Shift-Share Designs: Theory and Inference*

Rodrigo Adão† Michal Kolesár‡ Eduardo Morales§

June 20, 2018

Abstract

Since Bartik (1991), it has become popular in empirical studies to estimate regressions in which the variable of interest is a shift-share, such as when a regional labor market outcome is regressed on a weighted average of observed sectoral shocks, using the regional sector shares as weights. In this paper, we discuss inference in these regressions. We show that standard economic models imply that the regression residuals are likely to be correlated across regions with similar sector shares, independently of their geographic location. These correlations are ignored by inference procedures commonly used in these regressions, which can lead to severe undercoverage. In regressions studying the effect of randomly generated placebo sectoral shocks on actual labor market outcomes in U.S. commuting zones, we find that a 5% level significance test based on standard errors clustered at the state level rejects the null of no effect in up to 45% of the placebo interventions. We derive novel confidence intervals that correctly account for the potential correlation in the regression residuals.

*We thank Juan Manuel Castro Vicenzi for excellent research assistance, and seminar audiences at Carleton University, Yale University, Globalization & Inequality BFI conference, IDB, GTDW, Unil, EESP-FGV and PUC-Rio for very useful comments. All errors are our own.

†University of Chicago Booth School of Business. Email: radao@uchicago.edu

‡Princeton University. Email: mkolesar@princeton.edu

§Princeton University. Email: ecmorales@princeton.edu

1 Introduction

We study inference in shift-share designs—regression designs in which one studies the impact of a set of shocks, or "shifters", on units differentially exposed to them, and whose differential exposure depends on a set of weights, or "shares". More precisely, shift-share designs are regressions that have the form

$$Y_i = \beta X_i + Z_i'\delta + \epsilon_i, \qquad \text{where} \quad X_i \equiv \sum_{s=1}^{S} w_{is}\mathcal{X}_s, \qquad \text{and} \qquad \sum_{s=1}^{S} w_{is} = 1. \tag{1}$$

For example, in an investigation of the impact of sectoral demand shifters on regional employment changes, Y_i corresponds to the change in employment in region i, the shifter $\mathcal{X}_s$ is a measure of the change in demand for the good produced by sector s, and the share w_{is} may be measured as the initial share of region i's employment in sector s. Other observed characteristics of region i are captured by the vector Z_i, which includes the intercept, and ϵ_i is the regression residual.[1]

Shift-share specifications can be very appealing in many contexts: they are simple to apply and have the potential to both circumvent complicated endogeneity issues, and to provide estimates of treatment effects that are robust to different microfoundations. As a result, numerous influential studies, including Bartik (1991), Card (2001), Autor and Dorn (2013) and Autor, Dorn and Hanson (2013) have exploited these designs as their main specifications. At the same time, two types of concerns have been raised: first, the designs may not be appropriate in the presence of cross-regional general equilibrium effects, and second, it is unclear whether the estimand is interesting when the effects of the shifters $\mathcal{X}_s$ are heterogeneous across sectors and regions.[2] In this paper, we put these two concerns aside—we assume no cross-regional spillover effects, and, while we allow for shifter-effect heterogeneity and clarify the definition of the regression estimand in this case, we are agnostic about its policy relevance. We do this to focus on a different question: how to perform inference in shift-share regressions.

We find that usual standard error formulas may substantially understate the true variability of the estimates in shift-share regressions. We highlight this problem in two steps. First, we show that a simple multi-sector gravity trade model implies that the residuals ϵ_i in the shift-share specification in eq. (1) will have a shift-share structure similar to that of the regressor X_i, because, in addition to the shifter $\mathcal{X}_s$, regions are subject to other sector-level shocks, not all of which are observable. Consequently, the regression residuals for two regions with similar shares will be correlated, even if they are far apart geographically, because they will be exposed to similar combinations of unobserved shifters. In practice, researchers are typically careful to allow for correlations of residuals among regions that are geographically close to each other by reporting clustered standard errors, with clusters defined as

[1]For simplicity of exposition, we refer to the unit of observation i at which the outcome variable Y_i is measured as a region, and the unit of observation s at which the shifter $\mathcal{X}_s$ is measured as a sector. However, the setup applies to any setting in which the regressor X_i of interest admits the representation in eq. (1).

[2]These concerns are an instance of a broader discussion about whether regression analysis is useful to estimate policy-relevant parameters; see Heckman, Lance and Taber (1998), and Heckman and Vytlacil (2007a,b). For a discussion in the context of regional shift-share regressions, see Redding and Rossi-Hansberg (2017), Monte, Redding and Rossi-Hansberg (2018) and Adão, Arkolakis and Esposito (2018). Shift-share regressions have however also been used to estimate structural model parameters, see Diamond (2016), Adão (2016), Galle, Rodríguez-Clare and Yi (2017), Burstein et al. (2018b) and Bartelme (2018).

groups of regions that belong to the same administrative entity (e.g. clustering U.S. commuting zones by states). However, geographic clustering does not properly account for the shift-share correlation structure, and, as a result, generates standard errors that are too small.

Second, to illustrate the empirical importance of this problem, we conduct a placebo exercise. As outcomes, we use 2000–2007 changes in labor market outcomes for 722 commuting zones in the United States. We then build a shift-share regressor by combining actual sectoral employment shares in 1990 for each of 398 sectors that correspond to 4-digit SIC manufacturing industries with sector-level shifters that are randomly generated. We repeat this many times to construct many placebo samples. Since the shifters are randomly generated, their true effect is zero. Valid 5% level significance tests should therefore reject the null of no effect for at most 5% of the samples. We find that traditional standard errors—clustering on state as well as heteroscedasticity-robust, unclustered errors—are much smaller than the true standard deviation of the OLS estimator, and, as a result, lead to severe overrejection. Depending on the labor market outcome, the rejection rate can be as high as 55% for heteroscedasticity-robust and 45% for standard errors clustered on state, and it is never below 17%. In other words, suppose that instead of using real shifters, 100 researchers used randomly generated shocks and mistakenly thought that these are, for example, actual changes in trade flows, tariffs or the number of foreign workers employed in an industry. If these researchers were to use standard inference procedures, up to 55 of them would find a statistically significant effect of the randomly generated shocks on U.S. labor market outcomes between 2000 and 2007. The overrejection is even more severe when 2- and 3-digit SIC codes are used to define the sectors, so that the total number of sectors is smaller.

To correctly account for the correlations between residuals for regions with similar shares, we derive novel confidence intervals that are easy to construct, and remain valid under an arbitrary correlation structure of the residuals. The key assumption underlying their validity is that, conditionally on the covariates Z_i, the shifters are as good as randomly assigned and independent across sectors.[3] In the special case in which each region is fully specialized in one sector (i.e. for every i, $w_{is} = 1$ for some sector s), the procedure is identical to using the usual clustered standard error formula, with clusters defined as groups of regions specialized in the same sector.[4] Using our novel confidence intervals, we revisit two seminal papers that have employed shift-share specifications to reach their main conclusions, Bartik (1991) and Autor et al. (2013), and, as we describe in more detail below, find that, in some specifications, our confidence intervals are up to over twice as long as those based on the usual standard errors.

The starting point for our analysis is a simple multi-sector model, in which labor is the sole factor of production. Our model is a simplified version of that in Adão, Arkolakis and Esposito (2018) and its purpose is threefold. First, it helps us characterize the estimand β of the shift-share regression design in eq. (1) when the shifter effects are heterogeneous across sectors and regions. Second, we use it to guide the specification of the statistical model used to determine the properties of different

[3]This result is similar to that in Barrios et al. (2012), who consider cross-section regressions estimated at an individual level, with a variable of interest that varies at a state level. They show that standard errors clustered on state are valid when the variable of interest is as good as randomly assigned, and independent across states, even when the residuals are correlated across states.

[4]For an application that falls within this special case of our inference procedure, see Amiti and Weinstein (2011).

inference procedures in the context of shift-share regressions. Third, and most important, we use it to characterize the likely correlation structure of the residuals ϵ_i in eq. (1).

With these goals in mind, we shut down in our model cross-regional spillovers by assuming that individuals cannot move across regions, and account for cross-sector spillovers within each region by allowing individuals to freely choose between employment in any sector of the economy and being out of the labor force. We then characterize the general equilibrium impact of different sector-level shocks, which, in the model, correspond to changes in productivity, final demand, and foreign competition. We establish that the general equilibrium impact of sector-level shocks on changes in employment and wages may be written, to a first-order approximation around the initial equilibrium, as an additive function of multiple shift-share structures, where the shares correspond to sectoral employment shares in the initial equilibrium and the shifters correspond to the different sources of sector-level shocks in the model.[5]

We then focus on a set of small open regional economies and consider the problem of recovering the effect of one particular source of sector-level shocks—foreign competition, as measured by changes in world sectoral prices—on regional labor market outcomes. To make precise what we mean by "the effect on an outcome" and to help us express in general terms the implications of our economic model, we use a potential outcome framework. Let $Y_i(x_1, \ldots, x_S)$ denote the potential outcome in region i if the shifters were exogenously set to $x_1, \ldots, x_S$, so that the observed outcome would be $Y_i = Y_i(\mathcal{X}_1, \ldots, \mathcal{X}_S)$. Our economic model implies that the potential outcomes have the structure

$$Y_i(x_1, \ldots, x_S) = Y_i(0) + \sum_{i=1}^{S} w_{is} x_s \beta_{is}, \tag{2}$$

where $Y_i(0) = Y_i(0, \ldots, 0)$ indicates region i's outcome if all the sector-level shocks of interest were set to zero, and the shares w_{is} correspond to sectoral employment shares in the initial equilibrium.

Our economic model has two other important implications. First, even after conditioning on the observable measure of exposure of region i to a sector s (i.e. the share w_{is}), the general equilibrium impact of sectoral shocks on regional outcomes is a function of structural parameters that takes different values across sectors and regions; i.e. the parameters β_{is} vary across i and s. Second, the potential outcome $Y_i(0)$ accounts for the impact of sectoral shocks other than the shifters $\mathcal{X}_s$ of interest, and consequently, includes terms that also have a shift-share structure. In terms of the elements of the regression (1), our economic model thus implies that the regression residual generally has the structure $\epsilon_i = \sum_{s=1}^{S} w_{is} \mathcal{E}_s + \eta_i$, where w_{is} is the same share entering the construction of the regressor X_i, $\mathcal{E}_s$ is an unobserved shifter that varies at the same level as the shifter of interest $\mathcal{X}_s$, and η_i captures other unobserved determinants of the outcome Y_i. Consequently, whenever two regions have similar exposure to the observed shifters, they will also tend to have similar values of the residuals ϵ_i.

In the second part of the paper, we establish the asymptotic properties of the OLS estimator of

[5] We show in Appendix A.4 that this conclusion is robust to additionally allowing for sector-specific factors of production, as in the seminal paper by Jones (1971) and in the recent application in Kovak (2013). We show in Appendix A.5 that the expression for changes in regional employment is analogous to that generated by models with Roy-Fréchet workers with idiosyncratic sectoral productivities, as in Galle, Rodríguez-Clare and Yi (2017), Lee (2017) and Burstein, Morales and Vogel (2018a). We discuss in Appendix A.3 the differences in predictions that arise from allowing for labor mobility across regions as in Allen and Arkolakis (2016) and Redding (2016) (see Redding and Rossi-Hansberg (2017) for a review of this literature).

β in eq. (1). We show that the estimand β corresponds to a weighted average of the heterogeneous parameters β_{is}, and that the OLS estimator is consistent as the number of sectors and regions goes to infinity. As we discuss in more detail in Section 4.1.1 below, the estimand β does *not* in general equal a weighted average of the shifter effects $w_{is}\beta_{is}$. We also show that the distribution of the OLS estimator converges to a normal distribution with variance that depends on the cross-regional correlation of ϵ_i. When this correlation is zero, this asymptotic variance is identical to the usual asymptotic variance formula under independent, heteroscedastic errors ϵ_i. However, the asymptotic variance can be substantially larger if ϵ_i incorporates a shift-share aggregator of unobserved sectoral shocks, as in our economic model. We show that our new standard error formula consistently estimates this asymptotic variance. In our placebo exercise, we show that it correctly accounts for the correlations between the regression residuals, and therefore, unlike existing methods, yields tests with correct size.

To illustrate the practical application of our new standard errors, we revisit two applications of shift-share regressions that are popular in the literature. First, we revisit the empirical analysis in Autor, Dorn and Hanson (2013). We find that although the 95% confidence intervals that we propose are between 30% and over 100% larger than those computed using standard inference procedures, the effect of import flows from China on the reduction in employment and wages in the U.S. during the period 1999 to 2007 remains statistically significant. Our results indicate that, given the identification assumptions in Autor, Dorn and Hanson (2013), the negative effect of import flows from China on U.S. regional labor market outcomes could have been much larger than implied by the usual confidence intervals.

Second, we estimate the elasticity of average regional wages with respect to regional employment using the instrumental variables approach in Bartik (1991). Our confidence intervals are in this case virtually identical to those constructed using standard approaches. Intuitively, the sectoral shifter used in this empirical approach—the change in national employment by sector—soaks up all sectoral shocks affecting the labor market outcome of interest and, consequently, there is in this case no shift-share structure left in the regression residuals. To illustrate this point, we additionally estimate the same labor supply elasticity using the instrumental variable introduced in Autor, Dorn and Hanson (2013). The sector shifter in this case—changes in trade flows from China to developed countries other than the U.S.—leaves other sectoral shocks affecting U.S. commuting zones in the residual and, consequently, standard inference procedures generate confidence intervals that are too small.

Shift-share designs have been applied with great success to estimate the consequences of a variety of shocks to economic fundamentals. In seminal papers, Bartik (1991) and Blanchard and Katz (1992) explore shift-share strategies to analyze the impact on local labor markets of sector-level demand shocks measured as changes in national sectoral employment. More recently, this strategy has been applied to investigate the local labor market consequences of exposure to observable shocks of various sources, including international trade competition (Topalova, 2007, 2010; Kovak, 2013; Autor, Dorn and Hanson, 2013; Dix-Carneiro and Kovak, 2017; Pierce and Schott, 2017), credit supply (Greenstone, Mas and Nguyen, 2015), technological change (Acemoglu and Restrepo, 2017, 2018), and industry reallocation (Chodorow-Reich and Wieland, 2018). Shift-share regressors have been used to study the impact of sectoral shocks on alternative outcomes, such as political preferences (Autor et al., 2017a; Che et al., 2017; Colantone and Stanig, 2018), marriage patterns (Autor, Dorn and Hanson, 2018), crime levels (Dix-Carneiro, Soares and Ulyssea, 2017), innovation (Acemoglu and Linn, 2004; Autor et al., 2017b). Shift-share regressors also have been extensively used to estimate the impact of immigration on local labor markets, as in Card (2001) and many other papers following his approach; see reviews of this literature in Lewis and Peri (2015) and Dustmann, Schönberg and Stuhler (2016). Finally, recent papers have explored versions of shift-share strategies to estimate the effect on firms of shocks to outsourcing costs and foreign demand (Hummels et al., 2014; Aghion et al., 2018).[6]

Our paper is related to three other recent papers that study the statistical properties of shift-share specifications. First, Goldsmith-Pinkham, Sorkin and Swift (2018) focus on the case in which the shift-share regressor is used as an instrument in an instrumental variables regression. Within this setting, these authors argue that the full vector of shares $(w_{i1}, \ldots, w_{iS})$ can be used as an instrument for the endogenous treatment. This approach thus requires that the vector of shares be as good as randomly assigned conditional on the shifters, and independent across regions, or clusters of regions. Given our interest in exploring the impact of a set of shifters on an outcome of interest, rather than the impact of a set of shares, this approach is not attractive in our setting. That said, there may be otter settings in which this approach is more appealing. Second, Borusyak, Hull and Jaravel (2018), also focusing on the use of shift-share regressor as an instrumental variable, show that it is a valid instrument if the set of shifters is as good as randomly assigned conditional on the shares, and discuss consistency of the IV estimator as the number of sectors goes to infinity. Our approach to inference and the way we set up the potential outcomes framework follows their identification insight; this way of thinking about the shift-share design validity is also natural given our economic model. Third, Jaeger, Ruist and Stuhler (2018) study complications with this instrument when the shift-share regressor is correlated over time and there is a sluggish adjustment of the outcome variable of interest to changes in this regressor.

The rest of this paper is organized as follows. Section 2 introduces our economic model, and Section 3 maps the model implications into a potential outcomes framework. Section 4 establishes the asymptotic properties of the OLS estimator of β in eq. (1), and provides a consistent estimator of its standard error; we also show how the results extend to an instrumental variables setting. Section 5 presents the results of a placebo exercise in which we compare our novel inference procedures to those previously used in the literature on shift-share specifications. Section 6 revisits the conclusions from several prior applications of shift-share regression analysis, and Section 7 concludes. Proofs as well as additional results are collected in Appendices A to D.

2 Model

This section presents a stylized economic model featuring local labor markets, and discusses two key implications. Specifically, we illustrate that the general equilibrium impact of sectoral shifters

[6] There is also a large literature using a shift-share approach that treats the shifters as unobserved, and for this reason uses the shares directly as regressors. This approach has been applied to investigate the impact of technological shifters (Autor and Dorn, 2013; Bustos, Caprettini and Ponticelli, 2016), credit supply shifters (Huber, 2018), and immigration shifters (Card and Dinardo, 2000; Dustmann, Frattini and Preston, 2013; Monras, 2015). We treat the sectoral shares $\mathcal{X}_s$ in eq. (1) as observed and, leave the extension to the unobserved case to future work.

on regional labor market outcomes: (a) has a shift-share structure; and (b) is heterogeneous across regions and sectors, even conditional on the shares.

2.1 Environment

We consider a model with multiple sectors $s = 1, \dots, S$ and multiple regions $i, j = 1, \dots, J$. Regions are partitioned into $c = 1, \dots, C$ countries, and we denote the set of regions located in a country c by J_c. Region i has a population of M_i individuals who cannot move across regions.

Production. Each sector s in region i has a representative firm that produces a differentiated good. We assume that the quantity Q_{is} produced by sector s in region i is produced using labor with a linear technology A_{is},

$$Q_{is} = A_{is} L_{is}, \tag{3}$$

where L_{is} denotes the number of workers 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-region pair specific differentiated goods. Specifically, we assume that individuals have Cobb-Douglas preferences over sectoral composite goods,

$$C_j = \prod_{s=1}^{S} \left(C_{js} \right)^{\gamma_s}, \tag{4}$$

where C_j denotes 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 constant elasticity of substitution (CES) index of the sector s differentiated goods produced in different regions; i.e.

$$C_{js} = \left[\sum_{i=1}^{J} \left(c_{ijs} \right)^{\frac{\sigma_s - 1}{\sigma_s}} \right]^{\frac{\sigma_s}{\sigma_s - 1}}, \qquad \sigma_s \in (1, \infty), \tag{5}$$

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. Workers located in a region i have the choice of being employed in one of the sectors $s = 1, \dots, S$ of the economy or opting for non-employment, which we index as $s = 0$. Conditional on being employed, all workers 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 living in region j is

$$U(\iota \mid C_j) = \begin{cases} u(\iota) C_j & \text{if employed in any sector } s = 1, \ldots, S, \\ C_j & \text{if not employed } (s = 0). \end{cases} \tag{6}$$

We assume that $u(\iota)$ is distributed independently and identically across individuals ι according to a Pareto distribution with scale ν_i and shape ϕ, so that the cumulative distribution function of $u(\iota)$ is given by

$$F_u(u) = 1 - \left(\frac{u}{\nu_i}\right)^{-\phi}, \qquad u \in [\nu_i, \infty), \qquad \phi > 1. \tag{7}$$

If a worker living in region j chooses to be employed, she will earn wage ω_j (as workers are indifferent about the sector of employment and can move freely across sectors, region-specific wages must be equalized across sectors in equilibrium). If a worker chooses to not be employed, she receives a benefit b_j.[7] We denote the total number of employed workers in region j by L_j, and the employment rate in j as $e_j \equiv L_j / M_j$.

Market structure. Goods and labor markets are perfectly competitive.

Trade costs. For simplicity, we assume that there are no trade costs, which implies that the equilibrium price of the good produced in region i is the same in every other region j of the world economy; i.e. $p_{ijs} = p_{is}$ for $j = 1, \ldots, J$. Consequently, for every sector s there is a composite sectoral good that has identical price P_s in all regions; i.e.

$$(P_s)^{1-\sigma_s} = \sum_{s=1}^{S} (p_{is})^{1-\sigma_s}, \tag{8}$$

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

2.2 Equilibrium

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

Consumption. We first solve the expenditure minimization problem of an individual residing in region j. Given the sector-level utility in eq. (5) and the condition that $p_{ijs} = p_{is}$ for $j = 1, \ldots, J$, all regions j have identical spending shares x_{is} on goods from region i, given by

$$x_{is} = \left(\frac{p_{is}}{P_s}\right)^{1-\sigma_s}. \tag{9}$$

[7] We assume that these benefits are paid by a national government that imposes a flat tax equal to χ_c on all income earners in a country c. The budget constraint of the government is thus $\sum_{j \in J_c} \{\chi_c(\omega_j e_j + b_j(1 - e_j)) M_j\} = \sum_{j \in J_c} \{b_j (1 - e_j) M_j\}$. 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 described in the main text.

Labor supply. Every worker maximizes the utility function in eq. (6) in order to decide whether to be employed. Consequently, conditional on the wage ω_i and the non-employment benefit b_i, the employment rate in region i is $e_i = \Pr[u_i(\iota)\omega_i > b_i] = 1 - \Pr[u_i(\iota) < b_i/\omega_i]$. It therefore follows from eq. (7) that

$$e_i = v_i \omega_i^{\phi}, \qquad v_i \equiv (\nu_i / b_i)^{\phi}. \tag{10}$$

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

$$p_{is} = \frac{\omega_i}{A_{is}}. \tag{11}$$

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 $W \equiv \sum_i W_i$. Given preferences in eq. (4), 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 $x_{is}\gamma_s W$. 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. (9)

$$L_{is} = (\omega_i)^{-\sigma_s} (A_{is} P_s)^{\sigma_s - 1} \gamma_s W, \tag{12}$$

and total labor demand in region i is

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

Labor market clearing. Labor market clearing requires labor supply in eq. (10) to equal labor demand in eq. (13):

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

Equilibrium. Given technology parameters $\{A_{is}\}_{i=1,s=1}^{J,S}$, preference parameters $\{(\sigma_s, \gamma_s)\}_{s=1}^{S}$, total employment and labor supply parameters $\{(M_i, v_i)\}_{i=1}^{J}$, and normalizing world income to equal 1, $W = 1$, we can use eqs. (8), (11) and (14) to solve for the equilibrium wage in every world region, $\{\omega_i\}_{i=1}^{J}$, the equilibrium price of every sector-region specific good $\{p_{is}\}_{i=1,s=1}^{J,S}$, and the sectoral price indices $\{P_s\}_{s=1}^{S}$. Given these equilibrium wages and sectoral price indices, we can use eq. (13) to solve for the equilibrium level of employment in every region, $\{L_i\}_{i=J}^{J}$.

2.3 Labor market impact of economic shocks in a small open economy

We assume that, in every period, our model characterizes the labor market equilibrium in every region of the world economy. Across periods, we assume that the parameters $\{\sigma_s\}_{s=1}^{S}$, ϕ, and $\{M_i\}_{i=1}^{J}$ are fixed and that all changes in the labor market outcomes $\{(\omega_i, L_i)\}_{i=1}^{J}$ are generated by changes in technology $\{A_{is}\}_{i=1,s=1}^{J,S}$, sectoral preferences $\{\gamma_s\}_{s=1}^{S}$ and labor supply parameters $\{v_i\}_{i=1}^{J}$. Specifically, in every period, the values of these exogenous parameters correspond to realizations from an unknown joint distribution $F(\cdot)$:[8]

$$(\{A_{is}\}_{i=1,s=1}^{J,S}, \{\gamma_s\}_{s=1}^{S}, \{v_i\}_{i=1}^{J}) \sim F(\cdot). \tag{15}$$

We focus in this section on understanding how changes in these exogenous parameters affect the labor market equilibrium in a set of "small" regions whose share in world output is approximately zero for all sectors, i.e., $x_{is} \approx 0$ for $s = 1, \ldots, S$, with x_{is} defined in eq. (9). Without loss of generality, we assume that all small regions of interest belong to the same country c. Thus, the sectoral price index P_s of every sector s will not depend on the technology and labor supply parameters of the regions in c (i.e., $\{P_s\}_{s=1}^{S}$ does not depend on $\{A_{is}\}_{s=1,i\in J_c}^{S}$ and $\{v_i\}_{i\in J_c}$) and, from the perspective of any one of these regions, changes in sectoral prices operate as exogenous shocks. Furthermore, as illustrated in eqs. (13) and (14) these sectoral prices mediate the impact of all foreign technology and labor supply parameters on the labor market equilibrium of every region in country c.

We use $\hat{z} = \ln(z^t/z^0)$ to denote log-changes in a variable z between some initial period $t = 0$ and any other period t. Therefore, to a first-order approximation around the initial equilibrium, the equilibrium conditions in eqs. (12) to (14) imply that, for every region i in country c,

$$\hat{L}_{is} = \hat{\gamma}_s + (\sigma_s - 1)\left(\hat{A}_{is} + \hat{P}_s\right) - \sigma_s \hat{\omega}_i, \tag{16}$$

$$\hat{\omega}_i = \lambda_i \sum_{s=1}^{S} l_{is}^0 \left[\hat{\gamma}_s + (\sigma_s - 1)\left(\hat{A}_{is} + \hat{P}_s\right)\right] - \lambda_i \hat{v}_i, \tag{17}$$

$$\hat{L}_i = \phi\lambda_i \sum_{s=1}^{S} l_{is}^0 \left[\hat{\gamma}_s + (\sigma_s - 1)\left(\hat{A}_{is} + \hat{P}_s\right)\right] + (1 - \phi\lambda_i)\,\hat{v}_i, \tag{18}$$

where l_{is}^0 is the initial employment share of sector s in region i, $l_{is}^0 \equiv L_{is}^0/L_i$, and $\lambda_i \equiv \left[\phi + \sum_s l_{is}^0 \sigma_s\right]^{-1}$. Equations (16) and (17) illustrate that, for any region i, the change in employment in a particular sector s, $\hat{L}_{is}$, depends: (a) directly on sector s preferences, technology and price shocks, $\hat{\gamma}_s$, $\hat{A}_{is}$, and $\hat{P}_s$; and, (b) through changes in equilibrium wages, $\hat{\omega}_i$, on region i labor supply shocks, $\hat{v}_i$, and on preferences, technology and price shocks to any sector, $\{\hat{\gamma}_s, \hat{A}_{is}, \hat{P}_s\}_{s=1}^{S}$. The model described in Section 2.1 thus features cross-sectoral spillovers within each region: exogenous shocks to region i's labor demand in a sector k affect equilibrium wages and, consequently, also affect region i's labor demand in every other sector s. Importantly, as reflected in eq. (17), the impact of a sectoral shock on equilibrium wages in a region i increases in the initial share of region i's employment in sector s, l_{is}^0. Consequently, the expression for the total change in wages in eq. (17) includes several terms that have a shift-share structure: the "share" term is always the initial employment share in a sector l_{is}^0, and the "shift" term is either a preference, a technological, or a price shock in the corresponding sector: $\hat{\gamma}_s$, $\hat{A}_{is}$, or $\hat{P}_s$. Furthermore, as eq. (18) shows, the expression for the change in total employment in a region i inherits the same shift-share structure featuring in the expression for the total change in

[8]Since the labor market equilibrium depends on b_i and ν_i only through v_i, it suffices to specify a distribution for ν_i

wages in eq. (17).[9] We summarize the content of this paragraph in the following remark:

Remark 1. *According to the model in Section 2.1, the general equilibrium impact of sectoral shocks on regional labor market outcomes have a shift-share structure.*

For concreteness, we focus in the remainder of this section on the impact of changes in the price in a sector s on changes in wages and employment across regions in the country of interest c. Similar points would also apply if we were interested in the effect of sectoral shocks that affect the technology parameters A_{is} symmetrically in all regions i of country c. Specifically, without loss of generality, we can rewrite $\hat{A}_{is} = \hat{A}_s + \hat{\tilde{A}}_{is}$ and, according to eqs. (17) and (18), $d\hat{\omega}_i/d\hat{P}_s = d\hat{\omega}_i/d\hat{A}_s$ and $d\hat{L}_i/d\hat{P}_s = d\hat{L}_i/d\hat{A}_s$.

