$$\beta = \check{\beta}\frac{\sum_{i=1}^N \sum_{s=1}^S w_{is} E[\gamma Z_s \log(\sum_{k=1}^S w_{ik} e^{\gamma Z_k})]}{\gamma^2 \sum_{i=1}^N \sum_{s=1}^S w_{is}^2}, \tag{D.15}$$

as well as for the difference between this value of β and the parameter from the nonlinear model in eq. (D.5):

$$\beta - \check{\beta} = \check{\beta}\frac{\sum_{i=1}^N \sum_{s=1}^S w_{is} E[\gamma Z_s \log(\sum_{k=1}^S w_{ik} e^{\gamma Z_k})]}{\gamma^2 \sum_{i=1}^N \sum_{s=1}^S w_{is}^2} - \check{\beta}, \tag{D.16}$$

where $\{Z_s\}_{s=1}^S$ are *i.i.d* standard normal. As it is clear from this expression, the difference between β and $\check{\beta}$ depends on the shares $\{w_{is}\}_{i=1,s=1}^{N,S}$, the value of the γ (i.e. the standard deviation of $\mathcal{X}_s$ for every s, according to eq. (D.10)), and the value of $\check{\beta}$ itself.

The expression analogous to that in eq. (D.8) when the linear potential outcome framework in eq. (11) in Section 3.3 is assumed is the following,

$$Y_i(\mathcal{X}_1,\ldots,\mathcal{X}_S) - Y_i(\mathcal{X}_1',\ldots,\mathcal{X}_S') = \beta\Big(\sum_{s=1}^{S} w_{is}(\mathcal{X}_s - \mathcal{X}_s')\Big). \tag{D.17}$$

and the expression analogous to that in eq. (D.8) is

$$\bar{Y}(\mathcal{X}_1,\ldots,\mathcal{X}_S) - \bar{Y}(\mathcal{X}_1',\ldots,\mathcal{X}_S') = \beta\frac{1}{N}\sum_{i=1}^{N}\Big(\sum_{s=1}^{S} w_{is}(\mathcal{X}_s - \mathcal{X}_s')\Big). \tag{D.18}$$

D.5.3 Simulation

In this section, we construct a simulation exercise to quantify: (a) the difference between β and $\check{\beta}$, using eq. (D.16) to compute such difference; (b) the correlation coefficient between the i-specific treatment effects in eq. (D.8) and those in eq. (D.17); and, (c) the difference between the average treatment effect in eq. (D.9) and that in eq. (D.18).

In all simulations, we calibrate the labor supply elasticity to equal 2, $\sigma = 2$, and the inverse labor supply elasticity to equal 0.5, $\phi = 0.5$, implying that $\check{\beta} = 0.4$. To remain close to our baseline placebo exercise, we calibrate the shares $\{w_{is}\}_{i=1,s=1}^{N,S}$ using 1990 data on sector-region employment shares for 722 US CZs and 396 4-digit manufacturing sectors. Concerning the value of the variance of the sectoral shifters, we present results for five different values of $var(\mathcal{X}_s) = \gamma^2$ varying between $\gamma^2 = 0.5$ and $\gamma^2 = 10$. For each value of γ, we then generate $30,000$ samples indexed by m such that $\{\mathcal{X}_s^m\}_{s=1}^{396}$ are independently drawn according to eq. (D.10) and $\{Y_i^m\}_{i=1}^{722}$ are constructed according to eq. (D.7).

For each placebo sample m, we compute the OLS estimator $\hat{\beta}$ of the parameter β defined in eq. (D.12), confidence intervals for β according to the *Robust*, *Cluster*, *AKM* and *AKM0* inference procedures, the true linear approximation to the i-specific treatment effect and to the average treatment effect (i.e. the expressions in eqs. (D.17) and (D.18) with $\check{\beta}$ instead of β), the estimated linear approximation to the i-specific treatment effect and to the average treatment effect (i.e. the expressions in eqs. (D.17) and (D.18) with $\hat{\beta}$ instead of β), and the true i-specific treatment effects and their average (i.e. the expressions in eqs. (D.8) and (D.9) with $\check{\beta} = 0.4$).

A comparison of columns (2) and (3) in Table D.4 illustrates that the average across the placebo samples generated under the same value of γ of the OLS estimates of β, $\bar{\hat{\beta}} \equiv (30,000)^{-1}\sum_{m=1}^{30,000}\hat{\beta}^m$ (reported in column (3)) is very close to the true value of the parameter β (reported in column (2)). We compute this true value of β using the expression in eq. (D.15) and Monte Carlo integration based on 50,000 draws of $(Z_1,\ldots,Z_S)$ from the distribution in eq. (D.10). Thus, as expected, the average value of $\hat{\beta}^m$ is very close to its theoretical value.

Columns (4)–(7) of Table D.4 report different measures of the average treatment effect across simulated samples. Specifically, we compute in these three columns, in this order, the average across the 30,000 placebo samples of: (a) the true linear approximation to the average treatment effect (i.e. the expression in eq. (D.18) with the value β set to the expression in eq. (D.15)); the estimated linear approximation to the average treatment effect (i.e. the expression in eq. (D.18) with $\hat{\beta}^m$ instead of β); and the true average treatment effect (i.e. the expression in eq. (D.9)). When the variance of sector-level shocks is low ($\gamma^2 = 0.1$), the first row in Table D.4 shows that all these three averages are very close to each other. As the variance of sector-level shocks grows, the remaining rows in Table D.4 show that the bias in the linear approximations to the average treatment effect grows. Columns (6) and (7) of Table D.4 illustrate that not only the linear approximation to the average treatment effects worsen as γ^2 increases, but the average (across the 30,000 placebo samples) correlation coefficient between the i-specific linear treatment effects in eq. (D.17) (computed with $\check{\beta}$ instead of β) and the nonlinear ones in eq. (D.8) becomes much lower.

In summary, Table D.4 shows that, when the value of the variance of the sector-level shocks is small, the difference between β and $\check{\beta}$ reported in eq. (D.16) is small, and the linear approximations to the treatment effects in eqs. (D.17) and (D.18) remain very close to their non-linear counterparts in eqs. (D.8) and (D.9). Conversely, these approximations become much worse as the variance of the sector-level shocks increases.

In Table D.5, we study the performance of different inference methods in their capacity to provide information about the value of β in eq. (D.15) or about the parameter $\check{\beta}$. Columns (2)–(6) report the standard deviation of the OLS estimated coefficients $\hat{\beta}^m$ and the average estimated standard errors obtained with different inference procedures. Columns (7)–(10) report the rejection rate of the null hypothesis that $\beta = \check{\beta}$ and columns (11)–(14) report the rejection rate of the null hypothesis that β coincides with the expression in eq. (D.15). Results are similar for all levels of γ^2: robust and state-clustered standard errors significantly underestimate the standard deviation of the OLS estimator, while the *AKM* and *AKM0* are much closer to this standard deviation. In line with these results, when testing the null that β coincides with the expression in eq. (D.15) at the 5% significance level, columns (13)-(14) show that the rejection rates are close to 5% for *AKM* and *AKM0* inference procedures, but columns (11)–(12) show that the analogous rejection rates are around 50% for the *Robust* and *Cluster* inference procedures. Given the difference (reported in Table D.4) between the value of β in eq. (D.15) and the value of $\check{\beta}$, it is not surprising that, as illustrated in columns (7)–(10) of Table D.5, rejection rates for the null that β equals $\check{\beta}$ are larger than for the null that β equals the expression in eq. (D.15), no matter what inference procedure we use. However, it is remarkable that, when the *AKM0* inference procedure is used, these rejection rates remain quite close to 5% and always below 10%.

In summary, Table D.5 shows that, no matter what the value of the variance of the sector-level shocks is, the relative performance of the four different inference procedures that we consider in all our placebo simulations is consistent with what we have documented in Sections 2 and 6.1. *Robust* and *Cluster* lead to overrejection of the estimand of the OLS estimator, while *AKM* and *AKM0* maintain their good coverage properties for this estimand. Interestingly, even when the OLS estimated does not coincide with the structural parameter $\check{\beta}$, the *AKM0* inference procedure maintains good coverage for this structural parameter; the reason is that, as the variance of the sector-level shocks increases and the OLS estimand becomes more different from $\check{\beta}$, the length of the *AKM0* confidence interval also increases, and it does so at a rate such that it contains $\check{\beta}$ in a fraction of placebo samples that is always between 5% and 10%.

Table D.4: First-order approximation error: bias in $\hat{\beta}$ and in estimated average treatment effect

$var(\mathcal{X}_s)$	$eq.$ (D.15)	$\bar{\hat{\beta}}$	Avg. Treatment Effect			Correlation between linear & non-linear avg. treatment effect
			Linear		Non-linear	
			Estimated	True	True	
(1)	(2)	(3)	(4)	(5)	(6)	(7)
0.1	0.41	0.41	0.00	0.00	0.00	0.96
1	0.48	0.48	0.00	0.00	0.05	0.76
2	0.54	0.53	0.00	0.00	0.11	0.64
5	0.63	0.62	0.01	0.00	0.34	0.47
10	0.65	0.62	0.01	0.00	0.81	0.36

Notes: The sectoral shifters $\mathcal{X}_s$ are *i.i.d*, drawn from a normal distribution with mean zero and variance $var(\mathcal{X}_s)$. Column (1) indicates the different values of $var(\mathcal{X}_s)$ that we consider in our simulation exercise; for each value of $var(\mathcal{X}_s)$ listed in column (1), we generate 30,000 simulated samples. Given a set of draws of the shifters $(\mathcal{X}_1^m, \ldots, \mathcal{X}_s^m, \ldots, \mathcal{X}_S^m)$ for a simulated sample indexed by m, their true impact on the outcome of a region i is $\check{\beta} \log(\sum_s w_{is} \exp(\mathcal{X}_s^m))$ and the first-order approximation to this expression is $\beta \sum_s w_{is} \mathcal{X}_s^m$. We set $\check{\beta} = 0.4$ for all our simulation exercises. Given this value of $\check{\beta}$ and the value of $var(\mathcal{X}_s)$ in column (1), we report in column (2) the value of β, the estimand of the OLS estimator in a regression of Y_i on X_i computed according to the expression in eq. (D.15). We report in column (3) the average (across the simulated samples) value of this OLS estimator $\hat{\beta}^m$. Column (4) and (5) reports the average (across the simulated samples) value of the linearly approximated average treatment effect in eq. (D.18), with the only difference being whether the value of β in this expression is set to the value in eq. (D.15) or to the average of the OLS estimator $\hat{\beta}^m$. Column (6) reports the average (across the simulated samples) value of the true average treatment effect in eq. (D.9). Column (7) reports the median (across the simulated samples) value of the correlation coefficient between the true treatment effect in eq. (D.9) and that arising from the first-order approximation in eq. (D.18). See the description in Appendix D.5.3 for additional details.

Table D.5: First-order approximation error: impact on standard errors and rejection rates.

$var(\mathcal{X}_s)$	Estimate		Median eff. s.e.				Rejection rate of $H_0\colon \beta = \check{\beta}$				Rejection rate of $H_0\colon \beta = eq.$ (D.15)			
	Mean (1)	Std. dev (2)	Robust (3)	Cluster (4)	AKM (5)	AKM0 (6)	Robust (7)	Cluster (8)	AKM (9)	AKM0 (10)	Robust (11)	Cluster (12)	AKM (13)	AKM0 (14)
0.1	0.41	0.07	0.03	0.03	0.07	0.08	45.3%	38.7%	9.6%	3.9%	45.1%	38.4%	9.5%	4.0%
1	0.48	0.10	0.03	0.04	0.09	0.10	60.7%	56.3%	15.3%	5.9%	52.8%	47.9%	10.4%	2.8%
2	0.53	0.14	0.04	0.05	0.12	0.14	66.7%	62.6%	17.3%	7.7%	54.1%	49.1%	9.1%	2.5%
5	0.62	0.19	0.06	0.07	0.18	0.21	72.5%	68.4%	18.2%	9.1%	53.9%	48.6%	7.7%	2.4%
10	0.62	0.22	0.07	0.08	0.22	0.25	67.8%	63.0%	14.2%	6.8%	53.5%	47.4%	7.3%	3.0%

Notes: The sectoral shifters $\mathcal{X}_s$ are *i.i.d* drawn from a normal distribution with mean zero and variance $var(\mathcal{X}_s)$. Column (1) indicates the different values of $var(\mathcal{X}_s)$ that we consider in our simulation exercise; for each value of $var(\mathcal{X}_s)$ listed in column (1), we generate 30,000 simulated samples. This table indicates the median and standard deviation of the OLS estimates of β in eq. (1) across the placebo samples (columns (1) and (2)), the median effective standard error estimates (columns (3) to (6)), the percentage of placebo samples for which we reject the null hypothesis $H_0\colon \beta = \check{\beta}$ using a 5% significance level test (columns (7) to (10)), and the percentage of placebo samples for which we reject the null hypothesis that β coincides with the expression in eq. (D.15) using a 5% significance level test (columns (11) to (14)). *Robust* is the Eicker-Huber-White standard error; *Cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5; *AKM0* is the confidence interval in Remark 6. For each inference procedure, the median effective standard error is equal to the median length of the corresponding 95% confidence interval divided by 2×1.96.

D.6 Unobserved shift-share components with different shares

Equation (21) in Section 4.1 characterizes the source of the overrejection problem affecting traditional inference methods in shift-share specifications, showing that heteroskedasticity-robust and cluster-robust standard errors overreject whenever the correlation between residuals is positive. This positive correlation arises when the residual has a shift-share structure in eq. (22), the unobserved shifters may vary at the same level as the shift-share covariate of interest (e.g. sectors) or a different one (e.g. countries of origin of immigrants). In this section, we conduct a placebo simulation to illustrate the bias in both robust and state clustered standard errors that arises when the regression residual has a shift-share component.

We generate 30,000 placebo samples indexed by m with 722 US CZs and 396 4-digit SIC manufacturing industries. As in the baseline placebo exercise discussed in Sections 2 and 6.1, we compute the shift-share covariate of interest using the sectoral employment shares of US CZs in 1990 and sectoral shifters that are drawn independently from a normal distribution with mean equal zero and variance equal to five; i.e.

$$X_i^m = \sum_{s=1}^{396} w_{is} \mathcal{X}_s^m \quad \text{such that} \quad \mathcal{X}_s^m \sim N(0,5).$$

The difference between the simulation exercise we consider here and the baseline placebo simulation in Sections 2 and 6.1 is that the outcome variable is no longer taken from the observed data. Instead, this outcome variable varies across placebo samples and it is drawn randomly for each simulated sample m as

$$Y_i^m = \sum_{s=1}^{396} \tilde{w}_{is} \mathcal{A}_s^m \qquad \text{such that} \quad \mathcal{A}_s^m \sim N(0,5),$$

where $\tilde{w}_{is}$ are shares that may be different from (but possibly correlated with) the baseline sectoral employment shares in each CZ; i.e. $\tilde{w}_{is}$ may be different from w_{is}. Specifically, for all placebo samples, we generate a single set of alternative shares as

$$\tilde{w}_{is} = \frac{\exp\left(u_{is} + \ln(w_{is} + v_{is})\right)}{\sum_{k=1}^{396} \exp\left(u_{ik} + \ln(w_{ik} + v_{ik})\right)} \left(\sum_{k=1}^{396} w_{ik} \right) \tag{D.19}$$

where u_{is} and v_{is} drawn randomly such that $u_{is} \sim N(0, \sigma_u^2)$ and $v_{is} \sim U[0, \sigma_v]$.

Given a pair of values (σ_u, σ_v), for each placebo sample we compute: (a) the OLS estimator of the regression of Y_i^m on X_i^m and a constant; (b) effective standard errors according to the robust, state-clustered, *AKM* and *AKM0* inference procedures; (c) for each of these inference procedures, the outcome of a 5% significance level test of hypothesis of the null hypothesis $H_0\colon \beta = 0$. Each row of Table D.6 reports several summary statistics of the distribution of these quantities across the 30,000 placebo samples. Each row does so for placebo samples generated by different values of σ_u and σ_v.

The first row of Table D.6 considers the case in which $\sigma_v = \sigma_u = 0$. In this case, $w_{is} = \tilde{w}_{is}$ for every i and s and, thus, the correlation coefficient between the shares entering the covariate of interest and those entering the regression residual equal 1 (see column (3) in Table D.6). In this case, as in our baseline placebo, robust and state-cluster standard errors have rejection rates for a 5% significance

Table D.6: Bias in standard errors when regression residual is a shift-share term with shares correlated with those entering the shift-share covariate of interest

			Estimate		Median eff. s.e.				Rejection rate of $H_0\colon \beta = 0$			
σ_u^2	σ_v	$\rho_{w_{is},\tilde{w}_{is}}$	Mean	Std. dev	Robust	Cluster	AKM	AKM0	Robust	Cluster	AKM	AKM0
(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)
0	0	1.00	0.00	0.17	0.08	0.08	0.14	0.16	34.8%	31.0%	10.2%	3.7%
1	0	0.77	0.00	0.16	0.08	0.09	0.14	0.16	31.5%	27.3%	9.9%	4.0%
3	0	0.55	0.00	0.15	0.09	0.09	0.13	0.15	23.8%	22.7%	9.8%	4.1%
5	0	0.44	0.00	0.14	0.10	0.10	0.13	0.15	18.0%	17.4%	9.6%	4.3%
0	0.001	1.00	0.00	0.11	0.06	0.06	0.10	0.11	31.6%	28.7%	10.0%	3.7%
1	0.001	0.70	0.00	0.10	0.05	0.06	0.09	0.10	28.1%	26.5%	9.8%	4.2%
3	0.001	0.41	0.00	0.09	0.06	0.06	0.08	0.09	19.0%	19.3%	8.8%	4.1%
5	0.001	0.28	0.00	0.09	0.06	0.06	0.08	0.09	13.4%	14.5%	8.1%	4.4%
0	0.01	1.00	0.00	0.04	0.02	0.02	0.04	0.04	25.3%	23.4%	9.4%	3.6%
1	0.01	0.38	0.00	0.05	0.03	0.04	0.04	0.05	14.3%	14.2%	7.6%	3.8%
3	0.01	0.18	0.00	0.04	0.03	0.03	0.04	0.05	11.6%	12.3%	7.7%	4.4%
5	0.01	0.10	0.00	0.05	0.05	0.04	0.05	0.06	7.9%	9.0%	7.5%	4.2%

Notes: We impose that, for every simulated sample $m = 1, \ldots, 30000$, the outcome variable is $Y_i^m = \sum_s \tilde{w}_{is}\mathcal{A}_s^m$, with $\mathcal{A}_s^m$ drawn from a normal distribution with mean zero and variance equal to five. The shares $\{\tilde{w}_{is}\}_{i,s}$ vary across the cases described in each of the rows in the table above but, for each of these rows, are fixed across the 30,000 simulated samples. Specifically, given shares $\{w_{is}\}_{i,s}$ that capture the employment share in CZ i employed in sector s in 1990, we generate each $\tilde{w}_{is}$ according to the expression in eq. (D.19), with u_{is} and v_{is} drawn randomly according to the distributions $u_{is} \sim \mathcal{N}(0, \sigma_u^2)$ and $U[0, \sigma_v]$. The first two columns in the table above indicate the values of σ_u and σ_v used to generate $\{\tilde{w}_{is}\}_{i,s}$ in each case. As illustrated in the third column, the larger the value of either σ_u or σ_v, the lower the correlation coefficient $\rho_{w_{is},\tilde{w}_{is}}$ between w_{is} and $\tilde{w}_{is}$ across regions and sectors. Given the generated outcome variables $\{Y_i^m\}_i$ for each simulated sample m, we compute the OLS estimate of β in the regression $Y_i^m = \beta X_i^m + \epsilon_i^m$, with $X_i^m = \sum_s w_{is}\mathcal{X}_s^m$ and each $\mathcal{X}_s^m$ drawn randomly from a normal distribution with mean zero and variance equal to 5. We indicate the mean and standard deviation of the OLS estimates of β across the simulated samples (columns (4) and (5)), the median effective standard error estimates (columns (6) to (9)), and the percentage of placebo samples for which we reject the null hypothesis $H_0\colon \beta = 0$ using a 5% significance level test (columns (10) to (13)). *Robust* is the Eicker-Huber-White standard error; *Cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5; *AKM0* is the confidence interval in Remark 6.

level test that are around 30%–35%. In contrast, the *AKM* and *AKM0* inference procedures exhibit rejection rates that are 10% and 4%, respectively. The remaining rows of Table D.6 show that, as we increase the value of σ_v and σ_u, the correlation between w_{is} and $\tilde{w}_{is}$ declines, which attenuates the overrejection problem affecting testing procedures that rely on robust and state-clustered standard errors. However, the rejection rates of these two inference methods are still above 10% even when the correlation between w_{is} and $\tilde{w}_{is}$ is as low as 0.18. For all cases, the rejection rates of the *AKM* and *AKM0* testing procedures remain stable and close to 5%.

