Where are they coming from? Visualizing the movement of people to major urban centers in the late 19th century

Sula Sarkar, Minnesota Population Center, University of Minnesota
Rebecca Vick, Minnesota Population Center, University of Minnesota

Background

The availability of complete count historical microdata has permitted the linking of individuals between census years for longitudinal analysis. These linked datasets provide important research opportunities regarding historical population mobility. In our previous work, we estimated migration distance for the linked people using population weighted county centroids.[A] In this paper we visualize these linked migrants from their centroid[1] of origin to major urban hubs in late 19th century United States. The linked samples provide a variable called MILEMIG, which is the number of miles travelled by each migrant; we visually represent MILEMIG with the help of maps for this study. We also, with the help of these maps, analyze changes in the occupational character of the migrating individuals over time. Prior to presenting results, we describe our method of mapping the migrants with the help of line maps commonly known as spider diagrams. Next, we track migration patterns to major urban and industrial hubs in the late 19th and early 20th century United States through time from 1870-1880, 1880-1900, and 1880-1910.   Finally, we compare occupations of the linked individuals migrating to the major urban and industrial hubs from one census year to another and associate the change in occupational with their migration distances.

Data

Our study utilizes both the male and female linked records from 1870-1880, 1880-1900, and 1880-1910 IPUMS Linked Samples.[B] The linked records are based on the Integrated Public Use Microdata Series (IPUMS) data for United States 1870, 1900, and 1910, which are 1-in-100 national random sample.[C] The 1880 complete count historical microdata is obtained from the North Atlantic Population Project (NAPP).[D] For map files, we use National Historical Geographic Information Systems (NHGIS) historical county shapefiles for 1880, 1900, and 1910.[E]

Our analysis in this paper is based on the total number of linked cases in the IPUMS linked samples.[B] For 1870-1880, we have 58 total linked cases of which, 41% are females. The weighted migrating population for 1870-1880 would be 101,979 persons. For 1880-1900, the total linked cases are 201 persons and 28% of these are females. The weighted N for 1880-1900 is 252,687 migrants. 1880-1910 has 179 linked cases with only 19% females. Weighted N is 213,337 persons for 1880-1910. One of the greatest challenges of working with historical data is a dearth of readily available information thereby making the total number of linked cases often very small.

Methods

To map migration distance we need a starting (origin) and end (destination) point for each linked migrant. We use U.S. state and county of residence, as listed in the IPUMS STATE and COUNTY variables, to locate each person's place of residence in a given year. We do this on the county level because the county is the most precise level of geography available for place of residence for all IPUMS census records from 1870 to 1910. Using county level precision to map individual place of residence exemplifies the challenge of working with historical data, where detailed level geographical information is often very limited. The IPUMS samples also

---

[1] A centroid in a Geographical Information Systems (GIS) framework is the geographical center of a polygon.

have a variable known as "city" where individuals list their city of residence in a particular census year.[c] If a person lists a city variable in the linked census year, we use this as the destination point for each migrant.

We start by locating the origin and destination of the migrants in the 3 linked samples; 1870-1880, 1880-1900, and 1880-1910. As our destination cities for 1880, 1900, and 1910, we chose Chicago, Boston, New York, and Philadelphia as our eastern urban hubs and San Francisco as our western urban hub. All these cities have been centers of industrialization and migration during the late 19th and early 20th century. If a migrant has a destination city of interest to our study, we retain the record. The origin counties in 1870 and 1880, for each of the migrants moving into the above mentioned destination cities, change with each migrant. The spatial nature of their flow and the temporal change represented during this crucial period of industrialization can be visualized very efficiently with Geographical Information Systems (GIS) technology.

