1-20-2021

Measuring Local Job Distress

Timothy J. Bartik
W.E. Upjohn Institute for Employment Research, bartik@upjohn.org

Upjohn Institute working paper ; 20-335

Authors
Timothy J. Bartik, *W.E. Upjohn Institute for Employment Research*

Upjohn Author(s) ORCID Identifier
https://orcid.org/0000-0002-6238-8181

W.E. UPJOHN INSTITUTE
FOR EMPLOYMENT RESEARCH
Researching the causes and consequences of unemployment

MEASURING LOCAL JOB DISTRESS

Upjohn Institute Working Paper 20-335

Timothy J. Bartik
W.E. Upjohn Institute for Employment Research
Email: bartik@upjohn.org

January 2021

ABSTRACT

In this paper, estimates are presented on short-run effects of demand shocks on a local labor market's employment to population ratio (employment rate). Based on the estimates, commuting zones (CZs) better define a local labor market than counties, because both employment and employment rate effects exhibit large spillovers across counties within a CZ. In addition, the estimates suggest that demand shock effects vary, by an amount that is both statistically and substantively significant, with a CZ's prior overall employment rate.

JEL Classification Codes: R23

Key Words: Local labor markets; distressed regions; job creation benefits; local labor demand

Acknowledgments: I thank Shane Reed, Lillian Petrovic, Claire Black, and Allison Colosky for assistance.

How should we measure which local areas suffer from "job distress" and are most in need of jobs? This paper addresses that question.

Why care about measuring local job distress? Geographic variation in job distress rationalizes geographically targeting job creation. If jobs have higher benefits in some places—for example, because job creation boosts local employment rates more—then reallocating jobs to those places—such as through business incentives—may benefit the nation.

Local job creation is divided between increases in local employment to population ratios (employment rates) and increases in local population. Local jobs are directly filled by three sources: 1) already-employed residents, 2) nonemployed residents, and 3) in-migrants. Jobs filled by already-employed residents create vacancies, filled by the same three sources. The job vacancy chain is terminated when a vacancy is filled by a nonemployed resident or in-migrant.

As jobs are filled along a vacancy chain, the share that goes to the local nonemployed will vary with local labor's availability. But how should local labor's availability be measured? This paper addresses that question along two dimensions.

The first measurement issue: how to define local labor markets geographically. Typically, local labor markets are defined as areas that encompass most commuting flows, such as metro areas. But nearby jobs may be more accessible because of better transportation or information. This paper considers which geographic area's job creation has the most effect on local employment rates.

The second measurement issue is which labor should be considered available. Should we use the unemployment rate or the employment rate? Should we include persons of all ages and education levels, or only subgroups (prime-age workers, the less educated)?

This paper examines these questions using U.S. data. Key findings include the following:

- Local labor markets are better defined as “commuting zones” (CZs) than as counties, although the estimates cannot rule out some importance of counties.
- Local labor availability is better defined using broader measures of all nonemployed.

REVIEW OF PRIOR RESEARCH LITERATURE

This paper expands on Austin, Glaeser, and Summers (2018). They find that job creation has greater effects on local employment rates in distressed places. This finding implies that job creation in distressed places will boost earnings per capita more; social problems will be reduced more with greater employment rate effects; the opportunity cost of labor in distressed areas is lower, as inferred from the greater labor supply response (Bartik 2020); and redistributing jobs to distressed places will boost national employment by targeting more elastically supplied labor (Austin, Glaeser, and Summers 2018).

Austin, Glaeser, and Summers find that the prime-age male employment rate is more responsive to labor demand shocks in “consistent PUMAs” that have a lower prime-age male employment rate. PUMAs are “Public Use Microdata Areas”: areas of around 100,000 in population that are the smallest area for which the census will identify an individual’s location. “Consistent” PUMAs are larger areas that combine PUMAs to match across census years.

The current paper considers alternatives to Austin, Glaeser, and Summers (2018) in three modeling choices: geographic areas, dependent variable, and distress measure. First, regarding geography: consistent PUMAs are not consistent labor market areas.[1] Some consistent PUMAs are small neighborhoods, while others are very large multicounty groups. For the 1,066 consistent PUMAs that bridge the 2000 and 2010 census definitions, 456 are less than half the

[1] Austin, Glaeser, and Summers (2018) also look at effects for states. States are also not local labor market areas.

population of a single county, too small to be a local labor market. But 18 consistent PUMAs span 30 or more counties, and another 50 span between 10 and 29 counties, which are mostly too big to be plausible local labor markets.

The conventional wisdom among regional economists is that local labor markets should be areas that encompass most commuting flows. Due to changes in commuting, if a labor supply or demand shock occurs anywhere within a local labor market, the shock's effects on wages or employment rates will be spread throughout the area. Metropolitan and micropolitan areas have been created by the Census Bureau as areas that encompass most commuting flows. The U.S. Department of Agriculture has defined CZs, which also encompass most commuting flows, but assign each and every county in the United States, whether urban or rural, to a zone.

But recent research suggests that effects of local shocks might be more localized. For example, Manning and Petrongolo (2017) find that half of a labor demand shock's effects on unemployment flows occurs within a radius of about 9 miles, and 90 percent within a radius of about 19 miles. The former radius would encompass a little less than 250 square miles, the latter just under 1,200 square miles. The median county in the United States is 616 square miles.

With respect to Austin, Glaeser, and Summers' (2018) dependent variable: the change in the prime-age male employment rate does not capture effects on all ages or on women. A more inclusive measure would also divide job shocks between employment rate effects and migration effects. With respect to their distress indicator: allocating government aid to areas based on their prime-age male employment rate would be illegal, because this indicator differentiates by gender. In addition, experimenting with different distress indicators seems desirable.

I expand upon Austin, Glaeser, and Summers (2018) in three respects. First, my dependent variable is the overall employment rate. Second, I test how results vary when defining local labor markets as CZs versus counties. Third, I test how demand shock effects vary with different legally permissible definitions of labor market distress, including the unemployment rate and the employment rate, for either all persons aged 16 and over or various subgroups.

HOW THIS PAPER'S ELASTICITIES RELATE TO BENEFIT-COST RATIOS FOR LOCAL JOB CREATION

This paper's model, detailed in the next section, uses pooled time series cross-section data on either CZs or counties. The change from one year to the next in the ln(*employment rate*) for a CZ is related to the demand-induced change in the CZ's ln(*employment*). The change from one year to the next in the ln(*employment rate*) for a county is related to both the demand-induced change in ln(*employment*) for the county, and to the demand-induced change in ln(*employment*) for the CZ.[2]

These estimated elasticities of employment rates with respect to demand shocks to employment can be shown to be proportional to local benefit-cost ratios for job creation policies under plausible assumptions: local benefits are proportional to the change in average employment rates for various groups; local costs are proportional to the demand-induced change in employment. If we find that an area's baseline distress affects these elasticities by *x* percent, we can infer that they will alter the benefit-cost ratio by *x* percent.

