A More Generalized Way-
The Effect of Internal Migration on Local Labor Markets: using American Cities in 1980

Grace Wang[1]
University of Arizona

November 9, 2016

Abstract

This paper investigates the causal effect of internal migration on three local labor market outcomes: earnings, employment and the probability of migrating out. For the endogeneity concern that migrants may be attracted to cities with good economies, I construct predicted in-migration and out-migration rates as instruments for the net-migration rate in two steps. First, I use the variation in the shift-share measures and extreme weather events to predict the migrant flows to and from U.S. cities. Then I estimate migrants' settlement patterns by distance between the sending area and the receiving area. The predicted in-migration (out-migration) rate is the sum over all sending (receiving) areas of the products between predicted out-migration (in-migration) flows and the corresponding probabilities of migrating to (from) that destination. From the results, the net-migration had little effect on the earning outcomes except for the the hours worked in a usual week. Also, existing residents seem not to be crowded out by the in-migrants. However, net-migration flows into a metropolitan area significantly increases the employment status of existing residents. Relative to being part-time and unemployed, a one percentage increase in the net-migration rate led to a 0.335% higher probability of being full-time employed. The probability of being in part-time employment was decreased by 0.171% and the unemployment probability was reduced by 0.164% with a one percentage increase in the net-migration rate. These results imply that instead of direct wage reduction, workers have to work longer under the same wage payment in response to the internal migration.

[1]I thank my advisor Price V.Fishback and Stanley Reynolds for useful comments. All errors are mine.

1 Introduction

This paper studies how local labor markets respond to internal migration in United States in the 1980s. Even though most studies examine the effect on international immigration, internal migration is still the main source of in-migration into cities. In the 1980 Integrated Public Use Microdata Series, internal migration accounts for around 98% of the migrants in the sample. This substantial mobility in the United States in the past decades has renewed public debate about the effect of internal migration on local labor markets. I examine the effect of internal migration on the three aspects of the labor market: earnings, employment and probability of out-migration.

One challenge in studying this effect is the endogeneity problem in the immigration literature. Migrants are not randomly assigned to destinations. Instead, they tend to be attracted to cities with high wages or strong economies. Natural policy changes or catastrophes can provide a unique laboratory to explore the causal impact of migration on labor markets. Card (1990) applies a difference-in-difference method to examine how the labor inflow from the Mariel Boatlift changed labor market outcomes by increasing the labor supply. Similar work about migration from Hurricane Katrina includes McIntosh (2008), Imberman et al. (2012), and Groen and Polivka (2008). However, in most empirical situations, we do not have natural disasters to follow up with. Also, prospective migrants make their settlement choice based on the expected benefits and cost of migration. In this case, the endogeneity of migration should be considered. Boustan et al. (2010) construct predicted in-migration and out-migration rates as instruments for actual net-migration rates. In the first step, they use data on a set of "push factors" ("pull factors") to instrument for out-migration (in-migration) rates during the Great Depression. These factors contain variation in the generosity of New Deal policies, which takes advantage of particularly exogenous policy change during that period, and the presence of extreme weather conditions.

One contribution of this paper is to combine the methods used by Boustan et al. (2010) to construct instruments and using shift-share variables in Blanchard et al. (1992) as one of the "push" and "pull" factors that contribute to local out-migration and in-migration. This shift-share approach has been widely used in Wallis and Benjamin (1981), Blanchard et al. (1992), McGuire and Bartik (1992), Bartik (2002), Hoynes et al. (2012) and Bound and Holzer (2000). One condition in applying this method is that geographic areas differ sufficiently in their sectoral employment composition, otherwise, this constructed shift-share shock will not explain the variation of migration rates across divisions. This condition is satisfied in this paper by the strong relationship between shift-share instruments and in-migration and out-migration rates.

The method for constructing the instruments involves several steps. Since the procedure to construct the instrument for the out-migration rate is similar, I explain the instrument construction for in-migration rate in details as an example. First, I regress the out-migration rate of sending area k on "push factors" and region dummies, for all geographic areas. This regression is used to predict the out-migration rate of all migrants who leave the area k. The "push factors" contains shift-share shocks, following Blanchard et al. (1992) and extreme weather conditions. I select 42 two-digit SIC industry sections with complete data information to construct the shift-share measure, based on Blanchard et al. (1992). Also, I include the extreme weather conditions about average temperature, average precipitation, the months of extreme drought and wetness, and their interaction with the metropolitan dummy among the "push factors". The total estimated out-migration flow from area k is the product of predicted out-migration rate and local population of area k. In the next step, I predict the probabilities that a migrant leaving a source area k picks the destination location j based on its distance to the sending area k. The settlement patterns are area specific. I run regressions for each of the 269 sending areas and then predict the share of out-migrants from that city who go to each location. Afterwards, the predicted in-migration flow into the receiving area j is the sum over all sending areas k of the predicted out-migration flows in city k, multiplied by and the corresponding probabilities of migrating to the destination j. I then find the in-migration rates in the receiving area j through dividing the total predicted in-migration flow in j by its corresponding population size. The instrument for the actual out-migration rate is constructed in the similar way.

The IV results show that in-migration into a metropolitan area has significant effect on the employment status for existing residents. Relative to being part-time and unemployed, a one percentage increase in the net-migration rate led to a 0.335% higher probability of being in full-time employment. The probability of part-time employment was decreased by 0.171% and the probability of unemployment was reduced by 0.164% with a one percentage increase in net-migration rate. However, wage did not respond to the migration supply shocks in a statistically significant way. All the regression coefficients for wage are statistically insignificant. Besides, the coefficient of net migration on hourly wage is larger than that of weekly earnings and annual earnings. The effects of net-migration on working time are positive in all the IV specifications, but only the coefficient of hours worked is statistically significant at the 5% level. Altogether, it implies that instead of direct wage reduction, workers have to work longer under the same wage payment. Also, existing residents seem not to be crowded out by the in-migrants since there is no significant effect on the probability of migrating out.

The next section provides an overview of studies about the economic effects of immigration. Section 3 describes data construction about internal migration rates and economics outcomes and some preliminary statistical analysis. Section 4 specifies the reduced form model and Section 5 deals with the endogeneity concern and the instrument construction. Section 6 reports results from three main economic outcomes and possible explanations. Section 7 summarizes the main findings of this paper and some further work.

2 Previous Work on the Economic Effects of Immigration

Earlier studies on the economic effects of Immigration arrivals in the United States have mixed conclusions. Goldin (1994) documents that at the turn of twentieth century, mass migration from Europe led to a large reduction in the wages of native workers in cities which had high in-migration rates. Borjas et al. (1996) and Borjas (2003) also show that immigration has a pronounced negative effect on natives' wages. Although some studies have detected a wage response to the immigration across cities, many, using modern data, find no effect on wages at all. For example, Card (1990) finds that the Mariel Boatlift influx from Cuba, which led an increase of Miami labor force by 7%, had virtually no discernible effect on the wages or unemployment rates of less-skilled native-born workers relative to comparison cities. Card (2009) says that the discrepancy between these findings can be explained by recognizing the imperfect substitutability between natives and immigrants as well as the small degree of substitutability between high school graduates and dropouts. Unlike many immigrants, internal migrants are more likely to speak English fluently, be educated in American schools, and may be difficult to distinguish in appearance and accent. Therefore, internal migrants may be closer substitutes to existing residents and thus, their migration could lead to a strong negative effects on natives' wage by increasing labor supply. However, recently Frank (2009) finds that the arrival of East Germans into West Germany after the fall of the Berlin Wall had no significant effect on native wages, but did have a positive and short-lived effect on employment. Overall, these different estimates of the effects of immigration on native outcomes could reflect a true change over time in regard to the economic response to a local labor supply shock or simply the result of difference in feasible measures and research design.