As the expressions in eqs. (17) and (18) show, the response of regional wages and total employment to a particular change in the world price of a sector s is

$$d\hat{L}_i/d\hat{P}_s = l^0_{is}\beta_{L,is}, \qquad \text{with} \quad \beta_{L,is} \equiv \phi\lambda_i(\sigma_s - 1); \tag{19}$$

$$d\hat{\omega}_i/d\hat{P}_s = l^0_{is}\beta_{\omega,is}, \qquad \text{with} \quad \beta_{\omega,is} \equiv \lambda_i(\sigma_s - 1). \tag{20}$$

The general equilibrium impact of the world price of a sector s on both region-i's total employment and wages is thus heterogeneous across sectors and regions. This heterogeneity is not exclusively due to differences in the initial employment share l^0_{is}, as the parameters $\beta_{L,is}$ and $\beta_{\omega,is}$ will also generally vary across regions and sectors.[10] This is relevant as, while standard datasets will usually contain information on the initial employment shares for every sector and region, the parameters $\beta_{L,is}$ and $\beta_{\omega,is}$ are not generally known or directly observed and, thus, need to be estimated.

Remark 2. *Even conditional on the shares, the general equilibrium impact of sectoral shifters on regional labor market outcomes will generally be heterogeneous across regions and sectors.*

Remarks 1 and 2 are not specific to the model described in Section 2.1. The property that, for some given functions of structural parameters $\beta_{L,is}$ and $\beta_{\omega,is}$, one can write the impact of $\hat{P}_s$ on $\hat{\omega}_i$ and $\hat{L}_i$ as $l^0_{is}\beta_{L,is}$ and $l^0_{is}\beta_{\omega,is}$, respectively, also holds for models other than that described in Section 2.1. Specifically, we consider in Appendix A.4 a model à la Jones (1971) in which the production function in eq. (3) is generalized to be Cobb-Douglas in labor L_{is} and a sector-specific input K_{is} that is in fixed supply every region i. In Appendix A.5, we consider a Roy (1951) model in which workers have heterogeneous preferences for being employed in the different $s = 1, \ldots, S$ sectors. Both models are isomorphic to that described in Section 2.1, with the only difference being the mapping of $\beta_{L,is}$ and $\beta_{\omega,is}$ to structural parameters of the corresponding model.[11, 12]

[9] As illustrated in Appendix A.2, once we substitute the expression for wage changes in eq. (17) into eq. (16), the expression for the change in employment in a specific region i and sector s also has a term with a shift-share structure.

[10] The parameters $\beta_{L,is}$ and $\beta_{\omega,is}$ will vary neither across regions nor sectors if and only if all sectors share the same elasticity of substitution, i.e. $\sigma_s = \sigma_{s'}$ for all s, s'.

[11] According to the models described in Appendices A.4 and A.5, the assumptions on structural parameters needed so that the corresponding model predicts $\beta_{L,is}$ and $\beta_{\omega,is}$ to be constant across regions and sectors are even stronger than those required by the model described in Section 2.1. See Appendices A.4 and A.5 for details.

[12] The model in Appendix A.4 is very similar to that Kovak (2013). See Autor, Dorn and Hanson (2013, 2016) for an alternative modeling approach that generates predictions similar to those in eqs. (16), (17), (18), (19) and (20).

The model described in Section 2.1 assumes that each region's population M_i is fixed. We extend this framework in Appendix A.3 to allow for migration across regions within each country, and show that the elasticity of regional labor market outcomes with respect to $\hat{P}_s$ will depend not only on the elasticities described in eqs. (19) and (20) but also on how $\hat{P}_s$ impacts labor market outcomes in other regions within the same country. In the context of this model with migration, the elasticities described in eqs. (19) and (20) thus capture only partial effects that do not account for cross-regional spillovers.

3 From theory to inference

In Section 2.3, we focused on the impact of sectoral price shocks on labor market outcomes across the regions of a small open economy. More generally, we are interested in estimating the effect of sectoral shocks or shifters on an outcome of interest that varies at the regional level. To make precise what we mean by "the effect on an outcome", we use the potential outcome notation, writing $Y_i(x_1, \ldots, x_S)$ to denote the potential (counterfactual) outcome that would occur in a region i if the shocks to the S sectors were exogenously set to $\{x_s\}_{s=1}^S$. We assume that the potential outcomes are linear in the shocks,

$$Y_i(x_1, \ldots, x_S) = Y_i(0) + \sum_{i=1}^{S} w_{is} x_s \beta_{is}, \qquad \text{where} \qquad \sum_{s=1}^{S} w_{is} = 1, \tag{21}$$

and $Y_i(0) \equiv Y_i(0, \ldots, 0)$ denotes the potential outcome in region i when all shocks $\{x_s\}_{s=1}^S$ are set to zero. The equilibrium relationships in eqs. (17) and (18) are both examples of eq. (21). For instance, eq. (17) maps into eq. (21) if, for every region i and sector s, we define

$$Y_i = \hat{\omega}_i,\ w_{is} = l_{is}^0,\ x_s = \hat{P}_s,\ \beta_{is} = \beta_{\omega,is}, \quad Y_i(0) = \lambda_i \sum_{s=1}^{S} l_{is}^0 \left[\hat{\gamma}_s + (\sigma_s - 1)\hat{A}_{is}\right] - \lambda_i \hat{v}_i. \tag{22}$$

According to eq. (21), increasing x_s by one unit, and holding the shocks to the other sectors constant, leads to an increase in region i's outcome by $w_{is}\beta_{is}$ units. This is the treatment effect of x_s on $Y_i(x_1, \ldots, x_S)$. The actual (observed) outcome is given by $Y_i = Y_i(\mathcal{X}_1, \ldots, \mathcal{X}_S)$, where $\mathcal{X}_1, \ldots, \mathcal{X}_S$ denote the realized values of these sectoral shifters.

We assume that we observe data for N regions and S sectors on sectoral shocks $\mathcal{X}_s$, regional outcomes Y_i, region- and sector-specific shares w_{is}, and a vector of regional controls Z_i. For example, we may think of the N regions in our data as the set J_c of regions included in the country c defined in Section 2.3. We are interested in the properties of the OLS estimator $\hat{\beta}$ of the coefficient on the shift-share regressor $X_i = \sum_{s=1}^S w_{is}\mathcal{X}_s$ in a regression of Y_i onto X_i and an additional vector of covariates Z_i (we henceforth use the convention that the sector-level shocks $\mathcal{X}_s$ are written in script font style, and the region-level aggregates X_i are written in normal style). To help us focus on the key conceptual issues, we abstract away from the covariates in Z_i for now, and assume that $\mathcal{X}_s$ and Y_i have been demeaned, so that we can omit the intercept in a regression of Y_i on X_i. The OLS estimator of the coefficient on X_i in this simplified setting is given by

$$\hat{\beta} = \frac{\sum_{i=1}^{N} X_i Y_i}{\sum_{i=1}^{N} X_i^2}, \tag{23}$$

and we can write the regression equation as

$$Y_i = \beta X_i + \epsilon_i, \qquad \text{where} \quad X_i \equiv \sum_{s=1}^{S} w_{is} \mathcal{X}_s, \qquad \sum_{s=1}^{S} w_{is} = 1, \tag{24}$$

and β denotes the population analog of $\hat{\beta}$.

The definition of the estimand β and the properties of the estimator $\hat{\beta}$ will depend on: (a) what the population of interest is; (b) how we think about repeated sampling; and (c) which restrictions we impose on the data-generating process (DGP).

For (a), we define the population of interest to be the observed set of N regions, as opposed to modeling the regions as drawn from a large superpopulation of regions. Consequently, we are interested in the parameters $\{\beta_{is}\}_{i=1,s=1}^{N,S}$ and the treatment effects $\{w_{is}\beta_{is}\}_{i=1,s=1}^{N,S}$ themselves, rather than the distributions from which they are drawn, which would be the case if we were interested in a superpopulation of regions.[13] For (b), given our interest on estimating the *ceteris paribus* impact of a specific set of shocks $\mathcal{X}_1, \ldots, \mathcal{X}_S$, we consider repeated sampling of these shocks, holding the shares w_{is}, the parameters β_{is}, and the potential outcomes $Y_i(0)$ fixed. Finally, for (c), as we describe in detail in Section 4 below, we characterize the DGP by imposing restrictions on the joint distribution of the shifters $\mathcal{X}_s$, zero-shock potential outcomes $Y_i(0)$, and parameters β_{is}, conditional on the shares w_{is}. In particular, we follow Borusyak, Hull and Jaravel (2018) in assuming that the shifters $\{\mathcal{X}_s\}_{s=1}^{S}$ are good as randomly assigned conditional on the shares $\{w_{is}\}_{s=1,i=1}^{S,N}$.

Given our assumptions on the population of interest and on the type of repeated sampling, the estimand of interest β is defined as

$$\beta = \frac{\sum_{i=1}^{N} E[X_i Y_i \mid \mathcal{F}_0]}{\sum_{i=1}^{N} E[X_i^2 \mid \mathcal{F}_0]}, \qquad \text{with} \qquad \mathcal{F}_0 = \{Y_i(0), \beta_{is}, w_{is}\}_{i=1,s=1}^{N,S}, \tag{25}$$

and the regression error ϵ_i is then defined as the residual

$$\epsilon_i = Y_i - X_i\beta = Y_i(0) + \sum_{i=1}^{S} w_{is}\mathcal{X}_s(\beta_{is} - \beta), \tag{26}$$

where the second equality uses eq. (21). Thus, the statistical properties of the regression residual ϵ_i will depend on the properties of the potential outcome $Y_i(0)$, the shifters $\{\mathcal{X}_s\}_{s=1}^{S}$ and the shares $\{w_{is}\}_{i=1,s=1}^{N,S}$. A key implication of the model in Section 2.1 is that the zero-shock potential outcome

[13] Our definition of the population is consistent with the question of interest outlined in Section 2.3, in which we focus on a specific set of regions included in a particular country c. It is also consistent with several seminal papers that use a shift-share approach. For example, the first sentence in the abstract of Autor, Dorn and Hanson (2013) reads: "We analyze the effect of rising Chinese import competition between 1990 and 2007 *on U.S. local labor markets*". Similarly, the first sentence in the abstract of Dix-Carneiro and Kovak (2017) reads: "We study the evolution of trade liberalization's effects *on Brazilian local labor markets*" (emphases added).

$Y_i(0)$ in eq. (21) will generally incorporate unobserved covariates that have a shift-share structure analogous to that of the regressor of interest, X_i. Specifically, as illustrated in eq. (22), $Y_i(0)$ incorporates weighted averages of unobserved sectoral shocks (i.e. preference shocks $\hat{\gamma}_s$ and any sectoral component of the technology shocks $\hat{A}_{is}$), with weights l_{is}^0 identical to those used to aggregate the sectoral shock of interest $\hat{P}_s$ into the regressor X_i. Consequently, if two regions i and i' have similar shares, they will tend to have similar shift-shares X_i and $X_{i'}$ as well as similar potential outcomes $Y_i(0)$ and $Y_{i'}(0)$. It then follows from eq. (26), that the residuals ϵ_i and $\epsilon_{i'}$ will also tend to be correlated.[14] As we discuss in all remaining sections of the paper, this has important implications for the variance of $\hat{\beta}$. We summarize this discussion in the following remark.

Remark 3. *Correctly performing inference for the OLS estimator $\hat{\beta}$ of the coefficient on a shift-share regressor requires taking into account that the regression residuals will generally inherit the same shift-share structure.*

4 Asymptotic properties of shift-share regressions

In this section, we formulate the statistical assumptions that we impose on the DGP, present the asymptotic results that we derive using these assumptions, and use the model in Section 2 to provide an economic interpretation for these assumptions. To convey the main issues that arise in these regressions, we first consider in Section 4.1 the simple case in which there is a single regressor with a shift-share structure, and no other covariates, as in Section 3. We introduce covariates in Section 4.2. Section 4.3 considers further extensions of the basic setup. All proofs are relegated to Appendix B.

We use the notation introduced in Sections 2 and 3. To compactly state our assumptions and results, we use standard matrix and vector notation. In particular, for a (column) L-vector A_i that varies at the regional level, A denotes the $N \times L$ matrix with the ith row given by A_i'. For a vector $\mathcal{A}_s$ that varies at the sectoral level, $\mathcal{A}$ denotes the $S \times L$ matrix with the sth row given by $\mathcal{A}_s'$. If $L = 1$, then A and $\mathcal{A}$ are an N-vector, and an S-vector, respectively. Let W denote the $N \times S$ matrix of shares, so that its (i, s) element is given by w_{is}, and let B denote the $N \times S$ matrix with (i, s) element given by β_{is}.

4.1 No additional covariates

We study here the statistical properties of the OLS estimator in eq. (23). We assume that, conditionally on the matrix of shares W, the shocks are as good as randomly assigned in that they are independent of the potential outcomes $Y_i(x_1, \dots, x_S)$:

$$(Y(0), B) \perp\!\!\!\perp \mathcal{X} \mid W. \tag{27}$$

In the next subsection, we weaken this assumption by assuming that the shocks are as good as randomly assigned conditionally on some covariates.

[14] As we discuss in Section 4.2, when controls Z_i are included, this conclusion remains to hold unless the controls account for the impact on the outcome of *all* sectoral shocks other than $\{\mathcal{X}_s\}_{s=1}^S$ that affect the outcome.

As in the literature on inference in randomized controlled trials (see Imbens and Rubin, 2015, for a review), we will consider the statistical properties of $\hat{\beta}$ under randomization-style inference: we will condition on the realized values of the shares, and on the potential outcomes, and consider the properties of $\hat{\beta}$ under repeated sampling of the shocks $\mathcal{X}$. This approach leverages the random assignment assumption in eq. (27), and ensures that the standard errors that we derive will remain valid under *any* dependence structure between the potential outcomes, or the shares w_{is}, across regions.

We impose some regularity conditions on the DGP for $(Y(0), B, W, \mathcal{X})$ that generate the observed data (Y, X, W). We consider asymptotics with the number of sectors going to infinity, $S \to \infty$. Formally, the number of regions $N = N_S$ depends on S, but we typically drop the S subscript and keep the conditioning implicit. Unless stated otherwise, all limits are taken as $S \to \infty$.

Assumption 1.

(i) $\{(Y(0), B, W, \mathcal{X}) \in \mathbb{R}^{N_S} \times \mathbb{R}^{N_S \times S} \times \mathbb{R}^{N_S \times S} \times \mathbb{R}^{S}\}_{S=1}^{\infty}$ is a triangular array of random variables with $N = N_S \to \infty$ as $S \to \infty$ that satisfies eq. (27), and $E[Y_i(0) \mid W] = 0$. The observed data consists of the tuple (Y, X, W), with $Y_i = Y_i(\mathcal{X}_1, \dots, \mathcal{X}_S)$, such that eq. (21) holds.

(ii) Conditional on W, the shocks $\mathcal{X}_1, \dots, \mathcal{X}_S$ are mean zero, independent across s, with fourth moments that exist and are bounded uniformly over s.

(iii) $\frac{1}{N}\sum_{i=1}^{N} E[X_i^2 \mid W] = \frac{1}{N}\sum_{i=1}^{N}\sum_{s=1}^{S} \operatorname{var}(\mathcal{X}_s \mid W) w_{is}^2$ converges in probability to a strictly positive non-random limit.

Assumption 2.

(i) Conditional on W, the second moments of $Y_i(0)$ exist, and are bounded uniformly over i. The support of β_{is} is bounded.

(ii) $\max_s n_s/N \to 0$, where $n_s = \sum_{s=1}^{S} w_{is}$ denotes the total share of sector s.

By modeling the data as a triangular array, Assumption 1(i) allows the distribution of the data to change with the sample size.[15] The assumption that $Y_i(0)$ and $\mathcal{X}_s$ are mean zero is made to simplify the exposition in this section by allowing us to drop the intercept from the regression of Y_i on X_i, and will be relaxed in Section 4.2. The key assumption underlying our approach to inference is Assumption 1(ii). This is the only independence assumption that we need; $Y_i(0)$ and the shares w_{is} can be correlated in an arbitrary manner across i. Consequently, the residuals ϵ_i (defined in eq. (26)) are allowed to be correlated in an arbitrary manner, which in particular allows them to have a shift-share structure. Furthermore, we do not require $\mathcal{X}$, or any other variables, to be identically distributed—the sectors and regions may be heterogeneous. Assumption 1(iii) 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. The bounded support condition on β_{is} in part (i) of Assumption 2 is made to keep the proofs simple and can be relaxed.

[15] 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 $\{(Y_{iS}(0), \beta_{isS}, w_{isS}, \mathcal{X}_{sS}) : i = 1, \dots, N_S, s = 1, \dots, S\}$.

Finally, Assumption 2(ii) is a key regularity condition ensuring that the share of each sector is asymptotically negligible. It generalizes the standard consistency condition in the clustering literature that the largest cluster be asymptotically negligible. To see the connection, consider the special case with "concentrated sectors", in which each region i specializes in one sector; denote it by $s(i)$. Then $w_{is} = 1$ if $s = s(i)$ and $w_{is} = 0$ otherwise, and n_s denotes the number of regions that specialize in sector s. In this case, $X_i = \mathcal{X}_{s(i)}$, so that if eq. (27) holds, $\hat{\beta}$ is equivalent to an OLS estimator in a randomized controlled trial in which the treatment varies at a cluster level; here the sth cluster consists of regions that specialize in sector s. The condition $\max_s n_s/N \to 0$ then reduces to the assumption that the largest cluster be asymptotically negligible.

Proposition 1. *Suppose Assumptions 1 and 2 hold. Then*

$$\beta = \frac{\sum_{i=1}^{N}\sum_{s=1}^{S} \pi_{is}\beta_{is}}{\sum_{i=1}^{N}\sum_{s=1}^{S} \pi_{is}}, \qquad \text{and} \qquad \hat{\beta} = \beta + o_P(1), \tag{28}$$

where $\pi_{is} = w_{is}^2 \operatorname{var}(\mathcal{X}_s \mid W)$.

The proposition gives two results. First, it uses eq. (27) to show that the estimand in eq. (25) can be expressed as a weighted average of the region- and sector-specific parameters β_{is}, with weights that are increasing in the variance of the shock and the share w_{is}. Second, it uses the remaining assumptions to show that the OLS estimator $\hat{\beta}$ converges to this estimand as $S \to \infty$. In the special case with concentrated sectors, $\sum_{s=1}^{S} \pi_{is}\beta_{is} = \operatorname{var}(\mathcal{X}_{s(i)} \mid W)\beta_{i,s(i)}$, so that Proposition 1 reduces to the standard result from the randomized control trials literature with cluster-level randomization (here the "cluster" consists of all regions that specialize in the same sector) that the weights are proportional to the variance of the shock.

For the estimator in eq. (23) to be asymptotically normal, we need to strengthen Assumption 1(ii) and Assumption 2 slightly:

Assumption 3.

(i) $\max_s n_s^2 / \sum_{t=1}^{S} n_t^2 \to 0$.

(ii) Conditional on W, the eighth moments of $\mathcal{X}_s$ are bounded uniformly over s, and the fourth moments of $Y_i(0)$ are bounded uniformly over i.

Part (i) ensures that the contribution of each sector to the asymptotic variance, which, according to the standard error formula below, is of the order $O(n_s^2)$, is asymptotically negligible. For instance, while the estimator $\hat{\beta}$ is consistent for β when the largest sector share is of the order $O(N/\sqrt{S})$ and the remaining sector shares are of the order $O(N/S)$, Assumption 3 rules this case out; $\hat{\beta}$ will not generally be asymptotically normal in this case due to failure of the Lindeberg condition.

Proposition 2. *Suppose Assumptions 1, 2 and 3 hold, and suppose that*

$$\mathcal{V}_N = \frac{1}{\sum_{s=1}^{S} n_s^2} \operatorname{var}\left(\sum_{i=1}^{N} X_i \epsilon_i \mid Y(0), B, W\right)$$

converges in probability to a non-random limit. Then

$$\frac{N}{\sqrt{\sum_{s=1}^{S} n_s^2}}(\hat{\beta} - \beta) = \mathcal{N}\left(0, \frac{\mathcal{V}_N}{\left(\frac{1}{N}\sum_{i=1}^{N} X_i^2\right)^2}\right) + o_P(1).$$

The proposition shows that $\hat{\beta}$ is asymptotically normal, with a rate of convergence equal to $N/\sqrt{\sum_{s=1}^{S} n_s^2}$. If the sector sizes n_s are all of the same order, equal to N/S, then the rate of convergence is equal to $\sqrt{S}$. However, if the cluster sizes are unequal, the rate may be slower. The asymptotic variance formula has the usual "sandwich" form.

Since X_i is directly observed, to construct a consistent standard error estimate, it suffices to construct a consistent estimate of $\mathcal{V}_N$, the middle part of the sandwich. Suppose that the effects of the shocks are homogeneous across regions and sectors, so that $\beta_{is} = \beta$ for all i, s, and $\epsilon_i = Y_i(0)$.[16] Then it follows from eq. (27) and Assumption 1(ii) that

$$\mathcal{V}_N = \frac{\sum_{s=1}^{S} \operatorname{var}(\mathcal{X}_s \mid W) R_s^2}{\sum_{s=1}^{S} n_s^2}, \qquad R_s = \sum_{i=1}^{N} w_{is}\epsilon_i. \tag{29}$$

The contribution of sector s to $\mathcal{V}_N$ is given by R_s^2, which is of the order n_s^2. Replacing $\operatorname{var}(\mathcal{X}_s \mid W)$ by $\mathcal{X}_s^2$, and ϵ_i by the regression residual $\hat{\epsilon}_i = Y_i - X_i\hat{\beta}$, we obtain to the standard error estimate

$$\widehat{se}(\hat{\beta}) = \frac{\sqrt{\sum_{s=1}^{S} \mathcal{X}_s^2 \hat{R}_s^2}}{\sum_{i=1}^{N} X_i^2}, \qquad \hat{R}_s = \sum_{i=1}^{N} w_{is}\hat{\epsilon}_i. \tag{30}$$

In the case with concentrated sectors, $\sum_{s=1}^{S} \mathcal{X}_s^2 \hat{R}_s^2 = \sum_{s=1}^{S}(\sum_{i=1}^{N} \mathbb{I}\{s(i) = s\} X_i\hat{\epsilon}_i)^2$, so that the standard error formula in the preceding display reduces to the usual cluster-robust standard error, clustered on the sectors. In the general case, the standard error accounts for the fact that regions with similar sectoral composition will have similar errors—unless the regression error $\epsilon_i = Y_i(0)$ has no sectoral component (so there are no unobserved sector-level shocks), it will generally not be the case that $\operatorname{cov}(X_i\epsilon_i, X_j\epsilon_j) = 0$ for $i \neq j$. In contrast, the usual heteroscedasticity-robust standard error fails to account for this correlation. Furthermore, standard errors clustered by groups of regions defined by their geographical proximity to each other will also fail to account for this correlation unless all regions are fully specialized in a single sector and the sector of specialization is the same for regions belonging to the same geographically defined cluster.

Remark 4. *In the expression for $\mathcal{V}_N$ in eq. (29), the only expectation is taken over $\mathcal{X}_s$, we do not take any expectation over the shares w_{is} or the residuals ϵ_i. This is because, as mentioned earlier, our inference is conditional on the realized values of the shares and on the potential outcomes. In terms of the linear regression equation $Y_i = X_i\beta + \epsilon_i$, this effectively means that we consider the properties of $\hat{\beta}$ under repeated sampling of $X_i = \sum_s w_{is}\mathcal{X}_s$ conditional on the shares w_{is} and conditional on the residuals ϵ_i (as opposed to, say, considering properties of $\hat{\beta}$ under repeated sampling of the residuals ϵ_i conditional on the realized values of X_i). As a result,*

[16] The standard error formula that we provide remains valid under heterogeneous effects as long as some mild restrictions on the form of heterogeneity apply; see Appendix B.5 for discussion.

our standard errors allow for arbitrary dependence structure between the residuals ϵ_i.

4.1.1 Economic interpretation of assumptions and results

We now interpret the assumptions described in Section 4.1 in terms of the economic model described in Section 2. To fix ideas, we consider the case in which the outcome of interest is the change in average wages between any two periods ($Y_i = \hat{\omega}_i$), the sectoral shocks of interest are changes in sectoral prices ($\mathcal{X}_s = \hat{P}_s$), and all N regions in the sample are "small" (the spending shares x_{is} in eq. (9) are close to zero for all s and i). As discussed in eq. (22), the shares then correspond to initial employment shares ($w_{is} = l^0_{is}$), the parameter β_{is} corresponds to the combination of structural parameters $\beta_{L,is}$ defined in eq. (19), and the potential outcome $Y_i(0)$ is a sum of shift-share variables.

In terms of our economic model, Assumption 1(i) imposes that the potential outcomes $Y(0)$ are independent of the sectoral price changes $\{\hat{P}_s\}_{s=1}^S$. As illustrated in eq. (22), the potential outcomes $Y(0)$ depend, among other variables, on the changes in sectoral preferences $\{\hat{\gamma}_s\}_{s=1}^S$ and the changes in technology, $\{\hat{A}_{is}\}_{i=1,s=1}^{N,S}$. As illustrated in Appendix A.1, the sectoral price changes depend on the changes in: (a) sectoral preferences; and (b) technology parameters of all world regions that are large (i.e. affect the sectoral world prices). Consequently, the independence assumption in eq. (27) imposes that: (a) there are no shocks to sectoral preferences, $\hat{\gamma}_s = 0$ for all s; (b) the technological shocks in the N regions in the population of interest are independent of those in every large region. As we discuss in Section 4.2, adding controls to the regression equation allows us to relax these assumptions.