D.7 Heterogeneous treatment effects

We now present a placebo exercise to evaluate the performance of our inference procedures in the presence of heterogeneous treatment effects. For each placebo sample m, we construct the dependent variable as

$$Y_i^m = Y_i + \sum_s w_{is}\mathcal{X}_s^m \beta_{is} \quad \text{such that} \quad \beta_{is} = \lambda w_{is}.$$

In all placebo samples, Y_i is the change in the share of working-age population employed in CZ i and w_{is} is the share of sector s in total employment of CZ i. As before, in each placebo sample, we take

Table D.7: Heterogeneous treatment effects

		Estimate		Median eff. s.e.				Rejection rate of $H_0: \beta = \beta_0$			
λ (1)	β_0 (2)	Mean (3)	Std. dev (4)	Robust (5)	Cluster (6)	AKM (7)	AKM0 (8)	Robust (9)	Cluster (10)	AKM (11)	AKM0 (12)
0	0.00	0.00	1.98	0.73	0.92	1.91	2.22	0.48	0.38	0.07	0.04
1	0.14	0.15	1.98	0.73	0.92	1.91	2.22	0.48	0.38	0.07	0.04
3	0.43	0.45	1.98	0.74	0.92	1.91	2.22	0.48	0.38	0.07	0.04
5	0.72	0.74	1.98	0.74	0.93	1.91	2.23	0.48	0.37	0.08	0.04

Notes: This table indicates the median and standard deviation of the OLS estimates of β in eq. (1) across the placebo samples (columns (3) and (4)), the median effective standard error estimates (columns (5) to (8)), and the percentage of placebo samples for which we reject the null hypothesis $H_0\colon \beta = \beta_0$ using a 5% significance level test (columns (9) to (12)) where the true value of β_0 shown in column (2) is given in eq. (D.20). *Robust* is the Eicker-Huber-White standard error; *Cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5; *AKM0* is the confidence interval in Remark 6. For each inference procedure, the median effective standard error is equal to the median length of the corresponding 95% confidence interval divided by 2×1.96. Results are based on 30,000 placebo samples.

independent draws of the sector-level shifters from a normal distribution with a mean of zero and a variance of 5.

The parameter λ controls the degree of heterogeneity in the treatment effect of the sector-level shifters. When $\lambda = 0$, this placebo exercise is identical to our baseline placebo exercise in Section 2. We are interested in inference on the OLS estimand. By Proposition 1, it is given by

$$\beta_0 = \sum_{i,s} w_{is}^2 \beta_{is} \Big/ \sum_{i,s} w_{is}^2 = \lambda \sum_{i,s} w_{is}^3 \Big/ \sum_{i,s} w_{is}^2, \tag{D.20}$$

which is linear in λ.

Table D.7 presents the results of the placebo exercise for different values of λ. For all values of λ, the average OLS estimate in column (3) is similar to β_0. Results indicate that both the standard deviation of the OLS estimator and the performance of the inference procedures are not sensitive to the value of λ.

D.8 Other extensions

In Table D.8, we report results analogous to those in Table D.1 for outcome variables Y_i other than the employment rate in CZ i. The rejection rates that we obtain are very similar to those reported in Table D.1 and discussed in Section 6.2.

In Table D.9, we investigate the sensitivity of our results to an alternative definition of "region". We report results for a placebo exercise that is analogous to the baseline placebo exercise discussed in Sections 2 and 6.1 except for the use of counties instead of CZs as regions. We use the County Business Patterns 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, we only implement the placebo exercise for the outcome variables used in Panel A of Tables 1 and 2: employment rate; employment rate in manufacturing; and, employment rate in non-manufacturing. The results show that the rejection rates of all four inference procedures we consider are very similar to those obtained in the baseline placebo exercise, which are reported precisely in Panel A of Tables 1 and 2.

In Table D.10, we investigate the sensitivity of our results to an alternative definition of "sector". We report results for a placebo exercise that is analogous to the baseline placebo exercise discussed in Sections 2 and 6.1 except for the use of 331 occupations instead of 396 sectors as the unit of observation at which the shifters vary. The results in Table D.10 show that the overrejection problem affecting tradition inference procedures is even more severe when the shift-share covariate aggregates occupation-specific shifters than when it aggregates sectoral shifters. Actually, only the *AKM0* inference procedure yields rejection rates for the null hypothesis $H_0\colon \beta = 0$ that are below the 5% significance level of the test.

Table D.8: Controlling for the size of the residual sector in each CZ

	Estimate		Median eff. s.e.				Rejection rate of $H_0\colon \beta = 0$			
	Mean	Std. dev	Robust	Cluster	AKM	AKM0	Robust	Cluster	AKM	AKM0
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Shifters with zero mean										
Outcome variable: change in the share of working-age population in manufacturing										
No controls	−0.02	1.87	0.60	0.76	1.78	2.06	55.5%	44.2%	8.1%	4.2%
Control: $1 - \sum_s w_{is}$	0.00	1.03	0.56	0.63	0.97	1.12	30.1%	25.8%	10.0%	4.4%
Change in the share of working-age population in non-manufacturing										
No controls	0.00	0.94	0.58	0.67	0.89	1.04	23.0%	17.5%	8.1%	4.5%
Control: $1 - \sum_s w_{is}$	0.00	1.05	0.60	0.68	0.97	1.12	27.5%	22.6%	9.8%	5.4%
Outcome variable: change in average log-weekly wage of all employees										
No controls	0.05	2.67	1.02	1.34	2.58	3.00	47.0%	33.9%	7.8%	4.4%
Control: $1 - \sum_s w_{is}$	0.00	1.21	0.95	1.07	1.15	1.33	12.9%	8.9%	7.9%	4.8%
Outcome variable: change in average log-weekly wage of all employees in manufacturing										
No controls	0.02	2.94	1.69	2.11	2.75	3.19	27.0%	17.3%	9.3%	4.5%
Control: $1 - \sum_s w_{is}$	0.01	2.13	1.66	1.92	1.98	2.28	12.5%	8.0%	7.7%	4.5%
Outcome variable: change in average log-weekly wage of all employees in non-manufacturing										
No controls	0.00	2.62	1.05	1.33	2.56	2.98	44.5%	32.8%	7.6%	4.4%
Control: $1 - \sum_s w_{is}$	0.00	1.24	0.98	1.08	1.17	1.35	12.8%	9.5%	8.5%	4.7%
Panel B: Shifters with non-zero mean										
Outcome variable: change in the share of working-age population in manufacturing										
No controls	−3.92	1.12	0.57	0.81	1.34	1.51	98.7%	97.6%	80.6%	78.7%
Control: $1 - \sum_s w_{is}$	0.00	1.05	0.56	0.63	0.97	1.12	31.1%	26.4%	10.3%	4.6%
Outcome variable: change in the share of working-age population in non-manufacturing										
No controls	−0.75	0.71	0.48	0.64	0.76	0.86	37.2%	22.2%	14.2%	13.9%
Control: $1 - \sum_s w_{is}$	0.01	1.05	0.60	0.68	0.97	1.13	27.6%	22.5%	9.7%	5.2%
Outcome variable: change in average log-weekly wage of all employees										
No controls	−6.52	1.55	0.97	1.58	1.91	2.15	99.6%	98.3%	90.9%	90.5%
Control: $1 - \sum_s w_{is}$	0.01	1.22	0.95	1.08	1.15	1.33	13.4%	9.1%	8.1%	4.9%
Outcome variable: change in average log-weekly wage of all employees in manufacturing										
No controls	−5.38	1.88	1.54	2.29	1.94	2.17	89.3%	69.8%	75.1%	71.0%
Control: $1 - \sum_s w_{is}$	−0.02	2.13	1.66	1.91	1.98	2.28	12.5%	8.1%	7.8%	4.7%
Outcome variable: change in average log-weekly wage of all employees in non-manufacturing										
No controls	−6.31	1.54	0.99	1.58	1.90	2.15	99.4%	97.8%	89.0%	88.6%
Control: $1 - \sum_s w_{is}$	0.01	1.24	0.98	1.08	1.17	1.35	12.6%	9.4%	8.3%	4.7%

Notes: This table indicates the median and standard deviation of the OLS estimates of β in eq. (1) across the placebo samples (columns (1) and (2)), the median effective standard error estimates (columns (3) to (6)), and the percentage of placebo samples for which we reject the null hypothesis $H_0\colon \beta = 0$ using a 5% significance level test (columns (7) to (10)). *Robust* is the Eicker-Huber-White standard error; *Cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5; *AKM0* is the confidence interval in Remark 6. For each inference procedure, the median effective standard error is equal to the median length of the corresponding 95% confidence interval divided by 2×1.96. Results are based on 30,000 placebo samples. In Panel A, $(\mathcal{X}_1^m, \ldots, \mathcal{X}_{S-1}^m)$ is drawn i.i.d. from a normal distribution with zero mean and variance equal to 5 in each placebo sample. In Panel B, $(\mathcal{X}_1^m, \ldots, \mathcal{X}_{S-1}^m)$ is drawn i.i.d. from a normal distribution with mean equal to one and variance equal to 5 in each placebo sample. For each of the two panels, the first row presents results in which no control is accounted for in the estimating equation; the second row presents results in which we control for the size of the residual sector, $1 - \sum_s w_{is}$.

Table D.9: Magnitude of standard errors and rejection rates: county-level analysis

	Estimate		Median eff. s.e.				Rejection rate of $H_0\colon \beta = 0$			
	Mean (1)	Std. dev (2)	Robust (3)	Cluster (4)	AKM (5)	AKM0 (6)	Robust (7)	Cluster (8)	AKM (9)	AKM0 (10)
Panel A: Change in the share of working-age population										
employed (all)	0.00	0.65	0.24	0.30	0.61	0.67	47.3%	36.3%	8.0%	4.8%
employed (manuf.)	0.00	0.77	0.18	0.27	0.71	0.78	65.5%	51.4%	8.1%	4.6%
employed (non-manuf.)	0.00	0.37	0.21	0.22	0.35	0.39	27.9%	25.3%	8.8%	4.6%

Notes: This table indicates the median and standard deviation of the OLS estimates of β in eq. (1) across the placebo samples (columns (1) and (2)), the median effective standard error estimates (columns (3) to (6)), and the percentage of placebo samples for which we reject the null hypothesis $H_0\colon \beta = 0$ using a 5% significance level test (columns (7) to (10)). *Robust* is the Eicker-Huber-White standard error; *Cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5; *AKM0* is the confidence interval in Remark 6. For each inference procedure, the median effective standard error is equal to the median length of the corresponding 95% confidence interval divided by 2×1.96. Results are based on 30,000 placebo samples.

Table D.10: Magnitude of standard errors and rejection rates: occupation-specific shifters

	Estimate		Median eff. s.e.				Rejection rate of $H_0\colon \beta = 0$			
	Mean (1)	Std. dev (2)	Robust (3)	Cluster (4)	AKM (5)	AKM0 (6)	Robust (7)	Cluster (8)	AKM (9)	AKM0 (10)
Panel A: Change in the share of working-age population										
employed (all)	0.01	8.59	1.13	2.45	7.46	27.82	83.5%	62.4%	24.9%	4.0%
employed (manuf.)	0.02	8.13	0.80	1.82	6.55	25.50	89.7%	75.3%	32.9%	3.2%
employed (non-manuf.)	−0.01	4.03	0.96	1.76	3.06	9.86	65.1%	38.4%	17.9%	3.8%
Panel B: Change in average log weekly wage										
employed (all)	0.00	12.58	1.74	4.23	10.1	38.10	84.8%	62.6%	30.6%	3.4%
employed (manuf.)	−0.07	11.11	3.24	6.18	9.41	31.96	56.2%	27.2%	11.7%	4.9%
employed (non-manuf.)	0.01	12.60	1.77	4.19	9.96	37.96	84.9%	64.1%	31.8%	3.2%

Notes: This table indicates the median and standard deviation of the OLS estimates of β in eq. (1) across the placebo samples (columns (1) and (2)), the median effective standard error estimates (columns (3) to (6)), and the percentage of placebo samples for which we reject the null hypothesis $H_0\colon \beta = 0$ using a 5% significance level test (columns (7) to (10)). *Robust* is the Eicker-Huber-White standard error; *Cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5; *AKM0* is the confidence interval in Remark 6. For each inference procedure, the median effective standard error is equal to the median length of the corresponding 95% confidence interval divided by 2×1.96. Results are based on 30,000 placebo samples.

Appendix E Empirical applications: additional results

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

This section presents additional results that complement the estimates in Section 7.1 of the effect of Chinese import competition on US local labor markets following the approach in Autor, Dorn and Hanson (2013, ADH hereafter).

E.1.1 Placebo exercise: alternative distributions of shifters

The reduced-form and the first-stage specifications have a panel data structure discussed Section 5.2. Since the outcome data and the share matrix W is the same as in the placebo exercise in Appendix D.4, the results of that placebo exercise are informative about the finite-sample properties of the four inference procedures that we consider (robust standard errors, state-clustered standard errors, and the *AKM* and *AKM0* procedures) in the ADH empirical application. In this section, we investigate the robustness of the results in Appendix D.4 to alternative distributions of the sectoral shifters. In particular, instead of assuming that the shifters are i.i.d. according to a normal distribution, we consider distributions that are arguably closer to the distribution of the actual shifters employed in ADH (the growth in sectoral Chinese exports to high-income countries other than the US).

First, we consider a placebo exercise that differs from that in Appendix D.4 only in that the sectoral shifters are drawn independently from the empirical distribution of the shifters used in ADH. The results are presented in Panel A of Table E.1. As in the analysis in Appendix D.4, although the data generating process for our placebo exercise implies that $\beta = 0$, the rejection rates of a 5% significance level test of the null hypothesis $H_0 \colon \beta = 0$ are substantially above 5% when robust and state-clustered standard errors are used. The rejection rates implied by the *AKM* and *AKM0* procedures are much closer to 5%, with rejection rates are close to 10%.

Second, to get closer to the specification in ADH, we incorporate into our placebo specification the baseline set of controls that ADH use (see, e.g., column (6) of Table 3 in ADH). In particular, we draw the sectoral shifters from the empirical distribution of shifters used in ADH *after partialling out the baseline set of controls used in ADH*.[5] Panel B of Table E.1 reports the results. For the *Robust*, *Cluster* and *AKM* testing procedures, the rejection rates in Panel B are very similar to those in Panel A, while the *AKM0* rejection rate is much closer to the nominal level.

Next, we consider relaxing the assumption that the sectoral shifters are independent, or independent across clusters. This specification is motivated by the concern that the 1990–2000 and 2000–2007 sector-specific growth rates in Chinese exports to high-income countries other than the US were determined at least partly by a common factor that had possibly heterogeneous effects across sectors. We formalize this by modeling year-t imports from China of goods in sector s by high-income countries

[5] To partial out a set of controls (which vary by region) from the shifters (which vary by sector), we implement the following two-step procedure. First, we obtain the residual of a regression of the shift-share instrumental variable X_i used in ADH on the set of controls listed in column (6) of Table 3 in ADH; let $\ddot{X}_i$ denote this residual. We then draw the shifters from the empirical distribution of the residualized sectoral shifters $\mathcal{X}^{res}$, which correspond to the regression coefficients from regressing $\ddot{X}_i$ onto the vector of shares $(w_{i1}, \dots, w_{iS})$, i.e. $\mathcal{X}^{res} = (W'W)^{-1}W'\ddot{X}$.

Table E.1: Alternative distributions of sectoral shifters: placebo

	Estimate		Median eff. s.e.				Rejection rate of $H_0: \beta = 0$			
	Mean	Std. dev	Robust	Cluster	AKM	AKM0	Robust	Cluster	AKM	AKM0
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
Panel A: Empirical distribution of ADH (2013) shocks										
period: 1990–2000	0.11	0.49	0.16	0.19	0.38	0.85	48.5%	39.9%	10.7%	9.3%
period: 2000–2007	0.04	0.16	0.05	0.06	0.13	0.31	47.9%	39.4%	11.0%	9.5%
Panel B: Empirical distribution of residualized ADH (2013) shocks										
period: 1990–2000	0.00	0.14	0.05	0.06	0.12	0.21	46.5%	37.9%	10.4%	3.7%
period: 2000–2007	0.00	0.07	0.03	0.03	0.06	0.11	46.4%	37.7%	10.9%	3.7%

Notes: This table indicates the median and standard deviation of the OLS estimates of β in eq. (1) across the placebo samples (columns (1) and (2)), the median effective standard error estimates (columns (3) to (6)), and the percentage of placebo samples for which we reject the null hypothesis $H_0: \beta = 0$ using a 5% significance level test (columns (7) to (10)). *Robust* is the Eicker-Huber-White standard error; *Cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5; *AKM0* is the confidence interval in Remark 6. For each inference procedure, the median effective standard error is equal to the median length of the corresponding 95% confidence interval divided by 2×1.96. Results are based on 30,000 placebo samples. In Panel A, each $\mathcal{X}_s^m$ is drawn from the empirical distribution of shifters $\mathcal{X}_s$ observed in the data; i.e. from the empirical distribution of changes in sectoral exports from China to high-income countries other than the US. In Panel B, each $\mathcal{X}_s^m$ is drawn from the empirical distribution of residualized shifters $\mathcal{X}_s$ observed in the data; i.e. from the empirical distribution of the residuals of projecting the changes in sectoral exports from China to high-income countries other than the US on the full vector of baseline controls in ADH; i.e. those in column 6 of Table 3 in Autor, Dorn and Hanson (2013).

other than the US, IMP_{st}, as

$$IMP_{st} = X_{st}^{Ch} + \epsilon_{st}, \tag{E.1}$$

where X_{st}^{Ch} is a sectoral component of Chinese exports common to all destinations (i.e. it accounts for export supply factors), and ϵ_{st} is sector- and destination-specific component (i.e. it accounts for export demand factors). We impose the following factor structure on X_{st}^{Ch}:

$$X_{st}^{Ch} = \eta_s \bar{X}_t^{Ch} + e_{st}. \tag{E.2}$$

The term $\bar{X}_t^{Ch}$ captures unobserved factors that may potentially impact Chinese exports across all sectors (e.g. growth in Chinese labor productivity). The row-vector of sector-specific loadings η_s indicates how Chinese exports in each sector s react to changes in the common unobserved factors captured by $\bar{X}_t^{Ch}$ (e.g. how sensitive each sector s is to growth in Chinese labor productivity). Finally, e_{st} is a sector- and year-specific idiosyncratic component of Chinese exports. Note that, as long as the distribution of $\bar{X}_t^{Ch}$ is not degenerate, the shifter IMP_{st} will be correlated across any two sectors s and s' unless the loadings η_s and $\eta_{s'}$ are orthogonal. This correlation in shifters violates the independence assumption imposed by Assumption 2(i) in Section 4.1 in a way that is not accounted for by the clustering extension considered in Section 5.1. In the placebo simulations that follow, we explore the consequences of the violation of this assumption, as well as modifications of the AKM and AKM0 procedures that account for the potential factor structure in the shifters.

Combining eqs. (E.1) and (E.2) yields

$$IMP_{st} = \eta_s \bar{X}_t^{Ch} + \varepsilon_{st}, \qquad \text{with} \qquad \varepsilon_{st} = \epsilon_{st} + e_{st}. \tag{E.3}$$

Figure E.1: Histogram of estimates of $\{\eta_s\}_{s=1}^{S}$

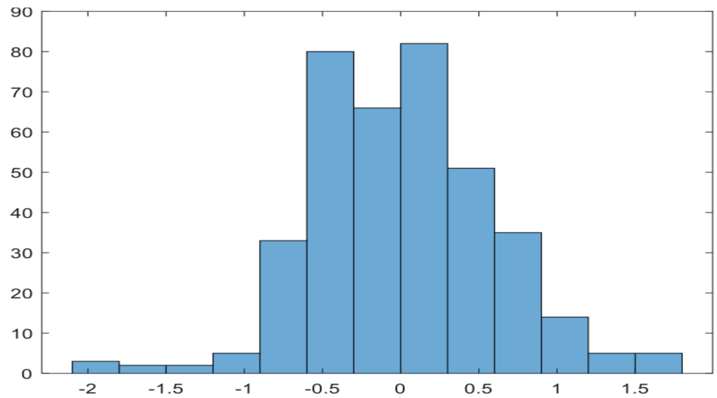
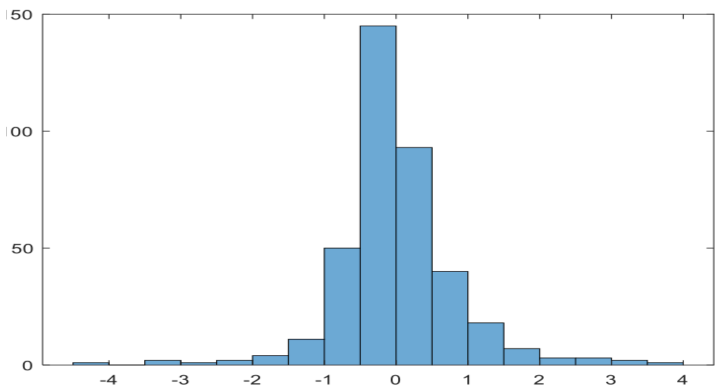


Figure E.2: Histogram of estimates of $\{u_{s,2007} - u_{s,1991}\}_{s=1}^{S}$

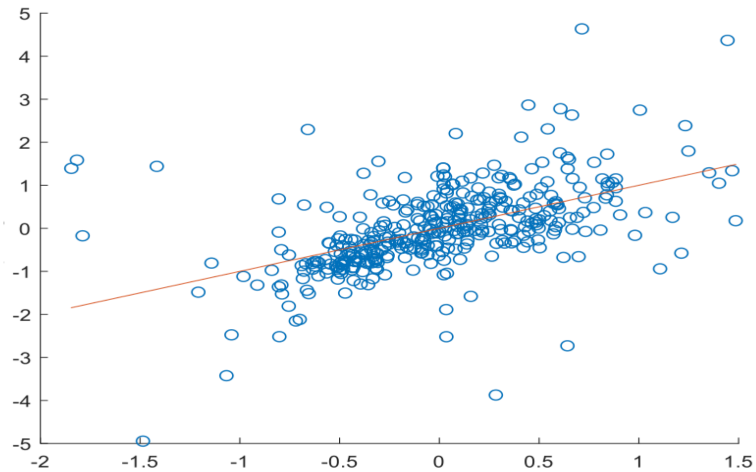


To remain as close as possible to the empirical application in ADH, we use annual data on sector-specific exports from China to other high-income countries between 1991 and 2007 (which corresponds to the variable IMP_{st} above) to estimate the common factor $\bar{X}_t^{Ch}$, the factor loadings $\{\eta_s\}_{s=1}^{S}$, and the residuals $\{\varepsilon_{st}\}_{s=1}^{S}$ for every year t and 4-digit SIC manufacturing sectors s using the interactive fixed effects estimator in Bai (2009), as implemented by Gomez (2017).