One concern about the relationship between immigration and wages in labor markets is that the local areas may adjust by other margins to the migrant supply shocks. In a basic model of the labor market, migration flows represent an outward shift in the labor supply. It would lead to an unambiguous decrease in wages, unless there is also an increased labor demand for locally produced goods and services that sufficiently offsets the increase in labor supply. Some related literature studies the relationship between immigration and subsequent native internal migration. Borjas et al. (1997) find that new arrivals to a city would induce some existing workforce to relocate, spreading the economic costs of immigration to other markets. Blanchard et al. (1992) point out that wage or employment adjustments to a labor supply shock might be tempered by the in-migration of capital or the out-migration of labor, with the free factor mobility between cities. Therefore, based on Borjas (2003) and Ottaviano and Peri (2012), the biased downward effect of immigration on wages might be larger at national level, compared to what local labor markets would suggest.

The empirical results about factor mobility responding to immigrant arrivals is also mixed. While some studies find that natives would respond to immigration by migrating out to other areas (see Filer (1992), Borjas (2006), Boustan et al. (2010)), other studies have not found a statistically significant out-migration response to immigrant arrivals (see Wright et al. (1997), Card and DiNardo (2000), Card (2001) and Kritz and Gurak (2001)). On the capital side, Lewis et al. (2004) and Lewis (2003) suggest that local labor markets would proceed a slower adoption of skill-intensive computer and production technology to adapt to low-skilled immigrants.

3 Data on Internal Migration and Labor Market Outcomes

In order to provide a comprehensive picture of local adjustment to internal migration, I investigate three economics activities: earnings, employment and out-migration of existing residents. I combine individual-level internal migration information into in-migration and out-migration rates at the metropolitan and balance-of-state level [1] from the 1980 Integrated Public Use Microdata Series (IPUMS) dataset. This dataset also contains information on individual-level earnings, employment and out-migration for city residents, together with a set of socioeconomic and demographic controls.

3.1 Internal Migration

The Integrated Public Use Microdata Series (IPUMS) in 1980 gathered information on the internal mobility in the United States from census questions in 1980 on both the respondents' current location in 1980 and their former residences in 1975. I use this information to aggregate the population flows to calculate 5-year internal migration rates at the metropolitan area level. Specifically, I calculate the number of migrants arriving in (or leaving) a city between 1975 and 1980 as a share of the existing population in year 1975[2]. Furthermore, to instrument in- and out-migration rates, I also define the portion of each state not contained in any of the large metropolitan areas in the sample of that state as the balance-of-state. For migration, the indicator move is defined as 1 if a person's current residence in 1980 is different from the resident location in 1975[3]. Since the migration sample accounts for only 46.45% of the overall census sample, I divide the

[1] The portion of each state not contained in any of the large metropolitan areas in the sample of that state is defined as the balance-of-state.

[2] In this paper, the population in year 1975 is estimated based on the population in 1980 from IPUMS-USA and adjusted by the national growth rate for the year 1975. From U.S. Bureau of the Census, total national population in year 1975 is 215464000 while in year 1980, the total population is 226542250, thus, the national adjusted growth rate for year 1975 is 0.95109853. While it is probably better to avoid using data from 1980, the IPUMS-USA in 1970 do not match 1980 IPUMS-USA data well. Another way is to use the population data by county level from the Census Bureau's Population Estimates Program on NBER website, considering the possible inconsistency in county and metropolitan areas, I use adjusted population level from IPUMS in 1980 instead.

[3] Meanwhile, I re-define the indicator move if the current metropolitan area or balance-of-state is the same as five years ago. For metropolitan areas, in the current location, there is no observation in New Haven(5482) while in the migration sample, there is no observation in New Haven-Meriden(5480). I combine both metropolitan areas into one, denoted as New Haven-Meriden(5480). Also, there is no observation in Norfolk-Virginia Beach-Newport News(5720) in current location while in the migration sample, there is no observation in Newport News-Hampton(5721) and Norfolk-Virginia Beach-Portsmouth(5722). Similarly, I combine these three metropolitan areas as Norfolk-Virginia Beach-Newport News(5720). Besides, based on the report Incompletely Identified Metropolitan Areas from IPUMS-USA, I adjusted the population by metropolitan areas for

aggregate flows by 0.4645 to get the absolute numbers of the migration population at metropolitan level. International migration constituted only 1.9% in the whole migration sample, thus I have not included this effect and foreign arrivals are excluded from the migration sample in my further analysis.

In- and out-migration flows are available for 220 metropolitan areas and 49 balance-of-states. I drop counties which belong to multiple cities as well as cities containing these counties[4], thus, the metropolitan areas in this paper contains counties which are uniquely belonged to one city, based on Historical Statistical Area Delineations from United States Census Bureau. The reason for doing this is that when calculating migration flows between cities from county level data, counties belonging to multiple cities cannot be precisely divided into the corresponding city. This selection criterion may affect the precision of economic activities in balance-of-states, but since balance-of-states are only used for estimating instruments for in- and out-migration rates, it won't affect the regression outcomes at city level. At this moment, I have not distinguished migration flows by race, gender, age and education level, instead, they are set as control variables in regression analysis. For migration rates calculation, it should be noted that I include movers of all ages, while some of them may be too young or old to participate in the labor market.

3.2 Economic Outcomes

I use individual records from the 1980 IPUMS to construct a sample of non-migrants residing in the 220 metropolitan areas of interest. In this paper, I limit my observations to individuals between the ages of 18 and 65 who were reported as in the labor force and did not serve as unpaid family workers or in the armed force.

I investigate the impact of internal migration on three earning measures: annual earnings, weekly wage, and hourly wage. I also examine the effect on weeks worked last year, whether working less than 26 weeks and hours worked for a usual week. These variables are computed conditional on men who reported being employed and not self-employed in 1980, working a positive number of weeks (hours) in the past year (week) and earning a positive income last year. The hourly wage is estimated by dividing the weekly wage by hours worked in a usual week [5]. To examine the migration effect on employment, I define the observation as full-time employed when the individual worked more than 35 hours in a usual week in 1980 [6]. Part-time

consistency since there are 18666 (0.164558% in the whole sample) observations in the sample which do not have metropolitan names while it is in the metro area but outside the principal cities.

[4]In total, there are 298 cities in 1980 IPUMS, while there are 25 cities that contain counties which belong to 2 or 3 or 4 cities. I also drop cities that mismatch with other dataset. All the remaining cities have complete datasets and totally, there are 220 metropolitan areas and 49 balance-of-states.

[5]This calculated variable is particularly subject to measurement error because the hours worked in a usual week is only reported in the census week.

[6]While current Affordable Care Act define working 30 hours or more per week in the United States as being full-time employed; at that time, generally, working 35 hours per week was regarded as being in full-time employment. Besides, the Fair Labor Standards Act (FLSA) does not define full-time employment or part-time employment. I have also 30 hours as a cutoff and the results are robust.

employment is defined when the worker is reported as employed who working less than 35 hours in a usual week. For the probability of out-migration, I set an indicator equal to one when the observation is defined as in-migrant based on the migration information. For analysis about existing residents, the estimates can be interpreted as a treatment effect of migration on the existing workforce rather than a compositional change in the labor force as new workers arrive. Moreover, I controlled age, education, race and gender at individual level, in case selective out-migration changed the composition of the remaining labor force.

3.3 Summary Statistics

From Table 1, the mean in- and out-migration rates are both around 17%, implying high mobility in the labor force in the United States from 1975 to 1980. On average, the economic outcomes during this period was strong with 82.6% full-time employed and only 5.4% unemployed. The earning outcomes about annual earnings, weekly wage and hourly wage have comparatively higher standard deviations. One cause is that at this moment, the data did not distinguish employment status.

Table 1: **Summary Statistics at the Metropolitan Area Level**

	Mean	SD
Dependent variable:		
Annual earnings	15391.52	11339.09
Weekly wage	303.5491	278.5971
Hourly wage	8.138166	22.5335
Weeks worked	47.33653	9.886862
Working less than 26 weeks	.0597228	.2369729
Hours worked per week	39.63495	10.38801
Full time	.8257579	.3793179
Part time	.1202349	.3252363
Idle	.0540072	.2260321
Left from original area in 1975, 1975-80	.1648627	.3710568
Explanatory variable (1975-80):		
In-migration rate	.179369	.0819072
Out-migration rate	.1739291	.0428636
Net-migration rate	.0054399	.0637722
Instruments (predicted rate):		
In-migration, all	.2278633	.0840531
Out-migration, all	.2504557	.0854195
In-migration, out of state	.1110989	.0475162
Out-migration, out of state	.1199903	.0428231
In-migration, out of region	.1446389	.0537653
Out-migration, out of region	.1576349	.0544276

Note: The sample for different dependent variables varies in earning (6 in total), employment (3 in total) and probability of migrating out (1 in total). Sample size corresponds to the number of observations in the regression.