Assumption 1(ii) imposes that the sectoral price changes $\{\hat{P}_s\}_{s=1}^S$ are independent across sectors. However, according to the model described in Section 2, any sector s price change, $\hat{P}_s$, will respond to changes in preferences and technology not only for sector s but also for any other sector $s' \neq s$ (see Appendix A.1). Our model thus predicts that sectoral price changes are not independent across sectors. In Section 5.2, we perform several simulations to study the importance of violations of the assumed sectoral independence of shocks imposed in Assumption 1(ii). In Section 4.3.1, we relax the independence restriction in Assumption 1(ii) and allow the shocks $\mathcal{X}$ to be correlated within groups of sectors, maintaining its independence across sectors belonging to different groups.[17]

Intuitively, Assumption 2(ii) imposes that no one sector dominates the others in terms of initial employment at the national level; i.e. $\sum_{i=1}^N L^0_{is} / \sum_{i=1} L_{is}$ is not too large for any one sector. As we illustrate in Section 5.2, this condition is satisfied for the U.S. when only manufacturing sectors are taken into account; it would not hold if the non-manufacturing sector is included as one of the S sectors incorporated into the analysis (unless the distribution of $\mathcal{X}_s$ for the non-manufacturing sector is degenerate at zero).[18]

[17] The assumption that the shifters $\mathcal{X}_s$ are independent across sectors or groups of sectors may be more accurate in settings in which the shifters $\mathcal{X}_s$ are not market equilibrium outcomes, such as tariffs and non-tariff trade barriers (Topalova, 2007, 2010; Kovak, 2013; Hakobyan and McLaren, 2016; Dix-Carneiro and Kovak, 2017, 2018), productivity (Table 10 in Autor, Dorn and Hanson, 2013), or tax rates (Zidar, 2018). In the context of the model in Section 2, the independence assumption would hold if the sectoral shocks of interest were sectoral technology shocks affecting equally all technology parameters A_{is} of the N regions in the population of interest.

[18] When analyzing the impact of international trade shocks on regional labor market outcomes, it is standard to either set the shock to the non-manufacturing sector to zero (Topalova, 2007, 2010; Autor, Dorn and Hanson, 2013; Hakobyan and McLaren, 2016) or to remove the non-manufacturing sector from the analysis and rescale the shares of all manufacturing sectors so that they add up to one (Kovak, 2013). Either of these approaches will satisfy the restriction in Assumption 2(ii).

Concerning the result in Proposition 1, it is important to remark that the estimand β does not equal a weighted average of the heterogeneous treatment effects: as discussed earlier, the effect of increasing the value of the sector s shock in one unit on the outcome variable is heterogeneous and equal to $w_{is}\beta_{is}$. Consider a set of region- and sector-specific weights $\{\xi_{is}\}_{i=1,s=1}^{N,S}$ chosen by the researcher (the weights may depend, for instance, on the policy that the researcher is interested in evaluating), and define the corresponding weighted average of the treatment effects as

$$\tau_{\xi} \equiv \frac{\sum_{i=1}^{N} \sum_{s=1}^{S} \xi_{is} w_{is} \beta_{is}}{\sum_{i=1}^{N} \sum_{s=1}^{S} \xi_{is}}.$$

If β_{is} is constant across i and s, then $\tau_{\xi} = \beta \overline{w}_{\xi}$, with $\overline{w}_{\xi} \equiv \sum_{i=1}^{N} \sum_{s=1}^{S} \xi_{is} w_{is} / \sum_{i=1}^{N} \sum_{s=1}^{S} \xi_{is}$. Given that $\overline{w}_{\xi}$ is a function of the observed shares in W and the weights $\{\xi_{is}\}_{i=1,s=1}^{N,S}$, the researcher can use $\hat{\beta}$ to compute a consistent estimate of τ_{ξ} as $\hat{\beta}\overline{w}_{\xi}$. Furthermore, in this case, the common parameter β such that $\beta_{is} = \beta$ for all i and s has the interpretation that it measures the total effect of increasing the shifters simultaneously in every sector by one unit. Conversely, when β_{is} varies across regions and sectors, then it is not clear how to exploit knowledge of the estimand β defined in eq. (28) to learn something about τ_{ξ}. This point has been made before in Monte, Redding and Rossi-Hansberg (2018).[19]

Finally, as Assumptions 1 and 2 impose no restriction on the cross-regional correlation of the potential outcomes $Y(0)$, the standard error estimate introduced in eq. (30) is consistent for the standard deviation of $\hat{\beta}$ no matter what the correlation in the labor supply shocks v_i (defined in eq. (10)) is: it remains consistent even if labor supply conditions are similar across regions that are geographically close to each other (e.g. due to labor market policies defined by a supra-regional entity).

4.2 Adding covariates

In many applications, assuming that the shocks $\mathcal{X}$ are as good as randomly assigned is unrealistic. However, they may be as good as randomly assigned conditional on a vector of covariates. In particular, we assume that there is a latent K-vector of covariates, $\mathcal{Z}_s$, measured at a sectoral level, such that conditional on $\{\mathcal{Z}_s\}_{s=1}^{S}$, the shocks $\mathcal{X}_s$ are as good as randomly assigned. We do not necessarily observe these covariates directly, but we do observe a region-level proxy

$$Z_i = \sum_{s=1}^{S} w_{is} \mathcal{Z}_s + U_i, \tag{31}$$

for each region i (as with the shocks $\mathcal{X}$, we use the notational convention that the sector-level variable $\mathcal{Z}_s$ is written in script font style, while the region-level variable Z_i is written in standard style). With

[19] An alternative estimation approach is to use the mapping between β_{is} and structural parameters implied by an economic model; e.g. eq. (20) expresses $\beta_{\omega,is}$ as a function of the demand elasticities $\{\sigma_s\}_{s=1}^{S}$ and the labor supply elasticity ϕ. By imposing this mapping and obtaining consistent estimates of these structural parameters, researchers may obtain consistent estimates of β_{is} for every i and s. This approach is however less robust to alternative modeling assumptions; e.g. the model in Appendix A.4 is isomorphic to that in Section 2.1 up to the mapping from $\beta_{L,is}$ and $\beta_{\omega,is}$ to structural parameters. Thus, both models are consistent with the potential outcomes framework in eq. (21), and consequently, both models justify an estimation approach based on a regression of the type in eq. (24). Conversely, they will imply different estimates if the mapping from β_{is} to structural parameters is exploited for identification.

this setup, we replace eq. (27) with the assumption that

$$(U, Y(0), B) \perp\!\!\!\perp \mathcal{X} \mid \mathcal{Z}, W. \tag{32}$$

where $\mathcal{Z}$ denotes the $S \times K$ matrix with sth row given by $\mathcal{Z}_s'$, and U denotes the N-vector with ith element given by U_i.

The condition in eq. (32) may be difficult to interpret directly because different variables are measured at different levels: the potential outcomes are measured at a regional level, and the covariates and shocks are measured at a sectoral level. It is therefore useful to consider a projection of the potential outcomes onto the sectoral space. For simplicity, consider the case with constant effects, $\beta_{is} = \beta$, and project $Y_i(0)$ onto the sector-level controls $\mathcal{Z}_s$, so that we can write $Y_i(0) = \sum_{s=1}^{S} w_{is}\mathcal{Z}_s'\kappa + \eta_i$. Then eq. (32) holds if the residuals η_i in this projection are independent of $\mathcal{X}$—if there are any other unobserved sector-level shocks that affect the outcomes (and are therefore in η_i), these must be unrelated to $(\mathcal{X}_s, \mathcal{Z}_s)$. The variable U_i in eq. (31) can be thought of in two ways. First, one can think of it as measurement error in Z when controlling for $\mathcal{Z}$: $U_{ik} = 0$ if the proxy Z_{ik} is perfect (it controls for $\mathcal{Z}_{1k}, \ldots, \mathcal{Z}_{Sk}$ without error). Alternatively, the kth covariate Z_{ik} may be included in the regression to increase precision—in this case $\mathcal{Z}_{sk} = 0$ for all s, and Z_{ik} is included because η_i and U_{ik} are correlated.

To ensure that it suffices to include the covariates in the regression linearly (instead of having to control for them non-parametrically), we assume that the expectation of $\mathcal{X}_s$ conditional on $\mathcal{Z}_s$ is linear in $\mathcal{Z}_s$,

$$E[\mathcal{X}_s \mid \mathcal{Z}] = \mathcal{Z}_s'\gamma, \tag{33}$$

where γ is a K-vector that equals 0 if and only if the scalar $\mathcal{X}_s$ is mean independent of the K-vector $\mathcal{Z}_s$. We now study the properties of the OLS estimator $\hat{\beta}$, the coefficient on X_i in a regression of Y_i onto X_i and Z_i. Let Z denote the $N \times K$ matrix with ith row given by Z_i', and let $\ddot{X} = X - Z(Z'Z)^{-1}Z'X$ denote an N-vector whose ith element is equal to the regressor X_i with the covariates in the vector Z_i partialled out (i.e. the residuals from regressing X onto Z). Then by the Frisch–Waugh–Lovell theorem, $\hat{\beta}$ can be written as

$$\hat{\beta} = \frac{\sum_{i=1}^{N} \ddot{X}_i Y_i}{\sum_{i=1}^{N} \ddot{X}_i^2} = \frac{\ddot{X}'Y}{\ddot{X}'\ddot{X}}, \tag{34}$$

and the OLS estimator of the coefficient on Z_i can be written as

$$\hat{\delta} = (Z'Z)^{-1}Z'(Y - X\hat{\beta}).$$

To state the consistency result, we first need to generalize Assumption 1 to allow for covariates:

Assumption 4.

(i) $\{(Y(0), B, W, U, \mathcal{X}, \mathcal{Z}) \in \mathbb{R}^{N_S} \times \mathbb{R}^{N_S \times S} \times \mathbb{R}^{N_S \times S} \times \mathbb{R}^{N_S \times K} \times \mathbb{R}^{S} \times \mathbb{R}^{S \times K}\}_{S=1}^{\infty}$ is a triangular array of random variables with $N = N_S \to \infty$ as $S \to \infty$ that satisfies eqs. (32) and (33). The observed data consists of the tuple (Y, X, Z, W), with $Y_i = Y_i(\mathcal{X}_1, \ldots, \mathcal{X}_S)$, such that eqs. (21) and (31) hold.

(ii) Conditional on W and $\mathcal{Z}$, the shocks $\mathcal{X}_1, \dots, \mathcal{X}_S$ are independent across s, with fourth moments that exist and are bounded uniformly over s.

(iii) $\frac{1}{N}\sum_{i=1}^{N}\sum_{s=1}^{S}\operatorname{var}(\mathcal{X}_s \mid W, \mathcal{Z})w_{is}^2$ converges in probability to a strictly positive non-random limit, and $Z'Z/N$ converges in probability to a positive definite non-random limit.

(iv) Conditional on W, the second moments of U_i and $\mathcal{Z}_s$ exist and are bounded uniformly over i and s.

Parts (i), (ii) and (iii) are straightforward generalizations of parts (i), (ii) and (iii) of Assumption 1. Part (iv) imposes very mild restrictions on U and $\mathcal{Z}$.

Proposition 3. *Suppose Assumptions 2 and 4 hold, and that $U_i'\gamma = 0$ for $i = 1, \dots, N$. Then*

$$\hat{\beta} = \beta + o_P(1), \qquad \beta = \frac{\sum_{i=1}^{N}\sum_{s=1}^{S}\pi_{is}\beta_{is}}{\sum_{i=1}^{N}\sum_{s=1}^{S}\pi_{is}}, \tag{35}$$

where $\pi_{is} = w_{is}^2 \operatorname{var}(\mathcal{X}_s \mid W, \mathcal{Z})$.

The result is very similar to Proposition 1; the only difference is that the weights π_{is} now reflect the variance of $\mathcal{X}_s$ that is conditional on $\mathcal{Z}$ and W, rather than just conditional on W. An important additional assumption is that $U_i'\gamma = 0$ for $i = 1, \dots, N$. Effectively, this requires that, for each covariate k, either $U_{ik} = 0$ for all i, so that Z_{ik} is a perfect proxy for the sector-level variables $\mathcal{Z}_{1k}, \dots, \mathcal{Z}_{Sk}$, or else $\gamma_k = 0$, so that the $\mathcal{Z}_{sk}$ is unrelated to the shock $\mathcal{X}_s$—in this case the proxy need not be perfect, since it is not necessary to control for $\mathcal{Z}_{sk}$ in the first place. If $U_i'\gamma \neq 0$, then there will be omitted variable bias due to inadequately controlling for the confounders $\mathcal{Z}$.

To state the asymptotic normality result, let

$$\delta = E[Z'Z]^{-1}E[Z'(Y - X\beta)]$$

denote the population regression coefficient on Z_i, so that we can write the regression model as

$$Y_i = X_i\beta + Z_i'\delta + \epsilon_i,$$

where the regression residual ϵ_i is defined as $\epsilon_i = Y_i - X_i\beta - Z_i'\delta$. 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 5.

(i) Conditional on W, the fourth moments of $\mathcal{Z}_s$, and U_i exist and are bounded uniformly over s and i.

(ii) $\frac{N}{\sqrt{\sum_s n_s^2}}(\check{\delta} - \delta) = O_P(1)$

Part (i) strengthens Assumption 4(iv). Part (ii) is a high-level assumption that implies $\hat{\delta}$ converges at least as fast as $\hat{\beta}$; otherwise the error in estimating δ could dominate the asymptotic variance of β.

Proposition 4. *Suppose Assumptions 2, 3, 4 and 5 hold, and that $U_i'\gamma = 0$. Suppose also that*

$$\mathcal{V}_N = \frac{1}{\sum_{s=1}^S n_s^2} \operatorname{var}\left(\sum_i (X_i - Z_i'\gamma)\epsilon_i \mid Y(0), B, U, \mathcal{Z}, W\right)$$

converges in probability to a non-random limit. Then

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

The result is very similar to that in Proposition 2; the only difference is that X_i in the definition of $\mathcal{V}_N$ is replaced by $X_i - Z_i'\gamma$, and that X_i is replaced by $\ddot{X}_i$ in the outer part of the "sandwich".

To construct a consistent standard error estimate, similarly to the case without covariates, it suffices to construct a consistent estimate of $\mathcal{V}_N$, the middle part of the sandwich. We derive the standard error formula under the assumption that $\beta_{is} = \beta$ for all i, s; in Appendix B.5, we discuss the restrictions under which it remains valid when the effects are heterogeneous. Under this assumption, $\epsilon_i = Y_i(0) - Z_i'\delta$, and it follows from eq. (32) and Assumption 4(ii) that, analogously to eq. (29),

$$\mathcal{V}_N = \frac{\sum_{s=1}^S \operatorname{var}(\tilde{\mathcal{X}}_s \mid W, \mathcal{Z}) R_s^2}{\sum_{s=1}^S n_s^2}, \qquad R_s = \sum_{i=1}^N w_{is}\epsilon_i, \quad \tilde{\mathcal{X}}_s = \mathcal{X}_s - \mathcal{Z}_s'\gamma.$$

A plug-in estimate of R_s can be constructed by replacing ϵ_i with the estimated regression residuals $\hat{\epsilon}_i = Y_i - X_i\hat{\beta} - Z_i\hat{\delta}$. To construct an estimate of the variance $\operatorname{var}(\tilde{\mathcal{X}}_s \mid W, \mathcal{Z})$, note that, under the assumption that $U_i'\gamma = 0$ for $i = 1, \dots, N$ and the assumption that W has full column rank—this requires that $N > S$, i.e. there are more regions than sectors—the sector-level variable $\tilde{\mathcal{X}}_s$ can be backed out from the region-level variable $X_i - Z_i'\gamma$ as the coefficients from the regression of $X_i - Z_i'\gamma$ onto the shares W_i,

$$\tilde{\mathcal{X}} = (W'W)^{-1}W'(X - Z'\gamma).$$

Therefore, if γ was known, we could estimate $\operatorname{var}(\tilde{\mathcal{X}}_s \mid W, \mathcal{Z}) = E[\tilde{\mathcal{X}}_s^2 \mid W, \mathcal{Z}]$ as the square of the sth element of the vector $(W'W)^{-1}W'(X - Z'\gamma)$. Since γ is unknown, this estimate is infeasible. However, it follows from the proof of Proposition 3 that $\hat{\gamma} = (Z'Z)^{-1}Z'X = \gamma + o_P(1)$, so that $\ddot{X}_i = X_i - Z_i'\hat{\gamma}$ is a consistent estimate of $X_i - Z_i'\gamma$, which suggests the feasible estimate

$$\widehat{\mathcal{X}} = (W'W)^{-1}W'\ddot{X}, \tag{36}$$

of $\tilde{\mathcal{X}}$, and the standard error estimate

$$\widehat{se}(\hat{\beta}) = \frac{\sqrt{\sum_{s=1}^S \widehat{\mathcal{X}}_s^2 \hat{R}_s^2}}{\sum_{i=1}^N \ddot{X}_i^2}, \qquad \hat{R}_s = \sum_{i=1}^N w_{is}\hat{\epsilon}_i. \tag{37}$$

The next remark summarizes these steps:

Remark 5. *To construct the standard error estimate in eq. (37):*

1. *Obtain the estimates $\hat{\beta}$ and $\hat{\delta}$ by regressing Y_i onto $X_i = \sum_s w_{is}\mathcal{X}_s$ and the controls Z_i. The estimate $\hat{\epsilon}_i$ corresponds to the estimated regression residuals.*

2. *Construct $\ddot{X}_i$, the residuals from regressing X_i onto Z_i. Compute $\widehat{\mathcal{X}}_s$, the regression coefficients from regressing $\ddot{X}$ onto W.*

Plug the estimates $\hat{\epsilon}_i$, $\ddot{X}_i$, and p $\widehat{\mathcal{X}}_s$ into the standard error formula in eq. (37).

Consider again the case with concentrated sectors. Suppose that $U_i = 0$ for all i, so that the regression of Y_i onto X_i and Z_i is identical to the regression of Y_i onto $\mathcal{X}_{s(i)}$ and $\mathcal{Z}_{s(i)}$. Then $\widehat{\mathcal{X}}_s = \mathcal{X}_s - \mathcal{Z}_s'\hat{\gamma}$, and the standard error formula in eq. (37) reduces to the usual cluster-robust standard error, with clustering on the sectors $s(i)$.

In the clustering literature, it has been shown that the cluster-robust standard error is generally biased due to estimation noise in estimating ϵ_i, which can lead to undercoverage, especially in cases with a few clusters (see, for example, Cameron and Miller, 2014 for a survey). Since the standard error in eq. (37) can be viewed as generalizing the cluster-robust formula, similar concerns arise in our setting. We therefore also consider a modification $\widehat{se}_{\beta_0}(\hat{\beta})$ of $\widehat{se}(\hat{\beta})$ that imposes the null hypothesis when estimating the regression residuals to reduce the estimation noise in estimating ϵ_i. In particular, to calculate the standard error $\widehat{se}_{\beta_0}(\hat{\beta})$ for testing the hypothesis $H_0\colon \beta = \beta_0$ against a two-sided alternative at significance level α, one replaces $\hat{\epsilon}_i$ with $\hat{\epsilon}_{\beta_0,i}$, the residual from regressing $Y_i - X_i\beta_0$ onto Z_i (that is, $\hat{\epsilon}_{\beta_0,i}$ is an estimate of the residuals with the null imposed). The null is rejected if the absolute value of the t-statistic $(\hat{\beta} - \beta_0)/\widehat{se}_{\beta_0}(\hat{\beta})$ exceeds $z_{1-\alpha/2}$, the $1 - \alpha/2$ quantile of a standard normal distribution (1.96 for $\alpha = 0.05$). To construct a confidence interval (CI) with coverage $1 - \alpha$, one collects all hypotheses β_0 that were not rejected. It follows from simple algebra that the endpoints of this CI are a solution to a quadratic equation, so that they are available in closed form—one doesn't numerically have to search for all the hypotheses that were not rejected. The next remark summarizes this procedure.

Remark 6 (Confidence interval with null imposed)**.** *To test the hypothesis $H_0\colon \beta = \beta_0$, or equivalently, to check whether β_0 lies in the confidence interval:*

1. *Obtain the estimate $\hat{\beta}$ by regressing Y_i onto $X_i = \sum_s w_{is}\mathcal{X}_s$ and the controls Z_i. Obtain the restricted regression residuals $\hat{\epsilon}_{\beta_0,i}$ as the residuals from regressing $Y_i - X_i\beta_0$ onto Z_i.*

2. *Construct $\ddot{X}_i$, the residuals from regressing X_i onto Z_i. Compute $\widehat{\mathcal{X}}_s$, the regression coefficients from regressing $\ddot{X}$ onto W (this step is identical to step 2 in Remark 5).*

Compute the standard error as

$$\widehat{se}_{\beta_0}(\hat{\beta}) = \frac{\sqrt{\sum_{s=1}^{S} \widehat{\mathcal{X}}_s^2 \hat{R}_{\beta_0,s}^2}}{\sum_{i=1}^{N} \ddot{X}_i^2}, \qquad \hat{R}_{\beta_0,s} = \sum_{i=1}^{N} w_{is}\hat{\epsilon}_{\beta_0,i}. \tag{38}$$

Reject the null if $|(\hat{\beta} - \beta_0)/\widehat{se}_{\beta_0}(\hat{\beta})| > z_{1-\alpha/2}$. A confidence set with coverage $1 - \alpha$ is given by all nulls that were not rejected, $CI_{1-\alpha} = \{\beta_0\colon |(\hat{\beta} - \beta_0)/\widehat{se}_{\beta_0}(\hat{\beta})| < z_{1-\alpha/2}\}$. This set is an interval with endpoints given *by*

$$\hat{\beta} - A \pm \sqrt{A^2 + \frac{\widehat{se}(\hat{\beta})^2}{Q/(\ddot{X}'\ddot{X})^2}}, \qquad A = \frac{\sum_{s=1}^{S} \widehat{\mathcal{X}}_s^2 \hat{R}_s \sum_i w_{is}\ddot{X}_i}{Q},$$

where $Q = (\ddot{X}'\ddot{X})^2/z_{1-\alpha/2}^2 - \sum_{s=1}^{S} \widehat{\mathcal{X}}_s^2 (\sum_i w_{is}\ddot{X}_i)^2$ *and* $\widehat{se}(\hat{\beta})$ *and* $\hat{R}_s$ *are given in eq.* (37).

Since in both $\hat{\epsilon}_i$ and $\hat{\epsilon}_{\beta_0,i}$ are consistent estimates of the residuals, both $\widehat{se}_{\beta_0}(\hat{\beta})$ and $\widehat{se}(\hat{\beta})$ are consistent estimates of the standard error and, consequently, yield tests and confidence intervals that are asymptotically valid. The next proposition formalizes this result.

Proposition 5. *Suppose that the assumptions of Proposition 4 hold, and that* $\beta_{is} = \beta$. *Suppose also that either* $\max_s \sum_i |((W'W)^{-1}W')_{si}|$ *is bounded, or else that* $U_i = 0$ *for* $i = 1, \dots, N$. *Define* $\widehat{\mathcal{X}}$ *as in eq.* (36), *and let* $\hat{R}_s = \sum_{i=1}^{N} w_{is}\tilde{\epsilon}_i$, *where* $\tilde{\epsilon}_i = Y_i - X_i\tilde{\beta} - Z_i'\tilde{\delta}$, *and* $\tilde{\beta}$ *and* $\tilde{\delta}$ *are consistent estimators of* δ *and* β. *Then*

$$\frac{\sum_{s=1}^{S} \widehat{\mathcal{X}}_s^2 \hat{R}_s^2}{\sum_{s=1}^{S} n_s^2} = \mathcal{V}_N + o_P(1). \tag{39}$$

The additional assumption of Proposition 5 is that either $\max_s \sum_i |((W'W)^{-1}W')_{si}|$ is bounded, or else $U_i = 0$ for all i. This assumption ensures that the estimation error in $\widehat{\mathcal{X}}_s$ that arises from having to back out the sector-level shocks $\mathcal{Z}_s$ from the covariates Z_i is not too large. If the sectors are concentrated, then $((W'W)^{-1}W')_{si} = \mathbb{I}\{s(i) = s\}/n_s$, so that $\max_s \sum_i |((W'W)^{-1}W')_{si}| = 1$, and the assumption always holds.

4.2.1 Economic interpretation of assumptions and results

The role that the vector of covariates Z_i plays in our framework is twofold. First, the kth element of the vector Z_i may proxy for the impact on region i of an unobserved sectoral shock $(\mathcal{Z}_{1k}, \dots, \mathcal{Z}_{Sk})$. In the context of the model in Section 2, regional labor market outcomes are potentially affected by several sectoral shocks: price shocks $\hat{P}_s$, preference shocks $\hat{\gamma}_s$, and any sectoral component of the region- and sector-specific technology shocks $\hat{A}_{is}$. As in Section 2.3, let us consider again the case in which the economic shock of interest is the price shock $\hat{P}_s$. Then, as discussed in Section 4.1.1, when the regression of Y_i on X_i does not include a vector of covariates Z_i, consistent estimation of the weighted average of the parameters $\{\beta_{is}\}_{i=1,s=1}^{N,S}$ described in eq. (35) requires assuming away the presence of preference shocks ($\hat{\gamma}_s = 0$ for all s), and imposing an independence restriction between the technology shocks $\hat{A}_{is}$ of the regions $i = 1, \dots, N$ included in the analysis and those of the regions impacting the sectoral prices. These restrictions can be relaxed once we allow for additional covariates in the shift-share regression of interest. As an example, if we can construct measures of changes in sectoral preferences (or perfect proxies for their impact in every region $i = 1, \dots, N$ in the sample), we can control for them in our regression and, thus, Proposition 3 holds even in the presence of sectoral preference shocks. Similarly, if we can build a proxy for the changes in technology in the N regions included in the sample, $\{\hat{A}_{is}\}_{i=1,s=1}^{N,S}$, then we can include them in the regression equation as part of the vector Z_i and, in this case, Proposition 3 holds even if these technology shocks are not independent of the sectoral price shocks of interest.

Second, each element of the vector Z_i may proxy for regional covariates that, although independent of the sectoral shocks of interest $\mathcal{X}$, have an effect on the outcome variable and, thus, enter the regression residual ϵ_i in eq. (24). Controlling for these covariates is not necessary for Proposition 3 to hold, but including them would increase the precision of $\hat{\beta}$. In the context of the model described in Section 2, the region-specific labor supply shock $\hat{v}_i$ is an example of these covariates, as long as $\{\hat{v}_i\}_{i=1}^N$ are independent of the analogous labor supply shocks impacting regions that are large enough to affect the sectoral price indices $\{\hat{P}_s\}_{s=1}^S$. If this independence condition does not hold, then it is important, as discussed above, to always control for these labor supply shocks in order to ensure consistency of $\hat{\beta}$.

4.3 Extensions

We now discuss two extensions of the basic setup: first, we weaken Assumption 4(ii), and allow $\mathcal{X}_s$ to be correlated with $\mathcal{X}_k$ if s and k are in the same "cluster" of sectors. Second, we consider using the shift-share regressor X_i as an instrument in an instrumental variables regression.