Figure E.1 reports the histogram of the estimates of $\{\eta_s\}_{s=1}^{S}$. There is considerable dispersion in the factor loadings across sectors. The estimates also reveal substantial variation across sectors and years in the idiosyncratic component of Chinese export growth ε_{st}; this can be seen in Figure E.2, which presents a histogram of the sector-specific changes in ε_{st} between 1991 and 2007. To provide a graphical illustration of the relative importance of the two terms entering the right-hand side of eq. (E.3), Figure E.3 provides a scatterplot of the variables $\{IMP_{s,2007} - IMP_{s,1991}\}_{s=1}^{S}$ against the estimates of the terms $\{\eta_s(\bar{X}_{2007}^{Ch} - \bar{X}_{1991}^{Ch})\}_{s=1}^{S}$; these terms explain only 27% of the cross-sectoral variation in export growth from China to high-income countries other than the US between 1991 and 2007.

Table E.2 reports the results of a placebo exercise illustrating the effects of the correlation in sectoral shifters implied by the estimated version of the model in eq. (E.3) on the finite-sample properties of the *AKM* and *AKM0* procedures. Specifically, we modify the baseline placebo exercise described

Figure E.3: Scatterplot of $\{IMP_{s,2007} - IMP_{s,1991}\}_{s=1}^{S}$ against $\{\eta_s(\bar{X}^{Ch}_{2007} - \bar{X}^{Ch}_{1991})\}_{s=1}^{S}$



Notes: Observed data on sector-specific export flows from China to high-income countries other than the US (i.e. $IMP_{s,2007} - IMP_{s,1991}$) appear in the vertical axis; estimates of $\eta_s(\bar{X}^{Ch}_{2007} - \bar{X}^{Ch}_{1991})$ appear in the horizontal axis. The R^2 of this regression is 0.273.

in Section 6.1 by instead generating the simulated sectoral shifters as

$$\mathcal{X}^m_s = \kappa\eta^m_s \Delta\hat{\bar{X}}_{Ch} + u^m_s, \qquad \text{with} \qquad \Delta\hat{\bar{X}}^{Ch} = \hat{\bar{X}}^{Ch}_{2007} - \hat{\bar{X}}^{Ch}_{1991} \tag{E.4}$$

where $\hat{\bar{X}}^{Ch}_t$ denotes the estimate of $\bar{X}^{Ch}_t$ for $t = 1991$ and $t = 2007$. The parameter κ controls the relative importance of the factor component in the simulated shifters. For each simulated sample m, the residuals u^m_s are drawn independently from a distribution that we vary across specifications. The term η^m_s is either fixed across the placebo samples m and set to equal to the estimate $\hat{\eta}_s$, or else drawn independently from the empirical distribution of $\hat{\eta}_s$. Whether the factor loadings η_s are fixed across the placebo samples or random (and independent across s) is important for the properties of the *AKM* and *AKM0* inference procedures. If the loadings are random and independent, the shifters $\mathcal{X}_s$ will also be independent across s, so that Assumption 2(i) in Section 4.1 holds, and we expect the *AKM* and *AKM0* inference procedures to have good asymptotic properties even if conditionally on the loadings, the interactive fixed effects structure in eq. (E.3) applies. On the other hand, if the loadings are fixed across simulation samples, the shifters will be correlated, so that the asymptotic results in Section 4 do not apply.

In Panels A and B in Table E.2, we fix $\eta^m_s = \hat{\eta}_s$ for every sector s and placebo sample m, with u^m_s drawn i.i.d. from mean-zero normal distribution with variance 5 in Panel A, and from the empirical the distribution of $\hat{\varepsilon}_{s,2007} - \hat{\varepsilon}_{s,1991}$ in Panel B, where $\hat{\varepsilon}_{st}$ is the interactive fixed effects estimate of the term ε_{st} in eq. (E.3). In the first three rows of each panel, when no controls are included, larger values of κ (which imply a larger weight on the interactive fixed effects component $\eta^m_s \Delta\hat{\bar{X}}_{Ch}$ in eq. (E.4)) imply larger rejection rates of the null $H_0\colon \beta = 0$ when we use either the *AKM* or the *AKM0* inference procedures. For $\kappa = 1$, which corresponds to the specification in ADH, the rejection rates for *AKM0* are close to the nominal rates, and *AKM* suffers from moderate overrejection. Importantly, this overrejection problem can be fixed by controlling for the term $\eta^m_s \Delta\hat{\bar{X}}_{Ch}$ as an additional covariate in our regression specification (see rows 4 to 6 in Panels A and B in Table E.2). This is in line with our theory, since conditioning on this control restores the independence assumption on the shifters. The takeaway form the results in Panels A and B in Table E.2 is thus that, if one thinks that the true data generating process for the sectoral shifters $\{\mathcal{X}_s\}_{s=1}^{S}$ corresponds to the model in eq. (E.4), then one should obtain a consistent estimate of $\eta_s \Delta \hat{\hat{X}}_{Ch}$ and control for it in the regression specification in order to ensure that the shifters are independent conditional on the controls, so that Assumption 2(i) holds once we condition on the control vector Z_i.

In Panel C of Table E.2, instead of holding the loadings fixed, we draw both η_s^m and ν_s^m in each placebo sample m from the empirical distribution of the interactive fixed effects estimates, independently across s. This makes the shifters independent across s, so that, as discussed above, Assumption 2(i) in Section 4.1 holds even without conditioning on $\eta_s \Delta \hat{\hat{X}}_{Ch}$. As a result, the rejection rates for the *AKM* and *AKM0* inference procedures reported in Panel C are similar to those reported in Table 2 in Section 6.1 and unaffected by the value of the parameter κ in eq. (E.4). In particular, the *AKM0* inference procedure yields always rejection rates that are very close to 5%.

E.1.2 Placebo exercise: accounting for controls in the first-stage regression

The placebo exercise described in Sections 2.2 and 6 use the outcome variables Y_i and the shares w_{is} used in Autor, Dorn and Hanson (2013) for the period 2000–2007. The placebo exercise discussed in Appendix D.4 gets closer to the reduced-form empirical specification in Autor, Dorn and Hanson (2013) by incorporating information on outcome variables and shares both for the period 1990–2000 and for the period 2000–2007. However, these two placebo exercises implement a specification that differs from that in Autor, Dorn and Hanson (2013) in that it includes no controls. As argued in Section 3.3, the overrejection problem affecting robust and state-clustered standard errors that is documented in the simulations is caused by cross-regional correlation in residuals across observations with similar shares. The inclusion of controls may improve the performance these methods, since the controls may soak up some (or even most) of the cross-regional correlation in the residuals.

In Table E.3, we introduce a placebo sample for the first-stage regression in Autor, Dorn and Hanson (2013). In Panel A, when we do not include any controls, both robust and state-clustered standard errors over-reject the null hypothesis $H_0\colon \beta_1 = 0$. In Panel B, we include as a control the shift-share instrumental variable used in Autor, Dorn and Hanson (2013), and the rejection rate for these procedures decreases to about 20%. Finally, in Panel C, we additionally include all controls used in the baseline specification in Autor, Dorn and Hanson (2013), and the *Robust* and *Cluster* rejection rates get closer to 14%. It can also be seen from Table E.3 that the rejection rates for the *AKM* and *AKM0* procedures are always very close to the 5% nominal level.

E.1.3 Additional empirical results

In Tables E.4 and E.5 we extend the results presented in Table 5 in Section 7.1. Specifically, Tables E.4 and E.5 present results not only for all workers (in Panel A), but also two subsets of workers: college graduates (in Panel B) and non-college graduates (in Panel C). Additionally, while the *AKM* and *AKM0* confidence intervals presented in Table 5 cluster observations belonging to the same 3-digit

Table E.2: Simulation for common China shock with heterogeneous sectoral exposure

κ	Control for $\eta_s^m \Delta\hat{\hat{X}}_{Ch}$	Estimate: Mean (1)	Estimate: Std. dev (2)	Median eff. s.e.: Robust (3)	Median eff. s.e.: Cluster (4)	Median eff. s.e.: AKM (5)	Median eff. s.e.: AKM0 (6)	Rejection rate of $H_0\colon \beta = 0$: Robust (7)	Rejection rate of $H_0\colon \beta = 0$: Cluster (8)	Rejection rate of $H_0\colon \beta = 0$: AKM (9)	Rejection rate of $H_0\colon \beta = 0$: AKM0 (10)
Panel A: $\eta_s^m = \hat{\eta}_s$ for all m and s; $u_s^m \sim \mathcal{N}(0,5)$											
0	No	0.00	0.17	0.08	0.09	0.14	0.17	35.4%	31.3%	10.3%	3.9%
1	No	0.00	0.15	0.07	0.07	0.12	0.14	38.2%	33.6%	12.4%	5.1%
3	No	0.00	0.09	0.04	0.04	0.06	0.07	42.2%	35.9%	17.5%	8.4%
0	Yes	0.00	0.16	0.08	0.08	0.14	0.16	34.9%	31.6%	10.4%	4.2%
1	Yes	0.00	0.16	0.08	0.08	0.14	0.16	35.1%	31.8%	10.5%	4.3%
3	Yes	0.00	0.16	0.08	0.08	0.14	0.16	34.9%	31.9%	10.5%	4.3%
Panel B: $\eta_s^m = \hat{\eta}_s$ for all m and s; $u_s^m \sim F_{emp}$											
0	No	0.00	0.43	0.20	0.21	0.35	0.49	36.5%	33.2%	12.1%	3.5%
1	No	0.00	0.26	0.11	0.12	0.18	0.21	42.3%	36.3%	17.3%	8.2%
3	No	0.00	0.10	0.04	0.05	0.07	0.08	43.9%	37.3%	18.8%	9.4%
0	Yes	0.00	0.43	0.19	0.21	0.34	0.46	36.7%	33.7%	12.7%	3.9%
1	Yes	0.00	0.43	0.19	0.21	0.34	0.46	36.0%	33.1%	12.3%	3.7%
3	Yes	0.00	0.43	0.19	0.21	0.34	0.46	36.3%	33.4%	12.3%	3.6%
Panel C: $(\eta_s^m, u_s^m) \sim F_{emp}$											
0	No	0.00	0.43	0.20	0.21	0.35	0.49	36.7%	33.1%	12.0%	3.5%
1	No	0.00	0.26	0.12	0.13	0.22	0.26	36.0%	32.1%	10.5%	3.8%
3	No	0.00	0.10	0.05	0.05	0.09	0.11	35.3%	31.4%	10.3%	3.7%
0	Yes	0.00	0.43	0.19	0.21	0.34	0.46	36.2%	33.1%	12.1%	3.5%
1	Yes	0.00	0.43	0.19	0.21	0.34	0.46	37.1%	33.5%	12.4%	3.9%
3	Yes	0.00	0.42	0.18	0.20	0.32	0.42	37.8%	34.4%	13.5%	5.2%

Notes: We impose the assumption that the year-specific sectoral shifters IMP_{st} are generated from the model in eq. (E.3). We compute the estimates of the parameters in this model using Gomez (2017), which implements the estimation approach in Bai (2009). To compute these estimates, we use annual data on exports from China to high-income countries other than the US, IMP_{st}, between 1991 and 2007 (i.e. the same sectoral exports used to construct the instrumental variable in Autor, Dorn and Hanson (2013)) for all sectors used in our baseline placebo exercise. We use these estimates to construct a treatment variable $X_i^m \equiv \sum_s w_{is} \mathcal{X}_s^m$, with each $\mathcal{X}_s^m$ defined as in eq. (E.4), for every simulated sample $m = 1, \ldots, 30{,}000$. The different panels impose different assumptions on the distribution of (η_s^m, u_s^m) across sectors and simulated samples. In Panels A and B in Table E.2, we fix $\eta_s^m = \hat{\eta}_s$ for every sector s and placebo sample m. The placebo simulations whose results we present in these two panels differ in the distribution from which u_s^m is drawn. In Panel A, we draw u_s^m independently across sectors and placebo samples either from a normal distribution with mean zero and variance equal to five. In Panel B, we draw u_s^m independently from the distribution of $\hat{\varepsilon}_{s,2007} - \hat{\varepsilon}_{s,1991}$ across sectors, where, for $t = 2007$ and $t = 1991$, $\hat{\varepsilon}_{st}$ is the estimate of the term ε_{st} in eq. (E.3) (in Panel B). The placebo exercises in Panel C of Table E.2 differs from that in Panel B in that, in the former, each η_s^m is independently drawn across sectors s and placebo samples m from the distribution of $\hat{\eta}_s$ across sectors, where $\hat{\eta}_s$ is our estimate of the term η_s in eq. (E.3). In all three panels, we compute the outcome variable as $Y_i^m = \sum_s w_{is} \mu_s^m$, with μ_s^m drawn randomly from a normal distribution with mean zero and variance equal to 5. Given the variables Y_i^m and X_i^m for each simulated sample m, we compute an estimate of β in the regression $Y_i = \beta X_i^m + \epsilon_i$ (whenever there is a 'No' in the second column) or in the regression $Y_i = \beta X_i^m + \gamma \sum_s \eta_s^m \Delta\hat{\hat{X}}_{Ch} + \epsilon_i$ (whenever there is a 'Yes' in the second column). We indicate the median and standard deviation of the OLS estimates of β across the simulated samples (columns (1) and (2)), the median effective standard error estimates (columns (3) to (6)), and the percentage of placebo samples for which we reject the null hypothesis $H_0\colon \beta = 0$ using a 5% significance level test (columns (7) to (10)). *Robust* is the Eicker-Huber-White standard error; *Cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5; *AKM0* is the confidence interval in Remark 6.

Table E.3: Placebo exercise for the first-stage regression in Autor, Dorn and Hanson (2013)

Estimate		Median eff. s.e.				Rejection rate $H_0: \beta = 0$			
Mean (1)	Std. dev (2)	Robust (3)	Cluster (4)	AKM (5)	AKM0 (6)	Robust (7)	Cluster (8)	AKM (9)	AKM0 (10)
Panel A: No controls									
0.01	1.73	0.72	0.81	1.63	1.88	41.5%	36.7%	6.5%	4.0%
Panel B: Controls: ADH IV									
0.01	1.01	0.63	0.63	0.93	1.06	20.6%	21.3%	7.8%	4.3%
Panel C: Controls: ADH IV and all controls included in Table 3, col. 6 of in Autor et al. (2013)									
0.00	0.68	0.51	0.51	0.64	0.72	14.4%	14.1%	5.6%	3.8%

Notes: This table indicates the median and standard deviation of the OLS estimates of β in eq. (1) across the placebo samples (columns (1) and (2)), the median effective standard error estimates (columns (3) to (6)), and the percentage of placebo samples for which we reject the null hypothesis $H_0: \beta = 0$ using a 5% significance level test (columns (7) to (10)). *Robust* is the Eicker-Huber-White standard error; *Cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5; *AKM0* is the confidence interval in Remark 6. For each inference procedure, the median effective standard error is equal to the median length of the corresponding 95% confidence interval divided by 2×1.96. Results are based on 30,000 placebo samples. In all three panels, each $\mathcal{X}_s^m$ is *i.i.d* drawn from a normal distribution with mean zero and variance equal to 5. In Panel A, we introduce no controls in the regression equation. In Panel B, we control for the instrumental variable used in Autor, Dorn and Hanson (2013); i.e. the shift-share aggregator of changes in sectoral exports from China to high-income countries other than the US. In Panel C, we control for the instrumental variable used in Autor, Dorn and Hanson (2013) and for the broadest set of controls used in that paper; i.e. the set of controls used in column 6 of Table 3 of Autor, Dorn and Hanson (2013).

sector in different periods (which we denote in Tables E.4 and E.5 as *AKM (3d cluster)* and *AKM0 (3d cluster)*), Tables E.4 and E.5 also present *AKM* and *AKM0* confidence intervals that only cluster on time (denoted as *AKM (4d cluster)* and *AKM0 (4d cluster)*), and *AKM* and *AKM0* that treat shifters as independent both across 4-digit sectors and across time periods (denoted as *AKM (indep.)* and *AKM0 (indep.)*)

There are several takeaways from the results in Tables E.4 and E.5. First, accounting for the possible correlation in the shifters has only a minimal impact on the *AKM* confidence intervals (i.e. the *AKM (indep.)*, *AKM (4d cluster)*, and *AKM (3d cluster)* confidence intervals are always very similar); the impact on the *AKM0* confidence intervals is a bit larger but also quite small. Second, while the *AKM* and *AKM0* confidence intervals are quite similar to the *Robust* and *Cluster* ones in the case of college graduates (Panel B), they are much larger for non-college graduates (Panel C). Finally, similarly to what we observed in Table 5 in Section 7.1, the *AKM0* confidence interval is not centered around the point estimate: it includes more values of the parameter to the left of the point estimate than it does to the right.

Table E.4: 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	Manuf.	Non-Manuf.	All	Manuf.	Non-Manuf.
	(1)	(2)	(3)	(4)	(5)	(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]	[-0.50,0.69]	[-0.83,-0.13]
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.)	[-0.79,-0.18]	[-0.52,-0.24]	[-0.33,0.10]	[-0.84,-0.12]	[-0.47,0.66]	[-0.88,-0.08]
AKM0 (indep.)	[-1.08,-0.25]	[-0.63,-0.26]	[-0.51,0.07]	[-1.08,-0.15]	[-0.91,0.58]	[-1.22,-0.15]
AKM (4d cluster)	[-0.79,-0.19]	[-0.52,-0.23]	[-0.33,0.10]	[-0.87,-0.09]	[-0.49,0.68]	[-0.90,-0.07]
AKM0 (4d cluster)	[-1.10,-0.26]	[-0.66,-0.25]	[-0.52,0.07]	[-1.16,-0.13]	[-0.99,0.59]	[-1.28,-0.14]
AKM (3d cluster)	[-0.81,-0.17]	[-0.52,-0.23]	[-0.35,0.12]	[-0.88,-0.07]	[-0.50,0.69]	[-0.93,-0.03]
AKM0 (3d cluster)	[-1.24,-0.24]	[-0.67,-0.25]	[-0.64,0.08]	[-1.27,-0.10]	[-1.16,0.61]	[-1.47,-0.11]
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]
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,-0.09]	[-0.50,-0.25]	[-0.03,0.24]	[-0.82,-0.13]	[-0.11,0.69]	[-0.83,-0.11]
AKM0 (indep.)	[-0.57,-0.11]	[-0.56,-0.24]	[-0.11,0.24]	[-1.00,-0.13]	[-0.35,0.68]	[-1.07,-0.14]
AKM (4d cluster)	[-0.45,-0.09]	[-0.51,-0.23]	[-0.04,0.25]	[-0.85,-0.10]	[-0.14,0.72]	[-0.85,-0.09]
AKM0 (4d cluster)	[-0.58,-0.11]	[-0.59,-0.23]	[-0.11,0.25]	[-1.08,-0.11]	[-0.41,0.70]	[-1.14,-0.13]
AKM (3d cluster)	[-0.45,-0.08]	[-0.52,-0.23]	[-0.04,0.25]	[-0.88,-0.08]	[-0.14,0.72]	[-0.89,-0.05]
AKM0 (3d cluster)	[-0.62,-0.09]	[-0.59,-0.20]	[-0.17,0.25]	[-1.20,-0.08]	[-0.46,0.73]	[-1.32,-0.09]
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]
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.)	[-1.18,-0.22]	[-0.55,-0.19]	[-0.68,0.01]	[-1.08,0.05]	[-0.70,0.57]	[-1.15,0.11]
AKM0 (indep.)	[-1.68,-0.34]	[-0.72,-0.23]	[-1.01,-0.06]	[-1.59,-0.06]	[-1.26,0.45]	[-1.78,-0.04]
AKM (4d cluster)	[-1.17,-0.23]	[-0.55,-0.18]	[-0.67,0.00]	[-1.09,0.06]	[-0.70,0.57]	[-1.14,0.11]
AKM0 (4d cluster)	[-1.69,-0.35]	[-0.74,-0.23]	[-1.01,-0.07]	[-1.64,-0.05]	[-1.30,0.45]	[-1.80,-0.04]
AKM (3d cluster)	[-1.22,-0.18]	[-0.55,-0.18]	[-0.71,0.04]	[-1.10,0.07]	[-0.72,0.60]	[-1.16,0.13]
AKM0 (3d cluster)	[-1.95,-0.32]	[-0.79,-0.23]	[-1.21,-0.04]	[-1.80,-0.04]	[-1.55,0.46]	[-2.02,-0.02]

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; i.e. those in column 6 of Table 3 in Autor, Dorn and Hanson (2013). 95% confidence intervals in square brackets. *Robust* is the Eicker-Huber-White standard error; *Cluster* is the standard error that clusters of CZs in the same state; *AKM (indep.)* is the standard error in Remark 5; *AKM (4d cluster)* is the standard error in eq. (40) with 4-digit SIC clusters; *AKM (3d cluster)* is the standard error in eq. (40) with 3-digit SIC clusters; *AKM0 (indep.)* is the confidence interval in Remark 6; *AKM0 (4d cluster)* is the confidence interval with 4-digit SIC clusters described in the last sentence of Section 5.1; and *AKM0 (3d cluster)* is the confidence interval with 3-digit SIC clusters described in the last sentence of Section 5.1.