The right shift of the labor supply curve from net-migration implies a reduction in earnings and an increase in employment. This prediction seem apparent even in the raw data about the in-migration rate[7]. The scatter plot in Figure 1 shows a significant negative relationship between the in-migration rate from 1975 to 1980 and the mean personal income in 1980 by city. Figure 2 depicts a significant positive link between the in-migration rate from 1975 to 1980 and the employment rate in 1980. In Figure 3, it shows that there is a significant positive relationship between the in-migration rate and the out-migration rate for all metropolitan cities between 1975 and 1980, using the out-migration rate proxy for the probability of migrating out of the metropolitan area.

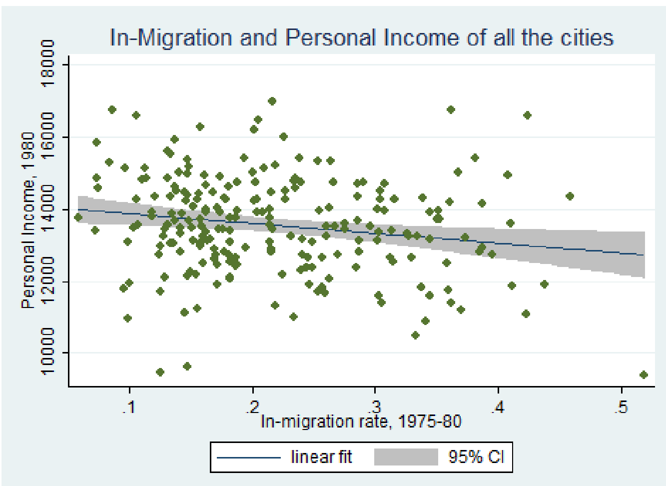

Figure 1: In-Migration Rate and Personal Income of all the cities, 1975-1980

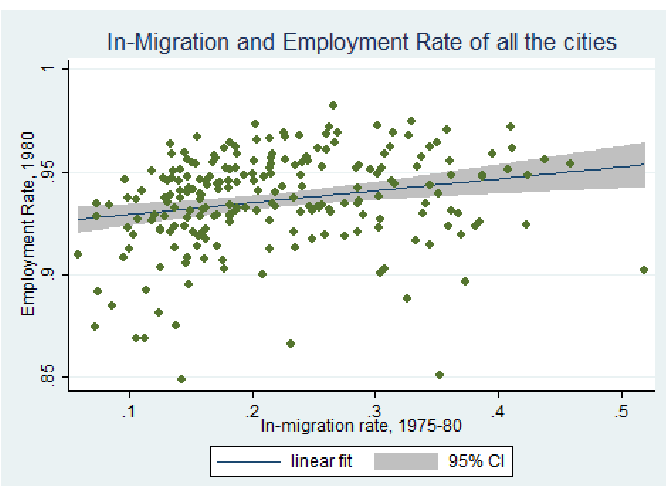

Figure 2: In-Migration Rate and Employment Rate of all the cities, 1975-1980

[7]These patterns also apply for the net-migration rate.

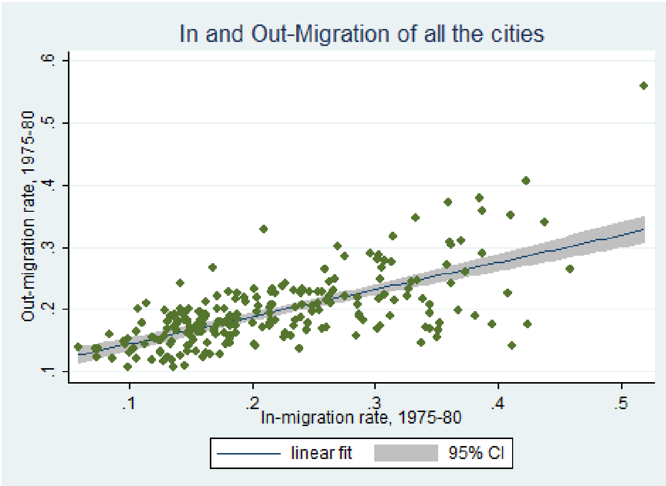

Figure 3: In and Out-Migration Rate of all the cities, 1975-1980

4 Main Model Specification

The main question of interest is the effect of internal migration into and out of a metropolitan area on earnings, employment and the probability of out-migration of existing residents. Y_{ijr80} represents an economic outcome for a non-migrant i who lives in metropolitan area j in region r in year 1980. $Y_{jr(75-70)}$ is a vector of information about economic conditions in 1970s because the information is not available in 1975. I posit Y_{ijr80} as a function of the migrant-induced change in labor supply:

$$Y_{ijr80} = \alpha + \beta n_{jr(80-75)} + \phi^{'} Y_{jr(75-70)} + \psi^{'} X_{ijr80} + \Pi_r + \epsilon_{ijr80}, \tag{1}$$

$n_{jr(80-75)}$ represents the net migration rate to area j from 1975 to 1980. There is a more flexible specification for the above estimation which allows in- and out-migration to have distinct effects on labor market outcomes.[8]

In this paper, for simplification, I restrict that in-migration and out-migration rates have equal and opposite effects on the economic outcomes, so β comes from the net effect of in-migration and out-migration rates. When instrumenting for the net-migration rate, I construct predicted in-migration and out-migration rates as instruments for both in- and out-migration rates[9].

[8] The equation can be expressed as $Y_{ijr80} = \alpha + \beta m_{jr(80-75)} + \gamma o_{jr(80-75)} + \phi^{'} Y_{jr(75-70)} + \psi^{'} X_{ijr80} + \Pi_r + \epsilon_{ijr80}$. $m_{jr(80-75)}$ represents the in-migration rate to area j and $o_{jr(80-75)}$ represents the out-migration rate from area j. In the earnings and employment equations, β is expected to be negative and γ is expected to be positive. For simplification, I restrict arrivals and departures to exert equal and opposite effects by including only the net migration rate on the right-hand side, that is, $n_{jr(80-75)} = m_{jr(80-75)} - o_{jr(80-75)}$, which is equivalent to equation (1) above.

[9] In this paper, the instruments are for actual net-migration rate since I constrain on the equal effects from in-migration and

I expect the equilibrium wage level in the area j to be affected by the migration flows. It will fall if the only effect comes through increased labor supply. It might rise if the dominant effect is a rise in labor demand associated with a rise in demand for local goods and services. Variable $Y_{jr(75-70)}$ measures the level of earnings or employment of individuals in area j in 1970s. Since IPUMS does not include the information about individual's earnings or employment status five years ago[10], to control for the pre-conditions, I use county-level estimates of wages and employment from prior years instead. I then aggregate them into each metropolitan area or balance-of-state level based on the file called Standard Metropolitan Statistical Areas and Components, issued on June 30, 1981, from Historical Statistical Area Delineations of United States Census Bureau. The county-level annual wage and employment in 1975 are from the Inter-university Consortium for Political and Social Research (ICPSR) about Historical, Demographic, Economic, and Social Data of the United States[11]. For this employment rate, I use data from the 1972 County Data Book which includes total civilian population that was employed in 1970 [12]. For payroll, I average over the payment per employee in manufacturing, retail trade, selected service, wholesale trade and mineral industries, weighted by the number of employees in each section, in 1972 from the 1977 County Data Book[13]. Meanwhile, for $Y_{jr(75-70)}$, I also control for the share of the city's population that were black, female, mean schooling years and mean age by metropolitan area in 1970. These measures are from 1972 County Data Book in Inter-university Consortium for Political and Social Research (ICPSR) about Historical, Demographic, Economic, and Social Data of the United States[14]. This dataset contains demographic information by county level in 1970, originally from U.S. Bureau of the Census in 1970. X_{ijr80} controls for a series of individual characteristics, including race, a cubic polynomial in age and a set of education dummies for schooling

out-migration rates.