Suppose that in some CZ *z*, the elasticity of the local employment rate with respect to a shock to employment in *z* is $H_z = \frac{\partial ln(E_z/P_z)}{\partial ln(E_z)}$ where E_z and P_z are employment and population in CZ z. Suppose the benefit *B* in dollars of some increase in the employment rate in CZ *z* equals

[2] Austin, Glaeser, and Summers (2018) examine the change in the prime-age male employment rate as a function of a shock to log employment, so the dependent variable is the change in the employment rate, not the change in the log rate.

some parameter b times the population P_z times $\mathrm{d}\frac{E_z}{P_z}$. Suppose total costs C in dollars of increasing employment in z can be written as some parameter c times $(\mathrm{d}E_z)$. Then the benefit-cost ratio of job creation policies in CZ z can be written as

(1) $$\frac{B}{C} = \frac{bP_z[\mathrm{d}(E_z/P_z)]}{c[\mathrm{d}E_z]} = \frac{b}{c} H_z.$$

If we estimate elasticities of a county's employment rate with respect to both shocks to county employment and CZ employment, then total benefits depend on how one weights benefits to the county, versus benefits elsewhere in the CZ, and also on whether the shock is to the county or elsewhere in the CZ. Suppose total benefits B are equal to benefit weight b_1 times the change in the county's employment rate times its population, plus some benefit weight b_2 times the product of the change in the employment rate and the population in the rest of the CZ, or

(2) $$B = b_1\left(\mathrm{d}\frac{E_1}{P_1}\right)P_1 + b_2\left(\mathrm{d}\frac{E_2}{P_2}\right)P_2$$

where E_1, P_1, E_2, and P_2 are employment and population in the county (Area 1) and elsewhere in the CZ (Area 2).

The employment rate in either Area 1 or Area 2 will change due to demand shocks to each area's employment and the CZ's employment. Suppose that we estimate an equation that shows elasticities of the county's employment rate with respect to both county employment and the CZ's employment, and these elasticities are G_a and G_z. Then the term $\left(\mathrm{d}\frac{E_1}{P_1}\right)P_1$ can be written as

(3) $$\left(\mathrm{d}\frac{E_1}{P_1}\right)P_1 = G_a\mathrm{d}(E_1) + \left(\frac{E_1}{E_1+E_2}\right)G_z(\mathrm{d}E_1 + \mathrm{d}E_2)$$

A similar equation can be written for Area 2:

(4) $$\left(\mathrm{d}\frac{E_2}{P_2}\right)P_2 = G_a\mathrm{d}(E_2) + \left(\frac{E_2}{E_1+E_2}\right)G_z(\mathrm{d}E_1 + \mathrm{d}E_2)$$

Costs of job creation are assumed to be the same everywhere:

(5) $$C = c(\mathrm{d}E_z) = c(\mathrm{d}E_1 + \mathrm{d}E_2)$$

The benefit-cost ratio can be calculated by plugging Equations (3) and (4) into Equation (2) and dividing by Equation (5). If we assume $b_1 = b_2 = b$; that is, we weight gains in employment rates equally throughout the CZ, then *regardless* of where in the CZ the jobs are created, the benefit-cost ratio is proportional to the sum of the two elasticities, or

(6) $$\frac{B}{C} = \frac{b}{c}(G_a + G_z)$$

Suppose alternatively that we set $b_2 = 0$: we only value Area 1 (the city?), and not Area 2 (the suburbs?). Then the benefit-cost ratio depends on where in the CZ the jobs are located if G_a is nonzero. If the jobs are created in Area 1, the benefit cost ratio is

(7) $$\frac{B}{C} = \frac{b_1}{c}\left[G_a + \frac{E_1}{(E_1+E_2)}G_z\right]$$

But if $b_2 = 0$, and the jobs are created in Area 2, the benefit-cost ratio is

(8) $$\frac{B}{C} = \frac{b_1}{c}\left[\frac{E_1}{(E_1+E_2)}\right]G_z$$

So, if we only value one part of the CZ, then demand shocks that occur elsewhere in the CZ are downweighted, as shown in Equation (8), by the proportion of employment in that one part of the CZ. But if the demand shock occurs in the valued part of the CZ, we add in differential effects of nearby jobs (Eq. 7).

EMPIRICAL MODEL

The baseline equation estimated is

(9) Change in ln(*employment to population ratio in area k from year t−1 to year t*)

= B_0 + *Year fixed effect* + B_d (*Demand shock to area k from year t−1 to year t*)

[+ B_m (*Demand shock to overlying area m from year t−1 to t*)]

As mentioned, when the dependent variable is for the CZ, the demand shock variable is for the CZ; when the dependent variable is for the county, one specification just considers demand shocks for the county, and another specification considers both the county and the CZ.

My preferred specification measures the demand shock as the predicted change in ln(*employment*) in the area k (or m) due to the area's industry mix and national growth in each industry, or

(10) *Demand shock to area k from year t−1 to year t* = $\ln(1+D_{kt})$

where

(11) D_{kt} = sum over all industries i of $\frac{E_{kit-1}}{E_{kt-1}} * \frac{E_{nit}-E_{nit-1}}{E_{nit-1}}$

In Equation (11), E_{kit-1} is employment in area k in industry i in year $t-1$, E_{kt-1} is total employment in area k in year $t-1$, E_{nit} is employment in the nation in industry i in year t. and E_{nit-1} is employment in the nation in industry i in year $t-1$. Equation (10) is the predicted growth rate of area k's ln(*employment*) if all industries in k grew at the same rate as their national counterpart from year $t-1$ to year t. Differentials in this predicted growth across areas will be mainly due to shocks to an area's export-base industries (Bartik 1991).

In addition to this approach, I also do estimates in which the demand-shock is measured as the growth in area k's log employment, or $\ln(E_{kt}) - \ln(E_{kt-1})$. I also explore the demand shock as an instrument for actual employment growth.

Why prefer the demand shock specification to the actual job growth specification or the instrumented growth specification? The actual job growth specification is potentially biased by local labor supply shocks. Labor supply shocks that increase local population may also increase employment; the resulting change in the employment rate may differ from what would occur because of demand shocks, which is what we should estimate to ascertain the effects of labor demand policies.

These labor supply shock effects stem from both direct labor supply effects and indirect product market effects. More population may lower the employment rate while raising employment, which is the opposite of the correlation resulting from a demand shock. If the population shock replaces people with low employment rates with people with high employment rates, this may increase local income, which will increase local labor demand. The resulting correlation between employment and employment rates may not reflect what would result from a direct demand shock to jobs.

Using industry mix demand shocks as instruments potentially corrects for this problem. But the predicted variation in ln(employment) due to the industry mix instrument represents the effect of the realized change in total employment. This reflects both the direct demand shock effects and multiplier effects on suppliers and retailers.