Table E.5: 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]
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.)	[-1.19,-0.36]	[-0.81,-0.38]	[-0.50,0.15]	[-1.30,-0.22]	[-0.76,1.06]	[-1.32,-0.20]
AKM0 (indep.)	[-1.40,-0.42]	[-0.89,-0.39]	[-0.65,0.11]	[-1.48,-0.23]	[-1.14,0.99]	[-1.58,-0.25]
AKM (4d cluster)	[-1.19,-0.36]	[-0.84,-0.36]	[-0.50,0.15]	[-1.35,-0.17]	[-0.80,1.10]	[-1.36,-0.17]
AKM0 (4d cluster)	[-1.46,-0.43]	[-0.96,-0.38]	[-0.66,0.12]	[-1.61,-0.21]	[-1.24,1.03]	[-1.69,-0.24]
AKM (3d cluster)	[-1.25,-0.30]	[-0.84,-0.35]	[-0.54,0.18]	[-1.37,-0.15]	[-0.81,1.11]	[-1.42,-0.10]
AKM0 (3d cluster)	[-1.69,-0.39]	[-1.01,-0.36]	[-0.84,0.14]	[-1.77,-0.17]	[-1.49,1.05]	[-1.97,-0.19]
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]
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.)	[-0.69,-0.16]	[-0.83,-0.36]	[-0.07,0.40]	[-1.30,-0.22]	[-0.22,1.14]	[-1.28,-0.20]
AKM0 (indep.)	[-0.78,-0.16]	[-0.87,-0.33]	[-0.14,0.40]	[-1.44,-0.19]	[-0.45,1.13]	[-1.47,-0.21]
AKM (4d cluster)	[-0.70,-0.15]	[-0.85,-0.33]	[-0.07,0.41]	[-1.34,-0.17]	[-0.27,1.18]	[-1.31,-0.18]
AKM0 (4d cluster)	[-0.82,-0.17]	[-0.93,-0.32]	[-0.15,0.42]	[-1.56,-0.17]	[-0.53,1.18]	[-1.57,-0.21]
AKM (3d cluster)	[-0.71,-0.13]	[-0.86,-0.32]	[-0.08,0.42]	[-1.37,-0.14]	[-0.25,1.17]	[-1.37,-0.11]
AKM0 (3d cluster)	[-0.90,-0.14]	[-0.96,-0.27]	[-0.23,0.42]	[-1.71,-0.13]	[-0.61,1.21]	[-1.82,-0.15]
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]
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.)	[-1.76,-0.47]	[-0.83,-0.33]	[-1.02,-0.04]	[-1.62,0.00]	[-1.09,0.89]	[-1.71,0.07]
AKM0 (indep.)	[-2.12,-0.58]	[-0.95,-0.37]	[-1.27,-0.11]	[-2.01,-0.10]	[-1.55,0.79]	[-2.20,-0.07]
AKM (4d cluster)	[-1.75,-0.47]	[-0.85,-0.32]	[-1.01,-0.05]	[-1.64,0.02]	[-1.10,0.90]	[-1.72,0.07]
AKM0 (4d cluster)	[-2.19,-0.59]	[-1.02,-0.36]	[-1.29,-0.12]	[-2.12,-0.09]	[-1.63,0.79]	[-2.28,-0.07]
AKM (3d cluster)	[-1.86,-0.36]	[-0.86,-0.30]	[-1.09,0.03]	[-1.68,0.06]	[-1.14,0.93]	[-1.78,0.14]
AKM0 (3d cluster)	[-2.62,-0.52]	[-1.13,-0.35]	[-1.59,-0.07]	[-2.43,-0.07]	[-2.00,0.79]	[-2.69,-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; i.e. those in column 6 of Table 3 in Autor, Dorn and Hanson (2013). 95% confidence intervals in square brackets. *Robust* is the Eicker-Huber-White standard error; *Cluster* is the standard error that clusters of CZs in the same state; *AKM (indep.)* is the standard error in eq. (39); *AKM (4d cluster)* is the standard error in eq. (39) with an adjustment analogous to that in eq. (40) with 4-digit SIC clusters; *AKM (3d cluster)* is the standard error in eq. (39) with an adjustment analogous to that in eq. (40) with d-digit SIC clusters; *AKM0 (indep.)* is the confidence interval built using the standard error in eq. (39) with the residual $(I - Z'(Z'Z)^{-1}Z')(Y_1 - Y_2\alpha_0)$ instead of the estimate $\hat{\epsilon}_\Delta = (I - Z'(Z'Z)^{-1}Z')(Y_1 - Y_2\hat{\alpha})$; *AKM0 (4d cluster)* and *AKM0 (3d cluster)* impose the same adjustment to the procedure in *AKM (4d cluster)* and *AKM (3d cluster)*, respectively.

E.2 Estimation of inverse labor supply elasticity

Shift-share IV regressions have been used extensively to estimate inverse local labor supply elasticities. Using the notation in Section 3, we can write the inverse labor supply in each region i as

$$\log \omega_i = \tilde{\phi} \log L_i - \tilde{\phi} \log v_i, \qquad \text{with} \qquad \tilde{\phi} \equiv \phi^{-1}, \tag{E.5}$$

and, consequently, we can relate log changes in wages and log changes employment rates (or number of employees) for each region i between any two time periods as

$$\hat{\omega}_i = \tilde{\phi}\hat{L}_i - \tilde{\phi}(\sum_{g=1}^{G} \tilde{w}_{ig}\hat{v}_g + \hat{v}_i). \tag{E.6}$$

E.2.1 Bias in OLS estimate of inverse labor supply elasticity

Using data on log changes in wages and employment rates for a set of regions, $\{(\hat{\omega}_i, \hat{L}_i)\}_i$, one may consider using OLS to compute an estimate of $\tilde{\phi}$. However, such estimator will be inconsistent. To show this formally, note that, up to a first-order approximation around the initial equilibrium, we can write the change in employment in any given region i as

$$\hat{L}_i = \sum_{s=1}^{S} l_{is}^0 \left[\theta_{is}\hat{\chi}_s + \lambda_i\hat{\mu}_s + \lambda_i\hat{\eta}_{is}\right] + (1-\lambda_i)\,(\sum_{g=1}^{G} \tilde{w}_{ig}\hat{v}_g + \hat{v}_i), \tag{E.7}$$

and the change in wages as

$$\hat{\omega}_i = \tilde{\phi} \sum_{s=1}^{S} l_{is}^0 (\theta_{is}\hat{\chi}_s + \lambda_i\hat{\mu}_s + \lambda_i\hat{\eta}_{is}) - \tilde{\phi}\lambda_i(\sum_{g=1}^{G} \tilde{w}_{ig}\hat{v}_g + \hat{v}_i). \tag{E.8}$$

Using eq. (E.6), the probability limit of the OLS estimator of $\tilde{\phi}$, $\hat{\tilde{\phi}}_{OLS}$, can be written as

$$plim(\hat{\tilde{\phi}}_{OLS}) = \frac{cov(\hat{\omega}_i, \hat{L}_i)}{var(\hat{L}_i)} = \tilde{\phi} + \frac{cov(-\tilde{\phi}(\sum_{g=1}^{G} \tilde{w}_{ig}\hat{v}_g + \hat{v}_i), \hat{L}_i)}{var(\hat{L}_i)}, \tag{E.9}$$

where $cov(-\tilde{\phi}(\sum_{g=1}^{G} \tilde{w}_{ig}\hat{v}_g + \hat{v}_i), \hat{L}_i)/var(\hat{L}_i)$ captures the asymptotic bias in $\hat{\tilde{\phi}}_{OLS}$ as an estimator of $\tilde{\phi}$. To characterize this term, we assume here that the of labor supply shocks $\{\hat{v}_g\}_g$ and $\{\hat{v}_i\}_i$ are independent of the vector of all labor demand shocks $(\{\hat{\chi}_s\}_s, \{\hat{\mu}_s\}_s, \{\hat{\eta}_{is}\}_{i,s})$ conditional on the matrix of weights $W \equiv \{l_{is}^0\}_{i,s}$ and the matrix of parameters $B \equiv (\{\beta_{is}\}_{i,s}, \{\lambda_i\}_i)$

$$(\{\hat{\chi}_s\}_s, \{\hat{\mu}_s\}_s, \{\hat{\eta}_{is}\}_{i,s}) \perp\!\!\!\perp (\{\hat{v}_g\}_g, \{\hat{v}_i\}_i) \mid (W, B). \tag{E.10}$$

Given this assumption and eq. (E.7), we can rewrite $plim(\hat{\tilde{\phi}}_{OLS})$ in eq. (E.9) as

$$plim(\hat{\tilde{\phi}}_{OLS}) = \tilde{\phi} + \frac{cov(-\tilde{\phi}(\sum_{g=1}^{G} \tilde{w}_{ig}\hat{v}_g + \hat{v}_i), (1-\lambda_i)(\sum_{g=1}^{G} \tilde{w}_{ig}\hat{v}_g + \hat{v}_i))}{var(\hat{L}_i)}$$

$$= \tilde{\phi} - \tilde{\phi}(1-\lambda)\frac{var(\sum_{g=1}^{G}\tilde{w}_{ig}\hat{v}_g + \hat{v}_i)}{var(\hat{L}_i)}, \tag{E.11}$$

where the second equality follows if we additionally assume that elasticity of labor demand in eq. (2) does not vary across sectors, $\sigma_s = \sigma$ for all s, so that $\lambda_i = \lambda$ for all i. As indicated in Section 3.2, in this case, $\lambda \equiv \phi[\phi + \sigma \sum_{s=1}^{S} l_{is}^0]^{-1}$. Thus, if $\sigma > 0$ (which guarantees that $\lambda < 1$) and $\tilde{\phi} > 0$, then the OLS will underestimate the inverse labor supply elasticity in the sense that $plim(\hat{\tilde{\phi}}_{OLS}) < \tilde{\phi}$.

E.2.2 Consistency of IV estimate of inverse labor supply elasticity

Using data for a set of regions and sectors on log changes in wages and employment rates $\{(\hat{\omega}_i, \hat{L}_i)\}_i$, initial employment shares $\{l_{is}^0\}_{i,s}$, and sectoral labor demand shifters $\{\hat{\chi}_s\}_s$, we can write the probability limit of the IV estimator of $\tilde{\phi}$ that uses $X_i \equiv \sum_{s=1}^{S} l_{is}^0 \hat{\chi}_s$ as IV, $\hat{\tilde{\phi}}_{IV}$, as

$$plim(\hat{\tilde{\phi}}_{IV}) = \frac{cov(\hat{\omega}_i, X_i)}{cov(\hat{L}_i, X_i)}.$$

Given the expressions for $\hat{L}_i$ and $\hat{\omega}_i$ in eqs. (E.7) and (E.8), respectively, and the independence assumption in eq. (E.10), we can rewrite

$$\begin{aligned}
plim(\hat{\tilde{\phi}}_{IV}) &= \frac{cov(\tilde{\phi}\sum_{s=1}^{S} l_{is}^0\left[\theta_{is}\hat{\chi}_s + \lambda_i\hat{\mu}_s + \lambda_i\hat{\eta}_{is}\right] - \tilde{\phi}\lambda_i(\sum_{g=1}^{G}\tilde{w}_{ig}\hat{v}_g + \hat{v}_i), X_i)}{cov(\sum_{s=1}^{S} l_{is}^0\left[\theta_{is}\hat{\chi}_s + \lambda_i\hat{\mu}_s + \lambda_i\hat{\eta}_{is}\right] + (1-\lambda_i)\left(\sum_{g=1}^{G}\tilde{w}_{ig}\hat{v}_g + \hat{v}_i\right), X_i)} \\
&= \tilde{\phi}\frac{cov(\sum_{s=1}^{S} l_{is}^0\left[\theta_{is}\hat{\chi}_s + \lambda_i\hat{\mu}_s + \lambda_i\hat{\eta}_{is}\right], X_i)}{cov(\sum_{s=1}^{S} l_{is}^0\left[\theta_{is}\hat{\chi}_s + \lambda_i\hat{\mu}_s + \lambda_i\hat{\eta}_{is}\right], X_i)} \\
&= \tilde{\phi}.
\end{aligned}$$

Therefore, under the distributional assumptions in eq. (E.10), the IV estimator that uses a shift-share instrument that aggregates sector-specific labor demand shifters is a consistent estimator of the inverse labor supply elasticity. Notice that the heterogeneity in θ_{is} does not affect the consistency of $\hat{\tilde{\phi}}$. However, the consistency of $\hat{\tilde{\phi}}$ will depend on the specific labor demand shock being employed by the researcher to construct its shift-share IV being independent of the specific labor supply shocks that have been prevalent in the set of regions belonging the population of interest.

E.2.3 Evaluation of leave-one-out IV through the lens of the model in Section 3

We describe in this section how one may use the model in Section 3 to frame the approach to the estimation of the inverse labor supply elasticity described in Section 7.2. This approach is described in general terms in Section 5.3.

In Section 7.2, we focus on the estimation of the inverse labor supply elasticity $\tilde{\phi}$ and we base the estimation of this parameter on the estimating equation

$$\hat{\omega}_i = \tilde{\phi}\hat{L}_i + \delta Z_i + \epsilon_i, \qquad \text{with} \qquad \tilde{\phi} = \phi^{-1}. \tag{E.12}$$

For simplicity, we assume here that we use no controls (i.e. $\delta = 0$) and that, thus, we can rewrite the estimating equation above as

$$\hat{\omega}_i = \tilde{\phi}\hat{L}_i + \epsilon_i, \qquad \text{with} \qquad \tilde{\phi} = \phi^{-1}. \tag{E.13}$$

The advantage of focusing on the version without controls is that, in this case, the model in Section 3 clarifies that $\epsilon_i = -\tilde{\phi}(\sum_{g=1}^{G} \tilde{w}_{ig}\hat{v}_g + \hat{v}_i)$, where $\{\hat{v}_g\}_g$ and $\{\hat{v}_i\}_i$ are labor supply shocks. Thus, in the version without controls, there is a clear mapping between the regression residual of the structural equation, ϵ_i, and the labor supply shocks in our economic model.

As discussed in Appendix E.2.1, the OLS estimator of $\tilde{\phi}$ will be biased. However, as discussed in Appendix E.2.2, one may obtain a consistent estimate of $\tilde{\phi}$ by computing an IV estimator that instruments for the log change in employment in region i, $\hat{L}_i$, using as an instrument a shift-share aggregator of labor demand shocks $\{\mathcal{X}_s\}_s$. In terms of the model in Section 3, $\mathcal{X}_s$ is any (possibly sector s-specific) function of the sector s-specific labor demand shocks χ_s and μ_s (see eqs. (2) and (3)). These sector-specific labor demand shocks are in many cases unobserved to the researcher. In these cases, following Bartik (1991) and the subsequent literature on the estimation of inverse local labor supply elasticities, it has become typical to estimate $\tilde{\phi}$ using as instruments one of two different IVs: either a shift-share aggregator of the growth in national employment in every sector s,

$$X_i = \sum_{s=1}^{S} l_{is}^0 \hat{L}_s, \qquad \text{with} \qquad \hat{L}_s = \sum_{j=1}^{N} \frac{L_{js}^0}{\sum_{j'=1}^{N} L_{j's}^0} \frac{L_{js}^t - L_{js}^0}{L_{js}^0}, \tag{E.14}$$

or a shift-share aggregator of the leave-one-out measure of the growth in national employment in sector s,

$$X_{i,-} = \sum_{s=1}^{S} l_{is}^0 \hat{L}_{s,-i}, \qquad \text{with} \qquad \hat{L}_{s,-i} = \sum_{j=1, j\neq i}^{N} \frac{L_{js}^0}{\sum_{j'=1,j'\neq i}^{N} L_{j's}^0} \frac{L_{js}^t - L_{js}^0}{L_{js}^0}. \tag{E.15}$$

We focus here on outlining the restrictions that one should impose on the sector-specific labor demand shifters $\{(\hat{\chi}_s, \hat{\mu}_s)\}_s$, region- and sector-specific labor demand shifters $\{\hat{\eta}_{is}\}_{i,s}$, group-specific labor supply shifters, $\{\hat{v}_g\}_g$, and region-specific labor supply shifters $\{\hat{v}_i\}_i$ (all of them introduced in the model in Section 3) so that the IV estimator that uses $X_{i,-}$ as an instrument yields a consistent estimate of $\tilde{\phi}$.

The variable $X_{i,-}$ in eq. (E.15) is a valid instrument as long as we can write

$$\hat{L}_{is} = \mathcal{X}_s + \psi_{is}, \tag{E.16}$$

and the following restrictions hold

$$E[\mathcal{X}_s | \hat{\omega}(0), \hat{L}(0), L^0] = E[\mathcal{X}_s], \qquad \text{for all } s, \tag{E.17}$$

$$E[l_{is}^0 \psi_{is} | \hat{\omega}_{-i}(0), \hat{L}_{-i}(0), L^0] = 0, \qquad \text{for all } i \text{ and } s, \tag{E.18}$$

$$E[l_{is}^0 \psi_{is} l_{js}^0 \psi_{js} | \hat{\omega}_{-i}(0), \hat{L}_{-i}(0), L^0] = 0, \qquad \text{for all } i \neq j \text{ and } s, \tag{E.19}$$

where $\hat{\omega}_{-i}(0)$ ($\hat{L}_{-i}(0)$) denotes the change in wages (employment shares) in every region other than i when the sectoral shock of interest equals 0 for all sectors (i.e. $\mathcal{X}_s = 0$ for all s), and L^0 is the vector all region- and sector-specific shares in the initial equilibrium (i.e. $L^0 = \{l^0_{is}\}_{i,s}$).

According to the model in Section 3, we can express the changes in employment in sector s in a region i as

$$\hat{L}_{is} = -\sigma_s \hat{\omega}_i + \rho_s \hat{\chi}_s + \hat{\mu}_s + \hat{\eta}_{is}.$$

Combining this expression with the expression for $\hat{\omega}_i$ in eq. (E.8) in Appendix E.2.1, we can rewrite the change in employment in sector s and region i approximately as

$$\hat{L}_{is} = -\sigma_s \tilde{\phi} \sum_{s'=1}^{S} l^0_{is'} \left[\theta_{is'} \hat{\chi}_{s'} + \lambda_i \hat{\mu}_{s'} + \lambda_i \hat{\eta}_{is'}\right] + \sigma_s \tilde{\phi} \lambda_i \left(\sum_{g=1}^{G} \tilde{\omega}_{ig} \hat{v}_g + \hat{v}_i\right) + \rho_s \hat{\chi}_s + \hat{\mu}_s + \hat{\eta}_{is}, \tag{E.20}$$

with $\lambda_i \equiv \phi \left[\phi + \sum_{s=1}^{S} l^0_{is} \sigma_s\right]^{-1}$, $\theta_{is} \equiv \rho_s \lambda_i$, and $\tilde{\phi} = \phi^{-1}$.

Without imposing any restrictions on the values of the labor demand and supply elasticities, the expression for $\hat{L}_{is}$ in eq. (E.20) will not satisfy the restrictions in eq. (E.16) to eq. (E.19). To illustrate this point, we can map the different terms in eq. (E.20) into those in eq. (E.16) as

$$\mathcal{X}_s = \rho_s \hat{\chi}_s + \hat{\mu}_s, \tag{E.21}$$

$$\psi_{is} = -\sigma_s \tilde{\phi} \sum_{s'=1}^{S} l^0_{is'} \left[\theta_{is'} \hat{\chi}_{s'} + \lambda_i \hat{\mu}_{s'} + \lambda_i \hat{\eta}_{is'}\right] + \sigma_s \tilde{\phi} \lambda_i \left(\sum_{g=1}^{G} \tilde{\omega}_{ig} \hat{v}_g + \hat{v}_i\right) + \hat{\eta}_{is}. \tag{E.22}$$

Under this definition of the labor demand shock $\mathcal{X}_s$, the potential outcomes $\hat{\omega}_i(0)$ and $\hat{L}_i(0)$ are

$$\hat{\omega}_i(0) = \tilde{\phi} \sum_{s=1}^{S} l^0_{is} \lambda_i \hat{\eta}_{is} - \tilde{\phi} \lambda_i \left(\sum_{g=1}^{G} \tilde{\omega}_{ig} \hat{v}_g + \hat{v}_i\right), \tag{E.23}$$

$$\hat{L}_i(0) = \sum_{s=1}^{S} l^0_{is} \lambda_i \hat{\eta}_{is} + (1 - \lambda_i) \left(\sum_{g=1}^{G} \tilde{\omega}_{ig} \hat{v}_g + \hat{v}_i\right). \tag{E.24}$$

Given the expressions in eqs. (E.22) to (E.24), the restriction on ψ_{is} in eq. (E.19) will not be satisfied: for any two regions i and i', ψ_{is} and $\psi_{i's}$ are a function of the same set of sectoral demand shocks $\{\hat{\chi}_s\}_s$ and $\{\hat{\mu}_s\}_s$ and, thus, ψ_{is} and $\psi_{i's}$ will generally be correlated with each other. Thus, unless additional restrictions are imposed, the IV estimator that uses the variable described in eq. (E.15) as instrument for $\hat{L}_i$ in eq. (E.13) will not be a consistent estimator of $\tilde{\phi}$.

