Online Appendix for: Local Ties in Spatial Equilibrium

Mike Zabek

October 22, 2019

Appendix A Data

The data comes primarily from the decennial census and ACS as collected by IPUMS at the University of Minnesota (Ruggles et al. (2010)). Data on the impact of trade on individual local labor markets comes from Autor, Dorn and Hanson (2013), and the vital statistics data comes from the NBER. I restrict my sample to prime-aged (16-64 inclusive) people not living in group quarters (barracks and dorms). I focus on workers who receive market wages by excluding unpaid family workers and workers who did not work for pay last year. Generally, I aggregate these data up to the Commuting Zone (Tolbert and Sizer (1996)) level and perform my analyses at this level, except in some cases where I focused on states to better match Vital Statistics and migration data.

IPUMS

The data from the U.S. Census comes via the IPUMS sample detailed in Ruggles et al. (2010). I use several PUMS samples: For 1970 I use the form 1 one percent sample at either the state or metro level, depending on whether the analysis uses states or commuting zones. For 1980, 1990, and 2000 I use the five percent samples. For 2008 I use the ACS 3 year estimates from 2006 to 2008. For the bulk of specifications I exclude people residing in group quarters, such as military barracks or dormitories. The only exception is the growth accounting by state that I performed in section two. In that case I include people residing in group quarters because this exclusion might cause me to lose young adults born 16 to 21 years earlier. For worker wages I exclude unpaid family workers and only include people who worked last year. In regressions using commuting zone data in 1970 I exclude 1990 commuting zone number 24600 because I suspect its geographic definition was mis-coded.[1]

I also compute "labor supply weights" following Autor and Dorn (2013) that weight each worker by their total hours worked last year, and I exclude the top and bottom 1 percent

[1] Using the commuting zone crosswalks from David Dorn suggests that its population in 1970 was ten times larger than its population in 1980.

of wages from the computation. All wages are deflated using the personal consumption expenditure chain type price index available from the Federal Reserve Bank of St. Louis via their FRED service. The reference year is 2007.

Birth Locations

The best available data I have access to concerns workers' states of birth. Unfortunately, this is the most detailed geography that the census bureau asks for, so it is impossible to determine precisely what local labor market a respondent was born in without using an outside data source. Consequently, I tally the proportion of residents of a local labor market who are living in the state of their birth. For large states with many local labor markets (California and Texas are examples) this should lead me to overshoot the proportion of residents living in the area of their birth. For labor markets that cut across state lines (New York for example) I would be understating the proportion of residents living in the same area they were born in since a resident could be in a different state, but the same commuting zone they were born in. On the one hand, imprecision in the measure of the proportion of residents born in the same local labor market is a concern. It is important to note that in a world where areas are not unique islands, the ideal geographic construct may be different in terms of work, family, and other considerations. For example, a worker may prefer to live further away from her parents compared with her work and a worker living in the same state but a different commuting zone as they were born might be almost as constrained as a worker living in the same commuting zone.

Another other issue with the variable is that most births are in hospitals and sometimes children will be born in a hospital in a different state from where their mother lives. Bartik (2009), for example, documents this using data from the PSID. In this situation, the question appears to ask for the state of the hospital, which is a poor proxy for the concepts I am examining. While this variable is far from perfect, its concordance with other measures of a respondent's local "ties" such as their tenure in their home should suggest that it is still meaningful for this application.

For all of these reasons, I include alternative specifications that use alternative measures of local ties. Generally these results are quite similar.

Local Labor Markets

I define a local area for this project as a Commuting Zone (CZ) defined by Tolbert and Sizer (1996). Commuting Zones are designed to reflect local labor markets where workers live and work, based on commuting data collected in the 1990 Census.[2] A given CZ can contain multiple states and states can contain multiple CZs. CZs are quite similar to Metropolitan Statistical Areas (MSAs) that are more commonly used, but CZs also include rural areas, covering the entire area of each of the 50 states. They are constructed to be an ideal analogue to the areas in traditional models of migration where workers live and work in the same area. To merge the IPUMS data I use in my specification I use the crosswalks created by David Dorn and available via his academic website. For historical charts, I exclude commuting zone 24600, which I believe may be improperly coded in 1970.