[10] IPUMS does include the information about change in income or employment status, but in weekly or yearly level. Another option is to use CPS dataset, but considering the limited sample size and coverage in CPS and as well as the comparability in the identification of metropolitan areas, CPS dataset is not a good choice for this analysis.

[11] Another alternative is to use Inter-university Consortium for Political and Social Research (ICPSR) about County Business Patterns from 1975 to 1980 which is used to construct the local exogenous demand shocks. One problem in applying this dataset here is that it withholds some data about employment size and payroll by industry for the disclosure concern. Instead, it denotes data suppression flag for employment size class and both employment and payroll data are replaced by zeros. The employment size class for each flag is as following: A 0-19; B 20-99; C 100-249; E 250-499; F 500-999; G 1.000-2.499; H 2,500-4.999; I 5.000-9,999; J 10,000-24.999; K 25,000-49,999; L 50,000-99,999; M 100,000 or More. I applied this dataset to calculate the employment rate and payroll by areas by replacing the minimum value in each range for the corresponding flag. Because of the measurement errors, the estimates are of little variations, thus I did not use this dataset for my further analysis.

[12] I choose to use the county level data book rather than the city level data book since there are fewer cities than the counties for each metropolitan area, which may lead larger measurement errors and inconsistency in metropolitan-level variables estimation.

[13] 1972 County Level Data Book has related data in 1967. I choose 1972 data from the 1977 county data book since it is closer to the pre-condition year 1975. Also, both county data books have median family income in 1969 by county level. Though this variable has similar summary statistics as the estimated payroll from 1977 county data book, since the main regression is in individual level and this family-level variable is less representative to account for the city-level payroll information.

[14] An ideal way is to get the information in year 1975 and IPUMS-CPS does have it. But there are two problems in using CPS dataset. First, the definition of county in 1975 and 1980 has changed: in 1970, the smallest identifiable geographic units are metropolitan areas which are the combinations of counties totaling at least 250,000 population; while in 1980 the minimum total population level to define a county is 100,000. Second, the sample size and coverage in CPS is comparatively small since there are only 23 states and 33 metropolitan areas in the whole dataset in year 1975. Another choice is IPUMS-USA in 1970. But there are 96 metropolitan areas that mis-match with the ones in 1980, which causes insufficient observations for instrument construction.

attendance. Finally, I also enter a set of regional dummies Π_r, 9 in total, to control for the geographical variation, following Boustan et al. (2010), Rosenbloom and Sundstrom (1999) and Wallis (1989).

Standard errors are clustered at the metropolitan level to allow for the correlation of economic shocks faced by individuals in the same labor markets. While Boustan et al. (2010) use the inverse of city population as a weight for each regression to ensure that each city contributes equally to the estimation, in this paper, I only cluster by city and use White correction for the standard errors. Because when adding interaction terms between the dummies for all metropolitan ares and its corresponding net-migration rate, there is no statistically significant relationship between the coefficients of net-migration rates by city and its city population size[15]. Therefore, city population size may not be a plausible weight and I only use White correction for the standard errors, clustered by city.

5 Endogeneity Concern and IV Construction

One concern for equation (1) is the endogeneity problem associated with the migration rate. Specifically, migrants may be attracted to areas with high wages or strong employment performance. This can be thought of as unmeasured and omitted city characteristics, which generates an increasing in average wages between 1975 and 1980, and will bias the in-migration rate (β) upward and the out-migration rate (γ) downward. This concern can be addressed by instrumenting for the in- and out-migration rates to an area j with variables that are uncorrelated with local economics conditions in area j in 1980.

To develop instruments for in- and out-migration rates, I make full use of the standard economic model of migration. Based on the model in Sjaastad (1970), prospective migrants in the source area k compare the expected benefits of moving to new locations j ($j \neq k$) and the costs of migration. Therefore, while positive economic shocks pulling migrants to destination j are clearly endogenous, local economic conditions pushing migrants to leave area k are arguably exogenous to labor market conditions in area j. In this paper, I define local economic conditions in sending areas k that push migrants to leave to destination j as "push factors" and local economic conditions in receiving areas j that pull migrants from the original area k as "pull factors". To be a valid "push factor" or "pull factor", its variation in area k should be uncorrelated with the one in area j. I adopt the shift-share measure, based on Blanchard et al. (1992), and local weather conditions as my "push factors" ("pull factors") to predict out-migration (in-migration) rates from sending (receiving) areas and construct my instrument for in-migration (out-migration) rates for all the receiving (sending) areas.

[15]From the result, there are indeed some significant difference in the out-migration rate on the economics outcomes across cities.

There is a two-step procedure to create the instrument for in-migration to a metropolitan area, following Boustan et al. (2010). First, I predict out-migration rates from source areas, indexed by k, using local economic conditions as "push factors". Multiplying the population in each sending area by the predict out-migration rates leads the predicted outflow from each source area k. In the next step, I predict the probability that migrants leaving source area k to settle in destination j on the basis of distance alone ($k \neq j$). For each sending area k, I then predict how many people out-migrated from sending area k to destination area j by multiplying the predicted outflow from k by the share who pick j. The total predicted migrant inflows to a destination area j is then the sum of these pair-wise predictions of people moving into j from all possible source areas. The results from the predicted migrant flow divided by the local population becomes the instrument for the actual in-migration rate in each receiving area. In next two subsections, I will explain the procedure in details about how to construct the predicted in-migration as an instrument for the actual in-migration rate as an example.

The instrument for the out-migration rate from area k is developed in a symmetric fashion. First, I predict the in-migration rate to each receiving area j as a function of local pull factors in j. I then convert the predicted in-migration rate into a predicted migration flow. In the next step, I use the distance between areas j and k to predict the share of in-migrants to area j that hail from each source area k. Using these shares, I calculate the predicted number of out-migrants flowing from source area k to destination j. Finally, I sum these predicted outflows across all possible destination areas j and convert this flow into a predicted out-migration rate for source area k. This value becomes the instrument for the out-migration rate for each city in the sample.

5.1 Predicting Out-migration Flow using "Push Factors"

There are two steps to get the predicted out-migration flows for each city. In the first step, I predict the out-migration rate from each of the sending areas. Therefore, I regress out-migration rates on the sample of 269 cities and balance-of-states in total, as a function of "push factors". Next, I multiply the predicted out-migration rate by its corresponding city population to get the predicted out-migration flows for all the cities and balance-of-states.

First, the out-migration rate ($o_{kr(80-75)}$) from source area k in region r is determined as a function of "push factors", controlling region dummies. The equation is as following:

$$o_{kr(80-75)} = \alpha + \Phi^{'} Z_{kr} + \lambda \hat{\eta}_k + \Pi_r + \varepsilon_{ijr80}, \tag{2}$$

where both Z_{kr} and $\hat{\eta}_k$ are the "push factors". Z_{kr} are measures of extreme weather conditions, and $\hat{\eta}_k$ is the constructed exogenous demand shock in the area k using a shift-share measure based on Blanchard et al. (1992), described below in section 5.1.1.

Next, for the predicted flow of migrants leaving area k, it is estimated as the product of the predicted out-migration rate for area k ($o_{kr(80-75)}$) and the population of area k in 1975[16]:

$$\hat{O}_k = o_{kr(80-75)} \times population_{kr75}, \tag{3}$$

5.1.1 Shift-share Measure

One condition to be a valid "push factor" is that local economic conditions pushing migrants to leave area k should be exogenous to labor market conditions in area j. In this paper, I use both shift-share measures and weather as the "push factors". While weather is a natural phenomenon[17], it mostly affects agriculture, which may be a weak "push factor" by itself. Generally, exogenous policy changes or natural disasters provide a unique laboratory to explore the causal impact of migration on local labor markets, such as in Card (1990) and Boustan et al. (2010). However, these exogenous natural experiments are uncommon. Blanchard et al. (1992) proxy the predicted growth rates of employment as an exogenous local demand shock to identify the effect from innovations on wage and unemployment. Blanchard et al. (1992), McGuire and Bartik (1992), Bartik (2002), Hoynes et al. (2012) and Bound and Holzer (2000) have applied this method similarly to construct an exogenous demand shock, using this shift-share measure.

