Urban Kin Propinquity Using Geocoded Complete Count Census Data, 1880

Matt Nelson
Senior Data Analyst
Institute for Social Research and Data Innovation
University of Minnesota

Declining kin propinquity in the United States between 1790 and 1940 has recently been established. Historically, urban kin propinquity rates were far lower than rural rates. While some evidence supports lower kin propinquity in urban areas compared to rural areas, previous estimates included some methodological shortcomings. Geocoded Census data corrects for some of the previous methodological issues. Using geocoded data for 39 cities in the 1880 Census, I attempt to provide an updated measure for urban kin propinquity. Numerous checks to insure accurate kin propinquity links include controlling for common surnames, identifying ideal household distance thresholds to identify kin propinquity, and consistent life course results for those with propinquitous kin. I hypothesize that results will be slightly higher than previous estimates and validate that previous sequential isonymic linking methods provided an accurate description of kin propinquity in urban areas historically.

New complete count census data allows researchers to study kin propinquity beyond the household in new ways, expanding our knowledge of the prevalence of nearby kin and their effects on demographic behavior. Recent research includes nearby mother-in-laws increased married women's net fertility by 2% in 1880 (Hacker & Roberts, 2017) and the decline of kin propinquity from 30% of households in 1790 to 6% by 1940. Kin propinquity was heavily influenced by the family life cycle where younger generations lived near siblings and parents, while older generations primarily lived near children (Nelson, forthcoming, Smith, 1989). These previous kin propinquity measures relied on the Census' sequential enumeration of households to measure kin associations. Increased urbanization explains a substantial part of the decline in kin propinquity, but sequential isonymic matching methods likely underreport urban kin propinquity. Because of this underreporting, it is unclear to what degree the decline in kin propinquity was due to increased urbanization or measurement error that does not capture kinship associations as well in urban areas. This paper aims to provide better reporting of kin propinquity in urban areas to better determine

how much of the decline in kin propinquity was associated with urbanization and improve kin propinquity measurement accuracy.

Researchers in the mid twentieth-century argued that urban kin networks were far stronger than previously thought (Litwak 1960, 1965, Shanas 1961, 1967, 1973). Rural kinship patterns related to geographic proximity had stronger mechanisms than in urban areas. A few examples include land inheritance patterns in agriculture (Ditz, 1986, Gjerde, 1997) and more economic opportunities via wage labor in urban areas allowing for children to leave the household (Ruggles, 2007, 2015). However, intergenerational exchange of resources, with parents providing financial assistance to younger generations, and younger generations caring for elderly parents, still provided a strong mechanism for urban kin propinquity. While urban kin propinquity theoretically was lower than rural areas, methodological error in urban areas could explain part of the long-run decline of kin propinquity in the United States.

Sequential measures of kin propinquity rely on the sequential ordering of census returns to capture kin associations within enumeration districts. By comparing surnames of households near each other, same surname households are considered potential kin (Figure 1). This methodological approach, while illuminating, is imperfect, particularly for urban areas. Sequential ordering does not arrange city blocks accurately, as a household on the opposite side of the street was not enumerated sequentially. Further, distances between households in urban areas compared to rural areas is far smaller, leading to potentially misleading results on kin propinquity with low distance thresholds used in previous research. One approach to alleviate these two particular issues is to use geocoded data. With geocoded data, surnames can be compared more efficiently and do not rely on the sequential nature of the Census. Previous studies of racial segregation used next-door neighbor analysis and geocoded data analysis, both of which expanded our understanding of racial

segregation historically (Grigoryeva & Ruef, 2015, J. Logan & Bellman 2016, T. Logan & Parman 2017).

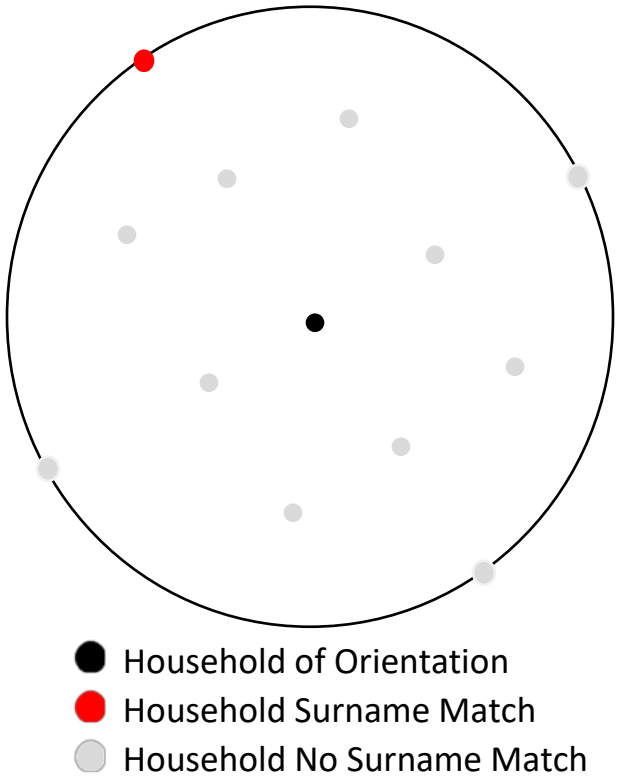
Figure 1: Example of Sequential Kin Propinquity