Appendix B Estimating Migration Elasticities

It is possible to directly measure the migration elasticities by measuring changes in population after the Bartik and China demand shocks. The equation below shows the basic empirical specification that I use to recover the migration elasticity, $\eta_{\text{Mig},j}$. The migration elasticity, $\eta_{\text{Mig},j}$, measures the effect of an increase in log incomes on log population, including the endogenous responses of other local prices. Since I intend to include the effects of these other local prices, like housing prices, I do not attempt to control for them. Following the reduced form results, I do control for decade fixed effects, γ_t, and the standard set of

[2] Different Commuting Zones exist following the 2000 census, however I keep with Autor, Dorn and Hanson (2013) and use the 1990 definitions. I do this to keep CZ definitions constant and I use 1990 because it reflects local areas at the beginning of the sample.

controls from the reduced form regressions, βX_{jt}. These ensure that the regressions are not being driven by different trends for areas where people are of different ages, different education levels, or places where more people are foreign born, for example. Following my earlier regressions, I allow heterogeneity across areas, j, by splitting areas into bins based on their levels of local ties, and also by including a continuous interaction with the level of local ties in each area.

$$\Delta \log \text{pop}_{jt} = \eta_{\text{Mig},j} \Delta \text{income}_{jt} + \gamma_t + \beta X_{jt} + \epsilon_{jt} \tag{1}$$

To isolate plausibly exogenous changes in local incomes, I use both the Bartik shifters in the 1980s and the Chinese import measures in the 1990s and early 2000s.[3] To maximize power, I stack the data for each of the three decades and estimate one set of parameters in the second stage. I allow the Bartik instruments to have different first stage effects from the trade instruments, but I assume the impact of the trade instruments is the same in each decade. [4]

I measure changes in incomes by combining information about changes in wages with information about the availability of jobs, as measured by the employment to population ratio. Wages are an imperfect measure of labor incomes because there appear to be significant frictions to their adjustments, particularly in periods when labor demand is falling. Workers and employees may be reluctant to accept declines in nominal wages, for example, and search frictions could also play a role.

In my empirical setup, labor incomes are the product of wages once one is employed times one's probability of being employed, as in Harris and Todaro (1970). Potential migrants

[4] Another point about the instruments is that the use of labor incomes abstracts from people's labor leisure choices. In my model, and in much of the literature on spatial equilibrium, an increase in labor incomes has an identical effect as an equivalent increase in local subsidies, because people work for a fixed number of hours in the place where they live. By the logic of the sufficient statistics derivation, however, the impacts on people's labor leisure choice should fall out from the first order welfare impacts of a local subsidy. Intuitively, people are roughly indifferent about working more or searching harder for a job. The most serious limitation for my empirical work appears to in terms of attracting population; local subsidies may be more or less appealing to migrants than increases in wages.

consider not only wages, but also the difficulty of finding and keeping a job. I use the employment to population ratio as a measure of this probability. Changes in log labor income, then, are changes in log wages, Δwage_{jt}, plus changes in the local employment to population ratio, $\Delta\text{emp ratio}_{jt}$.

$$\Delta\text{income}_{jt} = \Delta\text{wage}_{jt} + \Delta\text{emp ratio}_{jt}$$

The estimated migration elasticities – reported in Table 1 – are an order of magnitude lower in areas with higher levels of local ties. The migration elasticities in column one are 0.08 in high ties areas and 0.99 in low ties areas. These are statistically significantly different from one another at the ten percent level. The instrument also appears to be strong enough to support this inference, since the first stage since traditional Wald and Kleibergen-Paap corrected Wald F statistics are above traditional thresholds.