In this paper, I adopt this method, using the predicted growth rates of employment at the metropolitan and balance-of-state level, as well as the variation in local weather conditions to predict the out-migration rate from source areas. Another maintained assumption is that the extreme weather events in area k did not influence the labor market in area j, except through the resulting migration. The shift-share measure can be expressed as:

$$\hat{\eta}_k = \sum_j \gamma_{kj}\eta_j, \tag{4}$$

where $\hat{\eta}_k$ represents the predicted growth of employment level in area k, γ_{kj} represents the share of employment level by industry j in MSA k, based on the year 1975, and η_j represents the change in the log of total employment level in the same industry j nationally from 1975 to 1980. Data about the annual wage and the employment level as well as the share of employment level by industry in county and national level in year 1975 and 1980 are from Inter-university Consortium for Political and Social Research (ICPSR) about

[16] In this paper, the population in year 1975 is estimated based on the population in 1980 from IPUMS-USA and adjusted by the national growth rate for the year 1975. From U.S. Bureau of the Census, total national population in year 1975 is 215464000 while in year 1980, the total population is 226542250, thus, the national adjusted growth rate for year 1975 is 0.95109853.

[17] In order to avoid spatial correlation, I further control for region dummies in above equation.

County Business Patterns[18]. Since the observations in different counties vary in the number of their available two-digit SIC information, I select 42 SIC industry sections with complete data information[19].

5.2 Predicting Settlement Patterns using Distance

The predicted out-migrant flows are then assigned from source areas to destinations using the geographic distance between source-destination pairs. In this step, I predict the probability that migrants leaving source area k settle in destination j on the basis of distance[20] alone ($k \neq j$). Specifically, I use the geographic distance between area j and k to predict the probability that a migrant leaving area k would settle in destination j. Distance is a central determinant of a migrant's location choice and is invariant to contemporary economic conditions in either the sending or the receiving area (Borjas (2001); Levy and Wadycki (1974); Schwartz (1973)).

For 269 cities and balance-of-states, the probability of migrating from source area k to destinations j, denoted as P_{kj}, is the percentage of migrants from source area k to destinations j over all migrants leaving area k. For each of the sending areas, I regress this migration probability from source area k to all possible destinations j on their corresponding geographic distance in miles. These 269 regressions give me origin specific parameters for each location k, which is denoted as (θ_k, γ_k). The regression equation for source area k is as following:

$$P_{kj} = \alpha_k + \theta_k distance_{kj} + \gamma_k distance_{kj}^2 + \mu_k, \tag{5}$$

For total 269 sending areas, I use area specific parameters (θ_k, γ_k) to predict the migration flow probability P_{kj}, denoted as $\hat{P}_{kj}$.

To get the total predicted migrant flow to the destination area j, I then sum over all areas ($k \neq j$) of the predicted number of migrants leaving area k who are expected to settle in city j. The equation for the total predicted migrant flows into area j can be expressed as:

$$\hat{M}_j = \sum_{k=1...n(k \neq j)} \hat{O}_k \times \hat{P}_{kj}, \tag{6}$$

where $\hat{O}_k$ is the estimated out-migration flow from the first step and $\hat{P}_{kj}$ is the predicted probability of

[18] I appreciate Michael Matheis providing the dictionary and Stata code to convert the data in text or UDiv format into a usable Stata format.

[19] There are 71 SIC 2-digit SIC divisions in 1980 Inter-university Consortium for Political and Social Research (ICPSR) about County Business Patterns and there are no observations for SIC equal to 4000, 9900 and 4300.

[20] The measures of distances in miles between every county group in the United States is from the National Bureau of Economic Research website. I use the year 1990 dataset and the source is Gazetteer. For the distance between balance-of-state areas, I calculate the average distance from every non-metropolitan county group in state A to every non-metropolitan county group in state B to proxy it.

migrating from original area k to destination j. The total predicted migrant flows to the destination area j is then the sum of the product of these pair-wise predictions across all possible source areas. Next, I divide the predicted flow $\hat{M}_j$ from the above equation by city j's population in year 1975 to get the in-migration rate $\hat{m}_{jr80-75}$. This predicted in-migration rate $\hat{m}_{jr80-75}$ becomes the instrument for the actual in-migration rate in the receiving area j. For the instrument to be valid, the variables used to predict out-migration from sending areas must be uncorrelated with unobserved labor market conditions in destinations, except through migration. Therefore, I construct my instrument using predicted migrant flows, excluding cities from the same state of a destination city, in order to minimize concerns about the spatially correlated shocks[21].

From Table 1, the predicted in- and out-migration rates are both close to the true migration rates. In order to avoid spatial correlation, I estimate the predicting equations without the migration within the same state or region and these corresponding estimates are smaller than the ones when using all the migration observations.

5.3 Results from the Regressions Used to Construct the Instruments

In this paper, instruments for in- and out-migration rates are constructed based on two sets of regressions: estimating migration rates from local "push/pull factors" (shift-share shocks and weather conditions) and predicting settlement patterns by the distance between these two areas.

5.3.1 Determinants of In- and Out-Migration Rates

Table 2 shows results from equation (2), which estimates the net-migration rate or its subcomponents (in- and out-migration rates) from local economic conditions. Exogenous demand shocks are estimated based on Blanchard et al. (1992), using data from Inter-university Consortium for Political and Social Research (ICPSR) about County Business Patterns in 1975 and 1980. Weather condition data are constructed from NOAA weather dataset at the division-month level and use the division-area weight to combine the data into the county-monthly level[22]. I use the mean of county-monthly data from 1975 to 1980 to represent the weather condition at city and balance-of-state level.[23] Extreme drought and wetness add up both severe and extreme levels.

[21] Another approach is to restrict the attention to predicted migration from outside of a city's own census region. The results are similar to the ones using prediction from cities outside the same state.

[22] I appreciate Price Fishback for sharing this dataset. The dataset include monthly temperature and precipitation at county level as well as four severity index of a wet or dry spell. That is, Palmer Drought Severity Index (PDSI); Palmer Hydrological Drought Index (PHDI); Modified Palmer Drought Severity Index (PMDI); Palmer "Z" Index (ZNDX)-"Moisture Anomaly Index". I mainly use PDSI for all of the regressions reported in this paper, and the results using PHDI and PMDI are quite close to PDSI. Besides, these three index mostly have same summary statistics. ZNDX has quite different criterion and spells for indication of extreme weather, I have not used it for this paper.

[23] A better way is to use the county-area weight to aggregate the county level data into city and balance-of-state level data.

From the regression results, higher average temperature stimulated net- and in-migration to both urban and rural areas. While an increase in the number of months with severe and extreme drought led to higher in-migration, there is no significant difference between urban and rural areas. These results differ from Boustan et al. (2010), which shows that an increase in the months of severe wetness resulted in more net-migration in balance-of-state areas, many of which are specialized in agriculture. One explanation for this is that modern mobility relies more on residence comfort, even for rural areas. Somewhat surprisingly, the shift-share shock that generated more higher out-migration rate also led to more in-migration. This pattern can be explained if departures themselves prompted in-migration. On net, local shift-share shocks led to more net-migration. Overall, both local "push" or "pull" factors and favorable weather conditions attracted migrants to that area.

Table 2: Determinants of In- and Out-Migration, 1975-80

	Dependent Variable		
Right-Hand-Side Variable	Net-migration Rate	Out-migration Rate	In-migration Rate
Average temperature, 1975-80	0.00258*	-0.000448	0.00213*
	(0.00125)	(0.000945)	(0.000945)
Average precipitation, 1975-80	0.00611	-0.00767	-0.00156
	(0.00461)	(0.00634)	(0.00751)
Months of extreme drought, 1975-80	0.0199	0.00576	0.0257*
	(0.0150)	(0.0116)	(0.0119)
Months of extreme wetness, 1975-80	-0.00736	-0.00363	-0.0110
	(0.00658)	(0.00680)	(0.00735)
Drought months × metro area	-0.0168	-0.00511	-0.0220
	(0.0131)	(0.0131)	(0.0136)
Wet months × metro area	-0.00672	-0.00111	-0.00783
	(0.0103)	(0.0112)	(0.00825)
Exogenous Demand Shock	0.00518*	0.0100***	0.0152***
	(0.00245)	(0.00215)	(0.00296)
Metropolitan Dummy	0.0270	0.0511*	0.0780***
	(0.0234)	(0.0241)	(0.0164)
Observations	269	269	269
F	9.642	12.48	26.14

Note: Standard errors are in parentheses and clustered by state. * represents $p < 0.05$, ** represents $p < 0.01$, *** represents $p < 0.001$. Observations include 220 metropolitan areas and 49 balance-of-state areas; region dummies included.