4.3.1 Clusters of sectors

Suppose that the sectors $s = 1, \dots, S$ can be grouped into larger units, which we will refer to as "clusters", with $c(s) \in \{1, \dots, C\}$ denoting the cluster that sector s belongs to. For instance, s may be a four-digit industry code, while $c(s)$ is a three-digit or a two-digit code. With this structure, we replace Assumption 4(ii) with the assumption that, conditional on $\mathcal{Z}$ and W, the shocks $\mathcal{X}_s$ and $\mathcal{X}_t$ are independent if $c(s) \neq c(t)$.[20] Replace Assumption 2(ii) with the assumption that as $C \to \infty$, the largest cluster makes an asymptotically negligible contribution to the asymptotic variance,

$$\max_c \tilde{n}_c^2 / \sum_{d=1}^{C} \tilde{n}_d^2 \to 0,$$

where $\tilde{n}_c = \sum_s \mathbb{I}\{c(s) = c\} n_s$ is the total share of cluster c. Then, generalizing the arguments in Section 4.2, one can show that as $C \to \infty$, $\hat{\beta}$ remains asymptotically normal,

$$\frac{N}{\sqrt{\sum_{c=1}^{C} \tilde{n}_c^2}} (\hat{\beta} - \beta) = \mathcal{N}\left(0, \frac{\mathcal{V}_N}{\left(\frac{1}{N}\sum_i \ddot{X}_i^2\right)^2}\right) + o_P(1),$$

and, assuming that $\beta_{is} = \beta$, the term $\mathcal{V}_N$ is now given by

$$\mathcal{V}_N = \frac{\sum_{c=1}^{C} \sum_{s,t} \mathbb{I}\{c(s) = c(t) = c\} E[\tilde{\mathcal{X}}_s \tilde{\mathcal{X}}_t \mid W, \mathcal{Z}] R_s R_t}{\sum_{c=1}^{C} \tilde{n}_c^2}, \qquad R_s = \sum_{i=1}^{N} w_{is} \epsilon_i, \quad \tilde{\mathcal{X}}_s = \mathcal{X}_s - \mathcal{Z}_s' \gamma.$$

[20] Using the tools described in this section, one may also weaken Assumption 1(ii) in the context of a shift-share regression without covariates. We limit however the exposition to the more general case of the shift-share regression with covariates. Adjusting the standard error formulas to the setting without covariates requires only setting the K-vector $\mathcal{Z}_s$ to equal a vector of 1s for every sector s.

In other words, instead of treating $\tilde{\mathcal{X}}_s R_s$ as independent across s, the asymptotic variance formula clusters them. As a result, we replace the standard error estimate in eq. (37) with

$$\widehat{se}(\hat{\beta}) = \frac{\sqrt{\sum_{c=1}^{C} \sum_{s,t} \mathbb{I}\{c(s) = c(t) = c\} \widehat{\mathcal{X}}_s \hat{R}_s \widehat{\mathcal{X}}_t \hat{R}_t}}{\sum_{i=1}^{N} \ddot{X}_i^2}, \qquad \hat{R}_s = \sum_{i=1}^{N} w_{is} \hat{\epsilon}_i, \tag{40}$$

with $\widehat{\mathcal{X}}_s$ defined as in Remark 5. Confidence intervals with the null imposed can be constructed as in Remark 6, replacing $\hat{\epsilon}_i$ with $\hat{\epsilon}_{\beta_0,i}$ in the formula in eq. (40), and using this formula instead of that in eq. (38).

4.3.2 Instrumental variables regression

Consider the problem of estimating the effect of a regional variable Y_{2i}, that we refer to as a treatment variable, on an outcome variable Y_{1i}, using an instrumental variables (IV) regression, with the shift-share regressor $X_i = \sum_s w_{is} \mathcal{X}_s$ as an instrument, and a vector Z_i of regional controls. We assume there is a K-vector of latent covariates $\mathcal{Z}_s$, measured at a sectoral level, such that the regional controls Z_i have the structure in eq. (31) and that eq. (33) holds.

We assume that the effect of Y_{2i} onto Y_{1i} is linear and constant across regions, so that the potential outcome when Y_{2i} is exogenously set to y_2 is given by

$$Y_{i1}(y_2) = Y_{i1}(0) + y_2 \alpha,$$

where α is the treatment effect of Y_{2i} onto Y_{1i} for every region i. The observed outcome is thus given by $Y_{1i} = Y_{1i}(Y_{2i})$. The observed treatment level Y_{2i} may be correlated with the potential outcomes (i.e. endogenous), even when conditioning on a vector of covariates Z_i. In analogy with eq. (21), let

$$Y_{2i}(x_1, \dots, x_S) = Y_{2i}(0) + \sum_{i=1}^{S} w_{is} x_s \beta_{FS} \tag{41}$$

denote the potential treatment levels in region i that would occur if the region received shocks $x_1, \dots, x_S$. The observed treatment level is given by $Y_{2i} = Y_{2i}(\mathcal{X}_1, \dots, \mathcal{X}_S)$. For simplicity, we assume that β_{FS} does not vary across sectors or regions. Finally, we assume that conditional on $\mathcal{Z}$, the shocks $\mathcal{X}$ are as good as randomly assigned and satisfy the exclusion restriction, so that the following independence restriction holds:

$$(U, Y_1(0), Y_2(0)) \perp\!\!\!\perp \mathcal{X} \mid \mathcal{Z}, W. \tag{42}$$

Under this setup, both the reduced-form regression of Y_{i1} onto X_i and Z_i, and the first-stage regression of Y_{2i} onto X_i and Z_i fit into the setup of Section 4.2. Therefore, by generalizing the arguments in Section 4.2, it is straightforward to derive the joint asymptotic distribution of the OLS estimates

$$\hat{\beta}_{RF} = \frac{\sum_{i=1}^{N} \ddot{X}_i Y_{1i}}{\sum_{i=1}^{N} \ddot{X}_i^2}, \qquad \text{and} \qquad \hat{\beta}_{FS} = \frac{\sum_{i=1}^{N} \ddot{X}_i Y_{2i}}{\sum_{i=1}^{N} \ddot{X}_i^2},$$

of the reduced-form and first-stage coefficients on X_i, respectively. To state the result, define the reduced-form and first-stage regression errors

$$\epsilon_{1i} = Y_{1i} - Z_i'\delta_{RF} - X_i\beta_{RF}, \qquad \epsilon_{2i} = Y_{2i} - Z_i'\delta_{FS} - X_i\beta_{FS},$$

where $\delta_{RF} = E[Z'Z]^{-1}E[Z'(Y_1 - X\beta_{RF})]$ and $\delta_{FS} = E[Z'Z]^{-1}E[Z'(Y_2 - X\beta_{FS})]$, and it follows from eq. (42) that the population reduced-form coefficient on X_i is given by $\beta_{RF} = \beta_{FS}\alpha$. Then, under appropriate regularity conditions,

$$\frac{N}{\sqrt{\sum_{s=1}^{S} n_s^2}}\left(\begin{pmatrix}\hat{\beta}_{RF}\\ \hat{\beta}_{FS}\end{pmatrix} - \begin{pmatrix}\beta_{RF}\\ \beta_{FS}\end{pmatrix}\right) = \mathcal{N}\left(0, \frac{1}{(\frac{1}{N}\sum_i \ddot{X}_i^2)^2}\mathcal{V}_{IV,N}\right) + o_P(1),$$

where

$$\mathcal{V}_{IV,N} = \frac{1}{\sum_{s=1}^{S} n_s^2}\sum_{s=1}^{S}\operatorname{var}(\tilde{\mathcal{X}}_s \mid \mathcal{Z}, W)R_{IV,s}R_{IV,s}', \qquad R_{IV,s} = \sum_{i=1}^{N} w_{is}\epsilon_i, \qquad \tilde{\mathcal{X}}_s = \mathcal{X}_s - \mathcal{Z}_s'\gamma,$$

and $\epsilon_i = (\epsilon_{1i}, \epsilon_{2i})'$. Since the IV estimate of α is given by the ratio of the reduced form estimates,

$$\hat{\alpha} = \frac{\sum_{i=1}^{N}\ddot{X}_i Y_{1i}}{\sum_{i=1}^{N}\ddot{X}_i Y_{2i}} = \frac{\hat{\beta}_{RF}}{\hat{\beta}_{FS}}, \tag{43}$$

it follows by the delta method that, so long as $\beta_{FS} \neq 0$, so that the shift-share instrument is relevant,

$$\frac{N}{\sqrt{\sum_{s=1}^{S} n_s^2}}(\hat{\alpha} - \alpha) = \mathcal{N}\left(0, \frac{\frac{1}{\sum_{s=1}^{S} n_s^2}\sum_{s=1}^{S}\operatorname{var}(\tilde{\mathcal{X}}_s \mid \mathcal{Z}, W)R_s^2}{(\frac{1}{N}\sum_i \ddot{X}_i^2)^2\beta_{FS}^2}\right) + o_P(1), \qquad R_s = \sum_{i=1}^{N} w_{is}(\epsilon_{1i} - \epsilon_{2i}\alpha).$$

This suggests the standard error estimate

$$\widehat{se}(\hat{\alpha}) = \frac{\sqrt{\sum_{s=1}^{S}\widehat{\mathcal{X}}_s^2\hat{R}_s^2}}{|\hat{\beta}_{FS}|\sum_{i=1}^{N}\ddot{X}_i^2} = \frac{\sqrt{\sum_{s=1}^{S}\widehat{\mathcal{X}}_s^2\hat{R}_s^2}}{|\sum_{i=1}^{N}\ddot{X}_i Y_{2i}|}, \qquad \hat{R}_s = \sum_{i=1}^{N} w_{is}\hat{\epsilon}_{\Delta,i}, \tag{44}$$

with $\widehat{\mathcal{X}}_s$ constructed as in Remark 5, and $\hat{\epsilon}_\Delta = Y_1 - Y_2\hat{\alpha} - Z'(Z'Z)^{-1}Z'(Y_1 - Y_2\hat{\alpha})$ corresponds to the estimate of the residual in the structural equation, $\epsilon_{1i} - \epsilon_{2i}\alpha$.[21] The difference between the IV standard error formula in eq. (44) and the OLS version in eq. (37) is analogous to the difference between IV and OLS heteroscedasticity-robust standard errors: $\hat{\epsilon}_i$ is replaced in the numerator by the estimate of the structural residuals $\hat{\epsilon}_{\Delta,i}$, and the denominator is scaled by the first-stage coefficient. The IV analog of the standard error estimate with the null $H_0\colon \alpha = \alpha_0$ imposed instead estimates the residual as $(I - Z'(Z'Z)^{-1}Z')(Y_1 - Y_2\alpha_0)$, and the resulting confidence interval is a generalization of the Anderson and Rubin (1949) confidence interval (which assumes that the structural errors are independent). As a result, this confidence interval will remain valid even if the shift-share instrument

[21] Since the IV regression uses a single constructed instrument, $\hat{\epsilon}_\Delta$ is numerically equivalent to $\hat{\epsilon}_{1i} - \hat{\epsilon}_{2i}\hat{\alpha}$, where $\hat{\epsilon}_1$ and $\hat{\epsilon}_2$ are the reduced-form and first-stage residuals.

is weak.

Faced with the problem of estimating the treatment effect α in a setting in which the instrument has a shift-share structure, our approach to identification follows Borusyak, Hull and Jaravel (2018), who impose an assumption analogous to that in eq. (42), and also discuss the extension to a setting in which β_{FS} is allowed to vary across sectors and regions, and α is allowed to vary across regions. In contrast, Goldsmith-Pinkham, Sorkin and Swift (2018) suggest replacing the shift-share instrument X_i with the full vector of shares $(w_{i1}, \ldots, w_{iS})$. As pointed out by Tim Bartik in an online discussion, there are settings in which the shift-share instrument X_i satisfies the exclusion restriction, but the full vectors of shares $(w_{i1}, \ldots, w_{iS})$ does not, and is thus not a valid instrument.[22] Intuitively, this will be the case when the residual ϵ_Δ in the structural equation has a shift-share structure. Our independence restriction in eq. (42) allows for this possibility and, consequently, we adopt here the approach that has been standard since Bartik (1991), and use the the shift-share X_i as an instrumental variable.

5 Placebo exercise

In this section, we implement a placebo exercise to illustrate the finite-sample properties of the standard error estimators introduced in Section 4, and compare their behavior to that of standard error estimators typically used in applications of shift-share regressions. To this end, we use data on initial sectoral employment shares and changes in labor market outcomes observed for regional markets in the United States, and explore the properties of the OLS estimator of the coefficient on the shift-share covariate built from the combination of the actual employment shares with randomly generated sector-level shocks. Thus, consistently with the conceptualization of shift-share regressions in Section 4, we condition on observed data, Y and W, and combine it with randomly drawn shocks $\mathcal{X}$. We use this placebo analysis to explore the properties of: (a) the standard error estimates described in Section 4; (b) the Eicker-Hubert-White—or heteroskedasticity-robust—standard errors; and (c) cluster-robust standard errors with clusters defined as groups of regions geographically close to each other.

5.1 Data

In our baseline specification, we identify each region i with a U.S. Commuting Zone (CZ) and each sector s with either a 4-digit SIC manufacturing industry and an aggregated non-manufacturing sector. Specifically, the "shares" w_{is} correspond to employment shares in 1990 and the outcomes Y_i correspond to changes in employment rates and wages for different subsets of the population between 2000 and 2007. We index all manufacturing industries as $s = 1, \ldots, S-1$ and reserve the index $s = S$ for the non-manufacturing sector. To build this data, we follow the procedure described in the Online Appendix to Autor, Dorn and Hanson (2013). The County Business Patterns (CBP) is our main source of data on employment shares by sector and county in 1990. Our measures of changes in employment rates and average wages are based on data from the Census Integrated Public Use Micro Samples in

[22] See https://blogs.worldbank.org/impactevaluations/comment/5042#comment-5042. See also Borusyak, Hull and Jaravel (2018) for a discussion of different identification assumptions in this setting.

2000 and the American Community Survey for 2006 through 2008, which provide information at the Public Use Micro Area (PUMA) level.[23] We refer the reader to this Online Appendix to Autor et al. (2013) for details of the data construction.

The resulting baseline sample contains $N = 722$ regions and $S = 399$ sectors. To explore the robustness of our conclusions to the definition of regions and sectors in the data, we also perform our analysis using counties to define regions, and 3-digit and 2-digit SIC codes to define sectors.

5.2 Placebo exercise: impact of sector-level shocks on labor market outcomes

We consider a placebo exercise in which in each simulation draw $m = 1, \ldots, 1000$, we generate a random vector of sector-level shocks,

$$(\mathcal{X}_1^m, \ldots, \mathcal{X}_S^m) \sim \mathcal{N}(0, \Sigma), \tag{45}$$

where 0 denotes an S-vector of zeros and Σ is an $S \times S$ variance-covariance matrix given below. For each simulation draw m, we use the vector of generated sector-level "shifters" $(\mathcal{X}_1^m, \ldots, \mathcal{X}_S^m)$ and the data on the employment "shares" $\{w_{is}\}_{i=1,s=1}^{N,S}$ to construct a shift-share covariate $X_i^m = \sum_{s=1}^S w_{is}\mathcal{X}_s^m$ for every CZ i. Notice that in every simulation draw m, our computer-generated shocks $\{\mathcal{X}_s^m\}_{s=1}^S$ have no effect on the actual changes in labor market outcomes observed in U.S. CZs between 2000 and 2007. Thus, for the parameters $\{\beta_{is}\}_{i=1,s=1}^{N,S}$ introduced in eq. (21) we have $\beta_{is} = 0$ for all CZs i and sectors s. Consequently, the value of the parameter β defined in eq. (35) is also zero.

5.2.1 Baseline specification

Given data on the observed outcome Y_i and the generated shift-share covariate X_i^m, we compute for each simulation draw m the OLS estimate in eq. (34), the heteroskedasticity-robust standard error (which we label as *Robust* in our tables below), the standard error that clusters CZs in the same state (with label *St-cluster*), the standard error in eq. (37) (with label *AKM*), and the confidence interval and standard error in Remark 6 (with label *AKM0*). As this baseline specification includes no controls, we fix the matrix Z to be a column of ones when implementing the formulas in eqs. (34), (37) and (38).

In our baseline placebo specification, we impose two sets of restrictions on the variance-covariance matrix Σ. First, we set to 0 all elements in the Sth row and column of Σ; i.e. we assign to the aggregate non-manufacturing sector a shock of 0 in all the simulations.[24] Second, for all remaining elements of the matrix Σ, we set the element sk of Σ equal to $\sigma \mathbb{I}\{s = k\}$, with $\sigma = 5$.

[23] We download the CBP source data from the ICPSR website (`https://www.icpsr.umich.edu/icpsrweb/`), and apply the procedure in Autor, Dorn and Hanson (2013) to obtain employment by county and sector. The employment numbers in the CBP are often reported as an interval instead of an exact count; we use the cleaner files from Autor, Dorn and Hanson (2013) to estimate employment numbers within the indicated intervals. We also impose the sampling restrictions in Autor, Dorn and Hanson (2013) to construct our measures of employment rates and average wages. We match the resulting county- and PUMA-level information into CZs using the matching strategy in Dorn (2009), which has been previously applied, among others, in Autor and Dorn (2011, 2013). We thank the authors for making all codes publicly available through David Dorn's website: `https://www.ddorn.net/data.htm`.

[24] As discussed in footnote 18, this is standard practice in the literature that employs shift-share regressions with sectoral shocks. We explore the implications of this assumption in Section 5.2.2.

Table 1: Magnitude of standard errors.

	Estimates		Median effective std. error			
	Average (1)	Std. dev (2)	Robust (3)	St-cluster (4)	AKM (5)	AKM0 (6)
Panel A: Change in the share of working-age population						
employed	−0.03	1.95	0.73	0.92	1.87	2.17
employed in manufacturing	−0.05	1.83	0.60	0.76	1.76	2.04
employed in non-manufacturing	0.03	0.94	0.58	0.67	0.88	1.02
Panel B: Change in average log weekly wage						
employed	−0.03	2.60	1.00	1.32	2.52	2.93
employed in manufacturing	−0.02	2.94	1.67	2.09	2.71	3.16
employed in non-manufacturing	−0.05	2.57	1.04	1.31	2.50	2.91

Notes: For the outcome variable indicated in the first column and the inference procedure indicated in the first row, "median effective std. error" refers to the median length of the 95% confidence interval across the 1000 simulated datasets divided by 2×1.96. *Robust* is the Eicker-Huber-White standard error; *St-cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5; and *AKM0* is the confidence interval in Remark 6.

Table 1 presents the moments of the empirical distribution of the OLS estimates obtained with eq. (34), along with the median length of standard errors for different inference procedures. Since for *AKM0*, the standard error depends on the null being tested, the table reports "median effective standard error"—the median length of the 95% confidence interval divided by 2×1.96. For *Robust*, *St-Cluster*, and *AKM*, the "effective standard error" is the same as the actual standard error.

In line with our theoretical results, across simulations, the average of the estimated coefficient is approximately zero for all outcomes. Column (2) reports the standard deviation of the estimated coefficients. This dispersion is the target of our estimators of the standard error of the OLS estimator.[25] Columns (3)–(6) of Table 1 report the median effective standard error for *Robust*, *St-cluster*, *AKM*, and *AKM0*, respectively. Columns (3) and (4) show that *Robust* and *St-cluster* are downward biased relative to the standard deviation of the OLS estimator. On average across all outcomes, the median magnitude of the heteroskedasticity-robust and state-clustered standard errors are respectively 55% and 45% lower than its true value. In contrast, columns (5) and (6) show that our proposed methods perform much better. On average across all outcomes, the median standard error based on *AKM* is 5% lower than the true value. The median effective standard error of *AKM0* is slightly larger than the standard deviation of $\hat{\beta}$, by about 10% on average.[26]

Traditional inference methods assume that the regression residuals are independent across all regions, as in robust standard errors, or between geographically defined region groups, as in state-clustered standard errors. Given that shift-share regressors are correlated across regions with similar shares w_{is} (similar sector employment composition in our application), these methods generally lead

[25] Figure C.1 in Appendix C.2 reports the empirical distribution of the estimated coefficients when the dependent variable is the change in each CZ's employment rate. It distribution resembles a normal distribution centered around $\beta = 0$.

[26] Figure C.2 in Appendix C.2 presents histograms representing the empirical distribution of the effective standard errors, its mean and, for comparison, the standard deviation of the distribution of the OLS estimates $\{\hat{\beta}^m\}_{m=1}^{1000}$ for the placebo exercise based on change in the CZ's employment rate. Our new standard error estimators have a larger dispersion than the traditional ones, but are centered around the standard deviation of estimated parameters.

to a downward bias in the standard error estimate whenever the cross-region correlation in the regression residuals depends on the same shares used to construct the shift-share covariate; i.e. if regions with similar sector employment composition also tend to have similar regression residuals. Our proposed estimators allow for such a correlation and, for this reason, properly reflect the variability of the OLS estimator whether such correlation is present or not.

The difference between the confidence intervals based on the *AKM* and *AKM0* procedures and those based on either heteroskedasticity-robust or state-clustered standard errors indicates that the outcome variables used in the placebo have an important sectoral component. Consequently, whenever a researcher is running a shift-share regression in which the shifts are sectoral shocks, such as import, tariffs or technology shocks, both heteroskedasticity-robust and state-clustered standard errors will generally be biased downward. The only exception is the case in which either the shifters $\mathcal{X}$ or the controls $\mathcal{Z}$ fully account for all sectoral shocks affecting the outcome variable: only in this case the regression residual will have no remaining sectoral component. In contrast, the *AKM* and *AKM0* procedures will be valid, whether there is a sectoral component in the residual or not. We illustrate these points in Section 6.2, where we use different shift-share instruments to estimate the elasticity of regional employment to average wages.

The downward bias in the *Robust* and *St-cluster* standard errors translates into severe overrejection of the null hypothesis $H_0 : \beta = 0$. To show this, we report in Table 2 rejection rates for 5% significance level tests of this null hypothesis. Since the true value of β equals 0 by construction, a correctly behaved test statistic should generate a rejection rate of 5%. The results in Table 2 show that traditional standard error estimators yield much higher rejection rates. For example, when the outcome variable is the CZ's employment rate, the rejection rate for a 5% significance level for the null hypothesis $H_0 : \beta = 0$ is 48.8% and 37.2% when robust and state cluster standard errors are used, respectively. These rejection rates are very similar when the dependent variable is instead the change in the average log weekly wage.

The results in Table 2 thus indicate that, if we were to provide our 1000 simulated samples to 1000 researchers without disclosing to them the origin of the data (e.g. telling them instead that the sectoral shocks we provide to them are changes in trade flows, tariffs, or the number of foreign workers employed in an industry) and each of these researchers were to use standard inference procedures to perform a 5% significance level test for the null hypothesis $H_0 : \beta = 0$, at least a third of them would conclude that our computer generated shocks had a statistically significant effect on the evolution of employment rates in the United States between 2000 and 2007.

In contrast to the clear overrejection implied by traditional standard error estimators, our proposed inference procedures perform well. The *AKM* standard error has a rejection rate that is between 7.1% and 10.5% and the *AKM0* procedure has a rejection rate that is always between 4.1% and 5.5%. As discussed in Proposition 5 in Section 4.2, the standard error estimates in eqs. (37) and (38) are both consistent estimates of the true standard error of the OLS estimator of the coefficient on the shift-share covariate and, consequently, tests and confidence intervals that use either the *AKM* or the *AKM0* approach are asymptotically valid. The differences in rejection rates between *AKM* and *AKM0* in Table 2 are thus due to differences in finite-sample performance. In this context, the improved performance of the *AKM0* procedure, with rejection rates close to 5%, is not surprising. The improved

Table 2: Rejection rate of $H_0: \beta = 0$ at 5% significance level.

	Robust	St-cluster	AKM	AKM0
Panel A: Change in the share of working-age population				
Employed	48.8%	37.2%	7.2%	4.1%
Employed in manufacturing	55.8%	44.7%	7.1%	4.0%
Employed in non-manufacturing	23.0%	17.0%	8.9%	5.5%
Panel B: Change in average log weekly wage				
Employed	46.6%	32.8%	8.2%	4.5%
Employed in manufacturing	27.4%	18.1%	10.5%	4.6%
Employed in non-manufacturing	44.1%	32.3%	7.7%	5.0%

Notes: For the outcome variable indicated in the first column and the inference procedure indicated in the first row, this table indicates the percentage of the 1000 simulated datasets for which we reject the null hypothesis $H_0 : \beta = 0$ using a 5% significance level test. *Robust* is the Eicker-Huber-White standard error; *St-cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5; and *AKM0* is the test in Remark 6.

performance in small samples of standard errors computed under the null has been noted in other contexts; e.g., see Lazarus et al. (2018). Intuitively, imposing the null helps to reduce the finite-sample estimation noise when estimating the regression residuals.[27]

We summarize the conclusions from Table 1 and Table 2 in the following remark:

Remark 7. *In shift-share regressions, traditional inference methods suffer from a severe overrejection problem and yield confidence intervals that are too short and undercover; the inference procedures described in Remarks 5 and 6 yield tests with correct size and confidence intervals with the right coverage.*

Our placebo exercise follows closely our conceptualization of the shift-share regressions and, specifically, the notion that the sampling noise affecting the OLS estimator of the coefficient on the shift-share covariate arises from the fact that we only observe a particular realization of the "shifts" $(\mathcal{X}_1, \ldots, \mathcal{X}_S)$ and, thus, we do not observe the potential outcome for each region i from every possible realizations of these shocks (see Remark 4). As discussed in Section 4.3.2, this conceptualization differs from that in Goldsmith-Pinkham, Sorkin and Swift (2018), who instead treat the shares $(w_{i1}, \ldots, w_{iS})$ as independent across i and as good as randomly assigned. To explore the robustness of our inference procedures to this alternative characterization, we perform also an alternative placebo exercise in which we construct the shift-share covariate X_i^m for each simulation m by randomizing the vector of sectoral shares $\{w_{is}\}_{s=1}^S$ while holding constant across simulations the vector of sectoral shocks $(\mathcal{X}_1, \ldots, \mathcal{X}_S)$; i.e. $X_i^m = \sum_{s=1}^S w_{is}^m \mathcal{X}_s$.[28] Under this randomization scheme, all inference procedures considered in Table 2 have rejection rates close to 5%. Thus, while the consistency of the

[27] Appendix C investigates the sensitivity of our results to the definitions of geographic and exposure units. When we use counties (instead of CZs) as the regional unit of analysis, Table C.3 shows that rejection rates for all inference procedures are very similar to those in Table 2. We also investigate the performance of different inference procedures in an analogous placebo exercise using actual occupation employment shares in 1990 and randomly drawn shocks to 331 occupations. Table C.4 shows that the overrejection of traditional inference methods is more severe with shift-share regressors based on occupations; in this case, our inference procedure under the null yields the correct test size.

[28] Specifically, we generate a single random draw of $(\mathcal{X}_1, \ldots, \mathcal{X}_S)$ from the distribution in eq. (45) and we keep this draw constant across the 1000 simulated samples. Then, in each simulation draw m, we assign to each CZ i a vector of shares $(w_{i1}^m, \ldots, w_{iS}^m)$ drawn from the empirical distribution of the CZ shares $\{(w_{i1}, \ldots, w_{iS})\}_{i=1}^N$.

Table 3: Rejection rate of $H_0: \beta = 0$ at 5% significance level. Sectoral composition.