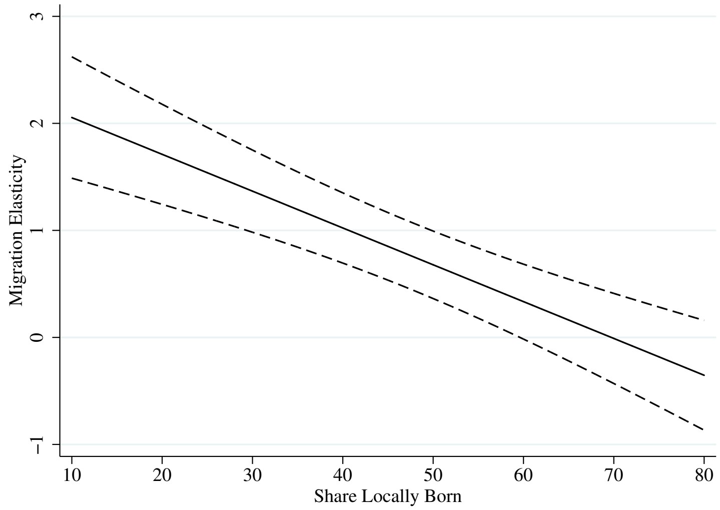
The slope of the continuous linear interaction in column two of Table 1 implies that migration elasticities decline by around 0.35 for every 10 percent increase in the share of locals. I also plot the estimate in Figure 1. To get an idea of the magnitudes, around 15 percent of people who live in Miami were born in Florida. So the estimated migration elasticity is around two in Miami. In areas with very high levels of local ties, the estimated migration elasticity reaches zero. The continuous linear interaction term is statistically different from zero at the one percent level, and the regression also passes conventional thresholds for having a sufficiently strong instrument.

Table 1: Estimated Migration Elasticities from Demand Shocks

	(1)	(2)	(3)
High Ties: Incomes	0.08		
	(0.24)		
Low Ties: Incomes	0.99		
	(0.39)		
High Ties Indicator	-0.95		
	(1.14)		
Main Effect of Incomes		2.40	1.35
		(0.34)	(0.58)
Interaction (x100)		-3.44	
		(0.64)	
Main Effect of ties		-0.14	
		(0.07)	
Year Fixed Effects	Y	Y	Y
Controls	Y	Y	Y
P-val: No diff	0.03	0.00	
First stage F: Wald	35	37	46
First stage F: K-P	12	11	14
Observations	2166	2166	2166

Note: Estimated migration elasticities are much lower in areas with high shares of locally born residents, or high local ties. This table displays the estimated coefficients from a regression using the two labor demand shocks – due to trade and national industry level changes – to instrument for log incomes in a regression predicting log population. So, the estimated relationship is an estimated migration elasticity. The regressions use the standard set of controls and standard errors clustered at the state level, as in Table 2 of the main text. The statistics at the bottom report a Wald test for no difference in the elasticities between low and high ties areas, the first stage partial Wald F statistic and Kleibergen-Paap corrected Wald F statistic.

Figure 1: Estimated Migration Elasticities



Note: Estimated migration elasticities are much lower in areas with high shares of locally born residents, or high local ties. This figure displays the migration elasticities implied by column two of Table 1 based on the methodology reported in that table. The dotted lines represent 95 percent confidence intervals of the values.

References

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

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

Bartik, Timothy. 2009. "What proportion of children stay in the same location as adults, and how does this vary across location and groups?" W.E. Upjohn Institute for Employment Research Upjohn Institute Working Paper 09-145.

Harris, John, and Michael Todaro. 1970. "Migration, Unemployment and Development: A Two-Sector Analysis." *American Economic Review*, 60(1): 126–142.

Ruggles, Steven, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. 2010. "Integrated Public Use Microdata Series: Version 5.0."

Tolbert, Charles M., and Molly Sizer. 1996. "U.S. Commuting Zones and Labor Market Areas: A 1990 Update." Rural Economy Division, Economic Research service, U.S. Department of Agriculture Staff Paper AGES-9614, Washington, DC.