5.3.2 Regression Results for Settlement Patterns using Distance

For the settlement patterns, I run 269 regressions, one for each city and balance-of-state as in equation (5). In general, the migration destination choice is strongly related to the geographic distance between the sending area and the receiving area. The median coefficient of the source area level regression on distance is -0.90984 with a standard deviation of 0.0092715, while the median coefficient of the sending area level regression on distance is -1.04578 with a standard deviation of 0.011174. These results suggest that a pair-wise distance increase by 1,000 miles decreases the share of migration pattern from area k to destination area j by around 1 percentage point[24]. When using all the metropolitan or balance-of-state areas, except for around 20 regressions, all the coefficients on linear distance term (nearly on both the linear and quadratic distance terms) in equation (5) was statistically significant at the 5% level. The number of these significance terms decrease when excluding areas within the same state or region. One potential explanation is that there are fewer observations in each regression.

5.4 First-stage Regression Results

The predicted in-migration and out-migration rates are the instruments for the actual net-migration rate in equation (1)[25]. Table 3 contains the results from first-stage regressions for all non-migrants about earning analysis. It relates actual net migration to the predicted in- and out-migration rates[26]. The results of the other two first-stage regressions about employment and out-migration probability are displayed in the appendix, which are quite close to the table below. In addition, the first-stage regressions contain a full set of region dummies and control variables included in the second stage. The preferred instrument excludes the predicted migration flows within the same state. This strategy reduces spatial correlation of unmeasured economic shocks. For comparison, the results using instruments that include all predicted migrant flows and that exclude migrants within the same region are also displayed below. In each case, predicted in-migration is positively related to actual net migration, while the relationship is negative between out-migration and net-migration rate.

The F-statistic for weak instruments test for the first-stage regression range between 7 to 21, thus we reject the weak instruments hypothesis. A one standard-deviation increase in the predicted in-migration or out-migration rate implies a 0.7-1.0 standard-deviation increase in the actual net-migration rate. The

[24]These results are based on regressions excluding migration flows from within the same state. When using all of migration observations or excluding migration flows within the same region, the results are close.

[25]The results stay similar if I use actual in-migration and out-migration rates in equation (1). In this case, the predicted in-migration and out-migration rates are the instruments for the actual in-migration and out-migration rates.

[26]While Table 3 only contain existing residents who were identified by staying at the house in migration sample, I have also added all the observations not in the migration sample for all the regressions. Results are all of the same signs and quite close to each other.

implied response to a one standard-deviation in predicted migration rate is larger when restricting the assigned migrant flows from either out-of-state or out-of-region scope. Moreover, based on the fact that the instruments are generated regressors(McKenzie and McAleer (1997)), I bootstrap for standard errors, clustered by the metropolitan and balance-of-state areas to account for the correlation across individuals within the same area. The coefficients and standard errors of predicted in- and out-migration rates from the bootstrap procedure are similar to the first-stage regression results.

Table 3: First-Stage Regressions for Earning: Relationship between Predicted and Actual Migration, 1975-80

	Assigned Migrant Flows		
	All	Out of State	Out of Region
OLS regression coefficients:			
Predicted in-migration rate	0.535***	1.027***	1.114***
	(0.144)	(0.181)	(0.247)
Predicted out-migration rate	-0.580***	-1.202***	-1.397***
	(0.164)	(0.182)	(0.245)
Observations	422686	422686	422686
F-statistic	6.93104	23.4734	18.2064
Implied effect of 1 SD:*			
Predicted in-migration rate	.7667685	.9355328	.8791634
Predicted out-migration rate	.8504967	1.1146	1.006792
Bootstrap procedure:†			
Predicted in-migration rate	.5913115***	1.065934***	1.165406***
	(.1589046)	(.2057895)	(.2876852)
Predicted out-migration rate	-.6266328***	-1.197755***	-1.380744***
	(.1798031)	(.1924475)	(.2679089)

Note: Standard errors are in parentheses and clustered by metropolitan area. * represents $p < 0.05$, ** represents $p < 0.01$, *** represents $p < 0.001$. Dependent variable is the net number of migrants between 1975 and 1980 as a percentage of the 1975 population. Right-hand-side variables include region dummies and all the other controls from the second-stage regression.

* Implied effect of a 1 SD increase is calculated by the estimated coefficient multiplying the standard deviation of predicted in or out-migration in shares of a standard deviation in the actual net migration rate.

† Bootstrap procedure accounts for the fact that predicted in- and out-migration rates are generated. Clustered by metropolitan area, I randomly selected individuals at the same number of population size for every metropolitan area with replacement for 100 times when performing the bootstrap procedure.

6 The Impact of In-Migration on Local Labor Markets

6.1 Earnings and Wage Outcomes

Table 4 investigates the relationship between net internal migration in a metropolitan area and three earnings and three work opportunity outcomes for existing non-migrants. The list includes annual earnings, weekly wages, hourly wages, weeks worked last year, whether working less than 26 weeks and usual hours worked during the previous week. Weekly wage is annual earnings divided by weeks worked last year, and hourly wage is the weekly wage divided by the usual hours worked during the previous week. The indicator working less than 26 weeks is equal to 1 for men who worked 26 or fewer weeks during the year. For comparison, both coefficients from an OLS and its corresponding IV specification are presented below. Because higher wages are likely to attract migrants to the area, the OLS specification may be biased in a positive direction. Indeed, the coefficients from OLS regressions for earnings are positive but not statistically significant. After instrumenting for the net-migration rate, the coefficients for earnings become negative, though they still remain statistically insignificant. While wage adjustments are statistically insignificant, the effect of net migration on earnings is larger for hourly wage than for weekly earnings and annual earnings.

This pattern does not apply for work opportunity outcomes. From the results, net migration seems to boost work opportunities. Compared with the IV results, the OLS regression results are downward biased toward zero and the signs of the coefficients vary from being negative and positive. But this effect of net-migration on working time become positive in IV specification while only the coefficient of hours worked is statistically significant and positive at the 5% level. This result contradict with Boustan et al. (2010), which shows that residents of metropolitan areas experiencing high in-migration worked fewer weeks over the year. One potential explanation for my results here is that net migration boost the quantity of labor demanded. In equilibrium, employees have to work longer under the same payment. Therefore, the labor market becomes more competitive. Overall, the effects of migration flows on existing residents at both earnings and work opportunity are not statistically significant, except for the hours worked per week.

6.2 Employment Outcomes

The earnings and work opportunity results suggest that net migration has little significant effect on existing residents. Next, I will investigate whether net migration affected the probability for an individual to be full-time, part-time employed or idle. I define full-time employment when the employee has worked 35 or more hours in a usual week, and define part-time employment when the employee has worked less than 35 hours per week. Being idle is defined when the employment status is unemployed. Table 5 presents results from a multinomial logistic (MNL) regression of employment status on migration rates with the base category of being idle, and separate regressions of employment status on net-migration rate using both OLS and IV specifications. Based on MNL results, net in-migration into a metropolitan area makes the existing residents 1.246% more likely to be full-time employed and 1.198% more likely to be part-time employed, relative to being unemployed. This pattern confirms that a shift in the labor supply from net migration

Table 4: Effect of Net Migration on Earnings, Wages and Work Time in 1980

Dependent Variable	Ordinary Least Squares	Instrument Variable
ln(annual earnings)	0.0577	-0.0992
	(0.0834)	(0.203)
ln(weekly wage)	0.0680	-0.133
	(0.0862)	(0.211)
ln(hourly wage)	0.0444	-0.371
	(0.0951)	(0.222)
ln(weeks worked)	-0.0295	0.0330
	(0.0223)	(0.0583)
Work less than 26 weeks	0.00474	-0.0144
	(0.0117)	(0.0305)
ln(hours worked per week)	0.0144	0.193^{*}
	(0.0398)	(0.0841)

Note: Number of observations is 422686. Standard errors are in parentheses and clustered by metropolitan area. * represents $p < 0.05$, ** represents $p < 0.01$, *** represents $p < 0.001$. Dependent variable is the net number of migrants between 1975 and 1980 as a percentage of the 1975 population and data in the above table are the coefficients on this dependent variable. The sample includes non-migrants employed in the dataset who reported positive earnings and working time.[27] Individual controls include an indicator of being female; race dummies of being White, Black, Asian or other; years of completed schooling; and a cubic polynomial in age. The regressions also have city-level controls for the share of the population that are female and Black in 1970, mean education level and age, and lagged annual earnings in 1972. Details about the data sources are described in the text.

leads to improvements in employment status.