In evaluating a demand shock policy, it is preferable to compare the change in employment rates with the directly induced job creation, which is what policy directly influences. This is particularly important in the specification that explains changes in county employment rates with both county and CZ right-hand-side variables. CZ demand shocks outside the county may spill over into the county via multiplier effects. The CZ and county demand shock effects tells us how demand-side policies at the county or CZ affects employment rates in the county, including via spillover effects.

But the instrumented estimates are also of interest. These estimates tell us, after multiplier effects, how the realized change in employment due to demand shocks, in both the county and CZ, affect the county's employment rate. These effects will reflect labor market spillovers beyond the multiplier effects. In contrast, the direct demand shock effects reflect spillovers due to both multipliers and commuting flows.

The disturbance might be correlated over time for a county or CZ, or across different counties in the same CZ. Therefore, I correct the standard errors for possible clustering at the CZ level.

This predicted growth instrument has been criticized (e.g., Goldsmith-Pinkham, Sorkin, and Swift 2020), the main critique being that an area's base period industry mix may be correlated with local labor supply shocks. But in the current paper, I am examining multiple short-run demand shocks due to industry mix and national trends. A local area's industry mix may sometimes predict high job growth, sometimes low. This reduces the potential correlation with supply shifters, which will have more persistent effects.

Furthermore, any instrument should not be judged from the standard of perfection, but rather on whether it improves matters relative to ordinary least squares. It is plausible that this predicted growth instrument captures demand shock effects better than unadjusted employment growth.

After estimating this baseline specification, I examine specifications where the demand shock is interacted with possible measures of local labor market distress. I restrict possible interaction terms to those that would be legal for targeting government policy; for example, I do not consider labor market distress indicators segregated by gender or race. The interaction terms considered are the level in year $t-1$ of the ln(*employment rate*), ln(*labor force participation rate*), and ln(*employment to labor force ratio*). These three interaction terms are considered for three groups: the entire population ages 16 and up; prime-age (ages 25–54); persons with less than a four-year college degree, ages 25–64. Thus, there are nine possible interaction terms (three types of labor outcomes times three groups). More elaborate regression-adjusted interaction terms could be considered, but policymakers are more likely to use these simpler methods. Furthermore, if controlling for age or education improves the predictive ability of interaction specifications, more elaborate adjustments can be considered in subsequent work.

DATA

The model is estimated using data for counties and CZs on 11 years of changes, from 2005–2006 through 2015–2016. The data used in estimation comes from the American Community Survey, County Business Patterns, and the Bureau of Economic Analysis.

The American Community Survey (ACS) is used to derive data on labor market outcomes, such as employment rates. The CZ data are derived from ACS microdata. The ACS data identify the individual's "public use microdata area" (PUMA) of residence, which as mentioned are areas of about 100,000 in population. The probability that the individual PUMA is in a CZ is combined with the ACS's person weights in making weighted estimates of a CZ's employment rates and other labor market outcomes. To increase the likelihood that these CZ assignments are accurate, I restrict the sample to CZs of 200,000 or greater population. See the Appendix for more details.

The labor market outcome data for counties is published ACS data. Such annual ACS-derived data are only available for counties of 65,000 or greater population.

The demand-shock instrument is constructed from the Upjohn Institute's WholeData. For each U.S. county, WholeData has complete information on industry employment for 1,079 six-digit NAICS industries for each year from 1998 to 2016. WholeData is derived from U.S. County Business Patterns data and overcomes data suppressions in CBP using an algorithm originally developed by Isserman and Westerveldt (2006). The change in actual employment uses Bureau of Economic Analysis data on wage and salary employment.

Estimates are done for three samples: sample 1 is county data only, sample 2 is CZ data only, and sample 3 uses changes in county employment rates as a dependent variable but explains these changes using both county and CZ data. Sample 1 ends up with 669 counties, sample 2 with 240 CZs, and sample 3 with 609 counties in 225 CZs. Sample 1 counties make up 79 percent of the U.S. population, sample 2 CZs are 89 percent of the U.S. population, and sample 3 counties are 77 percent of the U.S. population and are in CZs that collectively are 88 percent of U.S. population. The Appendix presents some descriptive statistics. In sample 3, the average county's employment, out of its CZ, is 33.7 percent; as described in a previous section, this county employment share plays a role in ascertaining the relative importance of CZ versus county demand shocks.

RESULTS, PART 1: WHICH GEOGRAPHIC UNIT DEMAND SHOCKS MATTER TO LOCAL EMPLOYMENT RATES, AND HOW DOES THIS VARY WITH DEMAND SHOCK MEASURES?

Table 1 shows baseline results for all three samples. For each sample, results are shown for the three different measures of demand shocks: the industry mix predictions, actual employment changes, and actual employment changes instrumented with industry mix predictions.

Turning first to samples 1 and 2, these two sample specifications look only at job shocks to the own geographic unit. Sample 1 looks at how county job shocks affect county employment rates. Sample 2 looks at how CZ job shocks affect CZ employment rates.

For both samples, all three of the demand shock measures show statistically significant effects of job shocks on employment rates. The instrumented demand shocks are significantly higher than the job shocks that just use actual employment change.[3] This suggests that the uninstrumented estimates are biased by supply side shocks to employment: actual job changes due to migration may both reduce employment rates and raise employment growth.

Also, for samples 1 and 2, the reduced form estimates and instrumented estimates are not much different. This suggests that the net short-run multiplier of demand shocks to job growth is not much different from one. This may reflect that multiplier effects take some time to develop, or that any multiplier effects on suppliers or retailers are to some extent offset by congestion effects of demand shocks, which raise local wages and prices, which will tend to reduce local employment.

Of most interest are the results from sample 3, which looks at how a county's employment rate is affected by job shocks at different geographies, the county or the CZ. My preferred estimates, the reduced form demand shock estimates, show that the elasticity of county employment rates with respect to CZ demand shocks is almost five times greater than the elasticity with respect to county demand shocks (0.5126 vs. 0.1184). Holding CZ demand shocks constant, the effect of county demand shocks on county employment rates is not statistically significant (t-statistic of 1.53 = 0.1184 / 0.0775), at conventional levels of statistical significance. But holding county demand shocks constant, CZ demand shocks are strongly statistically significant (t-statistics of 4.90 = 0.5126 / 0.1047).

[3] A Hausman test in sample1 for the difference between the instrumented estimate and the noninstrumented estimate yields a difference of 0.1593 (0.4296 − 0.2702), with a standard error of 0.0603. In sample 2, a similar Hausman test yields a difference of 0.2759 with a standard error of 0.0779.

In contrast, if we predict county employment rate changes using actual county job growth and actual CZ job growth, results are quite different. In this specification, county job shocks have about one-and-a-half times the effect of CZ job shocks (0.1989 vs. 0.1367).