However, under the restriction that $\sigma_s = 0$ for every sector s, the expression for $\hat{L}_{is}$ in eq. (E.20) will satisfy the restrictions in eq. (E.16) to eq. (E.19). In this case,

$$\mathcal{X}_s = \rho_s \hat{\chi}_s + \hat{\mu}_s, \tag{E.25}$$

$$\psi_{is} = \hat{\eta}_{is}, \tag{E.26}$$

and $\hat{\omega}_i(0)$ and $\hat{L}_i(0)$ correspond to the expressions in eq. (E.23) and eq. (E.24). Thus, if the sector-specific labor demand shocks $\{(\hat{\chi}_s, \hat{\mu}_s)\}_s$ are mean independent of the region-specific labor supply shocks $\{\hat{v}_g\}_g$ and $\{\hat{v}_i\}_i$ as well as of the region- and sector-specific labor demand shocks $\{\hat{\eta}_{is}\}_{i,s}$, the restriction in eq. (E.17) will hold. Additionally, under the additional assumption that η_{is} is mean zero and uncorrelated with η_{js} for every $i \neq j$ and s, the restrictions in eqs. (E.18) and (E.19) will hold. Thus, if these additional restrictions on the model in Section 3 hold, the IV estimator that uses the variable described in eq. (E.15) as instrument to estimate $\tilde{\phi}$ in eq. (E.13) will be consistent.

There are two alternative instrumental variables that do not use data on any specific labor demand shock and that lead to consistent estimates of the inverse labor supply elasticity $\tilde{\phi}$ under weaker restrictions than those needed for the instrument in eq. (E.15) to be valid.

First, conditional on a calibrated value of σ_s for every sector s, one may estimate $\tilde{\phi}$ using as an instrument for $\hat{L}_i$ the following leave-one-out estimator:

$$\tilde{X}_{i,-} = \sum_{s=1}^{S} l_{is}^0 \hat{\hat{L}}_{s,-i}, \qquad \text{with} \qquad \hat{\hat{L}}_{s,-i} = \sum_{j=1, j\neq i}^{N} \frac{L_{js}^0}{\sum_{j'=1,j'\neq i}^{N} L_{j's}^0} \hat{\hat{L}}_{js}, \quad \text{and} \quad \hat{\hat{L}}_{is} = \hat{L}_{is} - \sigma_s \hat{\omega}_i. \tag{E.27}$$

Combining the expression for $\hat{L}_{is}$ in eq. (E.20) and the expression for $\hat{\omega}_i$ in eq. (E.8), we can write

$$\hat{\hat{L}}_{is} = \rho_s \hat{\chi}_s + \hat{\mu}_s + \hat{\eta}_{is}. \tag{E.28}$$

Thus, we can define $\mathcal{X}_s$ and ψ_{is} as in eq. (E.25) and eq. (E.26). Consequently, as discussed above, eqs. (E.17) to (E.19) will hold if: (a) the sector-specific labor demand shocks $\{(\hat{\chi}_s, \hat{\mu}_s)\}_s$ are mean independent of the region-specific labor supply shocks $\{\hat{v}_g\}_g$ and $\{\hat{v}_i\}_i$ as well as of the region- and sector-specific labor demand shocks $\{\hat{\eta}_{is}\}_{i,s}$; and (b) η_{is} is mean zero and uncorrelated with η_{js} for every $i \neq j$ and s. Thus, under these two sets of assumptions, the IV estimator that uses the variable described in eq. (E.27) as instrument for $\hat{L}_i$ in eq. (E.13) will be a consistent estimator of $\tilde{\phi}$ no matter what the value of the labor demand elasticities $\{\sigma_s\}_s$ is.

Second, under the assumption that the labor demand elasticity is constant across sectors (i.e. $\sigma_s = \sigma$ for every s), the residual from projecting $\hat{L}_{is}$, as defined in eq. (E.20), on a set of region-specific fixed effects is equivalent to $\hat{\hat{L}}_{is}$, as defined in eq. (E.28). Therefore, once we define $\mathcal{X}_s$ and ψ_{is} as in eq. (E.25) and eq. (E.26), the IV estimator that uses the variable described in eq. (E.27) as an instrument for $\hat{L}_i$ in eq. (E.13) will be a consistent estimator of $\tilde{\phi}$ if two assumptions hold: (a) the sector-specific labor demand shocks $\{(\hat{\chi}_s, \hat{\mu}_s)\}_s$ are mean independent of the region-specific labor supply shocks $\{\hat{v}_g\}_g$ and $\{\hat{v}_i\}_i$ as well as of the region- and sector-specific labor demand shocks $\{\hat{\eta}_{is}\}_{i,s}$; and (b) η_{is} is mean zero and uncorrelated with η_{js} for every i, j, and s.

E.2.4 Placebo exercise

In this section, we implement a placebo exercise to evaluate the finite-sample properties of our suggested inference procedures when using the shift-share IVs introduced in Section 5.3. For each placebo sample $m = 1, \ldots, 30{,}000$, we construct sector- and region-specific shocks $X_{is}^m = \mathcal{X}_s^m + \psi_{is}^m$, where $\mathcal{X}_s^m$ and ψ_{is}^m are independently drawn from normal distributions with variances equal to 5 and 10, respectively. We then use data on employment shares of U.S. CZs by 4-digit manufacturing sectors, $\{w_{is}\}_{i=1,s=1}^{N,S}$ to compute

$$Y_{i2} = \sum_{s=1}^{S} w_{is} X_{is}^{m}, \qquad\qquad Y_{i1} = \rho \sum_{s=1}^{S} w_{is} \psi_{is}^{m} + \sum_{s=1}^{S} w_{is} A_{s}^{m},$$

where A_s^m is independently drawn from a normal distribution with variance equal to 20.

Our goal is to estimate the effect α of Y_{i2}^m on Y_{i1}^m,

$$Y_{i1}^m = Y_{i2}^m \alpha + \epsilon_i^m. \tag{E.29}$$

Note that, by the above construction, $\alpha = 0$. Therefore, the residual is $\epsilon_i^m = Y_{i1}^m = \rho \sum_{s=1}^{S} w_{is} \psi_{is}^m + \sum_{s=1}^{S} w_{is} A_s^m$, which indicates that there is a potential endogeneity problem stemming from the fact that ψ_{is}^m affects both Y_{i1}^m and Y_{i2}^m whenever $\rho \neq 0$.

We consider three different shift-share IVs. First, we consider the IV constructed directly with the shock $\mathcal{X}_s^m$:

$$X_i^m = \sum_{s=1}^{S} w_{is} \mathcal{X}_s^m.$$

Second, we consider an IV constructed with the aggregate growth in X_{is}:

$$\hat{X}_i^m = \sum_{s=1}^{S} w_{is} \hat{\mathcal{X}}_s^m \quad \text{such that} \quad \hat{\mathcal{X}}_s^m \equiv \sum_{i=1}^{N} \left(\frac{\check{w}_{is}}{\sum_{j=1}^{N} \check{w}_{js}} \right) X_{is}^m$$

where $\check{w}_{is} = L_{is}^0 / \sum_{j=1}^{N} L_{js}^0$ is the share of CZ i in the national employment of sector s in 1990. Third, we consider an IV constructed with leave-one-out aggregate growth in X_{is}:

$$\hat{X}_{i,-}^m = \sum_{s=1}^{S} w_{is} \hat{\mathcal{X}}_{s,-i}^m \quad \text{such that} \quad \hat{\mathcal{X}}_{s,-i}^m \equiv \sum_{j=1, j\neq i}^{N} \left(\frac{\check{w}_{js}}{\sum_{o=1, j\neq i}^{N} \check{w}_{os}} \right) X_{js}^m.$$

The instruments X_i^m and $\hat{X}_{i,-}^m$ are always valid in our setting. However, whenever $\rho \neq 0$, the instrument $\hat{X}_i^m$ is invalid since $\{\psi_{is}^m\}_{s=1}^S$ affect $\hat{X}_i^m$ and ϵ_i.

Table E.6 reports the results of this placebo exercise for different values of ρ. In Panel A, we report results using X_i^m as an instrument; we denote this instrument as the "infeasible" IV, as its construction requires observing the shifters $\{\mathcal{X}_s\}_s$. As expected, for all values of ρ, the median $\hat{\alpha}^m$ across placebo samples is zero. Because of the shift-share structure of ϵ_i, robust and state-clustered standard error estimators underestimate the variability of the estimates, while *AKM* and *AKM0* inference procedures yield good coverage. Panel B presents the results based on the feasible shift-share IV $\hat{X}_i^m$. When using this IV, higher levels of ρ yield higher average estimates of α. This follows from the endogeneity problem created by the fact that $\{\psi_{is}^m\}_s$ are part of both the dependent variable, Y_{i1}^m, and the instrument, $\hat{X}_i^m$. Finally, Panel C presents results based on the leave-one-out IV $\hat{X}_{i,-}^m$. This instrument is not affected by an endogeneity problem, as it does not use information on region i-specific shocks $\{\psi_{is}^m\}_s$ when constructing the region i-specific variable $\hat{X}_{i,-}^m$. Thus, the average of the IV estimates of α that use $\hat{X}_{i,-}^m$ as an instrument is also very close to zero for all values of ρ. The results in Panel C also show

Table E.6: Mismeasured shifter: impact on standard errors and rejection rates.

ρ	Estimate		Median eff. s.e.						Rejection rate of $H_0\colon \beta = 0$					
	Median	eff. s.e.	Robust	Cluster	AKM	AKM0	AKM Leave-one-out	AKM0 Leave-one-out	Robust	Cluster	AKM	AKM0	AKM Leave-one-out	AKM0 Leave-one-out
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)
Panel A: Unfeasible shift-share IV														
0	0.00	0.35	0.16	0.18	0.29	0.37	–	–	0.33	0.29	0.09	0.04	–	–
5	0.01	0.97	0.77	0.76	0.80	1.05	–	–	0.08	0.09	0.08	0.04	–	–
10	−0.02	1.86	1.53	1.50	1.52	2.01	–	–	0.07	0.08	0.08	0.04	–	–
Panel B: Shift-share IV with aggregate sector-level growth														
0	0.00	0.23	0.11	0.12	0.20	0.22	–	–	0.33	0.29	0.09	0.04	–	–
5	1.66	0.50	0.39	0.40	0.45	0.51	–	–	0.89	0.89	0.85	0.79	–	–
10	3.30	0.93	0.77	0.77	0.84	0.95	–	–	0.91	0.90	0.88	0.83	–	–
Panel C: Shift-share IV with aggregate sector-level growth (leave-one-out)														
0	0.00	0.36	0.17	0.18	0.30	0.40	0.31	0.32	0.32	0.29	0.09	0.04	0.08	0.03
5	−0.07	1.06	0.82	0.80	0.87	1.24	0.93	0.95	0.07	0.08	0.06	0.05	0.05	0.03
10	−0.16	2.03	1.62	1.59	1.66	2.37	1.78	1.80	0.06	0.07	0.06	0.05	0.05	0.03

Notes: This table reports the median and effective standard error estimates of the IV estimates of α in eq. (E.29) across the placebo samples (columns (1) and (2)), the median effective standard error estimates (columns (3) to (8)), and the percentage of placebo samples for which we reject the null hypothesis $H_0\colon \beta = 0$ using a 5% significance level test (columns (9) to (14)). For each value of ρ, we generate 30,000 simulated samples with sector-region shocks $X_{is}^m = \mathcal{X}_s^m + \psi_{is}^m$, where $\mathcal{X}_s^m \sim \mathcal{N}(0,5)$ and $\psi_{is}^m \sim \mathcal{N}(0,10)$. We then construct $Y_{i2} = \sum_{s=1}^{S} w_{is} X_{is}^m$ and $Y_{i1} = \rho \sum_s w_{is}\psi_{is}^m + \sum_s w_{is} A_s^m$, where $A_s^m \sim \mathcal{N}(0,20)$. *Robust* is the Eicker-Huber-White standard error; *Cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5; *AKM0* is the confidence interval in Remark 6. *AKM (leave-one-out)* and *AKM0 (leave-one-out)* are the versions of *AKM* and *AKM0* in Section 5.3. For each inference procedure, the median effective standard error is equal to the median length of the corresponding 95% confidence interval divided by 2×1.96.

Table E.7: Estimation of inverse labor supply elasticity: robustness with different control sets

	(1)	(2)	(3)	(4)
Panel A: Bartik IV, Not leave-one-out estimator				
$\hat{\beta}$	0.75	0.8	0.83	0.8
Robust	$[0.48, 1.03]$	$[0.64, 0.97]$	$[0.56, 1.1]$	$[0.64, 0.96]$
Cluster	$[0.44, 1.07]$	$[0.60, 1.01]$	$[0.55, 1.11]$	$[0.59, 1.02]$
AKM	$[0.59, 0.92]$	$[0.62, 0.98]$	$[0.60, 1.06]$	$[0.62, 0.98]$
AKM0	$[0.56, 0.95]$	$[0.59, 1.02]$	$[0.61, 1.21]$	$[0.59, 1.01]$
Panel B: Bartik IV, Leave-one-out estimator				
$\hat{\beta}$	0.76	0.82	0.83	0.81
Robust	$[0.48, 1.03]$	$[0.65, 0.98]$	$[0.56, 1.10]$	$[0.65, 0.98]$
Cluster	$[0.43, 1.08]$	$[0.60, 1.03]$	$[0.55, 1.11]$	$[0.59, 1.04]$
AKM	$[0.59, 0.92]$	$[0.62, 1.01]$	$[0.58, 1.08]$	$[0.63, 1.00]$
AKM0	$[0.57, 0.96]$	$[0.60, 1.07]$	$[0.59, 1.28]$	$[0.60, 1.04]$
AKM (*leave-one-out*)	$[0.59, 0.92]$	$[0.61, 1.02]$	$[0.58, 1.08]$	$[0.62, 1.01]$
AKM0 (*leave-one-out*)	$[0.56, 0.97]$	$[0.59, 1.09]$	$[0.59, 1.29]$	$[0.59, 1.06]$
Controls:				
Period dummies	Yes	Yes	Yes	Yes
Controls in Autor et al. (2013)	No	Yes	No	Yes
Controls in Amior and Manning (2018)	No	No	Yes	Yes

Notes: $N = 1,444$ (722 CZs $\times$ 2 time periods). The dependent variable is the log-change in mean weekly earnings in CZ i, and the regressor is the log-change in the employment rate in CZ i. Observations are weighted by the 1980 CZ share of national population. 95% confidence intervals in square brackets. *Robust* is the Eicker-Huber-White standard error; *Cluster* is the standard error that clusters of CZs in the same state; *AKM* is the standard error in eq. (40) with 3-digit SIC clusters; *AKM0* is the confidence interval with 3-digit SIC clusters described in Section 5.1; *AKM (leave-one-out)* is the standard error in section 5.3 with 3-digit SIC clusters; *AKM0 (leave-one-out)* is the confidence interval with 3-digit SIC clusters described in section 5.3. Baseline controls in Autor et al. (2013) are the controls in column 6 of Table 3 in ADH. Amenity controls in Amior and Manning (2018): binary indicator for presence of coastline, three temperature indicators, log population density in 1900, log distance to the closest CZ.

that the leave-one-out versions of the *AKM* and *AKM0* inference procedures (see Section 5.3) yield slightly larger median effective standard errors than the baseline versions of the *AKM* and *AKM0* procedures (see Section 4.3). In this particular application, the magnitude of the adjustment is modest: the implied rejection rates for the null hypothesis $H_0\colon \alpha = 0$ differ by less than 2 percentage points.

E.2.5 Additional results

Table E.7 reports estimates of the inverse labor supply elasticity with alternative sets of controls. Column (2) replicates the estimates of Panels A and B in column (3) of Table 6. Table E.7 shows that these results are robust to controlling (a) only for period dummies (column (1)); (b) for period dummies and the proxies for region-specific labor supply shocks included in Amior and Manning (2018) (column (3)); (c) for period dummies, the controls included in Autor, Dorn and Hanson (2013) and the proxies for region-specific labor supply shocks in Amior and Manning (2018) (column (4)).

Appendix F Effect of immigration on U.S. local labor markets

To complement the empirical applications discussed in Section 7, we present here the results of estimating of the impact of immigration on labor market outcomes in the US. To this end, we estimate the model

$$Y_{it} = \beta \Delta ImmShare_{it} + Z_{it}' \delta + \epsilon_{it}, \tag{F.1}$$

where, for observation or cell i, Y_{it} is the change in a labor market outcome for native workers between years t and $t-10$, $\Delta ImmShare_{it}$ is the change in the share of immigrants in total employment between years t and $t-10$, and Z_{it} is a control vector that includes fixed effects.

Following Dustmann, Schönberg and Stuhler (2016), one may classify different approaches to the estimation of β in eq. (F.1) on the basis of the definition of the cell i: in the *skill-cell approach*, i corresponds to an education-experience cell defined at the national level (e.g. Borjas, 2003); in the *spatial approach*, i corresponds to a region (e.g. Altonji and Card, 1991); in the *mixed approach*, i corresponds to the intersection of a region and an occupation, or a region and an education group (e.g. Card, 2001).

In the *spatial* and *mixed* approaches, since Altonji and Card (1991) and Card (2001), it has become common to instrument for the change in the immigrant share $\Delta ImmShare_{it}$ using a shift-share IV:

$$X_{it} = \sum_{g=1}^{G} ImmShare_{igt_0} \frac{\Delta Imm_{gt}}{Imm_{gt_0}}, \tag{F.2}$$

where g indexes countries (or groups of countries) of origin of immigrants, and t_0 is some pre-sample or beginning-of-the-sample time period. The variable $ImmShare_{igt_0}$ plays the role of the share w_{is} in eq. (1) and denotes the share of immigrants from origin g in total immigrant employment in cell i in year t_0; the ratio $\Delta Imm_{gt}/Imm_{gt_0}$ plays the role of the shifter $\mathcal{X}_s$ in eq. (1), with ΔImm_{gt} denoting the change in the total number of immigrants coming from origin g between years t and $t-10$, and Imm_{gt_0} denoting the total number of immigrants from region g at the national level in year t_0.

When estimating the parameter of interest β in eq. (F.1), the researcher must make a choice on the sample period or time frame of the analysis, and on the G countries (or areas) of origin used to construct the shift-share IV. In Appendix F.1, we discuss two different sample periods previously used in the literature, and present a list of areas of origin of immigrants for which information is available in each of the two sample periods. In Appendix F.2, we present placebo evidence that illustrates the finite-sample properties of the different inference procedures when applied to the two sample periods discussed in Appendix F.1 and when using different sets of countries of origin of immigrants to construct the shift-share IV in eq. (F.2). The main conclusion that arises from these placebo simulations is that restricting the set of countries of origin used in the construction of the shift-share IV to those with a relatively small value of $ImmShare_{igt_0}$ generally improves the finite-sample coverage of all different inference procedures. Consequently, in Appendix F.3, we present estimates of β in eq. (F.1) that use information on a restricted set of countries when building the shift-share IV in eq. (F.2). For the sake of comparison, in Appendix F.4, we present estimates that use information on all countries for which information on immigration flows into the US is available for the relevant sample period.

In Appendices F.3 and F.4, we present estimates of specifications that follow either the *spatial approach* or the *mixed approach*. In all specifications, information on all variables entering eqs. (F.1) and (F.2) comes from the Census Integrated Public Use Micro Samples for 1980–2000 and the American Community Survey for 2008–2012. In all regressions, the vector of controls Z_{it} includes period dummies and, when implementing the *mixed approach*, we also add occupation- or education-group-specific dummies to the vector Z_i. All tables referenced in this section are included at the end.

F.1 Sample periods and list of countries of origin of immigrants

The results we present use one of two time frames. The first one uses information on immigrant shares (i.e. the variable $ImmShare_{igt_0}$ in eq. (F.2)) measured in 1980, and information on the outcome variables, endogenous treatment, and shifters of interest (i.e. the variables Y_{it}, $\Delta ImmShare_{it}$ and ΔImm_{gt} in eqs. (F.1) and (F.2)) for the periods 1980–1990, 1990–2000, and 2000–2010. Table F.1 lists all countries or areas of origin that we consider for which information on the number of immigrants in the U.S. is available for all periods in this time frame (i.e. 1980, 1990, 2000, and 2010).

The second time frames uses information on immigrant shares measured in 1960, and information on the outcome variables, endogenous treatment, and shifters of interest for the period 1970–1980. Table F.2 lists all countries or areas of origin that we consider for which information on the number of immigrants in the U.S. is available for all periods in this time frame (i.e. 1960, 1970 and 1980).

In both Tables F.1 and F.2, we have marked in italics those countries or areas of origin that account for a relatively large share (larger than 3%) of the overall immigrant U.S. population in the corresponding base year (this base year is 1980 for Table F.1 and 1960 for Table F.2).

F.2 Placebo simulations

In Tables F.3 and F.4, we present the results of placebo exercises that illustrate the properties of different inference procedures for the parameter on the shift-share covariate in eq. (F.2) in regressions of labor market outcomes for native workers on this shift-share covariate. The only difference between the analysis in Table F.3 and the analysis in Table F.4 is in the set of areas of origin of immigrants used to construct the shift-share covariate in eq. (F.2). While the former uses information only on those countries of origin whose total share of immigrants in the corresponding baseline year t_0 (either 1960 or 1980, depending on the specification) is below 3% (i.e. it uses information only on those countries of origin g that satisfy $\sum_{i=1}^{N} ImmShare_{igt_0} / \sum_{i=1}^{N} \sum_{g'=1}^{G} ImmShare_{ig't_0} \leq 0.03$), the latter uses information on all areas of origin of immigrants listed in the tables described in Appendix F.1.