	Robust	St-cluster	AKM	AKM0	AKM	AKM0
Sector Shock Correlation:			Independent		3-digit cluster	
Panel A: Number of sectors						
2-digit ($S = 20$)	72.5%	59.2%	12.8%	6.1%	—	—
3-digit ($S = 136$)	53.6%	42.9%	7.1%	4.7%	—	—
4-digit ($S = 398$)	46.8%	36.8%	7.7%	4.6%	—	—
Panel B: Simulated shocks to non-manufacturing sector						
	93.0%	90.2%	78.0%	77.4%	—	—
Panel C: Simulated shocks with correlation within 3-digit SIC sectors						
$\rho = 0.00$	47.6%	36.6%	6.6%	4.2%	6.9%	4.1%
$\rho = 0.25$	48.5%	32.3%	7.0%	6.3%	5.0%	4.7%
$\rho = 0.50$	47.9%	31.8%	7.2%	8.1%	4.8%	5.5%
$\rho = 0.75$	53.7%	37.3%	9.3%	10.1%	4.9%	5.2%
$\rho = 1.00$	52.0%	36.5%	11.6%	13.2%	5.3%	5.3%

Notes: All estimates in this table use the total employment share in each CZ as the outcome variable Y_i. The inference procedure employed to compute the rejection rate in each of the columns is indicated in the first row. This rejection rate indicates the percentage of the 1000 simulated datasets for which we reject the null hypothesis $H_0 : \beta = 0$ using a 5% significance level test. *Robust* is the Eicker-Huber-White standard error; *St-cluster* is the standard error that clusters CZs in the same state; *AKM* (Independent) is the standard error in Remark 5; *AKM0* (Independent) is the test in Remark 6; *AKM* (3-digit cluster) is the standard error in eq. (40); and *AKM0* (3-digit cluster) is the confidence interval described in the last sentence of Section 4.3.1.

heteroskedasticity-robust and state-clustered standard error estimates depends crucially on the conceptualization of the shift-share regression, the inference procedures that we provide in Remarks 5 and 6 generate the right rejection rates when either the "shifts" or the "shares" are as good as randomly assigned.

5.2.2 Alternative number of sectors and correlated sectoral shocks

The inference procedures described in Remarks 5 and 6 generate tests and confidence intervals that are valid only if: (a) the number of sectors goes to infinity (see Assumption 4(i)); (b) all sectors are asymptotically "small" (see Assumption 3(i)); (c) all sectoral shocks are independent of each other conditional on the matrix of shares W and other sectoral shocks $\mathcal{Z}$ that we proxy for in the regression (see Assumption 4(ii)). We test how sensitive different inference procedures are to the number of sectors, the size of the largest sector and the correlation structure of the sectoral shocks in panels A, B, and C, respectively, of Table 3.

In Panel A, we perform the placebo exercise described in Section 5.2.1 with different sector definitions used to build the region- and sector-specific shares. The results in Panel A of Table 3 show that the overrejection problem affecting standard inference procedures is worse when the number of sectors decreases: the rejection rates for *Robust* and *St-cluster* standard errors reach 72.5% and 59.2%, respectively, when only the 20 2-digit SIC sectors are considered in the analysis. The rejection rates for *AKM* also increase to 12.8%, but those for *AKM0* remain quite close to the 5% significance level.

In Panel B, we redo the placebo exercise assigning a randomly generated sectoral shock also to the aggregate non-manufacturing sector. Across the CZs in our analysis, the non-manufacturing sector accounts on average for 78% of employment, with a minimum employment share of 38%. Our simulation indicates that all methods perform poorly in this case. So, in practice, it is important to have small sectors.

In Panel C, we allow the shocks $\mathcal{X}_s$ corresponding to industries that belong to the same 3-digit sector to be correlated. Specifically, we set again all elements in the Sth row and column of Σ to 0 and, for all remaining elements of Σ, we set the element sk of the matrix Σ equal to $(1-\rho)\sigma\,\mathbb{I}\{s=k\}+\rho\sigma\,\mathbb{I}\{c(s)=c(k)\}$, where, for every s, $c(s)$ indicates the 3-digit sector that industry s belongs to. When implementing either the standard inference procedures *Robust* and *St-cluster* or the version of the *AKM* and *AKM0* confidence intervals that assume no correlation across the sectoral shocks, we observe that within-cluster correlation in sectoral shocks has a small impact on the rejection rate. However, the inference procedures described in Section 4.3.1, which generalize *AKM* and *AKM0* to allow for correlation in sectoral shocks $\mathcal{X}_1,\ldots,\mathcal{X}_S$ across sectors belonging to the same cluster, yield the right rejection rate.

We summarize the conclusions from Table 3 in the following remark.

Remark 8. *In shift-share regressions, overrejection is more severe when there is a small number of large sectors. In this case, our methods significantly reduce the rejection rate, although they may still overreject relative to the nominal significance level when the number of sectors is very small.*

5.2.3 Confounding sector-level shocks: omitted variable bias and solutions

In Appendix C.1, we investigate the consequences of violations of Assumption 1(i) that requires observed sectoral shocks of interest $\mathcal{X}_1,\ldots,\mathcal{X}_S$ to be independent from other sectoral shocks affecting the outcome variable of interest. In particular, we study the impact that violations of this assumption have on the properties of the OLS estimator of the coefficient on the shift-share regressor of interest. We also consider the properties of two solutions to this problem: (i) the inclusion of regional controls as a proxy for sector-level unobserved shocks (discussed in Section 4.2), and (ii) the use of a shift-share instrumental variable constructed as a weighted average of exogenous sector-level shocks (discussed in Section 4.3.2).

Our simulations confirm that confounding sector-level shocks introduce bias in the OLS estimator of the coefficient on the shift-share regressor of interest. In such cases, region-level controls eliminate the bias only if they are a perfect proxy for the sector-level confounding shock. Otherwise, an IV approach is needed. We also verify that, whenever the corresponding estimator is consistent, the patterns discussed in Remark 7 apply to both OLS and IV estimators: traditional inference methods suffer from a severe overrejection problem and yield confidence intervals that undercover; the inference procedures described in Remarks 5 and 6 yield the correct test size and confidence intervals with the right coverage.

6 Empirical applications

We now apply our methodology to two empirical applications. First, the study of the effect of Chinese competition on labor market outcomes across U.S. Commuting Zones, as in Autor, Dorn and Hanson (2013). Second, the estimation of the elasticity of labor supply, as in Bartik (1991). In both of these applications, we use the data sources described in Section 5.1.

6.1 Effect of chinese exports on U.S. labor market outcomes

Autor, Dorn and Hanson (2013), henceforth referred to as ADH, explore the impact of exports from China on labor market outcomes across U.S. Commuting Zones. Specifically, they present estimates such as those in eq. (43) where Y_{i1} is the ten-year equivalent change in a labor-market outcome in CZ i in 1990–2000 and 2000–2007, w_{is} is the CZ i employment share in the 4-digit SIC sector s in the initial year of the corresponding period (either 1990 or 2000), Y_{i2} is a weighted average of the change in sectoral U.S. imports from China normalized by U.S. total employment in the corresponding sector, and X_i is analogous to Y_{i2} with the only difference that, instead of using U.S. imports from China as shifters, it uses imports from China to other high-income countries. As the K-vector of covariates Z_i, we will include the largest set of controls accounted for in ADH; see, e.g., column (6) of Table 3 of ADH.

Table 4 reports our replication of results presented in Tables 5 to 7 of ADH, along with 95% confidence intervals computed with different methodologies. Panel A presents the IV estimates; Panel B presents reduced-form estimates, and Panel C presents the first stage estimates. These corresponds to $\hat{\alpha}$, $\hat{\beta}_{RF}$ and $\hat{\beta}_{FS}$, respectively, in the notation introduced in Section 4.3.2. Our estimates closely replicate those in ADH.[29]

We focus on the comparison between confidence intervals obtained with different inference procedures. In all panels, state-clustered confidence intervals are very similar to the heteroskedasticity-robust ones. This suggests that there is not much correlation in residuals within states. In contrast, our proposed confidence intervals are wider than those implied by state-clustered standard errors. In the IV regression reported in Panel A, across all outcomes, the average increase in the length of the 95% confidence interval is 25% with the *AKM* procedure and 100% with the *AKM0* procedure. When the outcome variable is the change in the manufacturing employment rate, the length of the 95% confidence interval increases by 36% with the *AKM* procedure and by 136% with the *AKM0* procedure.[30] In light of the lack of impact of state-clustering on the 95% confidence interval, the wider intervals implied by our inference procedures indicate that cross-region residual correlation is mostly driven

[29]The slight difference between our estimates and those reported in ADH for both the point estimates $\hat{\alpha}$, $\hat{\beta}_{RF}$ and $\hat{\beta}_{FS}$ and the *Robust* and *Cluster* standard errors is due to differences in the measures of employment shares w_{is} and total U.S. employment by sector, which are not directly reported in the replication package shared by the authors. Note that information on these variables is generally not needed to replicate ADH once information on the shift-share variables Y_{i2} and X_i is provided (what the authors certainly do), but it is need in order to compute the confidence interval estimates introduced in Section 4.

[30]The *AKM* and *AKM0* estimates reported in Table 4 account for correlation in the sectoral shifters across periods and across 4-digit SIC sectors included in the same 3-digit SIC sector. Table D.2 shows that similar increases in the length of the 95% confidence intervals are implied by *AKM* and *AKM0* when we assume that sectoral shifters are: (a) independent across 4-digit SIC sectors and periods; (b) independent across 4-digit SIC sectors but possibly correlated across periods.

Table 4: Effect of Chinese exports on U.S. commuting zones—Autor et al. (2013).

	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: 2SLS Regression						
$\hat{\beta}$	-0.75	-0.58	-0.17	-0.70	0.20	-0.72
Robust	$[-1.09, -0.42]$	$[-0.77, -0.39]$	$[-0.48, 0.13]$	$[-1.18, -0.22]$	$[-0.79, 1.19]$	$[-1.23, -0.21]$
Cluster	$[-1.12, -0.39]$	$[-0.78, -0.38]$	$[-0.47, 0.12]$	$[-1.19, -0.21]$	$[-0.76, 1.16]$	$[-1.22, -0.22]$
AKM	$[-1.24, -0.26]$	$[-0.85, -0.31]$	$[-0.55, 0.20]$	$[-1.31, -0.09]$	$[-0.73, 1.13]$	$[-1.37, -0.07]$
AKM0	$[-2.02, -0.34]$	$[-1.25, -0.31]$	$[-0.96, 0.15]$	$[-1.95, -0.09]$	$[-1.45, 1.17]$	$[-2.23, -0.16]$
Panel B: OLS Reduced-Form Regression						
$\hat{\beta}$	-0.66	-0.50	-0.15	-0.61	0.18	-0.63
Robust	$[-0.97, -0.35]$	$[-0.65, -0.36]$	$[-0.43, 0.12]$	$[-1.05, -0.17]$	$[-0.67, 1.02]$	$[-1.11, -0.15]$
Cluster	$[-0.87, -0.44]$	$[-0.62, -0.39]$	$[-0.39, 0.09]$	$[-1.01, -0.21]$	$[-0.67, 1.02]$	$[-1.06, -0.20]$
AKM	$[-1.11, -0.20]$	$[-0.73, -0.28]$	$[-0.49, 0.18]$	$[-1.17, -0.05]$	$[-0.62, 0.97]$	$[-1.23, -0.02]$
AKM0	$[-1.72, -0.30]$	$[-0.94, -0.30]$	$[-0.89, 0.12]$	$[-1.66, -0.08]$	$[-1.45, 0.87]$	$[-1.97, -0.13]$
Panel C: 2SLS First-Stage						
$\hat{\beta}$			0.87			
Robust			$[0.63, 1.11]$			
Cluster			$[0.62, 1.12]$			
AKM			[0.71,1.03]			
AKM0			$[0.72, 1.17]$			

Notes: $N = 1,444$ (722 CZs $\times$ two time periods). Models are weighted by start of period CZ share of national population. All regressions include the full vector of baseline controls in ADH. 95% confidence intervals in square brackets. *Robust* is the Eicker-Huber-White standard error; *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 the last sentence of Section 4.3.1.

by similarity in sectoral compositions rather than by geographic proximity.

Panel B of Table 4 reports confidence intervals for the reduced-form specification. In this case, the increase in the confidence interval length is slightly larger: across outcomes, it increases on average by 54% for *AKM* and 129% for *AKM0*. The smaller relative increase in the confidence interval length for the IV estimate $\hat{\alpha}$ relative to its increase for the reduced-form estimate $\hat{\beta}_{RF}$ is a consequence of the fact that all inference procedures yield very similar confidence intervals for the first-stage estimate $\hat{\beta}_{FS}$, as reported in Panel C.

As discussed in Section 5, the differences between the *AKM* (or the *AKM0*) standard errors and state-clustered standard errors are related to the importance of the sector-level component in the regression residual. The results in Panel C suggest that, once we account for the changes in sectoral imports from China to other high-income countries, there is not much sectoral variation left in the first-stage regression residual; i.e., there are no other sectoral variables that are important to explain the changes in sectoral imports from China to the U.S.[31] To formally investigate this claim, Table D.3

[31]This is similar to the effect of state-specific regression covariates on state-clustered standard errors: if one includes enough covariates that vary at the state level, state-clustered standard errors in regressions in which the variable of interest also varies at the state level will be similar to heteroskedasticity-robust standard errors, since there is not much within-state

reports the rejection rates implied by a placebo exercise analogous to that described Section 5 when the outcome variable in the placebo exercise is the same as that in the first-stage specification reported in Panel C of Table 4. Panel A in Table D.3 shows that, when no covariates are included, traditional methods still suffer from severe overrejection problems and our methods yield the correct test size. However, as shown in Panels B and C in Table D.3, the problem is greatly attenuated when controlling for the instrumental variable and other covariates used in ADH. This indicates that the instrumental variable and additional covariates included in ADH soak most of the cross-CZ correlation in the CZs exposure to Chinese imports; i.e. most of the cross-CZ correlation in the ADH treatment variable.

Overall, Table 4 shows that, despite the wider confidence intervals obtained with our procedure, the qualitative conclusions in ADH with respect to the effect of U.S. imports from China on CZs labor market outcomes remain valid at usual significance levels. However, the increase in the length of the 95% confidence interval indicates that there is more uncertainty regarding the magnitude of the impact of Chinese import exposure on labor markets. In particular, the *AKM0* confidence interval is much wider than that based on state-clustered standard errors; it is asymmetric around the point estimate, and it indicates that the negative impact of the China shock could have been 2 to 3 times larger than the effect implied by the point estimates.[32]

6.2 Estimation of labor supply elasticity

Our second application focuses on the estimation of the labor supply elasticity. Following the model in Section 2, we consider the estimation of the parameter ϕ using eq. (10) as estimating equation:

$$\hat{e}_i = \phi\hat{\omega}_i + Z_i\delta + \hat{v}_i, \tag{46}$$

where, as in Section 2, we use $\hat{z} = \ln(z^t/z^0)$ to denote log-changes in a variable z between some initial period $t = 0$ and some other period t. In our empirical application, we define each region i as a U.S. CZ and use the same vector of covariates Z_i as in the application described in Section 6.1; i.e., the vector of controls listed in column (6) of Table 3 of ADH.

As illustrated through the model in Section 2, changes in local supply shocks, $\hat{v}_i$, will generally affect both changes in equilibrium local average wages and local employment rates. Thus, $\hat{\omega}_i$ and $\hat{v}_i$ are correlated and the OLS estimator of ϕ in eq. (46) will be biased. To circumvent this problem, a popular approach in the literature is the use of shift-share instrumental variables. In this section, we implement this strategy with two different sector-level shifters: (i) the national employment growth, as in Bartik (1991); and (ii) the increase in imports from China by a set of high-income countries that does not include the United States, as in Autor, Dorn and Hanson (2013).

Table 5 presents first-stage, reduced-form and IV estimates associated to the estimation of the

correlation left in the residuals.

[32]It follows from the formula in Remark 6 (see the expression for the quantity A, by which the confidence interval is recentered) that the asymmetry comes from the correlation between the regression residuals $\hat{R}_s$ and the shifters cubed, which is zero in large samples. In fact, it can be shown that *AKM* and *AM0* are asymptotically equivalent. The differences between the confidence intervals in Table 4 thus reflects differences in their finite-sample properties, which are not captured by the asymptotics. This notwithstanding, the placebo exercise presented in Section 5 shows that both inference procedures yield close to correct rejection rates in a sample with the same number of regions and sectors as that used to generate the estimates in Table 4.

Table 5: Estimation of labor supply elasticity.

	First-Stage $\hat{w}_i$ (1)	Reduced-Form $\hat{e}_i$ (2)	2SLS $\hat{e}_i$ (3)
Panel A: Bartik IV			
	30.08	37.99	1.26
Robust	$[21.05, 39.12]$	$[28.18, 47.80]$	$[0.97, 1.55]$
Cluster	$[18.28, 41.89]$	$[24.07, 51.91]$	$[0.89, 1.63]$
AKM	$[16.68, 43.49]$	$[22.77, 53.21]$	$[0.97, 1.56]$
AKM0	$[15.73, 44.81]$	$[21.68, 54.69]$	$[0.98, 1.72]$
Panel B: ADH IV			
	-0.61	-0.97	1.58
Robust	$[-1.05, -0.17]$	$[-1.43, -0.50]$	$[0.75, 2.42]$
Cluster	$[-1.01, -0.21]$	$[-1.27, -0.66]$	$[0.78, 2.39]$
AKM	$[-1.17, -0.05]$	$[-1.64, -0.29]$	$[0.53, 2.64]$
AKM0	$[-1.66, -0.08]$	$[-2.54, -0.45]$	$[0.88, 4.79]$

Notes: $N = 1,444$ (722 CZs $\times$ two time periods). Models are weighted by start of period CZ share of national population. All regressions include the full vector of baseline controls in ADH. 95% confidence intervals in square brackets. *Robust* is the Eicker-Huber-White standard error; *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 the last sentence of Section 4.3.1.

parameter ϕ in eq. (46). Panel A reports results using the Bartik (1991) instrumental variable with sectoral national employment growth as shifters, and Panel B reports results using the ADH instrumental variable with sectoral imports from China by high-income countries other than the US as shifters. In both cases, the estimates of the labor supply elasticity are similar: 1.26 in Panel A and 1.58 in Panel B.

We now compare the different IV confidence intervals in column (3). In Panel A, our proposed confidence intervals are wider than heteroskedasticity-robust 95% confidence intervals, but tighter than state-clustered 95% confidence intervals. For Panel B, our proposed confidence intervals are 30%–140% wider than those obtained with state-clustered and heteroskedasticity-robust standard errors. As discussed in Section 6.1, such differences are related to the sector-level component of the cross-regional correlation in residuals. Our results suggest that the national employment growth in the Bartik (1991) instrumental variable absorbs the bulk of this component, leaving little correlation left for our inference procedures to correct. In contrast, the ADH instrumental variable absorbs a lower fraction of the sectoral component of residuals, which implies that our procedure has a larger impact on the length of the 95% confidence interval.

7 Concluding remarks

This paper analyzes the statistical properties of shift-share empirical specifications. Our analysis shows that standard economic models predict changes in regional outcomes to depend on observed and unobserved sector-level shocks through several shift-share covariates. Our model thus implies that the residual in shift-share regressions is likely to be correlated across regions with similar sectoral composition, independently of their geographic location, due to the presence of unobserved sectoral shifters affecting the outcome of interest. Such a correlation is ignored by inference procedures typically used in shift-share regressions, such as when standard errors are clustered on geographic units. To illustrate the importance of this shortcoming, we implement a placebo exercise in which we study the effect of randomly generated sector-level shocks on actual changes in labor market outcomes across CZs in the United States. We find that traditional inference procedures severely overreject the null hypothesis of no effect. We derive two novel inference procedures that yield correct rejection rates.

It has become standard practice to report cluster-robust standard errors in regression analysis whenever the variable of interest varies at a more aggregate level than the unit of observation. This practice guards against potential correlation in the residuals that arises whenever the residuals contain unobserved shocks that also vary at a more aggregate level. In the same way, we recommend that researchers report confidence intervals in shift-share designs that allow for a shift-share structure in the residuals, such as one of the two confidence intervals that we propose.

References

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

_ **and Pascual Restrepo**, "Robots and jobs: Evidence from US labor markets," *MIT mimeo*, 2017.

_ **and** _ , "Demographics and Automation," *MIT mimeo*, 2018.

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

_ **, Costas Arkolakis, and Federico Esposito**, "Spatial linkages, Global Shocks, and Local Labor Markets: Theory and Evidence," *University of Chicago mimeo*, 2018.

Aghion, Philippe, Antonin Bergeaud, Matthieu Lequien, and Marc J. Melitz, "The impact of exports on innovation: Theory and evidence," Technical Report 2018.

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

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

Amiti, Mary and David E Weinstein, "Exports and financial shocks," *The Quarterly Journal of Economics*, 2011, *126* (4), 1841–1877.

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.

Anderson, Theodore W. and Herman Rubin, "Estimation of the Parameters of a Single Equation in a Complete System of Stochastic Equations," *The Annals of Mathematical Statistics*, 1949, *20* (1), 46–63.

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*, 1969, *16* (1), 159–178.

Autor, David H. and David Dorn, "This Job Is "Getting Old": Measuring Changes in Job Opportunities Using Occupational Age Structure," *American Economic Review*, 2011, *99* (2), 45–51.

_ **and** _ , "The Growth of Low-Skill Service Jobs and the Polarization of the US Labor Market," *American Economic Review*, August 2013, *103* (5), 1553–97.

_ **,** _ **, 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.

_ , _ , and _ , "The China Shock: Learning from Labor-Market Adjustment to Large Changes in Trade," *Annual Economic Review*, 2016, *8*, 205–240.

_ , _ , and _ , "When Work Disappears: Manufacturing Decline and the Falling Marriage-Market Value of Young Men," *American Economic Review: Insights*, 2018, *forthcoming.*

_ , _ , _ , and Kaveh Majlesi, "Importing Political Polarization? The Electoral Consequences of Rising Trade Exposure," *University of Zurich mimeo*, 2017.

_ , _ , _ , Gary Pisano, and Pian Shu, "Foreign Competition and Domestic Innovation: Evidence from U.S. Patents," *University of Zurich mimeo*, 2017.

Barrios, Thomas, Rebecca Diamond, Guido W. Imbens, and Michal Koles'ar, "Clustering, Spatial Correlation, and Randomization Inference," *Journal of the American Statistical Association*, 2012, *107* (498), 578–591.

Bartelme, Dominick, "Trade Costs and Economic Geography: Evidence from the U.S.," *University of Michigan mimeo*, 2018.

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

Blanchard, Olivier and Lawrence F Katz, "Regional Evolutions, Brooking Papers on Economic Activity," *Economic Studies Program, The Brookings Institution*, 1992, *23* (1), 76.

Borusyak, Kirill, Peter Hull, and Xavier Jaravel, "Quasi-experimental Shift-share Research Designs," Technical Report 2018. arXiv:1806.01221 [econ.EM].

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

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

Bustos, Paula, Bruno Caprettini, and Jacopo Ponticelli, "Agricultural productivity and structural transformation: Evidence from Brazil," *American Economic Review*, 2016, *106* (6), 1320–65.

Cameron, Colin A. and Douglas L. Miller, "A Practitioner's Guide to Cluster-Robust Inference," *Journal of Human Resources*, 2014, *50* (2), 317–372.

Card, David, "Immigrant inflows, native outflows, and the local labor market impacts of higher immigration," *Journal of Labor Economics*, 2001, *19* (1), 22–64.

_ and John Dinardo, "Do Immigrant Inflows Lead to Native Outflows?," *American Economic Review*, 2000, *90* (2), 360–367.

Che, Yi, Yi Lu, Justin R. Pierce, Peter K. Schott, and Tao Zhigang, "Did Trade Liberalization with China Influence U.S. Elections?," *Yale University mimeo*, 2017.

Chodorow-Reich, Gabriel and Johannes Wieland, "Secular Labor Reallocation and Business Cycles," *Harvard University mimeo*, 2018.

Colantone, Italo and Piero Stanig, "The Trade Origins of Economic Nationalism: Import Competition and Voting Behavior in Western Europe," *American Journal of Political Science*, 2018, *forthcoming.*

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

Dix-Carneiro, Rafael and Brian K. Kovak, "Trade Liberalization and Regional Dynamics," *American Economic Review*, October 2017, *107* (10), 2908–46.

— **and** — , "Margins of Labor Market Adjustment to Trade," *Duke University mimeo*, 2018.

— **, Rodrigo Soares, and Gabriel Ulyssea**, "Economic Shocks and Crime: Evidence From The Brazilian Trade Liberalization," *American Economic Journal: Applied Economics*, 2017, *forthcoming.*

Dorn, David, "Essays on Inequality, Spatial Interaction, and the Demand for Skills." PhD dissertation, University of St. Gallen no. 3613 2009.

Dustmann, Christian, Tommaso Frattini, and Ian P. Preston, "The Effect of Immigration along the Distribution of Wages," *The Review of Economic Studies*, 2013, *80* (1), 145–173.

— **, 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," *University of California Los Angeles mimeo*, 2018.

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

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift, "Bartik Instruments: What, When, Why, and How," Technical Report 24408, National Bureau of Economic Analysis, 2018.

Greenstone, Michael, Alex Mas, and Hoai-Luu Nguyen, "Do Credit Market Shocks Affect the Real Economy? Quasi-Experimental Evidence from the Great Recession and 'Normal' Economic Times," *Princeton University mimeo*, 2015.

Hakobyan, Sushanik and John McLaren, "Looking for Local Labor Market Effects of NAFTA," *Review of Economics and Statistics*, 2016, *98* (4), 728–741.

Heckman, James J. and Edward J. Vytlacil, "Econometric Evaluation of Social Programs, Part I: Causal Models, Structural Models and Econometric Policy Evaluation," in James J. Heckman and Edward E. Leamer, eds., *Handbook of Econometrics*, Vol. 6, Elsevier, 2007, pp. 4779–4874.

— **and** — , "Econometric Evaluation of Social Programs, Part II: Using the Marginal Treatment Effect to Organize Alternative Econometric Estimators to Evaluate Social Programs, and to Forecast their Effects in New Environments," in James J. Heckman and Edward E. Leamer, eds., *Handbook of Econometrics*, Vol. 6, Elsevier, 2007, pp. 4875–5143.

— **, Lochner Lance, and Christopher Taber**, "General-Equilibrium Treatment Effects: A Study of Tuition Policy," *American Economic Review*, 1998, *88* (2), 381–386.

Huber, Kilian, "Disentangling the Effects of a Banking Crisis: Evidence from German Firms and Counties," *American Economic Review*, March 2018, *108* (3), 868–98.

Hummels, David, Rasmus Jørgensen, Jakob Munch, and Chong Xiang, "The wage effects of offshoring: Evidence from Danish matched worker-firm data," *American Economic Review*, 2014, *104* (6), 1597–1629.

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," *CUNY University mimeo*, 2018.

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

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

Lazarus, Eben, Daniel J. Lewis, James H. Stock, and Mark W. Watson, "HAR Inference: Recommendations for Practice," *Harvard University mimeo*, 2018.

Lee, Eunhee, "Trade, Inequality, and the Endogenous Sorting ofHeterogeneous Workers," *University of Maryland mimeo*, 2017.

Lewis, Ethan and Giovanni Peri, "Immigration and the Economy of Cities and Regions," in Gilles Duranton, J. Vernon Henderson, and William C. Strange, eds., *Handbook of Regional and Urban Economics*, Vol. 5A, UK: North Holland, 2015, pp. 625–685.

Monras, Joan, "Immigration and Wage Dynamics: Evidence from the Mexican Peso Crisis," *CEMFI mimeo*, 2015.

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

Pierce, Justin R. and Peter K. Schott, "Trade Liberalization and Mortality: Evidence from US Counties," *Yale University mimeo*, 2017.

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.

Roy, Andrew D., "Some Thoughts on the Distribution of Earnings," *Oxford Economic Papers*, 1951, *3* (2), 135–146.

Topalova, Petia, "Trade Liberalization, Poverty and Inequality: Evidence from Indian Districts," in Ann Harrison, ed., *Globalization and Poverty*, University of Chicago Press, 2007, pp. 291–336.