Migrants may have been attracted to migrate to an area with higher employment rate, which implies higher probability of being hired. Therefore, coefficients from the MNL regression may be biased upward by endogenous location decisions. This pattern is confirmed by the biased-toward-zero coefficients from the OLS compared to the IV specification. While two of these coefficients from the OLS specification are statistically insignificant, the IV results imply that a one percentage point increase in the net-migration rate led to a 0.335% higher probability of being full-time employed at the 0.1% significant level. Besides, a one percentage increase in the net-migration rate will also decrease the probability of being part-time employed by 0.171% and the probability of being idle by 0.164% at 1.0% significant level.

Together with the positive effect on working time, a possible explanation for this is that instead of direct wage reduction, workers have to work longer under the same wage payment in response to the migrant inflows.

Table 5: Effect of Net Migration on Employment in 1980

Dependent Variable	Multinomial Logistic Regression*	Ordinary Least Squares†	Instrument Variable†
Full time	1.246**	0.0468	0.335***
	(0.481)	(0.0396)	(0.0722)
Part time	1.198**	0.000211	-0.171**
	(0.390)	(0.0253)	(0.0652)
Idle (Unemployed)	.	-0.0470*	-0.164**
	.	(0.0213)	(0.0594)

Note: Number of observations is 520190. Standard errors are in parentheses and clustered by metropolitan area. * represents $p < 0.05$, ** represents $p < 0.01$, *** represents $p < 0.001$. Dependent variable is the net number of migrants between 1975 and 1980 as a percentage of the 1975 population and data in the above table are the coefficients on this dependent variable. The sample includes non-migrants who have information of employment status or usually hours of work per week.[28] Individual controls include an indicator of being female; race dummies of being White, Black, Asian or other; years of completed schooling; and a cubic polynomial in age. The regressions also have city-level controls for the share of the population that are female and Black in 1970, mean education level and age, and lagged annual employment rate in 1970. Details about the data sources are described in the text.
* It is based on three categories: full-time employment (usual working hours per week is larger than 35), part-time employment (usual working hours per week is less than 35) and being idle (employment status is unemployed).
† Probability of being full-time, part-time employment or being idle is compared to the other two employment categories.

6.3 Out-migration of Existing Residents

Based on the above analysis, migration to a metropolitan area increases the work opportunity by increasing the working hour per week and the probability of being full-time employed. Though being more likely to be full-time employed, existing residents might migrate out in response to a more intensively competitive working environment which requires them to work more hour per week. In order to estimate this effect, I recovered all migration samples' residence in 1975 and set a dependent variable Y_{ijr80} equals to 1 if the migration sample migrated during 1975 to 1980. All city-level variables of residences in 1975, $n_{jr(80-75)}$ and $Y_{jr(75-70)}$, as well as individual characteristics X_{ijr80}, are the same as in former regressions[29]. Table 6 contains results about the effect of in-migration flows to a metropolitan area on the probability of a non-migrant to relocate his residence during 1975 to 1980. The coefficient from the OLS specification is statistically insignificantly and negative. While some unobserved city characteristics may attract more new arrivals and disencourage existing migrants to move out, this may lead a downward bias of the OLS specification. This is supported by the comparison between the OLS and IV results. The estimate from IV regression is larger than the OLS specification and stay positive, though still remaining statistically insignificant. This results accord with the literature about international migration on the probability of current residents migrating out.

Table 6: Effect of Net-Migration on Out-Migration in 1980

Dependent Variable	Ordinary Least Squares	Instrument Variable
Left standard metropolitan area, 1975-80	-0.0133	0.0752
	(0.0663)	(0.191)

Note: Number of observations is 1097240, which are all in migration sample. Standard errors are in parentheses and clustered by metropolitan area. * represents $p < 0.05$, ** represents $p < 0.01$, *** represents $p < 0.001$. Dependent variable is a dummy whether the individual had moved during the period from 1975 to 1980. Individual controls include an indicator of being female; race dummies of being White, Black, Asian or other; years of completed schooling; and a cubic polynomial in age. The regressions also have city-level controls for the share of the population that are female and Black in 1970, mean education level and age, and lagged annual earning in 1972 and employment rate in 1970. Details about the data sources are described in the text.

7 Conclusion

This paper focuses on the effects of internal migration on three local labor market outcomes: earning, employment and probability of migrating out. It mainly deals with the endogeneity concerns that migrants may be attracted to cities with good economies. I construct predicted in-migration and out-migration rates as instruments for the actual net-migration rate in two steps. I specify the process to construct the predicted in-migration rate and the method is symmetric for constructing the predicted out-migration rate. First, I use "push factors" to predict the out-migration rate from all 269 metropolitan areas and balance-of-states. The "push factors" include shift-share measures based on Blanchard et al. (1992) and extreme weather events. Then, I multiply the predicted out-migration rate by its local population size to get the out-migration flow

[29] I constructed dataset including individual characteristics X_{ijr80}, the corresponding dependent variable Y_{ijr80} and former residence in 1975, merged with the former regression dataset based on the residence name.

from sending area k. In the next step, I estimate the probability of settlement patterns from area k to remaining areas by regression on distance between these two areas. Afterwards, for the total in-migration flows to area j, I sum over all sending areas of the products between predicted out-migration flows from the sending areas to the destination j and the corresponding probabilities of their settlement patterns. The predicted in-migration rate is the result from dividing the in-migration flow by its local population size.

The comparison between OLS and IV regression results confirms that there are endogeneity problems in the OLS specification that biases the estimates toward zero. There are small and statistically insignificant effects on annual earnings, weekly wage and hourly wage, while the effect of migration on working time is positive but only statistically significant for hours worked in a usual week. Besides, the effect of net-migration on earnings is larger for the hourly wage than for the weekly and annual earnings. Moreover, net-migration flows into a metropolitan area significantly increases the employment status for existing residents. Relative to part-time employment and unemployment, a one percentage increase in the net-migration rate led to a 0.335% higher probability of being full-time employed. The probability of being in part-time employment was decreased by 0.171% and the unemployment probability was reduced by 0.164% with a one percentage increase in the net-migration rate. Together with the positive effect on working time, it implies that instead of direct wage reduction, workers have to work longer under the same wage payment. Furthermore, existing residents seem not to be crowded out by the in-migrants since there is no statistically significant effect of net-migration rate on the probability of migrating out.

However, the effect of migration on all three economic outcomes are based on the equilibrium state. Labor demand may simultaneously shift in response to the migration flows and this labor demand change by local area may offset the labor supply shift from migration. More further work which distinguish the equilibrium change caused by demand or supply shifts needs to be done for this concern. Besides, the industry composition may also affect the migration effect on economic outcomes. Since instruments have been constructed to deal with the endogeneity problem, including omitted local labor demand effects, this concern is negligible.