We use ArcGIS9.3 for our analysis. First we compute centroids based on the mean centers of the NHGIS historical county boundaries. Note that the county boundaries change through time and hence the mean centers of some counties also change through time. The output of the analysis is a point shapefile of county centroids with unique identification numbers. Next, we create a destination point file that locates the five historical cities. Finally, we use line diagrams, more commonly known as spider diagrams, to represent the movement of these migrants from their origin to their destination.
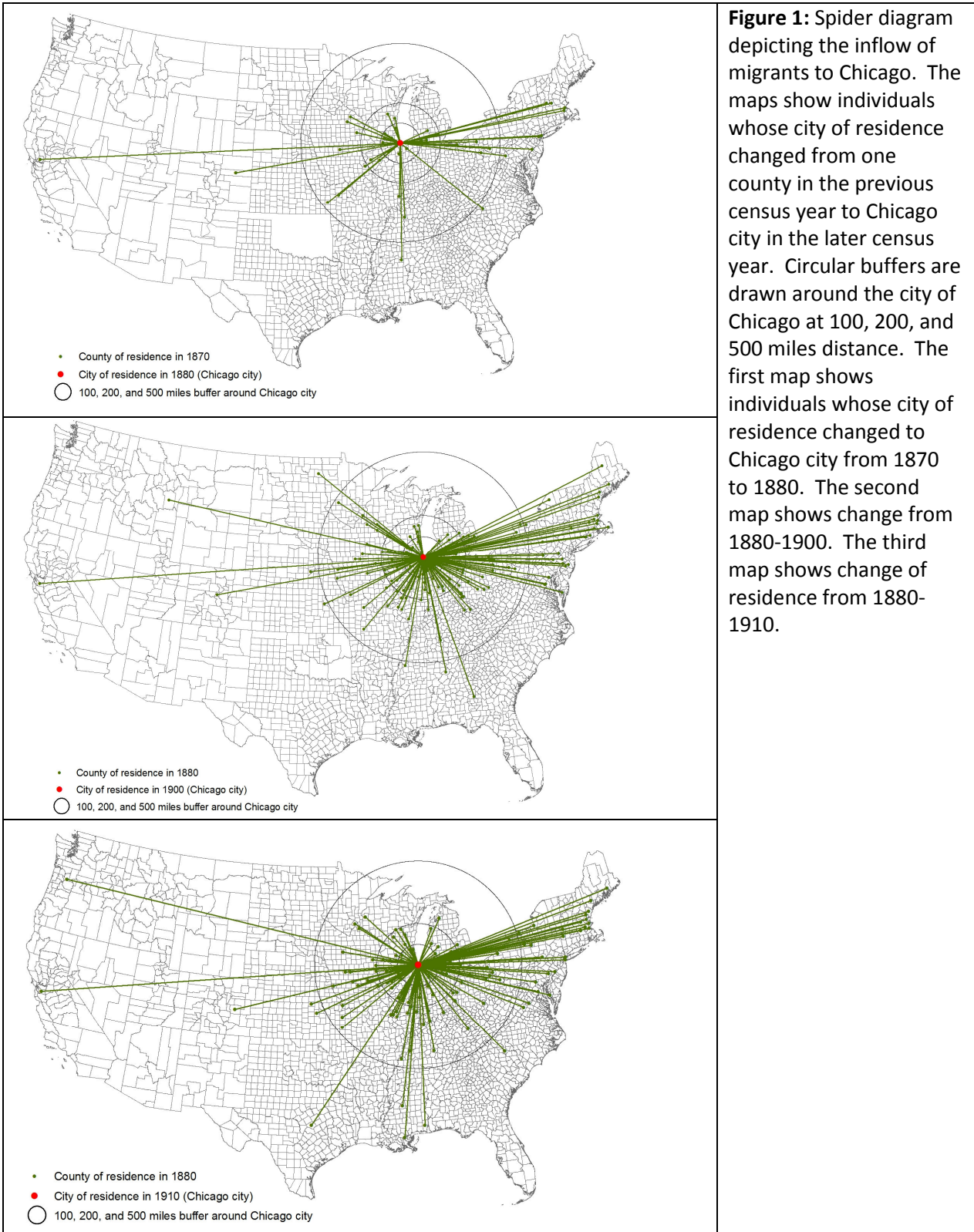
Once we constructed our spiders represented in Figure 1, we were able to visualize the spatial migration of incoming migrants to any particular city. It was clear that the distribution of points (or the origin of the migrants) was denser near the urban destination. Therefore, to evaluate the occupational characters of the migrants, we classified the migrants based on the distance they were travelling. We applied distance buffers around the destination cities and constructed 100, 200, and 500 miles buffer to analyze the characters of the migrants falling within and outside these buffer zones. The analysis remains the same for all the linked years, 1870-1880, 1880-1900, and 1880-1910.

Preliminary results and discussion

For the purposes of this proposal, we analyze Chicago city only for the three linked years. Spider diagrams for the migrants to Chicago from 1870-1880, 1880-1900, and 1900-1910 are represented in Figure 1. Note that in Figure 1, each line of the spider depicts one linked person. We list the total estimated population for each of these maps in the data section of our paper. Note that spider density in Figure 1 depends more on the number of linked cases than the total representative population. An initial look at the figure indicates an increase in migration to Chicago from the south over the years with a steady flow of migrants from the far east and the west. For our final paper, we will compare the migration pattern for all the five cities in our study over time.