We present results for four outcome variables: the change in employment ($\Delta \log E_i$) and average wages ($\Delta \log w_i$) across all native workers, and the change in average wages for high-skill and low-skill workers. For each of the four outcome variables, we consider several regressions in which we vary both the definition of a cell or unit of observation, and the sample period. The first four rows of each panel in Tables F.3 and F.4 implement a purely *spatial approach*, defining each unit of observation as a commuting zone (CZ) or as a metropolitan statistical area (MSA). The last four rows follow a *mixed approach*, defining each unit as the intersection of a CZ and either one of the fifty occupations defined in Burstein et al. (2018) (CZ-50 Occ.), one of seven aggregate occupations defined similarly to

Card (2001) (CZ-7 Occ.), or one of two education groups (CZ-Educ.). In terms of sample periods, we explore two alternatives. We either define the weights in 1980 and measure the outcome variable as the 1980–1990, 1990–2000, and 2000–2010 changes in log employment or log wages or, alternatively, we measure the weights in 1960 and measure the outcome variable as the 1970–1980 change in the variable of interest.

Tables F.3 and F.4 yield three key takeaways. First, robust standard errors are generally biased downward, leading frequently to an overrejection problem.

Second, when we construct the shift-share covariate in eq. (F.2) relying only on countries of origin with relatively small shares of U.S. immigrant population in the baseline year, state-clustered standard errors yield adequate rejection rates when the unit of observation is defined as the intersection of a CZ and fifty detailed occupation groups, shares are measured in 1980, and the outcome is defined as the subsequent three decadal changes. In all other cases, inference procedures based on state-clustered standard errors tend to overreject.

Third, the *AKM* and *AKM0* inference procedures perform much better when the shift-share covariate in eq. (F.2) is constructed using only countries of origin with relatively small shares of U.S. immigrant population in the baseline year, so that Assumptions 2(ii) and 2(iii) more plausibly hold. Furthermore, these inference procedures also tend to perform better in specifications that apply a *mixed approach* than in those that apply a purely *spatial approach*. One possible explanation for this pattern is that our asymptotics require that the number of observations $N \to \infty$; thus, the behavior of the *AKM* and *AKM0* inference procedures is generally better in samples with a larger number of observations, and the *mixed approach*, which intersects each region with several occupations or education groups, yields larger sample sizes. Importantly, while the *AKM* inference procedure may still lead to confidence intervals that are too short in several specifications, the *AKM0* inference procedure generally yields accurate rejection rates. However, confidence intervals based on the *AKM0* inference procedure may be very conservative for certain specifications.

F.3 Results with a restricted set of origin countries

All results presented in this section exploit information only on those countries of origin whose total share of immigrants in the corresponding baseline year t_0 (either 1960 or 1980, depending on the specification) is below 3%. More precisely, these results presented here are computed using an IV such as that in eq. (F.2) constructed excluding those countries of origin g for which $\sum_{i=1}^{N} ImmShare_{igt_0} / \sum_{i=1}^{N}\sum_{g'=1}^{G} ImmShare_{ig't_0} > 0.03$. We exclude large origin countries so that Assumptions 2(ii) and 2(iii) more plausibly hold. The simulations in Appendix F.2 also suggest that excluding large origin countries should lead to better finite-sample performance of the inference procedures that we propose.[6]

Table F.5 presents results for three different implementations of the *mixed approach*. In all three cases, the data comes from a three-period panel with $t = \{1990, 2000, 2010\}$ and $t_0 = 1980$. The implementations differ in the definition of a cell. In columns (1) to (4) of Table F.5, a cell corresponds

[6]Our theory currently does not provide guidance on the particular threshold that one should choose. While we find that the 3% threshold works well in the placebo exercises in this particular application, we leave the question of what threshold one should in general pick to ensure that Assumptions 2(ii) and 2(iii) plausibly hold to future research.

to the intersection of a CZ and one of the 50 occupations defined in Appendix F of Burstein et al. (2018). In columns (5) and (6), we define a cell as the intersection of a CZ and one of two education groups: high school-equivalent or college-equivalent educated workers (see Card, 2009). In columns (7) to (10), a cell corresponds to the intersection of a CZ and one of seven aggregate occupations (see Card, 2001).[7]

Although Table F.5 adopts occupational definitions that build on those in Burstein et al. (2018) and Card (2001), our specifications do not exactly match their definition of shares and shifters. Thus, our estimates should not be viewed as a test of the robustness of the results presented in these studies. Furthermore, no matter which definition of cell we use, when interpreting our estimates, one should bear in mind that, as discussed in Jaeger, Ruist and Stuhler (2018a), these may conflate the short- and the long-run responses to immigration shocks.

The magnitude and statistical significance of the estimates of β in eq. (F.1) is generally consistent across the specifications studied in Table F.5. In terms of the impact of immigration on native employment, we find that a one percentage point increase in the share of immigrants in total employment reduces the number of native workers employed by 1.19–1.49%, with all estimates of β being statistically different from zero at the 5% level for all four inference procedures that we consider. In terms of the impact of immigration on natives' average weekly wages, we find that the estimated impact of an increase in the immigrant share is not statistically different from zero at the 5% significance level according to the *AKM* and *AKM0* CIs; this is true for all three cell definitions and no matter whether we compute average wages for all workers, only for high-skill workers or only for low-skill workers. *Robust* and *Cluster* CIs also indicate that the effect of immigration on natives' average weekly wages is not statistically different from zero at the 5% significance level when each cell corresponds to the intersection of a CZ and an education group, but these standard inference procedures sometimes predict that immigration has a positive effect on the wages of high-skill workers when occupations are used to define the unit of analysis (see columns (3) and (9) in Table F.5).

While all inference procedures broadly agree in the statistical significance (at the 5% significance level) of the impact of immigration on natives' labor market outcomes, there is considerable heterogeneity across specifications in the length of the *AKM* and *AKM0* confidence intervals relative to those based on *Robust* and *Cluster* standard errors. In columns (1) to (4), which use detailed occupations to define cells, *AKM* and *AKM*0 CIs tend to be very similar (in some cases, even slightly smaller) to those based on state-clustered standard errors, although they are generally much larger than those based on robust standard errors. In contrast, for the other two cell definitions, the IV *AKM* and *AKM*0 CIs are on average, 200% and 356% wider than those based on state-clustered standard errors, and the reduced-form *AKM* and *AKM*0 CIs are on average 228% and 358% wider than those based on state-clustered standard errors. Similarly, the CIs for the first-stage coefficient, reported in Panel C, *AKM* and *AKM*0 CIs are more than twice as wide as *Robust* and *Cluster* CIs.[8]

[7] We group the 50 disaggregated occupations used in Burstein et al. (2018) into seven aggregate occupations: laborers, farm workers and low-skilled service workers; operatives and craft workers; clerical workers; sales workers; managers; professional and technical workers; and others.

[8] The results in Table F.5 are consistent with the placebo simulation results in Table F.3, which show that state-clustered standard errors lead to rejection rates that are very close to the nominal level when a cell is defined as the intersection of CZs and 50 occupations, but lead to overrejection for the other two cell definitions.

To understand why standard inference procedures may lead to overrejection of the null hypothesis of no effect in certain cases, recall from the discussion in Section 4.1 that robust and state-clustered standard errors may be biased downward even if there is no shock in the structural residual that varies exactly at the same level as the shifters of interest; a downward bias will arise so long as there is a shift-share component in the residual with shares that have a correlation structure similar to that of the shares used to construct the shift-share instrument. We present simulations that illustrate this point in Appendix D.6.

Tables F.6 and F.7 present results for different versions of the *spatial approach*, using CZs and MSAs as unit of observation, respectively. We present both estimates that measure the immigrant shares $ImmShare_{igt_0}$ in 1980 and use data on shifters and outcomes for the periods 1980–1990, 1990–2000, and 2000–2010, and estimates that measure the immigrant shares in 1960 and use data on outcomes only for the period 1970–1980. While the first sample definition mimics that in Table F.5, the second one is suggested in Jaeger, Ruist and Stuhler (2018b) as being more robust to potential bias in the estimates of β that arise from the combination of serial correlation in the shifters ΔImm_{gt} and the potentially slow adjustment of labor market outcomes to these immigration shocks.

The placebo simulation results for the different specifications considered in Tables F.6 and F.7 (see Table F.3) reveal that, due to the relatively small number of observations (i.e. small number of MSAs and CZs; small value of N) and, in the case of the specification that relies on immigrants shares measured in 1960, the relatively small number of countries of origin of immigrants (i.e. small value of G), only the *AKM0* inference procedure consistently yields rejection rates that are close to the nominal level of 5%. However, the *AKM0* procedure yields CIs with an implied median effective standard error that is much larger than the true standard deviation of the estimator. It is thus conservative. Thus, the placebo results suggest that, for most of the specifications considered in Tables F.6 and F.7, the *AKM0* CIs may be conservative and the *Robust*, *Cluster* and *AKM* CIs may be too small. It is thus not surprising that, for the different specifications considered in Tables F.6 and F.7, the *AKM0* CIs are much larger than than those implied by the other three inference procedures.[9]

Finally, Tables F.8 to F.10 report p-values for the null hypothesis of no effect for all specifications considered in Tables F.5 to F.7, respectively.

F.4 Results with all origin countries

Tables F.11 to F.16 present results analogous to those in Tables F.5 to F.10, respectively. While the latter set of tables, as described in Appendix F.3, use a shift-share instrument that excludes countries of origin that account for more than 3% of the overall immigrant population in the baseline year, the former uses all areas of origin of immigrants listed in Tables F.1 and F.2.

As the results of placebo simulations presented in Table F.4 show, using all countries of origin to construct the shift-share instrumental variable of interest results in the *Robust*, *Cluster* and *AKM* standard errors underestimating the sampling variability of the estimator of interest. Table F.4 also shows that not excluding any country of origin from the construction of the instrument in eq. (F.2)

[9] This is particularly noticeable for the IV results in Panel A; however, to interpret these CIs, one should bear in mind that, as the first-stage results in Panel C show, the shift-share IV is weak in these specifications. In the presence of weak IVs, only the *AKM0* confidence interval remains valid in general (see discussion in Section 4.3).

Table F.1: Origin countries (1980 weights)

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	

Notes: In italics, countries that are dropped from the sample when considering only countries whose share is below 3%; i.e. those countries in italics are countries of origin g such that $(\sum_{i=1}^{N} ImmShare_{ig,1980} / \sum_{i=1}^{N} \sum_{g'=1}^{G} ImmShare_{ig',1980}) > 0.03$.

Table F.2: Origin countries (1960 weights)

	France	Liechtenstein and Lux.	Scandinavia
Africa	Greece		Scotland
Albania			
	India		South America
Austria		*Mexico*	Spain
Belgium			Switzerland
		Netherlands	Syria
	Ireland	Oceania	
Canada	Israel/Palestine	*Other*	Turkey
Central America	*Italy*	Other Europe	
China	Japan	*Other USSR and Russia*	Wales
Cuba and West Indies		Philippines	
	Korea	Portugal	
Eastern Europe		Rest of Asia	
England	Lebanon		

Notes: In italics, countries that are dropped from the sample when considering only countries whose share is below 3%; i.e. those countries in italics are countries of origin g such that $(\sum_{i=1}^{N} ImmShare_{ig1960} / \sum_{i=1}^{N} \sum_{g'=1}^{G} ImmShare_{ig'1960}) > 0.03$.

results in the *AKM0* inference procedure being too conservative: the 95% confidence interval often has an infinite length and the rejection rates are generally much smaller than the 5% nominal rate.

Given the relatively poor performance of all inference procedures in the placebo simulations, one should use caution when extracting conclusions from the estimates presented in Tables F.11 to F.16.

Table F.3: Reduced-form placebo with origin countries below 3% of total immigrant share

		Median eff. s.e.				Rejection rate for $H_0\colon \beta = 0$ at 5%			
		$\Delta \log E_i$	$\Delta \log w_i$			$\Delta \log E_i$	$\Delta \log w_i$		
		All	All	High-Skill	Low-Skill	All	All	High-Skill	Low-Skill
Unit obs.	Weights	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Standard deviation of placebo estimate									
CZ	1980	1.14	0.46	0.30	1.04				
CZ	1960	0.96	0.81	0.80	1.51				
MSA	1980	1.94	0.49	0.74	0.83				
MSA	1960	1.99	0.25	0.33	0.71				
CZ-50 Occ.	1980	0.17	0.06	0.07	0.06				
CZ-7 Occ.	1980	1.59	0.40	0.47	0.42				
CZ-Educ.	1980	2.88	0.54	—	—				
Panel B: Robust standard error									
CZ	1980	0.49	0.21	0.20	0.25	37.82%	43.17%	12.45%	73.21%
CZ	1960	0.77	0.31	0.32	0.39	9.24%	47.70%	37.09%	68.84%
MSA	1980	1.82	0.33	0.35	0.37	8.12%	20.42%	43.80%	42.74%
MSA	1960	1.47	0.24	0.27	0.30	20.12%	2.73%	3.62%	48.74%
CZ-50 Occ.	1980	0.11	0.05	0.06	0.06	20.22%	12.46%	8.57%	4.51%
CZ-7 Occ.	1980	0.47	0.16	0.19	0.18	63.04%	44.33%	41.78%	45.89%
CZ-Educ.	1980	0.65	0.15	—	—	67.22%	66.12%	—	—
Panel C: State-clustered standard error									
CZ	1980	0.70	0.26	0.27	0.34	23.76%	30.73%	5.44%	62.41%
CZ	1960	1.03	0.46	0.48	0.60	2.01%	26.87%	12.68%	51.41%
MSA	1980	2.11	0.29	0.32	0.43	3.42%	24.52%	49.14%	36.30%
MSA	1960	1.62	0.20	0.23	0.35	17.72%	7.59%	12.52%	41.45%
CZ-50 Occ.	1980	0.14	0.07	0.06	0.06	9.21%	5.24%	5.01%	1.64%
CZ-7 Occ.	1980	0.64	0.23	0.26	0.25	49.70%	27.94%	27.90%	29.90%
CZ-Educ.	1980	0.98	0.23	—	—	50.31%	50.35%	—	—
Panel D: AKM standard error									
CZ	1980	0.91	0.40	0.24	0.95	12.01%	23.27%	5.40%	30.31%
CZ	1960	0.73	0.67	0.62	1.37	11.27%	15.35%	9.30%	20.06%
MSA	1980	1.60	0.39	0.62	0.73	25.50%	15.31%	22.26%	17.92%
MSA	1960	1.39	0.20	0.26	0.55	22.88%	11.71%	7.66%	23.57%
CZ-50 Occ.	1980	0.15	0.06	0.06	0.05	11.63%	10.77%	8.16%	10.02%
CZ-7 Occ.	1980	1.49	0.35	0.40	0.37	19.94%	14.79%	12.98%	20.63%
CZ-Educ.	1980	2.40	0.49	—	—	17.46%	25.09%	—	—
Panel E: AKM0 standard error									
CZ	1980	∞	∞	∞	∞	2.74%	1.63%	3.07%	1.00%
CZ	1960	3.04	3.13	2.67	6.31	3.10%	3.16%	5.04%	2.26%
MSA	1980	30.16	7.92	17.82	12.70	1.93%	1.25%	0.80%	1.41%
MSA	1960	∞	∞	∞	∞	2.85%	2.84%	2.30%	0.52%
CZ-50 Occ.	1980	0.25	0.09	0.10	0.08	4.43%	4.13%	3.63%	4.35%
CZ-7 Occ.	1980	2.89	0.66	0.76	0.72	4.67%	3.72%	4.08%	4.08%
CZ-Educ.	1980	8.10	1.75	—	—	4.65%	3.17%	—	—

Notes: $\Delta \log E_i$ denotes log change in native employment; $\Delta \log w_i$ denotes log change in average weekly wages of native workers. Whenever we use 1980 weights, we use observations for the time periods 1980–1990, 1990–2000, 2000–2010. Whenever we use 1960 weights, we use observations for the time period 1970–1980. We use information on 722 CZs when combined with both 1960 and 1980 weights, on 257 MSAs when combined with 1980 weights, and 217 MSAs when combined with 1960 weights. Models are weighted by the start-of-period share of national population. *Robust* is the Eicker-Huber-White standard error; *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. The median effective standard error is equal to the median length of the corresponding 95% confidence interval divided by 2×1.96. Results are based on 30,000 placebo samples.

Table F.4: Reduced-form placebo with all origin countries

		Median eff. s.e.				Rejection rate for $H_0: \beta = 0$ at 5%			
		$\Delta \log E_i$	$\Delta \log w_i$			$\Delta \log E_i$	$\Delta \log w_i$		
		All	All	High-Skill	Low-Skill	All	All	High-Skill	Low-Skill
Unit obs.	Weights	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Panel A: Standard deviation of placebo estimate									
CZ	1980	0.21	0.18	0.15	0.25				
CZ	1960	0.27	0.20	0.18	0.32				
MSA	1980	0.40	0.06	0.06	0.13				
MSA	1960	0.51	0.12	0.10	0.26				
CZ-50 Occ.	1980	0.11	0.04	0.02	0.04				
CZ-7 Occ.	1980	0.51	0.15	0.2	0.22				
CZ-Educ.	1980	0.60	0.15	—	—				
Panel B: Robust standard error									
CZ	1980	0.12	0.05	0.04	0.06	3.57%	73.01%	81.63%	38.91%
CZ	1960	0.13	0.05	0.05	0.06	15.86%	55.42%	57.59%	60.02%
MSA	1980	0.20	0.03	0.03	0.04	25.90%	18.07%	9.97%	53.39%
MSA	1960	0.34	0.08	0.08	0.09	9.21%	9.52%	2.18%	41.43%
CZ-50 Occ.	1980	0.04	0.02	0.02	0.02	62.52%	48.73%	9.07%	38.04%
CZ-7 Occ.	1980	0.15	0.05	0.04	0.07	75.86%	75.24%	86.16%	80.99%
CZ-Educ.	1980	0.13	0.04	—	—	82.08%	49.75%	—	—
Panel C: State-clustered standard error									
CZ	1980	0.09	0.07	0.05	0.10	3.46%	46.56%	68.91%	13.85%
CZ	1960	0.16	0.09	0.07	0.12	8.50%	20.66%	43.84%	29.18%
MSA	1980	0.28	0.04	0.03	0.07	10.98%	5.77%	7.30%	21.38%
MSA	1960	0.39	0.09	0.09	0.13	2.80%	7.02%	1.29%	33.24%
CZ-50 Occ.	1980	0.06	0.02	0.02	0.02	48.40%	34.88%	5.15%	30.04%
CZ-7 Occ.	1980	0.22	0.06	0.06	0.06	30.90%	60.28%	74.47%	75.69%
CZ-Educ.	1980	0.18	0.08	—	—	73.77%	9.82%	—	—
Panel D: AKM standard error									
CZ	1980	0.09	0.06	0.04	0.12	4.44%	56.32%	66.26%	10.19%
CZ	1960	0.14	0.11	0.08	0.19	17.27%	21.85%	37.77%	13.12%
MSA	1980	0.20	0.03	0.04	0.07	40.31%	30.36%	3.86%	31.97%
MSA	1960	0.33	0.08	0.07	0.20	23.68%	11.11%	1.43%	13.19%
CZ-50 Occ.	1980	0.07	0.03	0.02	0.03	32.38%	32.88%	25.60%	28.50%
CZ-7 Occ.	1980	0.21	0.06	0.07	0.08	54.43%	59.39%	61.74%	62.08%
CZ-Educ.	1980	0.23	0.07	—	—	53.00%	39.53%	—	—
Panel E: AKM0 standard error									
CZ	1980	∞	∞	∞	∞	2.52%	0.27%	0.33%	0.44%
CZ	1960	∞	∞	∞	∞	2.52%	0.32%	2.03%	0.30%
MSA	1980	∞	∞	∞	∞	0.54%	0.97%	2.47%	0.08%
MSA	1960	∞	∞	∞	∞	1.06%	0.12%	1.58%	0.06%
CZ-50 Occ.	1980	0.42	0.15	0.1	0.14	1.73%	2.49%	3.17%	2.72%
CZ-7 Occ.	1980	∞	∞	∞	∞	1.33%	0.73%	0.83%	0.56%
CZ-Educ.	1980	∞	∞	—	—	1.16%	1.20%	—	—

Notes: $\Delta \log E_i$ denotes log change in native employment; $\Delta \log w_i$ denotes log change in average weekly wages of native workers. Whenever we use 1980 weights, we use observations for the time periods 1980–1990, 1990–2000, 2000–2010. Whenever we use 1960 weights, we use observations for the time period 1970–1980. We use information on 722 CZs when combined with both 1960 and 1980 weights, on 257 MSAs when combined with 1980 weights, and 217 MSAs when combined with 1960 weights. Models are weighted by the start-of-period share of national population. *Robust* is the Eicker-Huber-White standard error; *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. The median effective standard error is equal to the median length of the corresponding 95% confidence interval divided by 2×1.96. Results are based on 30,000 placebo samples.