The results using demand shocks as instruments for actual employment growth suggest that county employment rate changes are driven more by CZ shocks than county shocks (0.3709 vs. 0.1782). The instrumented results are imprecise enough that the county or CZ-specific demand shock estimates are not statistically significant from the results using actual employment growth, but the combined county plus CZ-instrumented demand shock results are statistically significantly greater with instruments.[4]

The pattern of results with actual employment change, versus instrumented employment change, suggest that the actual employment change estimates are biased differently for the county shocks versus the CZ shocks. The actual employment change estimates of the effects of CZ shocks are biased downward by supply shocks: migration at the CZ level tends to increase job growth and lower employment rates. But the actual employment change estimates of the effects of county shocks are in part biased upward by supply shocks: migration within the CZ toward the county may often include persons with higher employment rates, who drive up both county employment and employment rates. Demand shock instruments are important for avoiding bias in these relative county versus CZ effects due to these disparate effects of county and CZ labor supply shocks.

[4] Hausman test results for the county shock alone yields a difference of −0.0201 with a standard error of 0.1481, and for the CZ shock yields a difference of 0.2342 with a standard error of 0.1584. But the sum of the county and CZ shocks yields a Hausman test difference of 0.2141 with a standard error of 0.0664. As these calculations suggest, correlations between demand shocks at the county and CZ level increase imprecision of the estimated effects of either shock holding the other constant.

The instrumented results, versus the reduced form results, are answering different questions. The instrumented demand shocks show the realized effect on county employment rates of the actual demand-induced growth that occurs at the county versus CZ level. The reduced form results show the effects by the location of the original demand shock. The results suggest that the direct effects of a demand shock at the CZ level is almost five times that at the county level, but when one includes indirect effects at the county and CZ level, the relative effects of the actual realized demand-induced job growth is only twice as great at the CZ versus county level.

To see why this difference occurs, it is helpful to look at the first stage of the instrumented regressions. This first stage is reported in Table 2. For sample 3, the results show that actual county job growth is highly responsive to both county and CZ demand shocks.[5] We would expect this pattern, as many of the multiplier effects on suppliers and retailers of a demand shock in one county would spill over into neighboring counties in the same CZ because of firms using supplies in the next counties, workers commuting across counties, and consumers buying goods and services in the next county over. In contrast, and also as one might expect, the CZ's employment change is not much predicted by the county demand shock, but is strongly predicted by the CZ demand shock. CZ employment change is affected by overall CZ demand conditions, not the distribution of demand shocks within the CZ.

Thus, this pattern of results suggests some answer to the question, Why are employment rate changes in a county more affected by CZ demand shocks than by county demand shocks? The answer, in part, is that CZ demand shocks spill over into the county via multiplier effects. But, in addition, the instrumented results show that even if we look at the realized job growth

[5] Also of note in samples 1 and 2 is that the first-stage *t*-statistics well exceed 10, suggesting that these are good instruments. In sample 3, the Cragg-Donald F-test is 145.3, the Kleinbergen-Paap F-test is 30.9.

due to demand shocks, a county's employment rate changes are more affected by a CZ's realized job growth than by a county's realized job growth. This pattern is probably due to commuting spreading labor market outcome changes across the CZ.

Thus, a central city county benefits from demand shocks to the suburbs for two reasons. First, a demand shock in the suburbs will induce some job growth in the central city county because of multiplier effects spillovers. Second, even the job growth that occurs in the suburbs will increase the employment rate in the central city county because of commuting effects.

Although a county's employment rate is more affected by CZ demand shocks than by county demand shocks, the point estimates on county demand shock effects still suggest that a county's workers benefit somewhat more from demand shocks in the county versus outside the county. As mentioned, the average county in sample 3 has employment that is 33.7 percent of CZ employment. As shown in a previous section, the benefit-cost ratio for the entire CZ of a job shock anywhere in the CZ will be proportional to 0.6310, the sum of the county plus CZ coefficients. But the benefit-cost ratio for the county of job growth to the county is proportional to 0.2913 [equals (0.1184 + (33.7% × 0.5126)], whereas the benefit-cost ratio for the county of job growth to the rest of the CZ, outside the county, is proportional to 0.1729 (equal to 33.7% × 0.5126). So, from the viewpoint of county residents, a demand shock to the county has a 69 percent greater effect on county employment rates than a demand shock to the rest of the CZ (1.69 = 0.2913 / 0.1729). However, that 69 percent differential is not statistically significantly different, as it rests on a coefficient that is not statistically significant. On the other hand, one cannot reject the hypothesis that the differential effects on county residents of county demand shocks, compared to noncounty CZ demand shocks, might be twice as great.

In the Appendix, I show that these results are robust instead to focusing on the largest CZs. One might think that in larger CZs, perhaps the within-CZ distribution of demand shocks is of greater importance. This does not appear to be the case.

RESULTS PART 2: MEASURING DISTRESS AT THE COMMUTING ZONE LEVEL

Based on the above results, local labor markets are better defined as CZs, compared to counties. Demand shocks at the CZ level are more important than demand shocks at the county level in determining county employment rates.

But if we want to measure local "distress," what measure is best? More specifically, if we want to predict how much a demand shock will increase employment rates, what measure of local labor market conditions "works best"?

The ideal measure of local labor market distress will be more statistically significant as a predictor of how much employment rates will increase because of a labor demand shock. Such a measure will better predict the benefits that occur because of a labor demand shock.

In addition, an ideal measure of local labor market distress will not vary as much from year to year, as any targeting from federal or state governments will have some lag. Volatile distress measures are less desirable.

As already mentioned, I explore nine possible distress indicators in nine different regressions. The dependent variable is always the change from year $t-1$ to year t in the natural logarithm of the employment to population ratio. The sample is sample 2, which is 240 CZs greater than 200,000 in population, with observations from 2005–2005 to 2015–2016. Each of the nine different regressions consider interacting one of the nine distress indicators with the industry mix predicted demand shock to the CZ.[6] The nine distress indicators are the CZ's value in year $t-1$ of the natural logarithm of one of three possible labor market outcomes (the employment rate, the labor force participation rate, the employment to labor force ratio) for one of three possible groups (everyone 16+, prime-age persons ages 25–54, persons with less than a college education ages 25–64).

The estimated interaction term coefficients are shown in Table 3. In all nine regressions, more distress, as judged by a lower value of the distress term, is associated with statistically significant effects of increasing the impact of demand shocks to the CZ on the CZ's employment rate. A CZ with lower employment rates, labor force participation rates, or lower ratios of employment to the labor force (a higher unemployment rate) will tend to have higher effects of demand shocks on boosting the local employment rate.

The interaction term has a higher absolute value of the *t*-statistic when distress is measured for everyone 16 and over rather than for only prime-age workers or only noncollege workers. Apparently, the employment or nonemployment of everyone in the local labor market has some relevance to how the local labor market responds to demand shocks.

The absolute *t*-statistic for the distress interaction term is highest for the employment to labor force or unemployment distress indicator, followed by the employment to population ratio, and then followed by the labor force participation rate. The *t*-statistic difference is modest between the unemployment indicator and the employment rate, but is much larger between the employment rate and the labor force participation rate.