Results Appendix

Table 7: First-Stage Regressions for Employment: Relationship between Predicted and Actual Migration, 1975-80

	Assigned Migrant Flows		
	All	Out of State	Out of Region
OLS regression coefficients:			
Predicted in-migration rate	0.540***	1.035***	1.137***
	(0.144)	(0.175)	(0.233)
Predicted out-migration rate	-0.600***	-1.231***	-1.450***
	(0.165)	(0.177)	(0.241)
Observations	520190	520190	520190
F-statistic	7.00771	27.0577	20.2656
Implied effect of 1 SD:*			
Predicted in-migration rate	.7695127	.9392448	.894187
Predicted out-migration rate	.8756075	1.139145	1.042803
Bootstrap procedure:†			
Predicted in-migration rate	.5894432***	1.061537***	1.162515***
	(.1591318)	(.1998948)	(.2717612)
Predicted out-migration rate	-.6546819***	-1.238404***	-1.439174***
	(.1831706)	(.1935773)	(.2662602)

Note: Standard errors are in parentheses and clustered by metropolitan area. * represents $p < 0.05$, ** represents $p < 0.01$, *** represents $p < 0.001$. Dependent variable is the net number of migrants between 1975 and 1980 as a percentage of the 1975 population. Right-hand-side variables include region dummies and all the other controls from the second-stage regression.

* Implied effect of a 1 SD increase is calculated by the estimated coefficient multiplying the standard deviation of predicted in or out-migration in shares of a standard deviation in the actual net migration rate.

† Bootstrap procedure accounts for the fact that predicted in- and out-migration rates are generated. Clustered by metropolitan area, I randomly selected individuals at the same number of population size for every metropolitan area with replacement for 100 times when performing the bootstrap procedure.

Table 8: First-Stage Regressions for Out-Migration: Relationship between Predicted and Actual Migration, 1975-80

	Assigned Migrant Flows		
	All	Out of State	Out of Region
OLS regression coefficients:			
Predicted in-migration rate	0.535***	1.057***	1.211***
	(0.148)	(0.181)	(0.241)
Predicted out-migration rate	-0.560**	-1.220***	-1.485***
	(0.176)	(0.190)	(0.248)
Observations	1097240	1097240	1097240
F-statistic	6.94565	21.0118	18.541
Implied effect of 1 SD:*			
Predicted in-migration rate	.7291008	.9210531	.9297664
Predicted out-migration rate	.7778818	1.077757	1.028992
Bootstrap procedure:†			
Predicted in-migration rate	.5851979***	1.076905***	1.228812***
	(.1621517)	(.2044918)	(.2819061)
Predicted out-migration rate	-.6129484***	-1.211831***	-1.449387***
	(.1930145)	(.1985274)	(.2713388)

Note: Standard errors are in parentheses and clustered by metropolitan area. * represents $p < 0.05$, ** represents $p < 0.01$, *** represents $p < 0.001$. Dependent variable is the net number of migrants between 1975 and 1980 as a percentage of the 1975 population. Right-hand-side variables include region dummies and all the other controls from the second-stage regression.

* Implied effect of a 1 SD increase is calculated by the estimated coefficient multiplying the standard deviation of predicted in or out-migration in shares of a standard deviation in the actual net migration rate.

† Bootstrap procedure accounts for the fact that predicted in- and out-migration rates are generated. Clustered by metropolitan area, I randomly selected individuals at the same number of population size for every metropolitan area with replacement for 100 times when performing the bootstrap procedure.

References

Bartik, T. J. (2002). Evaluating the impacts of local economic development policies on local economic outcomes: what has been done and what is doable?

Blanchard, O. J., Katz, L. F., Hall, R. E., and Eichengreen, B. (1992). Regional evolutions. *Brookings papers on economic activity*, 1992(1):1–75.

Borjas, G. J. (2001). Does immigration grease the wheels of the labor market? *Brookings papers on economic activity*, 2001(1):69–119.

Borjas, G. J. (2003). The labor demand curve is downward sloping: Reexamining the impact of immigration on the labor market. Technical report, National Bureau of Economic Research.

Borjas, G. J. (2006). Native internal migration and the labor market impact of immigration. *Journal of Human resources*, 41(2):221–258.

Borjas, G. J., Freeman, R. B., and Katz, L. F. (1996). Searching for the effect of immigration on the labor market. Technical report, National Bureau of Economic Research.

Borjas, G. J., Freeman, R. B., Katz, L. F., DiNardo, J., and Abowd, J. M. (1997). How much do immigration and trade affect labor market outcomes? *Brookings papers on economic activity*, 1997(1):1–90.

Bound, J. and Holzer, H. J. (2000). Demand shifts, population adjustments, and labor market outcomes during the 1980s. *Journal of labor Economics*, 18(1):20–54.

Boustan, L. P., Fishback, P. V., and Kantor, S. (2010). The effect of internal migration on local labor markets: American cities during the great depression. *Journal of Labor Economics*, 28(4):pp. 719–746.

Card, D. (1990). The impact of the mariel boatlift on the miami labor market. *Industrial and Labor Relations Review*, 43(2):pp. 245–257.

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

Card, D. (2009). Immigration and inequality. Technical report, National Bureau of Economic Research.

Card, D. and DiNardo, J. E. (2000). Do immigrant inflows lead to native outflows? Technical report, National bureau of economic research.

Filer, R. (1992). The effect of immigrant arrivals on migratory patterns of native workers. In *Immigration and the workforce: Economic consequences for the United States and source areas*, pages 245–270. University of Chicago Press.

Frank, D. H. (2009). The effect of migration on natives' employment outcomes: Evidence from the fall of the berlin wall.

Goldin, C. (1994). The political economy of immigration restriction in the united states, 1890 to 1921. In *The regulated economy: A historical approach to political economy*, pages 223–258. University of Chicago Press.

Groen, J. A. and Polivka, A. E. (2008). The effect of hurricane katrina on the labor market outcomes of evacuees. *The American Economic Review*, 98(2):pp. 43–48.

Hoynes, H., Miller, D. L., and Schaller, J. (2012). Who suffers during recessions? *The Journal of Economic Perspectives*, 26(3):27–47.

Imberman, S. A., Kugler, A. D., and Sacerdote, B. I. (2012). Katrina's children: Evidence on the structure of peer effects from hurricane evacuees. *The American Economic Review*, 102(5):pp. 2048–2082.

Kritz, M. M. and Gurak, D. T. (2001). The impact of immigration on the internal migration of natives and immigrants. *Demography*, 38(1):133–145.

Levy, M. B. and Wadycki, W. J. (1974). What is the opportunity cost of moving? reconsideration of the effects of distance on migration. *Economic Development and Cultural Change*, 22(2):198–214.

Lewis, E. et al. (2004). How did the miami labor market absorb the mariel immigrants? Technical report, Federal Reserve Bank of Philadelphia Philadelphia, PA.

Lewis, E. G. (2003). Local, open economies within the us: how do industries respond to immigration?

McGuire, T. J. and Bartik, T. J. (1992). Who benefits from state and local economic development policies?

McIntosh, M. F. (2008). Measuring the labor market impacts of hurricane katrina migration: Evidence from houston, texas. *The American Economic Review*, 98(2):pp. 54–57.

McKenzie, C. and McAleer, M. (1997). On efficient estimation and correct inference in models with generated regressors: A general approach. *Japanese Economic Review*, 48(4):368–389.

Ottaviano, G. I. and Peri, G. (2012). Rethinking the effect of immigration on wages. *Journal of the European economic association*, 10(1):152–197.

Rosenbloom, J. L. and Sundstrom, W. A. (1999). The sources of regional variation in the severity of the great depression: evidence from us manufacturing, 1919–1937. *The Journal of Economic History*, 59(03):714–747.

Schwartz, A. (1973). Interpreting the effect of distance on migration. *The Journal of Political Economy*, pages 1153–1169.

Sjaastad, L. A. (1970). The costs and returns of human migration. In *Regional Economics*, pages 115–133. Springer.

Wallis, J. J. (1989). Employment in the great depression: New data and hypotheses. *Explorations in Economic History*, 26(1):45–72.

Wallis, J. J. and Benjamin, D. K. (1981). Public relief and private employment in the great depression. *The Journal of Economic History*, 41(01):97–102.

Wright, R. A., Ellis, M., and Reibel, M. (1997). The linkage between immigration and internal migration in large metropolitan areas in the united states. *Economic Geography*, 73(2):234–254.