Table F.5: Effect of immigration: analysis by CZ-Occupations and CZ-Education groups (excluding large origin countries)

Outcome:	$\Delta \log E_i$	$\Delta \log w_i$			$\Delta \log E_i$	$\Delta \log w_i$	$\Delta \log E_i$	$\Delta \log w_i$		
Workers:	All	All	High-Skill	Low-Skill	All	All	All	All	High-Skill	Low-Skill
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	CZ-50 Occ. (1980 weights)				*CZ-Educ. (1980 weights)*		*CZ-7 Occ. (1980 weights)*			
Panel A: 2SLS										
$\hat{\beta}$	-1.19	0.05	0.26	-0.14	-1.49	0.18	-1.39	0.08	0.24	-0.14
Robust	$[-1.55, -0.83]$	$[-0.09, 0.20]$	$[0.11, 0.41]$	$[-0.32, 0.03]$	$[-2.09, -0.90]$	$[-0.20, 0.56]$	$[-1.89, -0.90]$	$[-0.17, 0.33]$	$[0.00, 0.47]$	$[-0.43, 0.15]$
Cluster	$[-1.89, -0.49]$	$[-0.35, 0.46]$	$[-0.14, 0.67]$	$[-0.69, 0.40]$	$[-2.14, -0.85]$	$[-0.04, 0.39]$	$[-1.87, -0.92]$	$[-0.11, 0.27]$	$[0.10, 0.38]$	$[-0.27, -0.01]$
AKM	$[-1.55, -0.83]$	$[-0.36, 0.47]$	$[-0.17, 0.69]$	$[-0.64, 0.35]$	$[-2.31, -0.68]$	$[-0.53, 0.88]$	$[-2.02, -0.76]$	$[-0.51, 0.67]$	$[-0.31, 0.79]$	$[-0.81, 0.53]$
AKM0	$[-1.66, -0.72]$	$[-0.53, 0.54]$	$[-0.32, 0.81]$	$[-0.92, 0.39]$	$[-3.00, -0.14]$	$[-0.90, 1.60]$	$[-2.35, -0.54]$	$[-0.75, 0.94]$	$[-0.51, 1.07]$	$[-1.12, 0.80]$
Panel B: Reduced-Form										
$\hat{\beta}$	-0.89	0.04	0.2	-0.11	-1.29	0.15	-1.05	0.06	0.18	-0.11
Robust	$[-1.17, -0.61]$	$[-0.07, 0.15]$	$[0.06, 0.33]$	$[-0.23, 0.02]$	$[-1.86, -0.73]$	$[-0.19, 0.50]$	$[-1.38, -0.73]$	$[-0.13, 0.25]$	$[-0.01, 0.37]$	$[-0.32, 0.11]$
Cluster	$[-1.37, -0.41]$	$[-0.27, 0.35]$	$[-0.16, 0.55]$	$[-0.47, 0.25]$	$[-1.69, -0.90]$	$[-0.02, 0.33]$	$[-1.29, -0.82]$	$[-0.08, 0.21]$	$[0.07, 0.29]$	$[-0.20, -0.01]$
AKM	$[-1.35, -0.43]$	$[-0.28, 0.36]$	$[-0.18, 0.57]$	$[-0.44, 0.23]$	$[-1.99, -0.60]$	$[-0.48, 0.79]$	$[-1.54, -0.57]$	$[-0.39, 0.52]$	$[-0.26, 0.62]$	$[-0.59, 0.38]$
AKM0	$[-1.55, -0.39]$	$[-0.27, 0.55]$	$[-0.16, 0.79]$	$[-0.44, 0.42]$	$[-2.25, -0.11]$	$[-0.53, 1.56]$	$[-1.62, -0.36]$	$[-0.42, 0.78]$	$[-0.29, 0.86]$	$[-0.62, 0.67]$
Panel C: First-Stage										
$\hat{\beta}$		0.75			0.87			0.76		
Robust		$[0.56, 0.93]$			$[0.55, 1.18]$			$[0.56, 0.95]$		
Cluster		$[0.62, 0.88]$			$[0.66, 1.08]$			$[0.60, 0.91]$		
AKM		$[0.36, 1.13]$			$[0.44, 1.30]$			$[0.42, 1.09]$		
AKM0		$[0.38, 1.37]$			$[0.34, 1.70]$			$[0.38, 1.24]$		

Notes: $\Delta \log E_i$ denotes log change in native employment; $\Delta \log w_i$ denotes log change in average weekly wages of native workers. The specifications *CZ-50 Occupations*, *CZ-2 Education Groups*, and *CZ-7 Occupations* differ in the definition of the unit of observation. In all three specifications, we use 1980 weights and three time periods, 1980–1990, 1990–2000, and 2000–2010. Thus, $N = 108,300$ (722 CZs $\times$ 50 occupations $\times$ 3 time periods) for the *CZ-50 Occupations* specification; $N = 4,332$ (722 CZs $\times$ 2 education groups $\times$ 3 time periods) for the *CZ-2 Education Groups* specification; and $N = 15,162$ (722 CZs $\times$ 7 occupations $\times$ 3 time periods) for the *CZ-7 Occupations* specification. Models are weighted by start-of-period occupation-region (or education group-region) share of national population. All regressions include occupation (or education group) and period dummies. 95% confidence intervals in square brackets. *Robust* is the Eicker-Huber-White standard error; *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. We exclude from the analysis those countries of origin whose immigrant share in 1980 is larger than 3%; i.e. $\sum_i ImmShare_{igt_0} / \sum_i \sum_{g'} ImmShare_{ig't_0} > 0.03$. See Table F.1 for a list of the origin countries included in the analysis.

Table F.6: Effect of immigration: analysis by CZ (excluding large origin countries)

Outcome:	$\Delta \log E_i$	$\Delta \log w_i$			$\Delta \log E_i$	$\Delta \log w_i$		
Workers:	All	All	High-Skill	Low-Skill	All	All	High-Skill	Low-Skill
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Commuting Zone (1980 weights)				*Commuting Zone (1960 weights)*			
Panel A: 2SLS Regression								
$\hat{\beta}$	-0.89	0.42	0.56	-0.10	-1.29	-0.71	-0.34	-1.14
Robust	$[-1.48, -0.30]$	$[-0.06, 0.90]$	$[0.16, 0.95]$	$[-0.68, 0.47]$	$[-3.69, 1.10]$	$[-1.10, -0.33]$	$[-1.01, 0.33]$	$[-1.66, -0.61]$
Cluster	$[-1.33, -0.45]$	$[0.15, 0.69]$	$[0.39, 0.72]$	$[-0.31, 0.11]$	$[-3.96, 1.37]$	$[-1.17, -0.26]$	$[-1.12, 0.44]$	$[-1.71, -0.56]$
AKM	$[-1.85, 0.08]$	$[-0.44, 1.28]$	$[-0.12, 1.24]$	$[-1.17, 0.97]$	$[-3.83, 1.24]$	$[-1.31, -0.12]$	$[-1.16, 0.48]$	$[-1.89, -0.38]$
AKM0	$[-\infty, -285.90] \cup [-4.12, \infty]$	$[-\infty, -230.52] \cup [-2.73, \infty]$	$[-\infty, -226.52] \cup [-1.46, \infty]$	$[-\infty, -158.12] \cup [-7.24, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$
Panel B: Reduced-Form Regression								
$\hat{\beta}$	-0.75	0.35	0.47	-0.08	-1.31	-0.72	-0.34	-1.15
Robust	$[-1.16, -0.33]$	$[-0.10, 0.81]$	$[0.07, 0.86]$	$[-0.55, 0.38]$	$[-3.39, 0.78]$	$[-1.06, -0.38]$	$[-0.94, 0.25]$	$[-1.61, -0.68]$
Cluster	$[-1.02, -0.47]$	$[0.12, 0.59]$	$[0.31, 0.62]$	$[-0.26, 0.09]$	$[-3.57, 0.96]$	$[-1.10, -0.35]$	$[-1.02, 0.33]$	$[-1.56, -0.73]$
AKM	$[-1.47, -0.02]$	$[-0.44, 1.15]$	$[-0.21, 1.14]$	$[-0.96, 0.79]$	$[-3.55, 0.94]$	$[-1.21, -0.23]$	$[-1.09, 0.40]$	$[-1.73, -0.57]$
AKM0	$[-1.71, 6.41]$	$[-0.73, 8.03]$	$[-0.44, 6.99]$	$[-1.29, 8.26]$	$[-\infty, \infty]$	$[-3.23, -1.51]$	$[-3.37, -1.87]$	$[-4.01, -2.12]$
Panel C: 2SLS First-Stage								
$\hat{\beta}$	0.84				1.01			
Robust	$[0.51, 1.18]$				$[0.46, 1.56]$			
Cluster	$[0.66, 1.03]$				$[0.65, 1.37]$			
AKM	$[0.37, 1.31]$				$[0.70, 1.32]$			
AKM0	$[0.00, 4.32]$				$[-\infty, \infty]$			

Notes: $\Delta \log E_i$ denotes log change in native employment; $\Delta \log w_i$ denotes log change in average weekly wages of native workers. In the specification *CZ (1980 weights)*, we use information on 722 CZs, 1980 weights and three time periods, 1980–1990, 1990–2000, 2000–2010; thus, $N = 2,166$ (722 CZs $\times$ 3 time periods). In the specification *CZ (1960 weights)*, we use information on 722 CZs, 1960 weights and one time period, 1970–1980; thus, $N = 722$ (722 CZs $\times$ 1 time period). 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; *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. We exclude from the analysis those countries of origin whose immigrant share in year t_0 is larger than 3%; i.e. $\sum_{i=1}^{N} ImmShare_{igt_0} / \sum_{i=1}^{N} \sum_{g'=1}^{G} ImmShare_{ig't_0} > 0.03$. See Tables F.1 and F.2 for a list of the origin countries included in the analysis.

Table F.7: Effect of immigration: analysis by MSA (excluding large origin countries)

Outcome:	$\Delta \log E_i$	$\Delta \log w_i$			$\Delta \log E_i$	$\Delta \log w_i$		
Workers:	All	All	High-Skill	Low-Skill	All	All	High-Skill	Low-Skill
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	MSA (1980 weights)				*MSA (1960 weights)*			
Panel A: 2SLS Regression								
$\hat{\beta}$	-1.87	0.50	0.63	-0.10	-1.55	-0.32	-0.17	-0.57
Robust	$[-3.17, -0.58]$	$[-0.06, 1.06]$	$[0.17, 1.09]$	$[-0.93, 0.73]$	$[-3.68, 0.57]$	$[-0.72, 0.08]$	$[-0.83, 0.49]$	$[-1.08, -0.06]$
Cluster	$[-2.97, -0.78]$	$[0.23, 0.77]$	$[0.48, 0.77]$	$[-0.47, 0.27]$	$[-3.94, 0.84]$	$[-0.74, 0.09]$	$[-0.88, 0.53]$	$[-1.06, -0.08]$
AKM	$[-3.77, 0.03]$	$[-0.44, 1.44]$	$[-0.15, 1.40]$	$[-1.49, 1.29]$	$[-5.61, 2.51]$	$[-0.90, 0.25]$	$[-1.07, 0.73]$	$[-1.26, 0.12]$
AKM0	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$
Panel B: Reduced-Form Regression								
$\hat{\beta}$	-1.47	0.39	0.49	-0.08	-1.48	-0.31	-0.16	-0.54
Robust	$[-2.06, -0.87]$	$[-0.17, 0.95]$	$[0.00, 0.98]$	$[-0.70, 0.54]$	$[-3.29, 0.34]$	$[-0.61, -0.01]$	$[-0.75, 0.42]$	$[-0.90, -0.19]$
Cluster	$[-1.94, -0.99]$	$[0.09, 0.69]$	$[0.27, 0.71]$	$[-0.35, 0.19]$	$[-3.45, 0.49]$	$[-0.63, 0.01]$	$[-0.79, 0.46]$	$[-0.86, -0.22]$
AKM	$[-2.43, -0.50]$	$[-0.50, 1.28]$	$[-0.28, 1.26]$	$[-1.13, 0.98]$	$[-4.79, 1.83]$	$[-0.73, 0.11]$	$[-0.96, 0.63]$	$[-0.98, -0.11]$
AKM0	$[-2.70, 8.87]$	$[-0.79, 9.41]$	$[-0.53, 8.41]$	$[-1.51, 10.46]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$
Panel C: 2SLS First-Stage								
$\hat{\beta}$	0.78				0.95			
Robust	$[0.29, 1.28]$				$[0.46, 1.45]$			
Cluster	$[0.51, 1.05]$				$[0.64, 1.27]$			
AKM	$[0.10, 1.47]$				$[0.49, 1.42]$			
AKM0	$[-0.61, 5.35]$				$[-\infty, \infty]$			

Notes: $\Delta \log E_i$ denotes log change in native employment; $\Delta \log w_i$ denotes log change in average weekly wages of native workers. In the specification *MSA (1980 weights)*, we use information on 257 MSAs, 1980 weights and three time periods, 1980–1990, 1990–2000, 2000–2010; thus, $N = 651$ (257 MSAs $\times$ 3 time periods). In the specification *MSA (1960 weights)*, we use information on 217 MSAs, 1960 weights and one time period, 1970–1980; thus, $N = 217$ (217 CZs $\times$ 1 time period). Models are weighted by start-of-period MSA share of national population. All regressions include period dummies. 95% confidence intervals in square brackets. *Robust* is the Eicker-Huber-White standard error; *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. We exclude from the analysis those countries of origin whose immigrant share in year t_0 is larger than 3%; i.e. $\sum_{i=1}^{N} ImmShare_{igt_0} / \sum_{i=1}^{N} \sum_{g'=1}^{G} ImmShare_{ig't_0} > 0.03$. See Tables F.1 and F.2 for a list of the origin countries included in the analysis.

Table F.8: Immigration: p-values by CZ-Occ. and CZ-Educ. (excluding large origin countries)

	$\Delta \log E_i$	$\Delta \log w_i$			$\Delta \log E_i$	$\Delta \log w_i$	$\Delta \log E_i$	$\Delta \log w_i$		
	All	All	High-Skill	Low-Skill	All	All	All	All	High-Skill	Low-Skill
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	CZ-50 Occupations				*CZ- Educ.*		*CZ-7 Occupations*			
Panel A: 2SLS Regression										
Robust	0.000	0.459	0.001	0.116	0.000	0.357	0.000	0.519	0.046	0.348
Cluster	0.001	0.790	0.205	0.611	0.000	0.102	0.000	0.412	0.001	0.042
AKM	0.000	0.793	0.235	0.576	0.000	0.620	0.000	0.788	0.395	0.684
AKM0	0.006	0.796	0.273	0.566	0.043	0.621	0.018	0.788	0.410	0.687
Panel B: Reduced-Form Regression										
Robust	0.000	0.473	0.004	0.091	0.000	0.379	0.000	0.528	0.057	0.330
Cluster	0.000	0.795	0.275	0.565	0.000	0.089	0.000	0.409	0.001	0.036
AKM	0.000	0.801	0.304	0.538	0.000	0.635	0.000	0.792	0.422	0.674
AKM0	0.006	0.796	0.273	0.566	0.043	0.621	0.018	0.788	0.410	0.687
Panel C: First-Stage										
Robust			0.000		0.000				0.000	
Cluster			0.000		0.000				0.000	
AKM			0.000		0.000				0.000	
AKM0			0.002		0.017				0.006	

Notes: $\Delta \log E_i$ denotes log change in native employment; $\Delta \log w_i$ denotes log change in average weekly wages of native workers. The specifications *CZ-50 Occupations*, *CZ-Educ.*, and *CZ-7 Occupations* differ in the definition of the unit of observation. In all three specifications, we use 1980 weights and three time periods, 1980–1990, 1990–2000, 2000–2010. Thus, $N = 108,300$ (722 CZs $\times$ 50 occupations $\times$ 3 time periods) for the *CZ-50 Occupations* specification; $N = 4,332$ (722 CZs $\times$ 2 education groups $\times$ 3 time periods) for the *CZ-Educ.* specification; and $N = 15,162$ (722 CZs $\times$ 7 occupations $\times$ 3 time periods) for the *CZ-7 Occupations* specification. Models are weighted by start-of-period occupation-region (or education group-region) share of national population. All regressions include occupation (or education group) and period dummies. 95% confidence intervals in square brackets. *Robust* is the Eicker-Huber-White standard error; *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. We exclude from the analysis those countries of origin whose immigrant share in year 1980 is larger than 3%; i.e. $\sum_{i=1}^{N} ImmShare_{igt_0} / \sum_{i=1}^{N} \sum_{g'=1}^{G} ImmShare_{ig't_0} > 0.03$. See Table F.1 for a list of the origin countries included in the analysis.

Table F.9: Effect of immigration: p-values by CZ (excluding large origin countries)

Outcome:	$\Delta \log E_i$	$\Delta \log w_i$			$\Delta \log E_i$	$\Delta \log w_i$		
Workers:	All	All	High-Skill	Low-Skill	All	All	High-Skill	Low-Skill
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Commuting Zone (1980 weights)				*Commuting Zone (1960 weights)*			
Panel A: 2SLS Regression								
Robust	0.003	0.083	0.005	0.731	0.291	0.000	0.317	0.000
Cluster	0.000	0.002	0.000	0.338	0.342	0.002	0.390	0.000
AKM	0.071	0.335	0.109	0.853	0.318	0.018	0.416	0.003
AKM0	0.216	0.360	0.174	0.855	0.393	0.179	0.472	0.142
Panel B: Reduced-Form Regression								
Robust	0.000	0.129	0.020	0.722	0.220	0.000	0.254	0.000
Cluster	0.000	0.003	0.000	0.330	0.258	0.000	0.316	0.000
AKM	0.044	0.384	0.174	0.850	0.254	0.004	0.366	0.000
AKM0	0.216	0.360	0.174	0.855	0.393	0.179	0.472	0.142
Panel C: First-Stage								
Robust			0.000				0.000	
Cluster			0.000				0.000	
AKM			0.001				0.000	
AKM0			0.050				0.062	

Notes: $\Delta \log E_i$ denotes log change in native employment; $\Delta \log w_i$ denotes log change in average weekly wages of native workers. In the specification *CZ (1980 weights)*, we use information on 722 CZs, 1980 weights and three time periods, 1980–1990, 1990–2000, 2000–2010; thus, $N = 2,166$ (722 CZs $\times$ 3 time periods). In the specification *CZ (1960 weights)*, we use information on 722 CZs, 1960 weights and one time period, 1970–1980; thus, $N = 722$ (722 CZs $\times$ 1 time period). 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; *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. We exclude from the analysis those countries of origin whose immigrant share in year t_0 is larger than 3%; i.e. $\sum_{i=1}^{N} ImmShare_{igt_0} / \sum_{i=1}^{N} \sum_{g'=1}^{G} ImmShare_{ig't_0} > 0.03$. See Tables F.1 and F.2 for a list of the origin countries included in the analysis.

Table F.10: Effect of immigration: p-values by MSA (excluding large origin countries)

Outcome:	$\Delta \log E_i$	$\Delta \log w_i$			$\Delta \log E_i$	$\Delta \log w_i$		
Workers:	All	All	High-Skill	Low-Skill	All	All	High-Skill	Low-Skill
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	MSA (1980 weights)				*MSA (1960 weights)*			
Panel A: 2SLS Regression								
Robust	0.005	0.080	0.008	0.812	0.152	0.115	0.610	0.028
Cluster	0.001	0.000	0.000	0.594	0.203	0.128	0.631	0.022
AKM	0.053	0.300	0.112	0.887	0.453	0.272	0.708	0.106
AKM0	0.146	0.361	0.203	0.887	0.532	0.376	0.732	0.284
Panel B: Reduced-Form Regression								
Robust	0.000	0.173	0.048	0.803	0.110	0.043	0.585	0.003
Cluster	0.000	0.012	0.000	0.564	0.141	0.059	0.607	0.001
AKM	0.003	0.388	0.213	0.884	0.381	0.150	0.687	0.015
AKM0	0.146	0.361	0.203	0.887	0.532	0.376	0.732	0.284
Panel C: First-Stage								
Robust			0.002				0.000	
Cluster			0.000				0.000	
AKM			0.025				0.000	
AKM0			0.096				0.126	

Notes: $\Delta \log E_i$ denotes log change in native employment; $\Delta \log w_i$ denotes log change in average weekly wages of native workers. In the specification *MSA (1980 weights)*, we use information on 257 MSAs, 1980 weights and three time periods, 1980–1990, 1990–2000, 2000–2010; thus, $N = 651$ (257 MSAs $\times$ 3 time periods). In the specification *MSA (1960 weights)*, we use information on 217 MSAs, 1960 weights and one time period, 1970–1980; thus, $N = 217$ (217 CZs $\times$ 1 time period). Models are weighted by start-of-period MSA share of national population. All regressions include period dummies. 95% confidence intervals in square brackets. *Robust* is the Eicker-Huber-White standard error; *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. We exclude from the analysis those countries of origin whose immigrant share in year t_0 is larger than 3%; i.e. $\sum_{i=1}^{N} ImmShare_{igt_0} / \sum_{i=1}^{N} \sum_{g'=1}^{G} ImmShare_{ig't_0} > 0.03$. See Tables F.1 and F.2 for a list of the origin countries included in the analysis.