[6] Each of the nine regressions also includes all the nonredundant distress indicators by themselves as predictors, that is all regressions include six variables for ln(labor force/population) and ln(employment/labor force) for the three groups, which implicitly incudes ln(employment/population) as well given that it is the sum of these other two outcome variables.

This pattern might suggest using the local unemployment rate as a distress indicator. However, the local unemployment rate is far more volatile over time than the local employment to population ratio. For example, if we use these data to look at the correlations for these 240 CZs across the nine different pairs of years that are exactly five years apart (e.g., starting with the year pair of 2005 and 2010, and ending with the year pair of 2013 and 2018), the average correlation is 0.935 for the ln(employment to population ratio), and 0.613 for the ln(employment to labor force ratio).[7] A CZ that is judged "distressed" using the employment to population ratio is quite likely to still be judged distressed five years later, which is not true if the unemployment rate is used as a distress indicator.

I conclude that if one wants to pick a single stable labor market distress indicator, the overall employment to population ratio for all persons age 16 and over is the best distress indicator.

Using these point estimates and descriptive statistics, we can look at how much the effects of a demand shock on local employment rates vary with the CZ's initial employment rate. Going from the 90th percentile of the employment rate distribution to the 10th percentile,[8] the estimated effect of a demand shock on the local employment to population ratio increases from an elasticity of 0.5655 to 0.7712, an increase of 36 percent (1.36 = 0.7712 / 0.5655).

Compared to prior estimates, this is somewhat less of a differential. Austin, Glaeser, and Summers (2018) find a differential of 44 percent between low and high employment rate areas. But this uses consistent PUMAs and looks at prime-age male employment rates. Bartik (2015)

[7] The Appendix looks at correlations for the employment rate and the employment to labor force ratio across all possible pairs of years. Correlations are consistently much higher for the employment rate than for the employment to labor force ratio across year pairs, particularly the year pairs that are more separated in time.

[8] This is from $\ln(E/P) = -0.4191$ to $\ln(E/P) = -0.6479$, or from $(E/P) = 65.8$ percent to $(E/P) = 52.3$ percent.

finds a differential of about two-thirds greater for initially high unemployment rate metro areas versus initially low unemployment rate metro areas. But Bartik's estimates are long-run effects: the relative percent differential in elasticities might increase in the long run because long-run elasticities of employment rates with respect to demand shocks are lower than short-run elasticities. The absolute elasticity differential, of about 0.20 (0.77 − 0.57), is greater than the long-run elasticity differential in Bartik (2015) of 0.14 (0.34 − 0.20).

CONCLUSION

This research leads to four key conclusions. First, in defining local labor markets, labor market definitions such as CZs, which encompass most commuting flows, seem preferable to smaller geographic areas such as counties. Second, despite that conclusion, the estimates are not precise enough to rule out effects of labor demand shocks at smaller geographic scales as being important for some policy purposes. Third, broader measures of economic distress, incorporating the entire population, are preferable to narrower ones. Fourth, in these short-run estimates, the estimated effects of a demand shock on local employment rates are considerably greater in areas with initially low employment rates.

The limitations of this research point to some needed future research. First, it would be desirable to obtain additional data that could give more precise estimates of how much smaller geographic unit's demand shocks matter to labor market outcomes for different groups. This is important to some pressing policy issues, such as how we should organize policy to address local labor market issues, and how we should better geographically target public policies that create jobs. Second, we need to further examine how long-run effects of demand shocks on employment rates vary with an area's characteristics, such as its preexisting employment rate.

The challenge in examining long-run effects of demand shocks is that we have effectively fewer observations in long-run estimates. Obtaining long-run estimates requires either a dynamic model with lags, or looking at longer-run changes, which limits available observations. However, public policy toward local labor markets clearly should consider how labor demand shocks affect labor market outcomes in both the short run and the long run.

Appendix

More Details on How Labor Market Outcomes Data were Calculated and Defined

As mentioned in the paper, the commuting zone (CZ) data on employment rates and other labor market outcomes is derived from microdata from the American Community Survey (ACS). I assign CZs based on the public use microdata area (PUMA) identifiers on the ACS files. PUMAs are defined based on 2000 U.S. census definitions in the ACS from 2005 to 2011, and based on the 2010 census definitions from 2012 to 2016. These can be mapped probabilistically to counties using Missouri Census Data Center (2020) (http://mcdc.missouri.edu/applications/geocorr.html) and then further mapped to the groups of counties called CZs using CZ definitions based on the 2010 census created by Penn State researchers (Fowler, Jensen, and Rhubart 2018).

The Missouri data report for each PUMA the percentage of its population that resides in different counties. Thus, the procedure assumes that the overall population percentage represents a reasonable estimate of the probability of residing in different counties for individual observations from the ACS. This probability weight is combined with the person weight also provided in the ACS to calculate weighted labor market outcome statistics for each CZ. This procedure is likely to be more problematic for smaller CZs; for example, if a CZ is 20,000 people, no more than one-fifth of any PUMA could be assigned to that CZ. To reduce measurement error due to probabilistic assignment of PUMAs to CZs, the estimates here restrict the sample to CZs with more than 200,000 in population in 2005.

The county level uses reported ACS statistics, which were downloaded from the now-defunct American Factfinder and are now available on other census data platforms. Why use American Factfinder (AFF) data? Largely because these data are available without probabilistic assignment for each individual year for all counties with more than 65,000 in population. The PUMA assignment algorithm would not give as accurate an assignment for counties down to this smaller sample size. In addition, the ACS data used in calculating these aggregate statistics has a 50 percent larger sample size than is publicly available in ACS microdata.

Why not use the AFF for CZs? Because AFF only reports annual data for areas of 65,000 and above, and CZs are not one of the designated areas by the Census Bureau, one has to construct CZ statistics from counties. Many CZs would contain one or more counties with less than 65,000 in population, and so annual data would not be available for the complete CZ.

The principal data analyzed was various labor market outcomes for local labor markets. These labor market outcomes include the employment to population ratio, the labor force participation rate, and the employment to labor force ratio. These three labor market outcomes were calculated for all persons 16 and above, all persons of prime-age (ages 25–54), and all persons with less than a four-year college degree. These variables were calculated for the civilian population.

The AFF calculations of civilian labor market outcomes require some algebra using the AFF data. AFF reports four data items for the three population groups we examine (everyone 16 and above, prime-age persons, persons with less than a college degree ages 25–64):

1) ratio of civilian employment E to the sum of civilian population C and military population (and employment) M, which I define as $e^{**} = E/(C + M)$;

2) total population P, which $= C + M$;

3) labor force participation rate including military, which I write as l^*, and which equals civilian labor force L plus military employment M, divided by total population P;

4) the civilian unemployment rate, which I write as u, and which is ratio of civilian unemployment U to civilian labor force L.