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

Zidar, Owen, "Tax Cuts For Whom? Heterogeneous Effects of Income Tax Changes on Growth and Employment," *Journal of Political Economy*, 2018, *forthcoming.*

Appendices

A Economic model: details and extensions

A.1 Sector-specific price index

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

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

where $\{\theta_{ss'}\}_{s,s'}$ are positive constants, and x_{js} is the share of the world production in sector s that corresponds to region j, $x^0_{js} \equiv X_{js} / \sum_{i=1}^{J} X_{is}$. Imposing that all regions j in a country c are small is equivalent to assuming that $x^0_{js} \approx 0$ for all $j \in J_c$ and for $s = 1, \ldots, S$. Therefore, when all regions $j \in J_c$ are small, we can rewrite the change in the sector-specific price index as

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

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

Proof of eq. (A.1). Equations (9) and (14) imply that

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

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

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

with $\hat{\eta}_s \equiv \sum_j x^0_{js} \left(\lambda_j \sum_k l^0_{jk} \left[\hat{\gamma}_k + (\sigma_k - 1)\hat{A}_{jk} \right] - \lambda_j \hat{v}_j - \hat{A}_{js} \right)$. As $\sigma_k > 1$ and $\phi > 0$, it will be true that

$$\sum_k |\tilde{\theta}_{sk}| = \sum_j x^0_{js} \frac{-1 + \sum_k l^0_{jk} \sigma_k}{\phi + \sum_k l^0_{jk} \sigma_k} < \sum_j x^0_{js} = 1,$$

and, therefore,

$$\boldsymbol{\theta} \equiv (I - \tilde{\boldsymbol{\theta}})^{-1} = \sum_{d=0}^{\infty} \tilde{\boldsymbol{\theta}}^d, \text{ with } \theta_{sk} > 0.$$

This immediately implies eq. (A.1). □

A.2 Expression for sector- and region-specific employment

Similarly to eq. (18), we can write the change in employment in sector s in a region i as

$$\hat{L}_{is} = \hat{\gamma}_s + (\sigma_s - 1)\left(\hat{A}_{is} + \hat{P}_s\right) - \sigma_s \lambda_i \sum_{s=1}^{S} l_{is}^0 \left[\hat{\gamma}_s + (\sigma_s - 1)\left(\hat{A}_{is} + \hat{P}_s\right)\right] + \sigma_s \lambda_i \hat{v}_i. \tag{A.3}$$

This expression shows that the change in employment in a particular sector- and region-specific pair also depends on the same set of shocks affecting each region's total employment and wages according to eqs. (17) and (18). However, as a comparison of eqs. (18) and (A.3) illustrates, the response of total regional employment $\hat{L}_i$ to, for example, a particular change in the world price of a sector s, $\hat{P}_s$, is different from the response of sectoral employment, $\hat{L}_{is}$, no matter whether we focus on its response to the same sector prices, $\hat{P}_s$, or its response to changes in prices in a different sector $k \neq s$, $\hat{P}_k$. Specifically,

$$d\hat{L}_{is}/d\hat{P}_s = (\sigma_s - 1) - l_{is}^0 \beta_{L,iss}, \qquad \text{with} \quad \beta_{L,iss} \equiv \sigma_s \lambda_i (\sigma_s - 1); \tag{A.4}$$

$$d\hat{L}_{is}/d\hat{P}_k = -l_{ik}^0 \beta_{L,isk}, \qquad \text{with} \quad \beta_{L,isk} \equiv \sigma_s \lambda_i (\sigma_k - 1), \qquad \forall\, s \neq k. \tag{A.5}$$

The expression in eq. (A.5) reflects the presence of cross-sectoral spillovers in the model described in Section 2.1: exogenous shocks to labor demand in one sector k affect employment in every other sector s. By doing simple algebra, the expressions in eqs. (A.4) and (A.5) also show that one can recover the total impact of $\hat{P}_s$ on regional employment $\hat{L}_i$ in eq. (19) as the appropriate weighted average of the elasticities in eqs. (A.4) and (A.5): $d\hat{L}_i/d\hat{P}_s = l_{is}^0 d\hat{L}_{is}/d\hat{P}_s + \sum_{k \neq s} l_{ik}^0 (d\hat{L}_{ik}/d\hat{P}_s)$.[33] Following this approach would however require estimating S parameters just to compute $d\hat{L}_i/d\hat{P}_s$ for a particular is pair; i.e. $\{\beta_{L,isk}\}_{k=1}^S$. Conversely, using the expression in eq. (18) as basis for analysis only requires estimating one single parameter: the parameter $\beta_{L,is}$ introduced in eq. (19).

A.3 Allowing for regional migration

We extend here the baseline environment described in Section 2.1 to allow for mobility of individuals across regions within a single country c. We still assume that the number of individuals living in each country c is fixed and equal to M_c.

Environment. The only difference with respect to the setting described in Section 2.1 is that the mass of individuals living in a region i, M_i, is no longer fixed. Instead, we assume that, before the realization of the shock $u(\iota)$ in eq. (6), 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

$$U(\iota) = \tilde{u}_i(\iota)\left(\bar{U}_i(\omega_i/P, b_i/P) - 1\right) \tag{A.6}$$

where $\bar{U}_i(\omega_i/P, b_i/P)$ is the expected utility of residing in region i, as determined by eqs. (6) and (7),

[33] Note that $\hat{L}_i = \sum_{k=1}^S l_{ik}^0 \hat{L}_{ik}$ and, therefore, $d\hat{L}_i/d\hat{P}_s = \sum_{k=1}^S l_{ik}^0 (d\hat{L}_{ik}/d\hat{P}_s) = l_{is}^0 d\hat{L}_{is}/d\hat{P}_s + \sum_{k \neq s} l_{ik}^0 (d\hat{L}_{ik}/d\hat{P}_s)$.

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:

$$\tilde{u}_i(\iota) \sim F_{\tilde{u}}(\tilde{u}) = e^{-\tilde{u}^{-\tilde{\phi}}}, \qquad \tilde{\phi} > 0. \tag{A.7}$$

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

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

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

To simplify the analysis, we assume that the unemployment benefit is identical in all regions and equal to the price index P; i.e. $b_i = P$ for all $i \in J$. Defining $v_i \equiv (\nu_i/b_i)^{\phi}$ as in eq. (10), the assumption that $b_i = P$ for all $i \in J$ implies that $v_i \equiv \nu_i/P$ and, thus,

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

and the share of national population in region i is

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

Given the distributional assumption in eq. (A.7), it holds that

$$M_i = \frac{v_i(\omega_i)^{\phi_m}}{\Phi_c}M_c \ \text{ such that } \ \Phi_c = \sum_{j \in J_c} v_j(\omega_j)^{\phi_m} \ \text{ and } \ \phi_m \equiv \tilde{\phi}\phi. \tag{A.8}$$

Combining eqs. (14) and (A.8), we obtain the following equilibrium equation

$$\frac{v_i(\omega_i)^{\phi_m}}{\sum_{j \in J_c} v_j(\omega_j)^{\phi_m}}M_c v_i(\omega_i)^{\phi} = \sum_s (\omega_i)^{-\sigma_s}\left(A_{is}P_s\right)^{\sigma_s - 1}\gamma_s W, \tag{A.9}$$

or, equivalently,

$$(\Phi_c)^{-1} M_c v_i (\omega_i)^{\phi+\phi_m} = \sum_s (\omega_i)^{-\sigma_s} (A_{is} P_s)^{\sigma_s - 1} \gamma_s W, \tag{A.10}$$

for every region i in every country c. As in the model described in Section 2, $W \equiv \sum_{i \in J} W_i$ and $W_i \equiv \sum W_i$ and we impose the normalization $W = 1$. Conditional on this normalization, the system of J labor market clearing conditions in eq. (A.9) and the system of S price indices in eqs. (8) and (11) jointly determine the set of equilibrium wages $\{\omega_i\}_{i \in J}$ as a function of the technology levels in all regions $\{A_{is}\}_{s=1, i \in J}^{S}$, sectoral preferences $\{\gamma_s\}_{s=1}^{S}$, labor supply shifters $\{v_i\}_{i \in J}$, and the parameters $\{\sigma_s\}_{s=1}^{S}$, ϕ, and ϕ_m.

Labor Market Impact of Economic Shocks in a Small Open Economy. Assuming that $\{M_c\}_{c=1}^{C}$, $\{\sigma_s\}_{s=1}^{S}$, and (ϕ, ϕ_m) are fixed and totally differentiating eq. (A.9) with respect to the remaining determinants of $\hat{\omega}_i$, we can express the changes in wages in every region i of the small economy c as

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

where $\tilde{\lambda}_i \equiv (\phi + \phi_m + \sum_s l_{is}^0 \sigma_s)^{-1}$ and l_{is}^0 is the share of workers living in region i that were employed in sector s at the initial period 0; i.e. $l_{is}^0 \equiv L_{is}^0 / L_i^0$. Eliminating sectoral differences in the constant elasticity of substitution in demand in eq. (5), $\sigma_s = \sigma \ \forall s$, the expression for the change in wages simplifies to:

$$\hat{\omega}_i = \tilde{\lambda} \hat{\Phi}_c + \tilde{\lambda} \sum_{s=1}^{S} l_{is}^0 \left[\hat{\gamma}_s + (\sigma - 1)(\hat{A}_{is} + \hat{P}_s) \right] - \tilde{\lambda} \hat{v}_i, \tag{A.11}$$

where $\tilde{\lambda} \equiv (\phi + \phi_m + \sum_s l_{is}^0 \sigma)^{-1}$ and

$$\begin{aligned} \hat{\Phi}_c &= \sum_{i \in J_c} m_i^0 \left(\phi_m \hat{\omega}_i + \hat{v}_i \right), \\ &= \tilde{\lambda} \phi_m \hat{\Phi}_c + \tilde{\lambda} \phi_m \sum_{i \in J_c} m_i^0 \sum_{s=1}^{S} l_{is}^0 \left[\hat{\gamma}_s + (\sigma - 1)(\hat{A}_{is} + \hat{P}_s) \right] - \tilde{\lambda} \phi_m \sum_{i \in J_c} m_i^0 \hat{v}_i, \\ &= \frac{\tilde{\lambda} \phi_m}{1 - \tilde{\lambda} \phi_m} \sum_{i \in J_c} m_i^0 \sum_{s=1}^{S} l_{is}^0 \left[\hat{\gamma}_s + (\sigma - 1)(\hat{A}_{is} + \hat{P}_s) \right] - \frac{\tilde{\lambda} \phi_m}{1 - \tilde{\lambda} \phi_m} \sum_{i \in J_c} m_i^0 \hat{v}_i, \end{aligned} \tag{A.12}$$

where m_i^0 is the share of individuals living in country c that had residence in region i at the initial period 0; i.e. $m_i^0 \equiv M_i^0 / M_c^0$, with $M_c^0 \equiv \sum_{i \in J_c} M_i^0$. As in the main text, let's consider the case in which all regions i in country c are small open economies. The variable $\hat{\Phi}_c$ thus depends on technology shocks and labor supply shocks in all regions i of country c, $\{(\hat{A}_{is}, \hat{v}_i)\}_{s=1, i \in J_c}^{S}$, and on the aggregate sectoral preference shocks and sectoral prices changes, $\{(\hat{\gamma}_s, \hat{P}_s)\}_{s=1}^{S}$. Furthermore, the sectoral price changes themselves depend on the aggregate sectoral preference shocks and on the technology shocks and labor supply shocks in every other region of the world, $\{(\hat{A}_{is}, \hat{v}_i)\}_{s=1, i \in J / J_c}^{S}$.

Once we impose the assumption that $\sigma_s = \sigma \ \forall s$, the expression in eq. (17) is identical to that in eq. (A.11) under the restrictions that $\hat{\Phi}_c = 0$ and $\tilde{\phi} = 0$. Specifically, focusing again on the impact of sectoral price shocks $\hat{P}_s$, note that

$$d\hat{\omega}_i/d\hat{P}_s = \tilde{\lambda}(d\hat{\Phi}_c/d\hat{P}_s) + l_{is}^0\tilde{\lambda}(\sigma-1), \qquad s=1,\ldots,S,\ i=1,\ldots,J_c, \tag{A.13}$$

and $\tilde{\lambda} = \lambda$ if $\tilde{\phi} = 0$. Therefore, the second term in eq. (A.13) has a shift-share structure, being the product of an observed weight l_{is}^0 and an unobserved parameter $\tilde{\lambda}(\sigma-1)$. However, even if one could consistently estimate the parameter $\tilde{\lambda}(\sigma-1)$ (e.g. using eq. (A.11) as estimating equation and controlling for $\hat{\Phi}_c$ through a country c dummy variable) the expression $l_{is}^0\tilde{\lambda}(\sigma-1)$ will *not* capture the impact of $\hat{P}_s$ on $\hat{\omega}_i$. It will only capture some partial impact that does not take into account the effect that $\hat{P}_s$ has on $\hat{\omega}_i$ through the term $\hat{\Phi}_c$.

To recover the total effect of $\hat{P}_s$ on $\hat{\omega}_i$ one would need to substitute for $\hat{\Phi}_c$ in eq. (A.11) using the expression in eq. (A.12). Interestingly, once we rewrite eq. (A.12) as

$$\hat{\Phi}_c = \frac{\tilde{\lambda}}{1-\tilde{\lambda}}\sum_{i\in J_c}\sum_{s=1}^{S}(m_i^0 l_{is}^0)\left[\hat{\gamma}_s + (\sigma-1)(\hat{A}_{is}+\hat{P}_s)\right] - \frac{\tilde{\lambda}}{1-\tilde{\lambda}}\sum_{i\in J_c} m_i^0\hat{v}_i,$$

one can observe that the term that depends on $\hat{P}_s$ in this equation also has a shift-share structure, where the "share" corresponding to the shifter $\hat{P}_s$ equals $\sum_{i\in J_c}(m_i^0 l_{is}^0)$. More precisely, from eqs. (A.11) and (A.12), one can rewrite the relationship between $\hat{\omega}_i$ and the vector of sectoral price shocks $\{\hat{P}_s\}_{s=1}^S$ as

$$\hat{\omega}_i = \tilde{\lambda}\frac{\tilde{\lambda}}{1-\tilde{\lambda}}\sum_{s=1}^{S}\tilde{l}_s^0(\sigma-1)\hat{P}_s + \tilde{\lambda}\sum_{s=1}^{S} l_{is}^0(\sigma-1)\hat{P}_s + \hat{\tilde{\varepsilon}}_i, \tag{A.14}$$

where $\tilde{l}_s^0 \equiv \sum_{i\in J_c}(m_i^0 l_{is}^0)$ and $\hat{\tilde{\varepsilon}}_i$ accounts for all remaining terms impacting $\hat{\omega}_i$ in eqs. (A.11) and (A.12). As the expression in eq. (A.14) shows, we can rewrite the effect of $\hat{P}_s$ on $\hat{\omega}_i$ in eq. (A.13) as

$$d\hat{\omega}_i/d\hat{P}_s = \tilde{l}_{is}^0\tilde{\beta}_{\omega,is} + l_{is}^0\beta_{\omega,is}, \qquad \text{with} \quad \tilde{\beta}_{\omega,is} \equiv \tilde{\lambda}\frac{\tilde{\lambda}}{1-\tilde{\lambda}}(\sigma-1) \quad \text{and} \quad \beta_{\omega,is} \equiv \tilde{\lambda}(\sigma-1), \tag{A.15}$$

for every sector $s = 1,\ldots,S$ and every region $i = 1,\ldots,J_c$.

Analogously, we can also construct an expression for the change in population in a location i as

$$\hat{M}_i = -\hat{\Phi}_c + \hat{v}_i + \phi_m\hat{\omega}_i, \tag{A.16}$$

and an expression for the change in employment in a location i as

$$\hat{L}_i = \hat{v}_i + \phi\hat{\omega}_i + \hat{M}_i, \tag{A.17}$$

where $\hat{\omega}_i$ is defined in eqs. (A.11) and (A.12). Therefore, estimating the impact of $\hat{P}_s$ on either $\hat{M}_i$ or $\hat{L}_i$ also requires accounting for the impact that it has on the country-c specific term $\hat{\Phi}_c$. If, on the contrary, a researcher builds an estimating equation for $\hat{M}_i$ using as basis the equilibrium condition in eq. (A.16) and controls for the term $\hat{\Phi}_c$ using a country-c specific dummy, then the resulting estimate will at best capture only a partial effect of $\hat{P}_s$ on $\hat{M}_i$. The same is true if the effect of $\hat{P}_s$ on $\hat{L}_i$ is estimated using the equilibrium condition in (A.17) as estimating equation and controlling for $\hat{\Phi}_c$ through a country-c fixed effect.

A.4 Sector-specific factors of production

We extend here the model described in Section 2.1 to incorporate other factors of production. In particular, we introduce a specific-factor in each sector, as in the seminal paper by Jones (1971). More recently, Kovak (2013) uses a specific-factors model to derives a shift-share specification.

Environment. The only difference with respect to the setting described in Section 2.1 is that the production function in eq. (3) is substituted for a Cobb-Douglas production function that combines labor and capital inputs:

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

We additionally 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 $s = 1, \ldots, S$, each region has an endowment of sector-specific capital $\bar{K}_{is}$.

Equilibrium. 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

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

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

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

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

$$L_{is} = \frac{1-\theta_{is}}{\theta_{is}} \bar{K}_{is} \left(\frac{p_{is} A_{is} \tilde{\theta}_{is}}{\omega_i}\right)^{\frac{1}{\theta_{is}}}, \tag{A.18}$$

and the total sales of the sector-s region-i good are

$$X_{is} = \frac{1}{1-\theta_{is}} \omega_i L_{is} = \frac{\bar{K}_{is}}{\theta_{is}} \left(p_{is} A_{is} \tilde{\theta}_{is}\right)^{\frac{1}{\theta_{is}}} \left(\omega_i\right)^{1-\frac{1}{\theta_{is}}}. \tag{A.19}$$

Given the normalization that sets world income to one, $W = 1$, the total expenditure in the sector-s region-i good is equal to $x_{is}\gamma_s$, with x_{is} defined as a function of the equilibrium prices p_{is} and P_s in eq. (9). Equating eqs. (9) and (A.19), we can solve for the equilibrium value of p_{is} as a function of the sector-s price index P_s:

$$p_{is} = \left[\frac{\bar{K}_{is}}{\theta_{is}} \left(A_{is}\tilde{\theta}_{is}\right)^{\frac{1}{\theta_{is}}} \left(\omega_i\right)^{1-\frac{1}{\theta_{is}}} \frac{P_s^{1-\sigma_s}}{\gamma_s} \right]^{-\theta_{is}\eta_{is}} \tag{A.20}$$

where $\eta_{is} \equiv (1+\theta_{is}(\sigma_s-1))^{-1} \in (0,1)$. Furthermore, combining the expression for p_{is} in eq. (A.20) and that for P_s in eq. (8), we can solve for the equilibrium value of the sectoral price P_s as a function of $\{\omega_i\}_{i\in J}$ and exogenous parameters.

Combining the expression in eq. (A.20) with eq. (A.18) we obtain an expression for labor demand in sector s of region i as a function of equilibrium wages ω_i, the sector-s price P_s and other exogenous determinants:

$$L_{is} = \kappa_{is}\gamma_s^{\eta_{is}} \left(A_{is}P_s\right)^{(\sigma_s-1)\eta_{is}} \left(\omega_i\right)^{-\sigma_s\eta_{is}}$$

where $\kappa_{is} \equiv (1-\theta_{is})(\bar{K}_{is}\tilde{\theta}_{is}^{\frac{1}{\theta_{is}}}/\theta_{is})^{1-\eta_{is}}$. Adding across sectors, we obtain aggregate labor demand in region i:

$$L_i = \sum_{s=1}^{S} \kappa_{is}\gamma_s^{\eta_{is}} \left(A_{is}P_s\right)^{(\sigma_s-1)\eta_{is}} \left(\omega_i\right)^{-\sigma_s\eta_{is}}$$

Finally, equalizing labor demand and labor supply in every region, we obtain the following J labor market clearing conditions:

$$M_i v_i(\omega_i)^{\phi} = \sum_{s=1}^{S} \kappa_{is}\gamma_s^{\eta_{is}} \left(A_{is}P_s\right)^{(\sigma_s-1)\eta_{is}} \left(\omega_i\right)^{-\sigma_s\eta_{is}}, \qquad j=1,\ldots,J. \tag{A.21}$$

Given these J equations and an expression for every sectoral price index P_s as function of the complete vector of wages in every region of the world, we can solve for these equilibrium wage levels $\{\omega_i\}_{i\in J}$.

Labor Market Impact of Economic Shocks in a Small Open Economy. Assuming that $\{M_i\}_{i\in J}$, $\{\sigma_s\}_{s=1}^{S}$, ϕ and $\{(\bar{K}_{is},\theta_{is})\}_{s=1,i\in J}^{S}$ are fixed and totally differentiating eq. (A.21) with respect to the remaining determinants of ω_i, we can express the changes in wages in every region i of the small open economy c as

$$\hat{\omega}_i = \tilde{\lambda}_i \sum_s l_{is}^0 \left[\eta_{is}\hat{\gamma}_s + (\sigma_s-1)\eta_{is}\left(\hat{A}_{is}+\hat{P}_s\right)\right] + \tilde{\lambda}_i\hat{v}_i, \tag{A.22}$$

where $\tilde{\lambda}_i \equiv \left(\phi + \sum_s l_{is}^0\sigma_s\eta_{is}\right)^{-1}$. Finally, plugging this expression into an equation relating changes in employment in region i to changes in wages in region i, $\hat{L}_i = \hat{v}_i + \phi\hat{\omega}_i$, we can express the changes in employment in every region i of the small open economy c as

$$\hat{L}_i = \phi\tilde{\lambda}_i \sum_s l_{is}^0 \left[\eta_{is}\hat{\gamma}_s + (\sigma_s-1)\eta_{is}\left(\hat{A}_{is}+\hat{P}_s\right)\right] + (1-\phi\lambda_i)\hat{v}_i. \tag{A.23}$$

Given eqs. (A.22) and (A.23) and focusing again on price shocks $\hat{P}_s$ for illustrative purposes, note that we can write, for every region i and sector s,

$$d\hat{L}_i/d\hat{P}_s = l_{is}^0 \beta_{L,is}, \qquad \text{with} \quad \beta_{L,is} \equiv \phi\tilde{\lambda}_i \eta_{is}(\sigma_s - 1); \tag{A.24}$$

$$d\hat{\omega}_i/d\hat{P}_s = l_{is}^0 \beta_{\omega,is}, \qquad \text{with} \quad \beta_{\omega,is} \equiv \tilde{\lambda}_i \eta_{is}(\sigma_s - 1). \tag{A.25}$$

Conditional on the values of the reduced-form parameters $\beta_{L,is}$ and $\beta_{\omega,is}$, these two expressions are identical to those in eqs. (19) and (20). Furthermore, the reduced-form parameters entering eqs. (A.24) and (A.25) will be identical to those entering eqs. (19) and (20) if and only if $\theta_{is} = 0$ for every i and s.

A.5 Sector-specific preferences

We extend here the model described in Section 2.1 to allow workers to have idiosyncratic preferences for being employed in the different $s = 1, \ldots, S$ sectors and for being non-employed $s = 0$.

Environment. The only difference with respect to the setting described in Section 2.1 is that the utility function in eqs. (6) and (7) 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_j from the consumption of goods, the utility of a worker ι living in region j is

$$U_{is} = u_s(\iota) C_i, \tag{A.26}$$

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

$$F_u(u) = e^{-v_{is}u^{-\phi}}, \qquad \phi > 1. \tag{A.27}$$

A similar modeling of workers' sorting patterns across sectors has been introduced in Galle, Rodríguez-Clare and Yi (2017) and Burstein, Morales and Vogel (2018a). See Adão (2016) for a framework that relaxes the distributional assumption in eq. (A.27). 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, which are financed as indicated in Section 2.1.

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

$$L_{is} = M_i \frac{v_{is}(\omega_{is})^\phi}{\Phi_i} \quad \text{with} \quad \Phi_i \equiv v_{i0} b_i^\phi + \sum_{s=1}^S v_{is}(\omega_{is})^\phi, \tag{A.28}$$

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

$$L_{i0} = M_i \frac{v_{i0}(b_i)^{\phi}}{\Phi_i}. \tag{A.29}$$

By combining the expression in eq. (A.28) and the labor demand in eq. (12), and imposing the normalization $W = 1$, the labor market clearing condition in every sector $s = 1, \ldots, S$ and region $i = 1, \ldots, J$ is

$$M_i \frac{v_{is}(\omega_{is})^{\phi}}{\Phi_i} = (\omega_{is})^{-\sigma_s} \left(A_{is} P_s\right)^{\sigma_s - 1} \gamma_s. \tag{A.30}$$

The system of $J \times S$ equations formed by this expression for every every sector $s = 1, \ldots, S$ and region $i = 1, \ldots, J$ yields the equilibrium wages $\{\omega_{is}\}_{s=1, j\in J}^{S}$.

In equilibrium, Burstein, Morales and Vogel (2018a) show that the average wage in every sector s of the region i is

$$\omega_i = \gamma \Phi_i^{\frac{1}{\phi}}, \tag{A.31}$$

where $\gamma = \Gamma\left(1 - 1/\phi\right)$ with $\Gamma(.)$ denoting the Gamma function.

Labor Market Impact of Economic Shocks in a Small Open Economy. Assuming that $\{M_i\}_{i\in J}$, $\{\sigma_s\}_{s=1}^{S}$, and ϕ are fixed and totally differentiating eq. (A.30) with respect to the remaining determinants of ω_{is}, we can express the changes in wages in every sector s and every region i of the small open economy c as

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

From the definition of Φ_i in eq. (A.28), we can derive the expression

$$\hat{\Phi}_i = \sum_{s=0}^{S} l_{is}^0 \hat{v}_{is} + \phi l_{i0}^0 \hat{b}_i + \phi \sum_{s=1}^{S} l_{is}^0 \hat{\omega}_{is},$$

and, plugging in the expression for $\hat{\omega}_{is}$ and grouping terms, we obtain

$$\hat{\Phi}_i = \hat{v}_i + \bar{\lambda}_i \phi \sum_{s=1}^{S} l_{is}^0 (\phi + \sigma_s)^{-1} \left[\hat{\gamma}_s + (\sigma_s - 1)(\hat{A}_{is} + \hat{P}_s)\right]$$

where $\bar{\lambda}_i \equiv (1 - \phi \sum_{s=1}^{S} l_{is}^0 (\phi + \sigma_s)^{-1})^{-1}$, and $\hat{v}_i \equiv \bar{\lambda}_i \phi l_{i0}^0 \hat{b}_i + \bar{\lambda}_i l_{i0}^0 \hat{v}_{i0} + \bar{\lambda}_i \sum_{s=1}^{S} l_{is}^0 \sigma_s (\phi + \sigma_s)^{-1} \hat{v}_{is}$. Thus, plugging this expression back in the expression for $\hat{\omega}_{is}$, we obtain that

$$\hat{\omega}_{is} = (\phi + \sigma_s)^{-1} \Big[\hat{v}_i - \hat{v}_{is} + \hat{\gamma}_s + (\sigma_s - 1)(\hat{A}_{is} + \hat{P}_s) + \bar{\lambda}_i \phi \sum_{k=1}^{S} l_{ik}^0 (\phi + \sigma_k)^{-1} \left[\hat{\gamma}_k + (\sigma_k - 1)(\hat{A}_{ik} + \hat{P}_k)\right]\Big].$$

We can similarly compute an expression for the change in total employment in region i as

$$\hat{L}_i = -\frac{l_{i0}^0}{1 - l_{i0}^0} \hat{l}_{i0} = \frac{l_{i0}^0}{1 - l_{i0}^0} (\hat{\Phi}_i - \phi \hat{b}_i)$$

$$= \hat{v}_i + \tilde{\lambda}_i \phi \sum_{s=1}^{S} l_{is}^0 (\phi + \sigma_s)^{-1} (\hat{\gamma}_s + (\sigma_s - 1)(\hat{A}_{is} + \hat{P}_s)) \tag{A.32}$$

where $\hat{v}_i = l_{i0}^0 (1 - l_{i0}^0)^{-1} (\hat{v}_i - \phi \hat{b}_i)$, and $\tilde{\lambda}_i \equiv l_{i0}^0 (1 - l_{i0}^0)^{-1} \bar{\lambda}_i$.