Table F.11: Effect of immigration: analysis by CZ-Occupations and CZ-Education groups (including all origin countries)

Outcome:	$\Delta \log E_i$	$\Delta \log w_i$			$\Delta \log E_i$	$\Delta \log w_i$	$\Delta \log E_i$	$\Delta \log w_i$		
Workers:	All	All	High-Skill	Low-Skill	All	All	All	All	High-Skill	Low-Skill
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	CZ-50 Occupations				*CZ- 2 Education Groups*		*CZ-7 Occupations*			
Panel A: 2SLS Regression										
$\hat{\beta}$	-0.73	-0.07	0.15	-0.24	-0.53	-0.01	-0.79	-0.08	0.08	-0.27
Robust	$[-1.04, -0.42]$	$[-0.22, 0.09]$	$[0.01, 0.29]$	$[-0.42, -0.06]$	$[-1.03, -0.02]$	$[-0.45, 0.44]$	$[-1.25, -0.33]$	$[-0.39, 0.22]$	$[-0.17, 0.33]$	$[-0.64, 0.09]$
Cluster	$[-1.15, -0.31]$	$[-0.49, 0.36]$	$[-0.23, 0.52]$	$[-0.80, 0.32]$	$[-1.03, -0.02]$	$[-0.24, 0.23]$	$[-1.34, -0.25]$	$[-0.33, 0.16]$	$[-0.09, 0.25]$	$[-0.47, -0.08]$
AKM	$[-1.22, -0.24]$	$[-0.42, 0.29]$	$[-0.18, 0.47]$	$[-0.66, 0.18]$	$[-1.85, 0.79]$	$[-0.81, 0.80]$	$[-1.67, 0.09]$	$[-0.74, 0.57]$	$[-0.52, 0.67]$	$[-0.99, 0.44]$
AKM0	$[-1.61, 0.24]$	$[-0.52, 0.94]$	$[-0.24, 1.11]$	$[-0.97, 0.68]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$
Panel B: Reduced-Form Regression										
$\hat{\beta}$	-0.19	-0.02	0.04	-0.06	-0.19	0.00	-0.25	-0.03	0.02	-0.09
Robust	$[-0.27, -0.11]$	$[-0.05, 0.02]$	$[0.00, 0.08]$	$[-0.10, -0.02]$	$[-0.41, 0.04]$	$[-0.16, 0.15]$	$[-0.41, -0.10]$	$[-0.12, 0.06]$	$[-0.06, 0.11]$	$[-0.18, 0.01]$
Cluster	$[-0.32, -0.06]$	$[-0.12, 0.09]$	$[-0.08, 0.15]$	$[-0.17, 0.05]$	$[-0.38, 0.01]$	$[-0.08, 0.08]$	$[-0.42, -0.09]$	$[-0.10, 0.04]$	$[-0.04, 0.09]$	$[-0.13, -0.05]$
AKM	$[-0.38, 0.01]$	$[-0.10, 0.07]$	$[-0.06, 0.13]$	$[-0.15, 0.02]$	$[-0.74, 0.37]$	$[-0.29, 0.28]$	$[-0.62, 0.11]$	$[-0.23, 0.17]$	$[-0.17, 0.22]$	$[-0.29, 0.11]$
AKM0	$[-1.01, 0.03]$	$[-0.08, 0.51]$	$[-0.04, 0.60]$	$[-0.14, 0.42]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-1.28, -0.26]$	$[-1.18, -0.20]$	$[-1.13, -0.37]$
Panel C: First-Stage										
$\hat{\beta}$	0.26				0.35		0.32			
Robust	$[0.19, 0.32]$				$[0.17, 0.53]$		$[0.20, 0.44]$			
Cluster	$[0.17, 0.35]$				$[0.24, 0.46]$		$[0.21, 0.43]$			
AKM	$[0.12, 0.39]$				$[0.10, 0.60]$		$[0.12, 0.52]$			
AKM0	$[0.11, 0.84]$				$[-\infty, \infty]$		$[-\infty, \infty]$			

Notes: $\Delta \log E_i$ denotes log change in native employment; $\Delta \log w_i$ denotes log change in average weekly wages of native workers. The specifications *CZ-50 Occupations*, *CZ-2 Education Groups*, and *CZ-7 Occupations* differ in the definition of the unit of observation. In all three specifications, we use 1980 weights and three time periods, 1980–1990, 1990–2000, 2000–2010. Thus, $N = 108,300$ (722 CZs $\times$ 50 occupations $\times$ 3 time periods) for the *CZ-50 Occupations* specification; $N = 4,332$ (722 CZs $\times$ 2 education groups $\times$ 3 time periods) for the *CZ-2 Education Groups* specification; and $N = 15,162$ (722 CZs $\times$ 7 occupations $\times$ 3 time periods) for the *CZ-7 Occupations* specification. Models are weighted by start-of-period occupation-region (or education group-region) share of national population. All regressions include occupation (or education group) and period dummies. 95% confidence intervals in square brackets. *Robust* is the Eicker-Huber-White standard error; *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. We include all countries of origin in the analysis. See Table F.1 for a list of the origin countries included in the analysis.

Table F.12: Effect of immigration: analysis by CZ (including all origin countries)

Outcome:	$\Delta \log E_i$	$\Delta \log w_i$			$\Delta \log E_i$	$\Delta \log w_i$		
Workers:	All	All	High-Skill	Low-Skill	All	All	High-Skill	Low-Skill
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Commuting Zone (1980 weights)				*Commuting Zone (1960 weights)*			
Panel A: 2SLS Regression								
$\hat{\beta}$	-0.49	0.13	0.27	-0.2	0.05	-0.25	0.09	-0.52
Robust	$[-1.12, 0.14]$	$[-0.37, 0.63]$	$[-0.09, 0.64]$	$[-0.85, 0.44]$	$[-0.96, 1.07]$	$[-0.52, 0.02]$	$[-0.16, 0.35]$	$[-0.86, -0.18]$
Cluster	$[-0.98, 0.01]$	$[-0.15, 0.41]$	$[0.08, 0.47]$	$[-0.49, 0.08]$	$[-0.93, 1.03]$	$[-0.59, 0.09]$	$[-0.16, 0.34]$	$[-0.92, -0.11]$
AKM	$[-1.74, 0.77]$	$[-0.88, 1.14]$	$[-0.53, 1.08]$	$[-1.42, 1.01]$	$[-2.39, 2.50]$	$[-1.14, 0.64]$	$[-0.76, 0.95]$	$[-1.64, 0.61]$
AKM0	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$
Panel B: Reduced-Form Regression								
$\hat{\beta}$	-0.19	0.05	0.11	-0.08	0.04	-0.17	0.06	-0.36
Robust	$[-0.39, 0.02]$	$[-0.16, 0.26]$	$[-0.07, 0.28]$	$[-0.30, 0.14]$	$[-0.66, 0.73]$	$[-0.40, 0.05]$	$[-0.12, 0.24]$	$[-0.67, -0.04]$
Cluster	$[-0.37, 0.00]$	$[-0.07, 0.17]$	$[0.01, 0.20]$	$[-0.17, 0.01]$	$[-0.63, 0.71]$	$[-0.48, 0.14]$	$[-0.11, 0.24]$	$[-0.79, 0.08]$
AKM	$[-0.71, 0.33]$	$[-0.36, 0.46]$	$[-0.24, 0.45]$	$[-0.52, 0.37]$	$[-1.64, 1.72]$	$[-0.82, 0.47]$	$[-0.51, 0.64]$	$[-1.20, 0.49]$
AKM0	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$
Panel C: 2SLS First-Stage								
$\hat{\beta}$	0.38				0.69			
Robust	$[0.20, 0.57]$				$[0.41, 0.97]$			
Cluster	$[0.27, 0.49]$				$[0.35, 1.03]$			
AKM	$[0.10, 0.67]$				$[0.53, 0.85]$			
AKM0	$[-\infty, \infty]$				$[-\infty, \infty]$			

Notes: $\Delta \log E_i$ denotes log change in native employment; $\Delta \log w_i$ denotes log change in average weekly wages of native workers. In the specification *CZ (1980 weights)*, we use information on 722 CZs, 1980 weights and three time periods, 1980–1990, 1990–2000, 2000–2010; thus, $N = 2,166$ (722 CZs $\times$ 3 time periods). In the specification *CZ (1960 weights)*, we use information on 722 CZs, 1960 weights and one time period, 1970–1980; thus, $N = 722$ (722 CZs $\times$ 1 time period). 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; *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. We include all countries of origin in the analysis. See Tables F.1 and F.2 for a list of the origin countries included in the analysis.

Table F.13: Effect of immigration: analysis by MSA (including all origin countries)

Outcome:	$\Delta \log E_i$	$\Delta \log w_i$			$\Delta \log E_i$	$\Delta \log w_i$		
Workers:	All	All	High-Skill	Low-Skill	All	All	High-Skill	Low-Skill
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	MSA (1980 weights)				*MSA (1960 weights)*			
Panel A: 2SLS Regression								
$\hat{\beta}$	-1.41	0.16	0.28	-0.21	-0.18	-0.14	0.18	-0.35
Robust	$[-2.62, -0.21]$	$[-0.38, 0.71]$	$[-0.11, 0.68]$	$[-1.06, 0.63]$	$[-1.11, 0.75]$	$[-0.30, 0.02]$	$[-0.05, 0.42]$	$[-0.56, -0.14]$
Cluster	$[-2.55, -0.28]$	$[-0.14, 0.46]$	$[0.06, 0.51]$	$[-0.62, 0.20]$	$[-1.21, 0.86]$	$[-0.31, 0.04]$	$[-0.08, 0.44]$	$[-0.52, -0.17]$
AKM	$[-3.24, 0.41]$	$[-0.91, 1.23]$	$[-0.61, 1.18]$	$[-1.64, 1.21]$	$[-5.04, 4.69]$	$[-0.92, 0.64]$	$[-0.90, 1.26]$	$[-1.26, 0.56]$
AKM0	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$
Panel B: Reduced-Form Regression								
$\hat{\beta}$	-0.41	0.05	0.08	-0.06	-0.12	-0.10	0.13	-0.24
Robust	$[-0.61, -0.20]$	$[-0.13, 0.23]$	$[-0.07, 0.24]$	$[-0.27, 0.15]$	$[-0.78, 0.54]$	$[-0.22, 0.03]$	$[-0.05, 0.31]$	$[-0.43, -0.05]$
Cluster	$[-0.62, -0.19]$	$[-0.06, 0.15]$	$[0.00, 0.17]$	$[-0.16, 0.03]$	$[-0.83, 0.59]$	$[-0.23, 0.04]$	$[-0.08, 0.33]$	$[-0.42, -0.06]$
AKM	$[-0.91, 0.10]$	$[-0.28, 0.37]$	$[-0.21, 0.37]$	$[-0.45, 0.33]$	$[-3.47, 3.23]$	$[-0.64, 0.45]$	$[-0.60, 0.85]$	$[-0.89, 0.41]$
AKM0	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$	$[-\infty, \infty]$
Panel C: First-Stage								
$\hat{\beta}$	0.29				0.69			
Robust	$[0.10, 0.48]$				$[0.35, 1.02]$			
Cluster	$[0.16, 0.41]$				$[0.34, 1.03]$			
AKM	$[0.03, 0.54]$				$[0.44, 0.93]$			
AKM0	$[-\infty, \infty]$				$[-\infty, \infty]$			

Notes: $\Delta \log E_i$ denotes log change in native employment; $\Delta \log w_i$ denotes log change in average weekly wages of native workers. In the specification *MSA (1980 weights)*, we use information on 257 MSAs, 1980 weights and three time periods, 1980–1990, 1990–2000, 2000–2010; thus, $N = 651$ (257 MSAs $\times$ 3 time periods). In the specification *MSA (1960 weights)*, we use information on 217 MSAs, 1960 weights and one time period, 1970–1980; thus, $N = 217$ (217 CZs $\times$ 1 time period). Models are weighted by start-of-period MSA share of national population. All regressions include period dummies. 95% confidence intervals in square brackets. *Robust* is the Eicker-Huber-White standard error; *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. We include all countries of origin in the analysis. See Tables F.1 and F.2 for a list of the origin countries included in the analysis.

Table F.14: Immigration: p-values by CZ-Occ. and CZ-Educ. (including all origin countries)

	$\Delta \log E_i$	$\Delta \log w_i$			$\Delta \log E_i$	$\Delta \log w_i$	$\Delta \log E_i$	$\Delta \log w_i$		
	All	All	High-Skill	Low-Skill	All	All	All	All	High-Skill	Low-Skill
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
	CZ-50 Occupations				*CZ- Educ.*		*CZ-7 Occupations*			
Panel A: 2SLS Regression										
Robust	0.000	0.402	0.042	0.011	0.040	0.978	0.001	0.586	0.543	0.139
Cluster	0.001	0.764	0.445	0.403	0.040	0.958	0.004	0.496	0.363	0.005
AKM	0.004	0.721	0.380	0.267	0.434	0.988	0.077	0.801	0.798	0.454
AKM0	0.074	0.734	0.359	0.330	0.532	0.988	0.274	0.808	0.793	0.504
Panel B: Reduced-Form Regression										
Robust	0.000	0.372	0.072	0.002	0.106	0.978	0.001	0.560	0.567	0.074
Cluster	0.005	0.752	0.515	0.282	0.068	0.957	0.002	0.447	0.422	0.000
AKM	0.059	0.704	0.442	0.161	0.513	0.988	0.170	0.793	0.805	0.395
AKM0	0.074	0.734	0.359	0.330	0.532	0.988	0.274	0.808	0.793	0.504
Panel C: First-Stage										
Robust			0.000		0.000			0.000		
Cluster			0.000		0.000			0.000		
AKM			0.000		0.005			0.001		
AKM0			0.021		0.134			0.067		

Notes: $\Delta \log E_i$ denotes log change in native employment; $\Delta \log w_i$ denotes log change in average weekly wages of native workers. The specifications *CZ-50 Occupations*, *CZ-Educ.*, and *CZ-7 Occupations* differ in the definition of the unit of observation. In all three specifications, we use 1980 weights and three time periods, 1980–1990, 1990–2000, 2000–2010. Thus, $N = 108,300$ (722 CZs $\times$ 50 occupations $\times$ 3 time periods) for the *CZ-50 Occupations* specification; $N = 4,332$ (722 CZs $\times$ 2 education groups $\times$ 3 time periods) for the *CZ-Educ.* specification; and $N = 15,162$ (722 CZs $\times$ 7 occupations $\times$ 3 time periods) for the *CZ-7 Occupations* specification. Models are weighted by start-of-period occupation-region (or education group-region) share of national population. All regressions include occupation (or education group) and period dummies. 95% confidence intervals in square brackets. *Robust* is the Eicker-Huber-White standard error; *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. We all origin countries in the analysis. See Table F.1 for a list of the origin countries included in the analysis.

Table F.15: Effect of immigration: p-values by CZ (including all origin countries)

Outcome:	$\Delta \log E_i$	$\Delta \log w_i$			$\Delta \log E_i$	$\Delta \log w_i$		
Workers:	All	All	High-Skill	Low-Skill	All	All	High-Skill	Low-Skill
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Commuting Zone (1980 weights)				*Commuting Zone (1960 weights)*			
Panel A: 2SLS Regression								
Robust	0.130	0.607	0.142	0.536	0.917	0.072	0.478	0.003
Cluster	0.055	0.353	0.005	0.154	0.914	0.152	0.464	0.013
AKM	0.449	0.799	0.504	0.741	0.965	0.581	0.830	0.367
AKM0	0.578	0.797	0.522	0.759	0.965	0.619	0.830	0.463
Panel B: Reduced-Form Regression								
Robust	0.075	0.640	0.252	0.485	0.916	0.136	0.483	0.025
Cluster	0.047	0.404	0.032	0.091	0.913	0.274	0.473	0.108
AKM	0.481	0.808	0.553	0.730	0.965	0.600	0.827	0.408
AKM0	0.578	0.797	0.522	0.759	0.965	0.619	0.830	0.463
Panel C: 2SLS First-Stage								
Robust			0.000				0.000	
Cluster			0.000				0.000	
AKM			0.008				0.000	
AKM0			0.179				0.156	

Notes: $\Delta \log E_i$ denotes log change in native employment; $\Delta \log w_i$ denotes log change in average weekly wages of native workers. In the specification *CZ (1980 weights)*, we use information on 722 CZs, 1980 weights and three time periods, 1980–1990, 1990–2000, 2000–2010; thus, $N = 2,166$ (722 CZs $\times$ 3 time periods). In the specification *CZ (1960 weights)*, we use information on 722 CZs, 1960 weights and one time period, 1970–1980; thus, $N = 722$ (722 CZs $\times$ 1 time period). 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; *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. We include all origin countries in the analysis. See Tables F.1 and F.2 for a list of the origin countries included in the analysis.

Table F.16: Effect of immigration: p-values by MSA (including all origin countries)

Outcome:	$\Delta \log E_i$	$\Delta \log w_i$			$\Delta \log E_i$	$\Delta \log w_i$		
Workers:	All	All	High-Skill	Low-Skill	All	All	High-Skill	Low-Skill
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
	Commuting Zone (1980 weights)				*Commuting Zone (1960 weights)*			
Panel A: 2SLS Regression								
Robust	0.021	0.562	0.159	0.621	0.709	0.096	0.131	0.001
Cluster	0.015	0.294	0.014	0.306	0.737	0.118	0.168	0.000
AKM	0.130	0.768	0.534	0.769	0.943	0.727	0.739	0.450
AKM0	0.343	0.771	0.554	0.778	0.944	0.753	0.730	0.562
Panel B: Reduced-Form Regression								
Robust	0.000	0.616	0.302	0.571	0.717	0.127	0.172	0.012
Cluster	0.000	0.375	0.061	0.203	0.737	0.168	0.237	0.010
AKM	0.115	0.782	0.584	0.757	0.943	0.731	0.735	0.467
AKM0	0.343	0.771	0.554	0.778	0.944	0.753	0.730	0.562
Panel C: First-Stage								
Robust			0.003				0.000	
Cluster			0.000				0.000	
AKM			0.029				0.000	
AKM0			0.165				0.200	

Notes: $\Delta \log E_i$ denotes log change in native employment; $\Delta \log w_i$ denotes log change in average weekly wages of native workers. In the specification *CZ (1980 weights)*, we use information on 722 CZs, 1980 weights and three time periods, 1980–1990, 1990–2000, 2000–2010; thus, $N = 2,166$ (722 CZs $\times$ 3 time periods). In the specification *CZ (1960 weights)*, we use information on 722 CZs, 1960 weights and one time period, 1970–1980; thus, $N = 722$ (722 CZs $\times$ 1 time period). 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; *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. We include all origin countries in the analysis. See Tables F.1 and F.2 for a list of the origin countries included in the analysis.

References

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

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," *Journal of Political Economy*, 2018, *forthcoming*.

Altonji, Joseph G and David Card, "The Effects of Immigration on the Labor Market Outcomes of Less-skilled Natives," in John M. Abowd and Richard B. Freeman, eds., *Immigration, Trade, and the Labor Market*, Chicago: University of Chicago Press, 1991, pp. 201–234.

Amior, Michael and Alan Manning, "The Persistence of Local Joblessness," *American Economic Review*, 2018, *108* (7), 1942–1970.

Anderson, James E., "A Theoretical Foundation for the Gravity Equation," *The American Economic Review*, 1979, *69* (1), 106–116.

_ and Eric van Wincoop, "Gravity with Gravitas: A Solution to the Border Puzzle," *The American Economic Review*, 2003, *93* (1), 170–192.

Arkolakis, Costas, Arnaud Costinot, and Andrés Rodríguez-Clare, "New Trade Models, Same Old Gains?," *American Economic Review*, 2012, *102* (1), 94–130.

Armington, Paul S., "A Theory of Demand for Products Distinguished by Place of Production," *Staff Papers (International Monetary Fund)*, 1969, *16* (1), 159–178.

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.

Bai, Jushan, "Panel Data Model with Interactive Fixed Effects," *Econometrica*, 2009, *77* (4), 1229–1279.

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

Borjas, George J., "The Labor Demand Curve is Downward Sloping: Reexamining the Impact of Immigration on the Labor Market," *Quarterly Journal of Economics*, November 2003, *118* (4), 1335–1374.

Burstein, Ariel, Eduardo Morales, and Jonathan Vogel, "Changes in Between-Group Inequality: Computers, Occupations, and International Trade," *American Economic Journal: Macroeconomics*, 2019, *11* (2), 348–400.

_ , Gordon Hanson, Lin Tian, and Jonathan Vogel, "Tradability and the Labor-Market Impact of Immigration: Theory and Evidence from the U.S.," 2018. Unpublished manuscript, University of California Los Angeles.

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

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

Chao, John C., Norman R. Swanson, Jerry A. Hausman, Whitney K. Newey, and Tiemen Woutersen, "Asymptotic Distribution of JIVE in a Heteroskedastic IV Regression with Many Instruments," *Econometric Theory*, 2012, *12* (1), 42–86.

Dustmann, Christian, Uta Schönberg, and Jan Stuhler, "The Impact of Immigration: Why Do Studies Reach Such Different Results?," *Journal of Economic Perspectives*, 2016, *30* (4), 31–56.

Fajgelbaum, Pablo, Eduardo Morales, Juan Carlos Suárez Serrato, and Owen Zidar, "State Taxes and Spatial Misallocation," *Review of Economic Studies*, 2019, *86* (1), 333–376.

Galle, Simon, Andrés Rodríguez-Clare, and Moises Yi, "Slicing the Pie: Quantifying the Aggregate and Distributional Effects of Trade," October 2018. Unpublished manuscript, University of California Berkeley.

Gomez, Matthieu, "REGIFE: Stata module to estimate linear models with interactive fixed effects," 2017. Statistical Software Components, Boston College, Department of Economics.

Imbens, Guido W. and Donald B. Rubin, *Causal Inference for Statistics, Social, and Biomedical Sciences: an Introduction*, New York, NY: Cambridge University Press, 2015.

Jaeger, David A, Joakim Ruist, and Jan Stuhler, "Shift-share instruments and the impact of immigration," Technical Report 24285, National Bureau of Economic Research February 2018.

Jaeger, David A., Joakim Ruist, and Jan Stuhler, "Shift-Share Instruments and the Impact of Immigration," February 2018. Unpublished manuscript, CUNY University.

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*, Amsterdam: 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.

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.