Table 1 lists the change in occupation for each of these migrants.[2] More than 90% of the migrants moving into Chicago list a change in occupation and the high percentages are steady over the 3 years in our study. Table 2 goes a bit further by looking at the percentage of migrants engaged in the non-agricultural sector.[3] We look at the miles migrated by the individual and break it up into four categories. Although most of the migrants take up a new job, there are several migrants who work in the professional and personal sector in the city and work in blue collar jobs as janitors, barbers, or in white collar jobs as nurses, or lawyers. Note that these blue and white collar jobs are excluded from our analysis on the non-agricultural sector.

---

[2] Note that there are several occupational classifications in the IPUMS datasets. We use the "OCC" variable that reports the person's primary occupation, coded into a contemporary census classification scheme.[c] More information on the codes can be found at: http://usa.ipums.org/usa/volii/88occup.shtml

Figure 1: Spider diagram depicting the inflow of migrants to Chicago. The maps show individuals whose city of residence changed from one county in the previous census year to Chicago city in the later census year. Circular buffers are drawn around the city of Chicago at 100, 200, and 500 miles distance. The first map shows individuals whose city of residence changed to Chicago city from 1870 to 1880. The second map shows change from 1880-1900. The third map shows change of residence from 1880-1910.

The tables also reflect the high percentage of migrants with non-occupational responses and missing occupations. Further work is needed on the linked females; we plan to alter the occupation of the married women to that of the household head. Furthermore, the answer to our research questions will be stronger

when we compare Chicago in the late 19th century with the other cities of Boston, New York, Philadelphia, and San Francisco.

Linked year	N (number of linked cases)	% whose occupation changed from agricultural to non-agricultural occupation[3]
1870-1880	58	16%
1880-1900	201	20%
1900-1910	179	*
*Large number of unknown occupations and non- occupational responses.		
Table 1: Change in occupation from agricultural to non-agricultural for the linked migrants from one census year to the linked census year for the city of Chicago. N is the total number of linked cases and does not represent the whole population.		

Linked year	% engaged in the non-agricultural sector[3] migrating 100 miles or less	% engaged in the non-agricultural sector[3] migrating 100 - 200 miles	% engaged in the non-agricultural sector[3] migrating 200 - 500 miles	% engaged in the non-agricultural sector[3] migrating 500 miles or more
1870-1880	25	*	63	61
1880-1900	66	71	34	32
1880-1910	13	*	*	*
*Large number of unknown occupations and non- occupational responses				
Table 2: Change in occupation for the linked migrants from one census year to the linked census year. Percent change in occupation is listed with respect to circular buffers with a radius of 100, 200, and 500 miles from the city of Chicago.				

Preliminary conclusions

Existing research on historical migration is based on statistical data on a group of migrating individuals. This paper illustrates a new way to look at each person and follow the same person through time. One of the major drawbacks of working with longitudinal and historical datasets is the small number of persons that can be linked and the large number of unknown attributes in the datasets. Nevertheless, working with historical cities and analyzing the spatial and temporal pattern of migration provides a novel approach to looking at this age old research. Our next step will expand on the finding in this preliminary study by looking closely at the occupations and comparing the eastern cities with the western city of San Francisco.

References

A. Sula Sarkar and Rebecca Vick, Using population weighted county centroids to calculate migration distance for historical linked census data: A case study of male migrants in the New York and New Jersey area, 1880-1910, Population Association of America, 2011
B. Downloaded from: http://usa.ipums.org/usa/linked_data_samples.shtml
C. Steven Ruggles, J. Trent Alexander, Katie Genadek, Ronald Goeken, Matthew B. Schroeder, and Matthew Sobek. Integrated Public Use Microdata Series: Version 5.0 [Machine-readable database]. Minneapolis: University of Minnesota, 2010
D. Minnesota Population Center. North Atlantic Population Project: Complete Count Microdata. Version 2.0 [Machine-readable database]. Minneapolis, MN: Minnesota Population Center, 2008.
E. Minnesota Population Center. *National Historical Geographic Information System: Pre-release Version 0.1.* Minneapolis, MN: University of Minnesota 2004.

---

[3] Non-agricultural occupation includes persons engaged in trade and transportation as well as the manufacturing, mechanical, and mining industries. Note that it does not include professional and personal services.