Using these four data items, it is possible to calculate the civilian employment rate, labor force participation rate, and employment to labor force ratio.

The labor market outcome data used in the CZ sample and the county sample are calculated to be as similar as possible. For CZs, I include allocated observations, but drop the military, drop people in institutional quarters such as prisons and nursing homes, but persons in group quarters in dorms and homeless shelters. For AFF, it appears to be the case that the 2005 ACS data exclude all those in group quarters, but 2006 and subsequent years include all those in group quarters. Therefore, there are slight differences in different years in the county and CZ data in what group quarters persons are included.

Appendix Tables Presenting Descriptive Statistics and Additional Results

Appendix Tables A1–A4 present various descriptive statistics for the data used in the paper's regressions.

Appendix Table A5 reestimates the specification in which county employment rates are explained by both county and CZ demand shocks, but further restricts the sample to only include large CZs, with more than 1 million in population as of 2005. The sample reported in the main paper includes all CZs with more than 200,000 in population.

Appendix Table A6 reports all possible correlations across year pairs, for 240 CZs, for the ln(employment to population ratio) and ln(employment to labor force ratio). This is for ACS annual data that go from 2005 to 2018.

References

Austin, Benjamin, Edward Glaeser, and Lawrence H. Summers. 2018. "Saving the Heartland: Place-based Policies in 21st Century America." *Brookings Papers on Economic Activity* Spring (2018).

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

———. 2015. "How Effects of Local Labor Demand Shocks Vary with the Initial Local Unemployment Rate." *Growth and Change* 46(4): 529–557.

———. 2020. "Using Place-Based Jobs Policies to Help Distressed Communities." *Journal of Economic Perspectives* 34(3): 99–127.

Fowler, Christopher. S., Leif Jensen, and Danielle C. Rhubart. 2018. *Assessing U.S. Labor Market Delineations for Containment, Economic Core, and Wage Correlation*. Technical report. https://doi.org/10.17605/OSF.IO/T4HPU

Goldsmith-Pinkham, Paul, Isaac Sorkin, and Henry Swift. 2020. "Bartik Instruments: What, When, Why, and How." *American Economic Review* 110(8): 2586–2624.

Isserman, Andrew M., and James Westervelt. 2006. "1.5 Million Missing Numbers: Overcoming Employment Suppression in County Business Patterns Data." *International Regional Science Review* 29(3): 311–335.

Manning, Alan, and Barbara Petrongolo. 2017. "How Local Are Labor Markets? Evidence from a Spatial Job Search Model." *American Economic Review* 107(10): 2877–2907.

Missouri Census Data Center. 2020 "Geocorr Applications: Geographic Correspondence Engine." Columbia, MO: University of Missouri, Office of Social and Economic Data Analysis. http://mcdc.missouri.edu/applications/geocorr.html (accessed August 17, 2020).

Table 1 Baseline Regression Results for Job Shock Effects in Commuting Zones (CZs) and Counties: Elasticity of Employment Rate with Respect to Job Shock

	Predicted demand shock	Job shock	Instrumented job shock
Sample 1 (Counties)			
County job shock effect	0.3633	0.2702	0.4296
	(0.0567)	(0.0224)	(0.0643)
Sample 2 (CZs)			
CZ job shock effect	0.6633	0.3596	0.6209
	(0.0957)	(0.0346)	(0.0853)
Sample 3 (counties, in CZs >200K)			
County job shock effect	0.1184	0.1989	0.1788
	(0.0775)	(0.0347)	(0.1521)
CZ job shock effect	0.5126	0.1367	0.3709
	(0.1047)	(0.0439)	(0.1643)
Sum(CZ+county job shock effects)	0.6310	0.3356	0.5498
	(0.0867)	(0.0284)	(0.0722)

NOTE: Standard errors in parentheses below estimated coefficients. These results come from nine separate regressions: the three different samples times the three different types of shocks: demand shocks vs. overall job shocks vs. job shocks instrumented with demand shocks. Dependent variable in all cases is change in ln(employment to population ratio) from year $t-1$ to year t for everyone 16 and above in a particular county (sample 1 or sample 3) or CZ (sample 2). All regressions include controls for ln(labor force participation rate) and ln(employment to labor force ratio) in year $t-1$ for all 16 and above, all prime age, and all noncollege ages 25–64. All standard errors allow for clustering, at CZ level. All estimates based on sample that runs from 2005–2006 to 2015–2016. Sample 1 is 669 counties; Sample 2 is 240 CZs; Sample 3 is 609 counties in 225 CZs.

Table 2 First-Stage of Demand-Instrumented Change in ln(wage and salary employment)

	Effects of demand shock on:	
	County employment	Commuting zone (CZ) employment
Sample 1		
County instrument	0.8457	
	(0.0538)	
Sample 2		
CZ instrument		1.0684
		(0.0749)
Sample 3		
County instrument	0.5530	0.0526
	(0.0664)	(0.0270)
CZ instrument	0.6014	1.0920
	(0.0898)	(0.0725)

NOTE: These estimates are from the first stage of the three 2SLS regressions reported in Table 1. The dependent variable in the first stage is the change from year $t-1$ to year t in the natural logarithm of county or CZ employment. All other right-hand-side variables are also included. Standard errors are clustered at CZ level. For sample 3, industry mix instruments are defined at both county and CZ level and attempt to predict both county and CZ change in ln(wage and salary employment). All estimates are based on a sample that runs from 2005–2006 to 2015–2016. Sample 1 is 669 counties. Sample 2 is 240 CZs. Sample 3 is 609 counties in 225 CZs.

Table 3 Interaction of Lagged Labor Market Outcome Variables for Different Groups with Demand Shock, Commuting Zone (CZ) Model

		All 16+	Prime age	Noncollege (ages 25–64)
ln(Emp/Pop)	coefficient	−0.8988	−1.1047	−0.9447
	st. error	(0.2068)	(0.3104)	(0.2527)
	t-stat	−4.35	−3.56	−3.74
ln(LabForce/Pop)	coefficient	−0.8897	−1.0065	−0.8283
	st. error	(0.2491)	(0.4059)	(0.3115)
	t-stat	−3.57	−2.48	−2.66
ln(Emp/LabForce)	coefficient	−4.1643	−4.0231	−4.0113
	st. error	(0.8964)	(0.9836)	(0.8773)
	t-stat	−4.65	−4.09	−4.57

NOTES: These results are for the CZ sample of CZs 200K or above in population in 2005, 240 of them. The years are 11 years, from 2005–2006 to 2015–2015. Dependent variable is always change from year $t-1$ to year t in ln(Emp/Pop) for all 16+. The regression includes year dummies, and controls for both lagged ln(LabForce/Pop) and ln(Emp/LabForce) for all three of these demographic groups. The regressions also include demand shock to CZ from year $t-1$ to year t. The interaction term is between this demand shock, and an interaction term measuring a particular lagged (year $t-1$) labor market outcome for any of three groups. These results come from nine different regressions, which differ only in what interaction term is used. The interaction term is the year $t-1$ value of natural log of either (Emp/Pop), (LabForce/Pop), or (Emp/LabForce) for either everyone ages 16+, prime-age, or noncollege grads (ages 25–64). The first column indicates which lagged labor market outcome variable is used as an interaction. The results shown here are the coefficient on that interaction term, the standard error, and the *t*-statistic. Standard errors reflect clustering by CZ.