In addition, the change in the average in region i is

$$\hat{\omega}_i = \frac{1}{\phi} \hat{v}_i + \bar{\lambda}_i \sum_{s=1}^{S} l_{is}^0 (\phi + \sigma_s)^{-1} \left[\hat{\gamma}_s + (\sigma_s - 1)(\hat{A}_{is} + \hat{P}_s) \right] \tag{A.33}$$

Given eq. (A.32) and eq. (A.33), the effect of price shocks $\hat{P}_s$ is

$$d\hat{L}_i / d\hat{P}_s = l_{is}^0 \beta_{L,is}, \qquad \text{with} \quad \beta_{L,is} \equiv \phi \tilde{\lambda}_i (\phi + \sigma_s)^{-1} (\sigma_s - 1) \tag{A.34}$$

$$d\hat{\omega}_i / d\hat{P}_s = l_{is}^0 \beta_{\omega,is}, \qquad \text{with} \quad \beta_{\omega,is} \equiv \tilde{\lambda}_i (\phi + \sigma_s)^{-1} (\sigma_s - 1). \tag{A.35}$$

Conditional on the value of the reduced-form parameters $\beta_{L,is}$ and $\beta_{\omega,is}$, these expressions are identical to those in equations eqs. (19) and (20).

B Proofs and additional details for Section 4

Since Propositions 1 and 2 are special cases or Propositions 3 and 4, we only prove Propositions 3, 4 and 5. Before proving these results in Appendices B.2, B.2 and B.4, we collect some auxiliary Lemmata used in the proofs in Appendix B.1. Finally, Appendix B.5 discusses inference when the effects β_{is} are heterogeneous. Throughout this appendix, we use the following notation. 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 $A_S \leq C B_S$. We denote the σ-field generated by $(Y(0), B, W, U, \mathcal{Z})$ by $\mathcal{F}_0 = \sigma(Y(0), B, W, U, \mathcal{Z})$. Define $\bar{w}_{st} = \sum_{i=1}^{N} w_{is} w_{it}$, $\tilde{\mathcal{X}}_s = \mathcal{X}_s - \mathcal{Z}_s' \gamma$, and $\sigma_s^2 = \text{var}(\mathcal{X}_s \mid \mathcal{F}_0) = \text{var}(\mathcal{X}_s \mid \mathcal{Z}, W)$.

B.1 Auxiliary results

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

Proof. By Hölder's inequality,

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

which yields the result. □

Lemma 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 bounded uniformly over S and i. Then $N^{-2}\sum_{s=1}^{S} E\left[(\sum_{i=1}^{N} w_{is}A_{Si})^2 \mid W\right] \to 0$, provided Assumption 2(ii) holds.*

Proof. By Cauchy-Schwarz inequality,

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

The result follows from the fact that $N^{-2}\sum_{s=1}^{S} n_s^2 \leq \max_s n_s/N$, which converges to zero by Assumption 2(ii). □

Lemma 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 such that $E[A_{Si}^4 \mid W]$, $E[B_{Si}^4 \mid W]$, and $E[\mathcal{A}_{Ss}^2 \mid W]$ exist and are bounded uniformly over S, i and s. Then $(\sum_s n_s^2)^{-1}\sum_{i,j,s} w_{is}w_{js}A_{Si}B_{Sj}\mathcal{A}_{Ss} = O_P(1)$.*

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

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

The result then follows by Markov inequality and the dominated convergence theorem. □

B.2 Proof of Proposition 3

Let E_W denote expectation conditional on W. We first show that

$$\frac{1}{N}X'Z = \frac{1}{N}\sum_{i,s} w_{is}\mathcal{Z}_s'\gamma Z_i + o_P(1) \tag{B.1}$$
$$\frac{1}{N}X'X = \frac{1}{N}\sum_{s}\sigma_s^2\overline{w}_{ss} + \frac{1}{N}\sum_{s,t}\mathcal{Z}_s'\gamma\mathcal{Z}_t'\gamma\overline{w}_{st} + o_P(1) \tag{B.2}$$
$$\frac{1}{N}Z'Y = \frac{1}{N}\sum_{i} Z_iY_i(0) + \frac{1}{N}\sum_{i,t} Z_iw_{it}\mathcal{Z}_t'\gamma\beta_{it} + o_P(1) \tag{B.3}$$
$$\frac{1}{N}X'Y = \frac{1}{N}\sum_{i,s,t} w_{is}w_{it}(\mathcal{Z}_s'\gamma)(\mathcal{Z}_t'\gamma)\beta_{it} + \frac{1}{N}\sum_{i,s} w_{is}^2\sigma_s^2\beta_{is} + \frac{1}{N}\sum_{i,s} w_{is}(\mathcal{Z}_s'\gamma)Y_i(0) + o_P(1). \tag{B.4}$$

Consider (B.1). We have

$$\frac{1}{N}X'Z = \frac{1}{N}\sum_{s}\mathcal{X}_s\sum_{i} w_{is}Z_i = \frac{1}{N}\sum_{s}\tilde{\mathcal{X}}_s\sum_{i} w_{is}Z_i + \frac{1}{N}\sum_{i,s} w_{is}\mathcal{Z}_s'\gamma Z_i.$$

It therefore suffices to show that

$$\frac{1}{N}\sum_{s}\tilde{\mathcal{X}}_s\sum_{i}w_{is}Z_i = o_P(1). \tag{B.5}$$

The left-hand side has mean zero conditional on W, with the variance of the kth row given by

$$\operatorname{var}\left(\frac{1}{N}\sum_{i,s}w_{is}\tilde{\mathcal{X}}_s Z_{ik} \mid W\right) = \frac{1}{N^2}\sum_{s}E_W\sigma_s^2\left(\sum_{i}w_{is}Z_{ik}\right)^2 \preceq \frac{1}{N^2}\sum_{s}E_W\left(\sum_{i}w_{is}Z_{ik}\right)^2.$$

By Lemma 1, Assumption 4(iv), and the C_r-inequality, $E_W[Z_{ik}^2] = E_W[(\sum_s w_{is}\mathcal{Z}_{sk}+U_{ik})^2]$ is bounded, so that by Lemma 2, the right-hand side converges to zero. Equation (B.5) then follows by Markov inequality and the dominated convergence theorem.

Next, consider eq. (B.2). We have

$$\frac{1}{N}X'X = \frac{1}{N}\sum_{i,s,t}\mathcal{X}_s\mathcal{X}_t w_{is}w_{it} = \frac{2}{N}\sum_{s<t}\tilde{\mathcal{X}}_s\tilde{\mathcal{X}}_t\overline{w}_{st} + \frac{1}{N}\sum_{i,s}(\mathcal{X}_s^2 - E[\mathcal{X}_s^2 \mid \mathcal{Z}_s, W])w_{is}^2 \\ + \frac{2}{N}\sum_{s\neq t}\mathcal{Z}_s'\gamma\tilde{\mathcal{X}}_t\overline{w}_{st} + \frac{1}{N}\sum_{s}\sigma_s^2\overline{w}_{ss} + \frac{1}{N}\sum_{s,t}\mathcal{Z}_s'\gamma\mathcal{Z}_t'\gamma\overline{w}_{st}. \tag{B.6}$$

We will show that the first three summands are of the order $o_P(1)$. All three summands are mean zero since they are mean zero conditional on $\mathcal{F}_0$, so by Markov inequality and the dominated convergence theorem, it suffices to show that their variances, conditional on W, converge to zero. To that end,

$$\operatorname{var}\left(\frac{2}{N}\sum_{s<t}\tilde{\mathcal{X}}_s\tilde{\mathcal{X}}_t\overline{w}_{st} \mid W\right) = \frac{4}{N^2}\sum_{s<t}E_W[\sigma_s^2\sigma_t^2]\overline{w}_{st}^2 \preceq \frac{1}{N^2}\sum_{s,t}\overline{w}_{st}^2 \\ \leq \frac{1}{N^2}\sum_{i,j,s}w_{is}w_{js} = \frac{1}{N^2}\sum_{s}n_s^2 \to 0. \tag{B.7}$$

where the last inequality follows from $\sum_s w_{is}w_{js} \leq \sum_s w_{is} = 1$, and the convergence to 0 follows by Assumption 2(ii). The variance of the second summand can be bounded by

$$\operatorname{var}\left(\frac{1}{N}\sum_{i,s}(\mathcal{X}_s^2 - E[\mathcal{X}_s^2 \mid \mathcal{Z}_s, W])w_{is}^2 \mid W\right) \preceq \frac{1}{N^2}\sum_{s}\left(\sum_{i}w_{is}^2\right)^2 \leq \frac{1}{N^2}\sum_{s}n_s^2,$$

which converges to zero by Assumption 2(ii). Finally, variance of the third summand in eq. (B.6) can be bounded by

$$\operatorname{var}\left(\frac{2}{N}\sum_{i,s\neq t}\mathcal{Z}_s'\gamma\tilde{\mathcal{X}}_t w_{is}w_{it} \mid W\right) \leq \frac{4}{N^2}\sum_{t}E_W\sigma_t^2\left(\sum_{s,i}|\mathcal{Z}_s'\gamma|w_{is}w_{it}\right)^2 \\ \preceq \frac{1}{N^2}\sum_{s}E_W\left(\sum_{i}w_{is}\sum_{t}w_{it}|\mathcal{Z}_t'\gamma|\right)^2.$$

By Lemma 1, the second moment of $\sum_t w_{it}|\mathcal{Z}_t'\gamma|$ is bounded, so by Lemma 2, the right-hand side converges to zero.

Next, consider eq. (B.3). We can decompose

$$\frac{1}{N}Z'Y = \frac{1}{N}\sum_{i,s} Z_i w_{is}\tilde{\mathcal{X}}_s\beta_{is} + \frac{1}{N}\sum_i Z_i Y_i(0) + \frac{1}{N}\sum_{i,t} Z_i w_{it}\mathcal{Z}_t'\gamma\beta_{it}.$$

We will show that the first summand is $o_P(1)$. Since it has mean zero, by Markov inequality, it suffices to show that the variance of each row k conditional on W converges to zero. Now,

$$\operatorname{var}\left(\frac{1}{N}\sum_{i,t} Z_{ik}w_{it}\tilde{\mathcal{X}}_t\beta_{it} \mid W\right) = \frac{1}{N^2}\sum_s E_W\sigma_s^2\left(\sum_i Z_{ik}w_{is}\beta_{is}\right)^2$$
$$\preceq \frac{1}{N^2}\sum_s E_W\left(\sum_i w_{is}|Z_{ik}|\right)^2 \to 0,$$

where the convergence follows by Lemma 2, since as observed above, $E_W[|Z_{ik}|^2]$ is bounded. Finally, consider eq. (B.4). Decompose

$$\begin{aligned}\frac{1}{N}\sum_i X_iY_i = \frac{1}{N}\sum_s \tilde{\mathcal{X}}_s\sum_i w_{is}Y_i(0) + \frac{1}{N}\sum_{i,s<t} w_{is}w_{it}\tilde{\mathcal{X}}_s\tilde{\mathcal{X}}_t\beta_{it} \\ + \frac{1}{N}\sum_{i,s>t} w_{is}w_{it}\tilde{\mathcal{X}}_s\tilde{\mathcal{X}}_t\beta_{it} + \frac{1}{N}\sum_{s\neq t}(\mathcal{Z}_s'\gamma)\tilde{\mathcal{X}}_t\sum_i w_{is}w_{it}\beta_{it} \\ + \frac{1}{N}\sum_{s\neq t}\tilde{\mathcal{X}}_s(\mathcal{Z}_t'\gamma)\sum_i w_{is}w_{it}\beta_{it} + \frac{1}{N}\sum_{i,s} w_{is}^2(\mathcal{X}_s^2 - E[\mathcal{X}_s^2 \mid \mathcal{Z}_s, W])\beta_{is} \\ + \frac{1}{N}\sum_{i,s,t} w_{is}w_{it}(\mathcal{Z}_s'\gamma)(\mathcal{Z}_t'\gamma)\beta_{it} + \frac{1}{N}\sum_{i,s} w_{is}^2\sigma_s^2\beta_{is} + \frac{1}{N}\sum_{i,s} w_{is}(\mathcal{Z}_s'\gamma)Y_i(0).\end{aligned}$$

We will show that all summands except for the last three are $o_P(1)$. Since they are all mean zero conditional on $\mathcal{F}_0$, it suffices to show that their variances conditional on W converge to zero. The variance of the first summand is bounded by

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

by Lemma 2. The variance of the second summand is bounded by

$$\operatorname{var}\left(\frac{1}{N}\sum_{i,s<t} w_{is}w_{it}\tilde{\mathcal{X}}_s\tilde{\mathcal{X}}_t\beta_{it} \mid W\right) = \frac{1}{N^2}\sum_{s<t} E_W\sigma_s\sigma_t\left(\sum_i w_{is}w_{it}\beta_{it}\right)^2 \preceq \frac{1}{N^2}\sum_{s<t}\overline{w}_{st}^2 \to 0,$$

where the convergence to zero follows by arguments analogous to those in (B.7). The variance of the third summand converges to zero by analogous arguments. Variance of the fourth summand satisfies

$$\operatorname{var}\left(\frac{1}{N}\sum_{s\neq t}(\mathcal{Z}_s'\gamma)\tilde{\mathcal{X}}_t\sum_i w_{is}w_{it}\beta_{it}\mid W\right)\leq\frac{1}{N}\sum_s E_W\sigma_s^2\left(\sum_t|(\mathcal{Z}_t'\gamma)|\sum_i w_{is}w_{it}|\beta_{is}|\right)^2$$
$$\preceq\frac{1}{N}\sum_s E_W\left(\sum_i w_{is}\sum_t w_{it}|(\mathcal{Z}_t'\gamma)|\right)^2,$$

which converges to by Lemma 2, since by Lemma 1, the second moment of $\sum_t w_{it}|(\mathcal{Z}_t'\gamma)|$ is bounded. Variance of the fifth summand converges to zero by analogous arguments. Finally, variance of the sixth summand satisfies

$$\operatorname{var}\left(\frac{1}{N}\sum_{i,s}w_{is}^2(\mathcal{X}_s^2-E[\mathcal{X}_s^2\mid\mathcal{Z}_s,W])\beta_{is}\mid W\right)\preceq\frac{1}{N^2}\sum_s E_W\left(\sum_i w_{is}^2\beta_{is}\right)^2\preceq\frac{1}{N^2}\sum_s n_s^2\to 0,$$

which yields (B.4). We now use eqs. (B.1), (B.2), (B.3) and (B.4) to derive the result. Since $U_i'\gamma=0$, Equation (B.1) implies $Z'X/N=Z'Z\gamma/N+o_P(1)$. Consequently, since by Assumption 4(iii), $(Z'Z/N)^{-1}=o_P(1)$,

$$\frac{1}{N}\ddot{X}'\ddot{X}=\frac{1}{N}X'X-\frac{1}{N}X'Z(Z'Z)^{-1}Z'X=\frac{1}{N}\sum_s\sigma_s^2\overline{w}_{ss}+o_P(1)=\frac{1}{N}\sum_{i,s}\pi_{is}+o_P(1),\tag{B.8}$$

and, since $Z'Y/N=o_P(1)$,

$$\frac{1}{N}\ddot{X}'Y=\frac{1}{N}X'Y-\gamma'\frac{1}{N}Z'Y+o_P(1)=\frac{1}{N}\sum_{i,s}\pi_{is}\beta_{is}+o_P(1).$$

Combining Assumption 4(iii) with the preceding two displays then yields the result.

B.3 Proof of Proposition 4

Let $r_N=1/\sum_s n_s^2$, and let E_W denote expectation conditional on W. Note that $\gamma'U_i=0$ implies $Z\gamma=W\mathcal{Z}\gamma$. Therefore, $\ddot{X}$ admits the decomposition

$$\ddot{X}=(I-Z(Z'Z)^{-1}Z')X=(I-Z(Z'Z)^{-1}Z')(X-Z\gamma)=(I-Z(Z'Z)^{-1}Z')W\tilde{\mathcal{X}}.$$

Using this decomposition, we obtain

$$\begin{aligned}r_N^{1/2}(\ddot{X}'\ddot{X})(\hat{\beta}-\beta)&=r_N^{1/2}\ddot{X}'(Y-X'\beta)=r_N^{1/2}\tilde{\mathcal{X}}'W'(Y-X\beta-Z\check{\delta})\\&=r_N^{1/2}\tilde{\mathcal{X}}'W'(Y-X\beta-Z\delta)-r_N^{1/2}\tilde{\mathcal{X}}'W'Z(\check{\delta}-\delta)\\&=r_N^{1/2}\sum_{s,i}\tilde{\mathcal{X}}_s w_{is}\epsilon_i-\frac{\tilde{\mathcal{X}}'W'Z}{N}(r_NN^2)^{1/2}(\check{\delta}-\delta)=r_N^{1/2}\sum_{s,i}\tilde{\mathcal{X}}_s w_{is}\epsilon_i+o_P(1).\end{aligned}$$

where the last line follows by Assumption 5(ii) and (B.5). It follows from eq. (B.8) and Assumption 4(iii) that $(\ddot{X}'\ddot{X}/N)^{-1} = (1 + o_P(1))(N^{-1}\sum_{i,s}\pi_{is})^{-1}$, so 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

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

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

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

Then we can write $\epsilon_i = V_i + \sum_t w_{it}\tilde{\mathcal{X}}_t(\beta_{it} - \beta)$, and, using the fact that $0 = \sum_{i,s}\pi_{is}(\beta_{is} - \beta) = \sum_s \sigma_s^2 b_{ss}$, we can decompose

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

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 $r_N^2 \sum_{s=1}^S E_W[\mathcal{Y}_s^4] \to 0$ 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

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

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

Note that since $\sum_s |\sum_t w_{it}\mathcal{Z}_t'\gamma(\beta_{it} - \beta)|^4 \preceq \sum_s |\sum_t w_{it}\mathcal{Z}_t'\gamma|^4$, it follows from Lemma 1, Assumption 3(ii), Assumption 5(i), 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 2, $\sum_s E_W[a_s^4] \preceq \sum_s n_s^4$, so that

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

by Assumption 3(i), since $\sum_s n_s^4 \leq \max_s n_s^2 / r_N$. Second, since β_{is} is bounded by Assumption 2(i), $b_{ss} \preceq \sum_i w_{is}^2 \leq n_s$, so that

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

Third, by similar arguments

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

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

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

we have

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

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 $\operatorname{var}(D_j \mid W) = E_W[\operatorname{var}(D_j \mid \mathcal{F}_0)] \to 0$. 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 3(i). 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 the fact that since $\sum_s \overline{w}_{st} = n_t$ and $\overline{w}_{st} \leq n_s$,

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

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

B.4 Proof of Proposition 5

Let $\hat{\theta} = (\hat{\beta}, \hat{\delta}')'$, $\theta = (\beta, \delta)$, $M_i = (X_i, Z_i')'$, $r_N = 1/\sum_{s=1}^{S} n_s^2$, and let

$$\hat{\mathcal{V}}_N = r_N \sum_s \widehat{\mathcal{X}}_s \hat{R}_s^2.$$

Since $\mathcal{V}_N = r_N \sum_s \sigma_s^2 R_s^2$, we can decompose this estimator as

$$\hat{\mathcal{V}}_N = r_N \sum_s (\widehat{\mathcal{X}}_s^2 - \tilde{\mathcal{X}}_s^2) \hat{R}_s^2 + r_N \sum_s \tilde{\mathcal{X}}_s^2 (\hat{R}_s^2 - R_s^2) + r_N \sum_s (\tilde{\mathcal{X}}_s^2 - \sigma_s^2) R_s^2 + \mathcal{V}_N. \tag{B.9}$$

We'll show that the first three terms are $o_P(1)$. Since $\hat{\epsilon}_i = \epsilon_i + M_i'(\theta - \hat{\theta})$, with $\epsilon_i = Y_i(0) - Z_i'\delta$, we can decompose

$$\hat{R}_s^2 = \sum_{i,j} w_{is} w_{js} \hat{\epsilon}_i \hat{\epsilon}_j = R_s^2 + 2 \sum_{i,j} w_{js} w_{is} M_i'(\theta - \hat{\theta}) \epsilon_j + \sum_{i,j} w_{is} w_{js} M_i'(\theta - \hat{\theta}) M_j'(\theta - \hat{\theta}). \tag{B.10}$$

Therefore, the second term in eq. (B.9) satisfies

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

where the second line follows from Lemma 3. Second, the variance of the third term in eq. (B.9) can be bounded by

$$\text{var}(r_N \sum_s (\mathcal{X}_s^2 - \sigma_s^2) R_s^2 \mid W) = r_N^2 \sum_s E[(\mathcal{X}_s^2 - \sigma_s^2)^2 R_s^4 \mid W] \preceq r_N^2 \sum_s E[R_s^4 \mid W] \preceq r_N^2 \sum_s n_s^4 \to 0$$

since $r_N^2 \sum_s n_s^4 \leq \max_s n_s^2 / \sum_t n_t^2 \to 0$ by Assumption 3(i). Since $E[r_N \sum_s (\mathcal{X}_s^2 - \sigma_s^2) R_s^2 \mid W] = E[r_N \sum_s E[(\mathcal{X}_s^2 - \sigma_s^2) \mid \mathcal{F}_0] R_s^2 \mid W] = 0$, it follows by Markov inequality and the dominated convergence theorem that $r_N \sum_s (\mathcal{X}_s^2 - \sigma_s^2) R_s^2 = o_P(1)$.

It remains to show that the first term in eq. (B.9) is $o_P(1)$. Let $\hat{\gamma} = (Z'Z)^{-1} Z'X$. Since $W\mathcal{X} = X$ and $Z = W\mathcal{Z} + U$, it follows that

$$\begin{aligned}
\widehat{\mathcal{X}} &= (W'W)^{-1} W' \ddot{X} = (W'W)^{-1} W'(X - Z(Z'Z)^{-1} Z'X) = \mathcal{X} - (W'W)^{-1} W'Z(Z'Z)^{-1} Z'X \\
&= \mathcal{X} - (W'W)^{-1} W'Z(\hat{\gamma} - \gamma) - (W'W)^{-1} W'Z\gamma \\
&= \tilde{\mathcal{X}} - (W'W)^{-1} W'Z(\hat{\gamma} - \gamma) \\
&= \tilde{\mathcal{X}} - \mathcal{Z}(\hat{\gamma} - \gamma) - (W'W)^{-1} W'U(\hat{\gamma} - \gamma).
\end{aligned}$$

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. Consequently,

$$\begin{aligned}
r_N \sum_s (\widehat{\mathcal{X}}_s^2 - \tilde{\mathcal{X}}_s^2) \hat{R}_s^2 &= (\hat{\gamma} - \gamma)' r_N \sum_s \mathcal{Z}_s \hat{R}_s^2 - (\hat{\gamma} - \gamma)' r_N \sum_s \mathcal{U}_s \hat{R}_s^2 \\
&= (\hat{\gamma} - \gamma)' O_P(1) - (\hat{\gamma} - \gamma)' O_P(1) \\
&= o_P(1),
\end{aligned}$$

where the second line follows by applying Lemma 3 after using the expansion in eq. (B.10), and the third line follows since by eq. (B.1) and Assumption 4(iii), $\hat{\gamma} = \gamma + o_P(1)$.

B.5 Inference under heterogeneous effects

For valid (but perhaps conservative) inference under heterogeneous effects, we need to ensure that that that when $\beta_{is} \neq \beta$, eq. (39) holds with inequality, that is,

$$\frac{\sum_{s=1}^{S} \widehat{\mathcal{X}}_s^2 \hat{R}_s^2}{\sum_{s=1}^{S} n_s^2} \geq \mathcal{V}_N + o_P(1). \tag{B.11}$$

To discuss conditions under which this is the case, observe that the "middle-sandwidch" in the asymptotic variance sandwich formula, $\mathcal{V}_N$, as defined in Proposition 4, can be decomposed into three terms:

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

where, as before $R_s = \sum_s w_{is}\epsilon_i$, and $\epsilon_i = Y_i(0) - Z_i'\delta + \sum_s \mathcal{X}_s w_{is}(\beta_{is} - \beta)$. Under homogeneous effects, R_s is non-random conditional on $\mathcal{F}_0$, and the second and third term are equal to zero, since in this case $E[\tilde{\mathcal{X}}_s R_s \mid \mathcal{F}_0] = E[\tilde{\mathcal{X}}_s \mid \mathcal{F}_0]R_s = 0$, and $E[\tilde{\mathcal{X}}_s R_s \tilde{\mathcal{X}}_t R_t \mid \mathcal{F}_0] = R_s R_t E[\tilde{\mathcal{X}}_s \tilde{\mathcal{X}}_t \mid \mathcal{F}_0] = 0$ if $s \neq t$. Therefore, only the first term remains, and the standard error estimator consistently estimates this term by Proposition 5.

It can be shown that the proposition remains valid under regularity conditions if the effects β_{is} are heterogeneous, so that to ensure valid inference under heterogeneous effects, one needs to ensure that the sum of the second and third term is weakly negative. This is the case under several different settings. We now discuss two of them.

First observe that since $E[\tilde{\mathcal{X}}_s R_s \mid \mathcal{F}_0] = E[\tilde{X}_s \sum_{t=1}^{S} \mathcal{X}_t w_{it}(\beta_{it} - \beta) \mid \mathcal{F}_0] = \sigma_s^2 w_{is}(\beta_{is} - \beta)$, the second term equals

$$-\frac{\sum_{s=1}^{S}\left(\sum_{i=1}^{N} \pi_{is}(\beta_{is} - \bar{\beta})\right)^2}{\sum_{s=1}^{S} n_s^2},$$

where $\pi_{is} = w_{is}^2\sigma_s^2$ as in the statement Proposition 3. The term is always negative, and it reflects the variability of the treatment effect. It makes the variance estimate that we propose conservative if the third term equals zero. This is analogous to the result that the robust standard error estimator is conservative in randomized trials, and that the cluster-robust standard error estimator is conservative in cluster-randomized trials (see, for example Imbens and Rubin, 2015, Chapter 6). The third term reflects the correlation between $\mathcal{X}_s R_s$ and $\mathcal{X}_t R_t$, and it has no analog in cluster-randomized trials. Indeed, the term can be written as

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

In the example with "concentrated sectors", which is the analog of the cluster-randomized setup if there are no covariates, the term is thus 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. Another sufficient condition for validity of inference is that β_{is} and β_{jt} are uncorrelated if $t \neq s$, in which case it follows from the display above that the third term converges to zero. Numerical work, not reported here, indicates that the correlation between β_{is} and β_{jt} needs to be quite high and depend on the weights w_{is} in order for the third term to dominate the second term. We therefore expect our inference to remain valid for empirically relevant distributions of the effects β_{is}.