Table A1 Employment and Population for Counties and Commuting Zones (CZs) in Samples

	Mean	Standard deviation	Median	10th percentile	90th percentile
Sample 1: Counties of 65,000 population or higher (669 counties)					
Population in 2005	349,751	587,891	169,740	83,962	763,016
Employment in 2005	144,736	260,047	61,149	26,177	347,225
Sample 2: CZs of 200,000 population or higher (240 CZss)					
Population in 2005	1,100,868	1,726,099	540,938	235,917	2,424,440
Employment in 2005	435,679	711,127	187,138	73,111	1,034,898
Sample 3: Counties of 65,000 population or higher in CZs of 200,000 population or higher (609 counties in 225 CZs)					
County population in 2005	372,790	610,412	188,117	90,198	789,110
County employment in 2005	154,854	270,132	67,059	26,492	377,242
CZ population in 2005	1,157,178	1,768,553	558,420	247,063	2,541,753
CZ employment in 2005	459,157	728,473	201,715	78,662	1,095,388

NOTE: Population figures from census. Employment figures from WholeData, which is almost entirely private employment data originally derived from County Business Patterns.

Table A2 Summary Statistics on Key Variables, 2005–2016, 669 Counties

Variable	Mean	Standard deviation	Median	10th percentile	90th percentile
Δ ln(E/P) from year t−1 to t	−0.0041	0.0382	−0.0010	−0.0536	0.0395
ln(1+demand shock from year t−1 to t)	0.0068	0.0248	0.0161	−0.0303	0.0270
Δ ln(WS emp.) from year t−1 to t	0.0060	0.0261	0.0079	−0.0245	0.0342
ln(16+ E/P in year t−1)	−0.5332	0.1065	−0.5216	−0.6692	−0.4085
ln(16+ LF/P in year t−1)	−0.0823	0.0330	−0.0758	−0.1267	−0.0460
ln(16+ E/LF in year t−1)	−0.4509	0.0905	−0.4391	−0.5653	−0.3474
ln(prime-age E/P in year t−1)	−0.2713	0.0725	−0.2616	−0.3670	−0.1880
ln(prime-age LF/P in year t−1)	−0.0696	0.0312	−0.0638	−0.1129	−0.0351
ln(prime-age E/LF in year t−1)	−0.2016	0.0560	−0.1934	−0.2737	−0.1400
ln(noncollege 25–-64 E/P in year t−1)	−0.3767	0.0856	−0.3677	−0.4890	−0.2741
ln(noncollege 25–64 LF/P in year t−1)	−0.0815	0.0355	−0.0754	−0.1311	−0.0417
ln(noncollege 25–64 E/LF in year t−1)	−0.2952	0.0690	−0.2872	−0.3867	−0.2145

NOTE: See notes to Table 2 in paper. "E" is employment, "P" is population, "LF" is labor force, and "WS" is wage and salary. These are descriptive statistics for "Sample 1" in the paper, which is counties with 65,000 population in all years from 2005 on. The only difference in sources: the ACS data are taken from American Factfinder, not IPUMS microdata. Total observations is 669 × 11 = 7,359.

Table A3 Summary Statistics on Key Variables, 2005–2016, 240 Commuting Zones

Variable	Mean	Standard deviation	Median	10th percentile	90th percentile
Δ ln(E/P) from year $t-1$ to t	−0.0030	0.0298	0.0005	−0.0416	0.0305
ln(1+demand shock from year $t-1$ to t)	0.0064	0.0242	0.0161	−0.0287	0.0255
Δ ln(WS emp.) from year $t-1$ to t	0.0043	0.0221	0.0075	−0.0222	0.0275
ln(16+ E/P in year $t-1$)	−0.5373	0.0949	−0.5342	−0.6479	−0.4191
ln(16+ LF/P in year $t-1$)	−0.4554	0.0809	−0.4520	−0.5484	−0.3577
ln(16+ E/LF in year $t-1$)	−0.0819	0.0283	−0.0771	−0.1217	−0.0500
ln(prime-age E/P in year $t-1$)	−0.2686	0.0659	−0.2642	−0.3451	−0.1924
ln(prime-age LF/P in year $t-1$)	−0.1990	0.0514	−0.1939	−0.2601	−0.1439
ln(prime-age E/LF in year $t-1$)	−0.0696	0.0266	−0.0649	−0.1057	−0.0397
ln(noncollege 25–64 E/P in year $t-1$)	−0.3750	0.0815	−0.3712	−0.4679	−0.2781
ln(noncollege 25–64 LF/P in year $t-1$)	−0.2947	0.0679	−0.2890	−0.3741	−0.2176
ln(noncollege 25–64 E/LF in year $t-1$)	−0.0803	0.0300	−0.0746	−0.1243	−0.0465

NOTE: Data are for 240 commuting zones (CZs) with greater than 200K population in 2005, from years 2005 to 2016. Total observations is 240 × 11 = 2,640. These "Sample 2" descriptive statistics are very similar to d-stats for Sample 1 (counties) and Sample 3 (counties with CZ data as well), which are reported in Appendix Tables A2 and A4. The dependent variable, the change in the natural log of the employment to population ratio, are data from 2005–2006 to 2015–2016, and come from American Community Survey. The key demand shock variable is derived from the Upjohn Institute's WholeData, which overcomes suppressions in the County Business Patterns data. The underlying county data are at the six-digit NAICS level. The demand shock is national growth in proportional terms, weighted by each CZ's share of employment in year $t-1$ in each industry. The natural log of 1+ this demand shock term is the predicted change in the natural log of employment due to national demand shocks to the CZ's specialized industries. The actual change in the log of CZ wage and salary employment is also sometimes used as a right-hand-side variable, and is taken from the U.S. Bureau of Economic Analysis. The other variables are included as control variables always in the initial baseline run, and are later used as possible interaction variables with shocks to employment. These variables are calculated for year $t-1$ (years 2005–2015) from the American Community Survey.

Table A4 Summary Statistics on 609 Counties in 225 Commuting Zones, 2005–2016

Variable	Mean	Standard deviation	Median	10th percentile	90th percentile
Panel A: County Statistics					
Δ ln(emp/pop) from year *t*−1 to *t*	−0.0041	0.0375	−0.0008	−0.0527	0.0381
ln(1+demand shock from year *t*−1 to *t*)	0.0068	0.0247	0.0161	−0.0301	0.0270
Δ ln(WS emp.) from year i−1 to *t*	0.0061	0.0260	0.0080	−0.0245	0.0343
ln(16+ E/P in year *t*−1)	−0.5297	0.1057	−0.5180	−0.6625	−0.4070
ln(16+ LF/P in year *t*−1)	−0.4476	0.0898	−0.4354	−0.5605	−0.3460
ln(16+ E/LF in year *t*−1)	−0.0821	0.0328	−0.0758	−0.1267	−0.0460
ln(prime-age E/P in year *t*−1)	−0.2690	0.0719	−0.2591	−0.3627	−0.1872
ln(prime-age LF/P in year *t*−1)	−0.1996	0.0554	−0.1912	−0.2697	−0.1389
ln(prime-age E/LF in year *t*−1)	−0.0694	0.0309	−0.0635	−0.1118	−0.0352
ln(non-college 25–64 E/P in year *t*−1)	−0.3747	0.0847	−0.3660	−0.4837	−0.2739
ln(non-college 25–64 LF/P in year *t*−1)	−0.2931	0.0682	−0.2850	−0.3828	−0.2139
ln(non-college 25–64 E/LF in year *t*−1)	−0.0816	0.0353	−0.0754	−0.1309	−0.0423
Panel B: Commuting Zone Statistics					
ln(1+demand shock from year *t*−1 to *t*)	0.0070	0.0240	0.0168	−0.0275	0.0257
Δ ln(WS employment) from year *t*−1 to *t*	0.0053	0.0224	0.0087	−0.0221	0.0290
ln(16+ E/P in year *t*−1)	−0.5150	0.0810	−0.5057	−0.6202	−0.4171
ln(16+ LF/P in year *t*−1)	−0.4321	0.0687	−0.4229	−0.5218	−0.3532
ln(16+ E/LF in year *t*−1)	−0.0830	0.0274	−0.0773	−0.1220	−0.0529
ln(prime-age E/P in year *t*−1)	−0.2575	0.0511	−0.2527	−0.3218	−0.1987
ln(prime-age LF/P in year *t*−1)	−0.1870	0.0374	−0.1828	−0.2338	−0.1466
ln(prime-age E/LF in year *t*−1)	−0.0705	0.0255	−0.0656	−0.1053	−0.0422
ln(noncollege 25–64 E/P in year *t*−1)	−0.3647	0.0636	−0.3608	−0.4475	−0.2882
ln(noncollege 25–64 LF/P in year *t*−1)	−0.2808	0.0504	−0.2756	−0.3458	−0.2222
ln(noncollege 25–64 E/LF in year *t*−1)	−0.0839	0.0302	−0.0781	−0.1265	−0.0506

NOTE: See notes to Table 1, and to prior appendix tables. The counties and commuting zones (CZs) are all counties with 65,000 or more population in all years, that are also in CZs that have 200,000 population or more in 2005. County data on labor market outcomes comes from the ACS via American Factfinder. The CZ data on labor market outcomes comes from the ACS via microdata from IPUMS. The number of observations is 609 × 11 = 6,699 for all these data, so the same CZ is represented more than once for some counties; 225 CZs are in the data, so we would get slightly different statistics just using these 225 CZs × 11 = 2,375 observations as the data for the CZs.

Table A5 Effects of Demand Shocks, Job Growth, and Instrumented Job Growth, for Counties in Large Commuting Zones (>1 million population)

	Predicted demand shock	Job shock	Instrumented job shock
County	0.1063	0.1377	0.1876
	(0.0844)	(0.0339)	(0.1410)
Commuting zone	0.6691	0.2114	0.3578
	(0.1443)	(0.0526)	(0.1688)
Sum(Commuting zone +county)	0.7754	0.3491	0.5454
	(0.1383)	(0.0439)	(0.0914)

NOTE: This sample is of counties with more than 65K population in commuting zones (CZs) with more than 1 million population. The sample ends up being 353 counties in 70 CZs. As sample is 11 years, total number of observations in regressions are 353 × 11= 3,883 observations. All estimates include a complete set of year dummies and similar controls as in baseline regressions in paper. Robust standard errors are clustered at CZ level.

Table A6 Correlations between Pairs of Commuting Zone Years in ln(employment/population) Ratio and ln(employment/labor force) Ratio

LN(E/P)	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
2005	1													
2006	0.945	1												
2007	0.951	0.964	1											
2008	0.951	0.948	0.957	1										
2009	0.909	0.923	0.938	0.959	1									
2010	0.900	0.920	0.932	0.956	0.963	1								
2011	0.906	0.910	0.925	0.949	0.960	0.965	1							
2012	0.906	0.917	0.932	0.948	0.962	0.959	0.971	1						
2013	0.900	0.915	0.932	0.941	0.949	0.955	0.960	0.970	1					
2014	0.906	0.918	0.930	0.947	0.949	0.947	0.959	0.964	0.972	1				
2015	0.914	0.924	0.938	0.945	0.939	0.940	0.953	0.958	0.969	0.971	1			
2016	0.922	0.928	0.937	0.942	0.933	0.929	0.944	0.951	0.961	0.961	0.972	1		
2017	0.917	0.924	0.940	0.933	0.923	0.920	0.927	0.941	0.951	0.954	0.965	0.970	1	
2018	0.907	0.922	0.926	0.931	0.922	0.923	0.924	0.940	0.953	0.957	0.962	0.969	0.970	1

LN(E/L)	2005	2006	2007	2008	2009	2010	2011	2012	2013	2014	2015	2016	2017	2018
2005	1													
2006	0.788	1												
2007	0.667	0.753	1											
2008	0.536	0.604	0.719	1										
2009	0.486	0.518	0.665	0.832	1									
2010	0.469	0.483	0.626	0.818	0.894	1								
2011	0.524	0.473	0.592	0.761	0.826	0.860	1							
2012	0.562	0.527	0.631	0.779	0.786	0.833	0.861	1						
2013	0.566	0.557	0.594	0.713	0.738	0.757	0.810	0.852	1					
2014	0.623	0.569	0.572	0.661	0.680	0.687	0.758	0.818	0.826	1				
2015	0.606	0.581	0.570	0.627	0.609	0.624	0.690	0.758	0.800	0.852	1			
2016	0.639	0.552	0.517	0.535	0.505	0.509	0.606	0.661	0.733	0.739	0.800	1		
2017	0.613	0.525	0.527	0.497	0.474	0.474	0.578	0.634	0.682	0.737	0.770	0.826	1	
2018	0.613	0.575	0.522	0.551	0.525	0.528	0.645	0.656	0.691	0.771	0.747	0.785	0.755	1

NOTE: Based on 240 commuting zones with greater than 200K population in 2005. The statistics are simple unweighted correlations between two different years. Average of ln(E/P) correlations is 0.941, average of ln(E/L) is 0.658. Five-year correlations (2005 and 2010 up through 2013 and 2018) range from 0.900 to 0.953 for ln(E/P), from 0.469 through 0.691 for ln(E/L).