C Placebo Exercise

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

In this appendix, we investigate the consequences of violations of Assumption 1(i) that requires observed sectoral shocks of interest $\mathcal{X}_1, \ldots, \mathcal{X}_S$ to be independent from other sectoral shocks affecting the outcome variable of interest. We study in this section the impact that violations of this assumption have on the properties of the OLS estimator of the coefficient on the shift-share regressor of interest. We also consider the properties of two solutions to this problem: (i) the inclusion of regional controls as a proxy for sector-level unobserved shocks (discussed in Section 4.2), and (ii) the use of a shift-share instrumental variable constructed as a weighted average of exogenous sector-level shocks (discussed in Section 4.3.2).

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

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

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

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

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

where Y_i^{obs} is the observed 2000–2007 change in the employment rate in CZ i, and 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. The parameter δ thus modulates the magnitude of the confounding effect of the unobserved shock $\mathcal{X}_s^b$. We explore the impact of δ by simulating data both with $\delta = 0$ and with $\delta = 6$.

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

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

The parameter σ_u thus modulates the measurement error in X_i^b as a proxy for the impact of the unobserved sectoral shocks $(\mathcal{X}_1^b, \ldots, \mathcal{X}_S^b)$ in CZ i. We explore the impact of σ_u by simulating data both with $\sigma_u = 0$ and with $\sigma_u = 6$.[34]

For each set of parameters (δ, σ_u) that we explore 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. (23). Second, we consider the OLS estimator in a regression in which we include X_i^b as a proxy for the vector of unobserved confounding sectoral shocks $(\mathcal{X}_1^b, \ldots, \mathcal{X}_S^b)$; i.e. the estimator in eq. (34). 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. (43). For each of these three estimators, we compute four estimates of its standard error: *Robust*, *St-cluster*, *AKM* and *AKM0*. All results are reported in Table C.1.

When there is no confounding sectoral shock ($\delta = 0$), Panel A shows that all three estimators yield an average coefficient close to zero. Panels B and C report results in the presence of confounding sectoral shocks ($\delta > 0$), in which case the OLS estimator in a simple regression of Y_i on X_i^a that does not include any additional covariates is positively biased ($\hat{\beta} = 4.23$). The introduction of the regional control only yields unbiased estimates when it is a good proxy for the underlying confounding sectoral shocks (i.e. if $\sigma_u = 0$). In contrast, the IV estimate always yields an average estimated coefficient of zero due to the orthogonality between the sector-level instrumental variable and the sector-level unobserved confounding shock.

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

[34]Using the notation in Section 4.2, the simulated variable $\mathcal{X}_s^a$ corresponds to $\mathcal{X}_s$, the simulated variable $\mathcal{X}_s^b$ is a column in the matrix $\mathcal{Z}$ (which also includes a column of ones), u_i corresponds to U_i, and X_i^b to Z_i. The value of the parameter γ in eq. (33) is thus equal to $\tilde{\rho}$.

Table C.1: Magnitude of standard errors. Confounding effects.

	Estimates		Median effective std. error			
	Average	Std. dev	Robust	St-cluster	AKM	AKM0
Panel A: No confouding effect ($\delta = 0$)						
OLS without controls	-0.02	1.27	0.48	0.59	1.21	1.40
OLS with regional control	-0.04	1.75	0.67	0.83	1.68	1.94
2SLS	-0.04	1.77	0.68	0.85	1.72	1.99
Panel B: Confouding effect ($\delta = 6$) and perfect regional control ($\sigma_u = 0$)						
OLS without controls	4.23	1.47	0.58	0.70	1.37	1.58
OLS with regional control	-0.05	1.75	0.67	0.82	1.67	1.93
2SLS	-0.05	1.79	0.68	0.84	1.72	1.98
Panel C: Confouding effect ($\delta = 6$) and imperfect regional control ($\sigma_u = 2$)						
OLS without controls	4.26	1.45	0.58	0.69	1.36	1.57
OLS with regional control	4.16	1.44	0.58	0.69	1.38	1.58
2SLS	-0.15	2.45	0.93	1.08	2.06	2.58

Notes: All estimates in this table use the total employment share in each CZ as the outcome variable Y_i. For the inference procedure indicated in the first row, "median effective std. error" refers to the median length of the 95% confidence interval across the 1000 simulated datasets divided by 2×1.96. *Robust* is the Eicker-Huber-White standard error; *St-cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in eq. (37); *AKM0* is the confidence interval Remark 6.

Table C.2: Rejection rate of $H_0\colon \beta = 0$ with significance level of 5%, Confounding Effects.

	Estimates		Rejection rate of $H_0\colon \beta = 0$ at 5%			
	Average	Std. Dev	Robust	St-cluster	AKM	AKM0
Panel A: No confouding effect ($\delta = 0$)						
OLS without controls	-0.02	1.27	49.1%	36.9%	7.6%	4.0%
OLS with regional control	-0.04	1.75	48.8%	38.8%	6.8%	3.1%
IV	-0.04	1.77	48.1%	39.3%	6.8%	3.4%
Panel B: Confouding effect ($\delta = 6$) and perfect regional control ($\sigma_u = 0$)						
OLS without controls	4.23	1.47	97.9%	96.8%	82.2%	74.7%
OLS with regional control	-0.05	1.75	48.3%	38.5%	7.3%	3.4%
IV	-0.05	1.79	46.7%	37.2%	7.1%	4.0%
Panel C: Confouding effect ($\delta = 6$) and imperfect regional control ($\sigma_u = 2$)						
OLS without controls	4.26	1.45	98.6%	97.6%	83.4%	76.0%
OLS with regional control	4.16	1.44	98.4%	96.9%	81.5%	74.0%
IV	-0.15	2.45	40.6%	33.3%	9.7%	5.7%

Notes: All estimates in this table use the total employment share in each CZ as the outcome variable Y_i. For the inference procedure indicated in the first row, this table indicates the percentage of the 1000 simulated datasets for which we reject the null hypothesis $H_0 : \beta = 0$ using a 5% significance level test. *Robust* is the Eicker-Huber-White standard error; *St-cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in eq. (37); *AKM0* is the test in Remark 6.

C.2 Additional Results

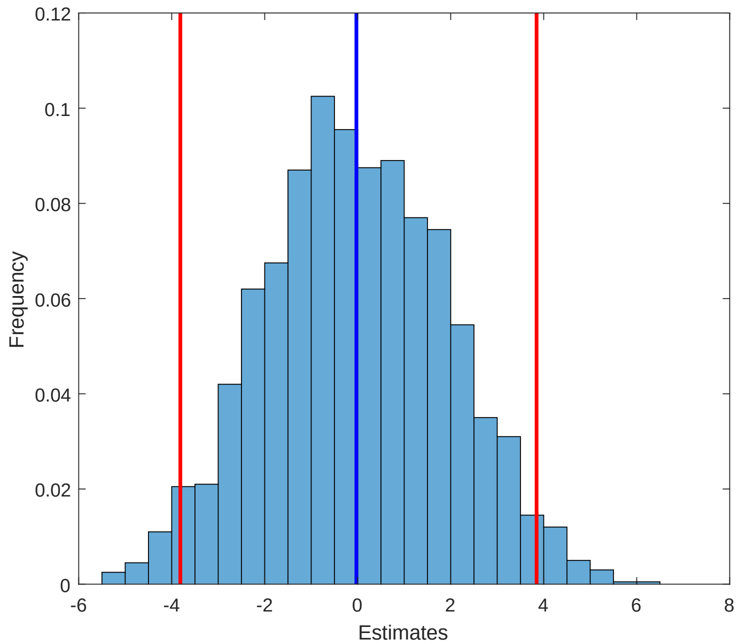

Figure C.1: Empirical distribution of estimated coefficients in the placebo exercise.

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 $m = 1, \ldots, 1000$ simulations we perform. The dependent variable Y is the 2000–2007 change in the employment rate.

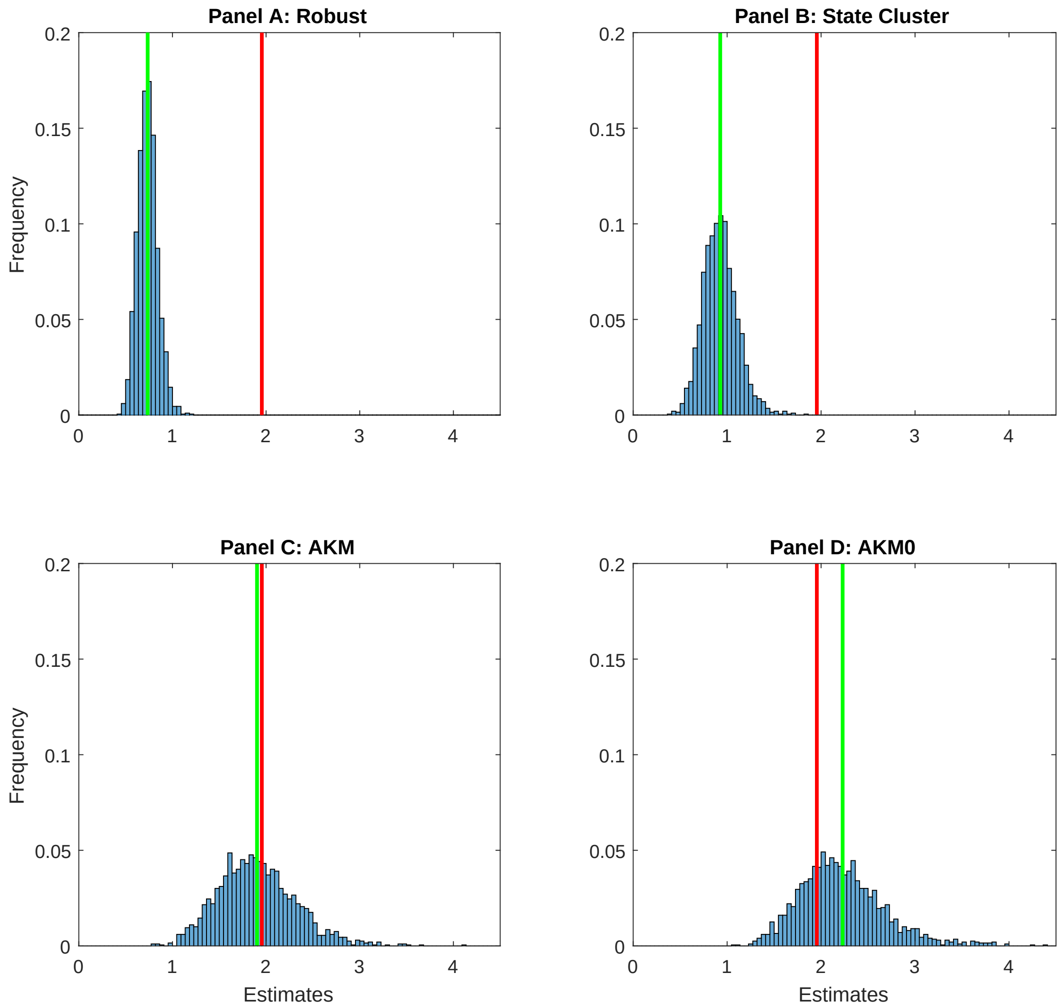

Figure C.2: Empirical distribution of effective standard errors for different standard error estimators in the placebo exercise.

Notes: In each of the four panels, the red line indicates the standard deviation of the empirical distribution of estimated coefficients represented in Figure C.1 (i.e. 1.95); the green line indicates average of the estimated effective standard error (95% confidence interval divided by 2×1.96) and the light blue bars represent the distribution of the effective standard errors across the 1000 simulations. The dependent variable Y is the 2000–2007 change in the employment rate in all four panels.

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

	Robust	St-cluster	AKM	AKM0
Panel A: Change in the share of working-age population				
employed	47.4%	38.0%	7.7%	4.4%
employed in manufacturing	65.2%	52.3%	7.7%	3.5%
employed in non-manufacturing	28.5%	26.8%	8.0%	5.2%

Notes: The content of this table is analogous to that in Table 2. The only difference is that 3107 counties (instead of CZs) are used as regional unit of analysis.

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

	Robust	St-cluster	AKM	AKM0
Panel A: Change in the share of working-age population				
employed	84.0%	63.8%	23.7%	4.0%
employed in manufacturing	90.2%	76.2%	32.8%	3.5%
employed in non-manufacturing	63.9%	38.7%	17.9%	3.6%
Panel B: Change in average log weekly wage				
employed	84.6%	64.8%	30.3%	3.8%
employed in manufacturing	55.7%	28.3%	12.9%	5.7%
employed in non-manufacturing	84.3%	66.4%	30.4%	3.5%

Notes: The content of this table is analogous to that in Table 2. The only difference is that the placebo exercise is based on random shocks for 331 occupations.

Table C.5: Magnitude of standard errors. Sectoral composition.

	Estimates		Median effective std. error					
	Average	St Dev	Robust	St-cluster	AKM	AKM0	AKM	AKM0
Sector Correlation:					Independent		3-digit cluster	
Panel A: Number of sectors								
2-digit ($S = 20$)	0.04	3.16	0.67	0.97	2.98	5.45	—	—
3-digit ($S = 136$)	0.04	2.21	0.72	0.93	2.15	2.68	—	—
4-digit ($S = 398$)	0.01	1.94	0.73	0.91	1.85	2.17	—	—
Panel B: Simulated Shocks to non-manufacturing sector								
	0.08	4.21	0.58	0.75	1.16	1.31	—	—
Panel C: Simulated shocks with correlation within 3-digit SIC sectors								
$\rho = 0.00$	0.00	1.95	0.73	0.92	1.86	2.16	1.85	2.20
$\rho = 0.25$	-0.06	2.06	0.77	1.07	2.00	2.14	2.11	2.33
$\rho = 0.50$	-0.06	2.10	0.76	1.06	1.99	2.13	2.22	2.46
$\rho = 0.75$	-0.03	2.24	0.76	1.08	1.98	2.08	2.30	2.55
$\rho = 1.00$	0.00	2.31	0.76	1.08	1.96	2.04	2.39	2.65

Notes: All estimates in this table use the total employment share in each CZ as the outcome variable Y. For inference procedure indicated in the first row, "median effective std. error" refers to the median length of the 95% confidence interval across the 1000 simulated datasets divided by 2×1.96. *Robust* is the Eicker-Huber-White standard error; *St-cluster* is the standard error that clusters CZs in the same state; *AKM* (Independent) is the standard error in Remark 5; *AKM0* (Independent) is the confidence interval in Remark 5; *AKM* (3-digit cluster) is the standard error in eq. (40); and *AKM0* (3-digit cluster) is the confidence interval described in the last sentence of Section 4.3.1.

D Empirical application: additional results

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

	Change in the employment share			Change in avg. log weekly wage		
	All (1)	Manuf. (2)	Non-Manuf. (3)	All (4)	Manuf. (5)	Non-Manuf. (6)
Panel A: All Workers						
$\hat{\beta}$	-0.66	-0.50	-0.15	-0.61	0.18	-0.63
Robust	[-0.97,-0.35]	[-0.65,-0.36]	[-0.43,0.12]	[-1.05,-0.17]	[-0.67,1.02]	[-1.11,-0.15]
Cluster	[-0.87,-0.44]	[-0.62,-0.39]	[-0.39,0.09]	[-1.01,-0.21]	[-0.67,1.02]	[-1.06,-0.20]
AKM (indep.)	[-1.08,-0.23]	[-0.71,-0.3]	[-0.46,0.15]	[-1.1,-0.11]	[-0.61,0.96]	[-1.17,-0.09]
AKM_0 (indep.)	[-1.47,-0.32]	[-0.85,-0.32]	[-0.72,0.10]	[-1.41,-0.14]	[-1.21,0.85]	[-1.62,-0.18]
AKM (4d cluster)	[-1.09,-0.22]	[-0.73,-0.28]	[-0.46,0.16]	[-1.15,-0.07]	[-0.62,0.97]	[-1.20,-0.06]
AKM_0 (4d cluster)	[-1.53,-0.32]	[-0.92,-0.32]	[-0.73,0.11]	[-1.55,-0.11]	[-1.27,0.87]	[-1.73,-0.17]
AKM (3d cluster)	[-1.11,-0.20]	[-0.73,-0.28]	[-0.49,0.18]	[-1.17,-0.05]	[-0.62,0.97]	[-1.23,-0.02]
AKM_0 (3d cluster)	[-1.72,-0.30]	[-0.94,-0.30]	[-0.89,0.12]	[-1.66,-0.08]	[-1.45,0.87]	[-1.97,-0.13]
Panel B: College Graduates						
$\hat{\beta}$	-0.34	-0.49	0.15	-0.64	0.42	-0.63
Robust	[-0.55,-0.13]	[-0.65,-0.32]	[-0.06,0.35]	[-1.12,-0.15]	[-0.14,0.97]	[-1.14,-0.13]
Cluster	[-0.52,-0.17]	[-0.65,-0.33]	[-0.07,0.36]	[-1.10,-0.17]	[-0.18,1.01]	[-1.09,-0.18]
AKM (indep.)	[-0.60,-0.09]	[-0.68,-0.30]	[-0.05,0.34]	[-1.11,-0.16]	[-0.15,0.98]	[-1.13,-0.14]
AKM_0 (indep.)	[-0.75,-0.10]	[-0.74,-0.28]	[-0.17,0.33]	[-1.35,-0.15]	[-0.48,0.96]	[-1.45,-0.18]
AKM (4d cluster)	[-0.60,-0.08]	[-0.70,-0.28]	[-0.06,0.35]	[-1.17,-0.11]	[-0.18,1.01]	[-1.16,-0.11]
AKM_0 (4d cluster)	[-0.80,-0.11]	[-0.81,-0.28]	[-0.16,0.35]	[-1.48,-0.11]	[-0.54,1.00]	[-1.55,-0.16]
AKM (3d cluster)	[-0.61,-0.07]	[-0.71,-0.27]	[-0.06,0.35]	[-1.19,-0.08]	[-0.17,1.00]	[-1.21,-0.06]
AKM_0 (3d cluster)	[-0.85,-0.09]	[-0.82,-0.22]	[-0.24,0.35]	[-1.63,-0.07]	[-0.57,1.05]	[-1.79,-0.13]
Panel C: Non-College Graduates						
$\hat{\beta}$	-0.94	-0.49	-0.45	0.18	-0.63	-0.64
Robust	[-1.38,-0.49]	[-0.65,-0.33]	[-0.82,-0.07]	[-0.67,1.02]	[-1.11,-0.15]	[-1.12,-0.15]
Cluster	[-1.26,-0.62]	[-0.63,-0.35]	[-0.76,-0.13]	[-0.67,1.02]	[-1.06,-0.20]	[-1.10,-0.17]
AKM (indep.)	[-1.61,-0.26]	[-0.74,-0.23]	[-0.93,0.04]	[-0.61,0.96]	[-1.17,-0.09]	[-1.11,-0.16]
AKM_0 (indep.)	[-2.27,-0.42]	[-0.97,-0.28]	[-1.38,-0.06]	[-1.21,0.85]	[-1.62,-0.18]	[-1.35,-0.15]
AKM (4d cluster)	[-1.62,-0.25]	[-0.76,-0.22]	[-0.94,0.04]	[-0.62,0.97]	[-1.20,-0.06]	[-1.17,-0.11]
AKM_0 (4d cluster)	[-2.36,-0.42]	[-1.03,-0.28]	[-1.42,-0.06]	[-1.27,0.87]	[-1.73,-0.17]	[-1.48,-0.11]
AKM (3d cluster)	[-1.67,-0.20]	[-0.75,-0.22]	[-0.99,0.09]	[-0.62,0.97]	[-1.23,-0.02]	[-1.19,-0.08]
AKM_0 (3d cluster)	[-2.71,-0.38]	[-1.10,-0.28]	[-1.69,-0.03]	[-1.45,0.87]	[-1.97,-0.13]	[-1.63,-0.07]

Notes: $N = 1,444$ (722 CZs $\times$ two time periods). Models are weighted by start of period CZ share of national population. All regressions include the full vector of baseline controls in ADH. *Robust* is the Eicker-Huber-White standard error; *Cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5 (for independent shocks), or that in eq. (40) (for correlated shocks). *AKM0* is the test in Remark 6 (for independent shocks) or that described in the last sentence of Section 4.3.1 (for correlated shocks). The table reports confidence intervals computed under the assumption of independent shocks, 4-digit SIC correlated shocks, and 3-digit SIC correlated shocks.

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

	Change in the employment share			Change in avg. log weekly wage		
	All (1)	Manuf. (2)	Non-Manuf. (3)	All (4)	Manuf. (5)	Non-Manuf. (6)
Panel A: All Workers						
$\hat{\beta}$	-0.75	-0.58	-0.17	-0.70	0.20	-0.72
Robust	[-1.09,-0.42]	[-0.77,-0.39]	[-0.48,0.13]	[-1.18,-0.22]	[-0.79,1.19]	[-1.23,-0.21]
Cluster	[-1.12,-0.39]	[-0.78,-0.38]	[-0.47,0.12]	[-1.19,-0.21]	[-0.76,1.16]	[-1.22,-0.22]
AKM (indep.)	[-1.18,-0.33]	[-0.81,-0.35]	[-0.51,0.16]	[-1.24,-0.16]	[-0.72,1.13]	[-1.28,-0.16]
AKM_0 (indep.)	[-1.41,-0.38]	[-0.89,-0.36]	[-0.67,0.13]	[-1.42,-0.16]	[-1.11,1.07]	[-1.54,-0.22]
AKM (4d cluster)	[-1.19,-0.32]	[-0.84,-0.32]	[-0.51,0.16]	[-1.29,-0.11]	[-0.74,1.14]	[-1.31,-0.13]
AKM_0 (4d cluster)	[-1.66,-0.38]	[-1.12,-0.33]	[-0.73,0.14]	[-1.72,-0.13]	[-1.21,1.16]	[-1.83,-0.20]
AKM (3d cluster)	[-1.24,-0.26]	[-0.85,-0.31]	[-0.55,0.20]	[-1.31,-0.09]	[-0.73,1.13]	[-1.37,-0.07]
AKM_0 (3d cluster)	[-2.02,-0.34]	[-1.25,-0.31]	[-0.96,0.15]	[-1.95,-0.09]	[-1.45,1.17]	[-2.23,-0.16]
Panel B: College Graduates						
$\hat{\beta}$	-0.39	-0.56	0.17	-0.73	0.48	-0.73
Robust	[-0.62,-0.16]	[-0.79,-0.33]	[-0.08,0.42]	[-1.28,-0.18]	[-0.19,1.15]	[-1.29,-0.17]
Cluster	[-0.64,-0.15]	[-0.81,-0.31]	[-0.08,0.41]	[-1.33,-0.14]	[-0.19,1.14]	[-1.30,-0.16]
AKM (indep.)	[-0.67,-0.12]	[-0.81,-0.32]	[-0.08,0.41]	[-1.27,-0.19]	[-0.22,1.18]	[-1.27,-0.19]
AKM_0 (indep.)	[-0.76,-0.12]	[-0.85,-0.28]	[-0.16,0.41]	[-1.42,-0.16]	[-0.46,1.17]	[-1.46,-0.21]
AKM (4d cluster)	[-0.68,-0.11]	[-0.83,-0.29]	[-0.08,0.42]	[-1.33,-0.13]	[-0.25,1.21]	[-1.30,-0.16]
AKM_0 (4d cluster)	[-0.89,-0.12]	[-1.05,-0.27]	[-0.16,0.47]	[-1.71,-0.12]	[-0.52,1.32]	[-1.72,-0.19]
AKM (3d cluster)	[-0.70,-0.09]	[-0.85,-0.27]	[-0.09,0.42]	[-1.36,-0.11]	[-0.23,1.19]	[-1.35,-0.11]
AKM_0 (3d cluster)	[-1.03,-0.09]	[-1.13,-0.21]	[-0.24,0.46]	[-1.94,-0.08]	[-0.58,1.39]	[-2.07,-0.15]
Panel C: Non-College Graduates						
$\hat{\beta}$	-1.08	-0.56	-0.51	0.20	-0.72	-0.73
Robust	[-1.56,-0.59]	[-0.74,-0.38]	[-0.93,-0.10]	[-0.79,1.19]	[-1.23,-0.21]	[-1.28,-0.18]
Cluster	[-1.60,-0.55]	[-0.75,-0.37]	[-0.95,-0.08]	[-0.76,1.16]	[-1.22,-0.22]	[-1.33,-0.14]
AKM (indep.)	[-1.74,-0.41]	[-0.82,-0.30]	[-1.02,-0.01]	[-0.72,1.13]	[-1.28,-0.16]	[-1.27,-0.19]
AKM_0 (indep.)	[-2.13,-0.52]	[-0.95,-0.33]	[-1.29,-0.08]	[-1.11,1.07]	[-1.54,-0.22]	[-1.42,-0.16]
AKM (4d cluster)	[-1.75,-0.40]	[-0.84,-0.28]	[-1.02,-0.01]	[-0.74,1.14]	[-1.31,-0.13]	[-1.33,-0.13]
AKM_0 (4d cluster)	[-2.49,-0.51]	[-1.17,-0.31]	[-1.44,-0.07]	[-1.21,1.16]	[-1.83,-0.20]	[-1.71,-0.12]
AKM (3d cluster)	[-1.86,-0.29]	[-0.86,-0.26]	[-1.09,0.06]	[-0.73,1.13]	[-1.37,-0.07]	[-1.36,-0.11]
AKM_0 (3d cluster)	[-3.11,-0.45]	[-1.36,-0.30]	[-1.86,-0.03]	[-1.45,1.17]	[-2.23,-0.16]	[-1.94,-0.08]

Notes: $N = 1,444$ (722 CZs $\times$ two time periods). Models are weighted by start of period CZ share of national population. All regressions include the full vector of baseline controls in ADH. *Robust* is the Eicker-Huber-White standard error; *Cluster* is the standard error that clusters CZs in the same state; *AKM* is the standard error in Remark 5 (for independent shocks), or that in eq. (40) (for correlated shocks). *AKM0* is the test in Remark 6 (for independent shocks) or that described in the last sentence of Section 4.3.1 (for correlated shocks). The table reports confidence intervals computed under the assumption of independent shocks, 4-digit SIC correlated shocks, and 3-digit SIC correlated shocks.

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

Estimates		Rejection rate of $H_0\colon \beta = 0$ at 5%			
Average	St Dev	Robust	St-cluster	AKM	AKM0
Panel A: Baseline simulation without controls					
0.02	1.72	42.0%	37.4%	5.7%	3.6%
Panel B: Controlling for ADH IV					
-0.08	1.32	30.0%	22.8%	7.5%	4.3%
Panel C: Controlling for ADH IV and baseline controls					
0.00	0.73	13.3%	12.3%	7.5%	4.5%

Notes: Dependent variable is the "shift-share" regressor in ADH contructed from the interaction of CZ's employment share in 4-digit SIC manufacturing industries and the normalized U.S. imports from China in the same industries. The inference procedure employed to compute the average length of the 95% confidence interval in each of the columns is indicated in the first row. *Robust* is the Eicker-Huber-White standard error estimate; *St-cluster* is the standard error clusters CZs in the same state; *AKM* is the standard error in Remark 5; *AKM0* is the test in Remark 6.