SERIAL	PERNUM	NAMEFRST	NAMELAST	KIN LINK
16638343	1	JAMES	CAMERFORD	0
16638343	2	MINNIE	CAMERFORD	0
16638343	3	PAUL	CAMERFORD	0
16638343	4	NINA	CAMERFORD	0
16638344	1	PETER	RASMUSSEN	0
16638344	2	CATHERINE	RASMUSSEN	0
16638344	3	IOALA	RASMUSSEN	0
16638344	4	HELEN	RASMUSSEN	0
16638344	5	GLEN	RASMUSSEN	0
16638344	6	LILLIAN	RASMUSSEN	0
16638345	1	CARROLL	GREEN	0
16638345	2	MARRIEN	GREEN	0
16638345	3	MARION	GREEN	0
16638346	1	EDWIN	JOHNSON	1
16638346	2	ADOLPH	JOHNSON	1
16638346	3	LILLIA	JOHNSON	1
16638347	1	RONALD	JOHNSON	1
16638347	2	ESTHER	JOHNSON	1
16638348	1	ANDREW L	JOHNSON	1
16638348	2	LENA	JOHNSON	1
16638348	3	MOYEL E	JOHNSON	1
16638348	4	HASLEY A	JOHNSON	1
16638348	5	JOAN B	JOHNSON	1
16638349	1	LEO	HALLING	0
16638349	2	LAURA	HALLING	0
16638349	3	LAURALL	HALLING	0

The census data comes from the Integrated Public Use Microdata Series (IPUMS) from the Institute of Social Research and Data Innovation (ISRDI) to measure kin propinquity (Ruggles et al, 2017). The geocoded data of 39 cities from the 1880 Census was created by the Urban Transition Historical GIS Project (UTP, J. Logan et al, 2011). The geocoded data was based on the IPUMS data, so the datasets can be merged directly. Because of the current geocoded data that

is publicly available, I only analyze 1880. Approximately 97% of the cases I use to analyze kin propinquity have geocoded addresses, approximately 6.7 million persons in 1880.[1] With the geocoded data, surnames from the household of orientation will be compared with surrounding households. I then measure the probability that the surname randomly occurs between the households. If the probability of a surname randomly occurring is greater than 1%, I disregard the match as a potential false kinship association. If the probability is less than 1%, I keep the match as potential kin propinquity.[2] Figure 2 shows a link made between two households with the surname using the geocoded data. If we consider all households within an equal distance, there are

Figure 2: Example of Geocoded Matching



[1] Previous sequential kin propinquity measures have only analyzed individuals living in non-group quarters. To compare the results of geocoded kin propinquity to sequential matching, I use the same universe of analysis. Only exact matching surnames are currently measured.

[2] This probability threshold was used in previous papers measuring kin propinquity using the sequential method.

12 other households. If this particular surname makes up 0.08% of the surnames, the remaining surnames compose 99.92% of the surnames, meaning within one household, there is a 0.08% chance that the surname occurs randomly.[3] Therefore, the probability of this particular surname occurring randomly within this distance is 0.96%. Given that the chance of the surname randomly occurring is less than 1%, these households would be consider propinquitous kin. To measure kin propinquity, I use distance thresholds of 0.25 mile, 0.5 mile, 1 mile, and 3.38 miles.[4]

I hypothesize that urban kin propinquity rates will be higher than those estimated using the sequential method, but kin propinquity rates will still be lower than rural and agricultural kin propinquity rates. These results will inform the sequential method, showing either that the sequential method overall is a relatively accurate measure of kin propinquity trends, or that the method needs to be revised to better capture kin propinquity in urban areas. Further, it will expand our knowledge on kinship networks within urban areas historically.

I have run preliminary testing on kin propinquity in Minneapolis, MN (Figure 3). Kin propinquity rates using the sequential method were 2.8% to 3.7% depending on the distance threshold. When using geocoded data, kin propinquity rates increased to 5.7% at 0.25 miles away to as high as 8.3% 3.38 miles away. The average distance of households within a quarter of a mile was approximately 0.06 miles away. However, when looking at household distance, the average number of households away (125) is highly skewed because of outliers. When using the median distance, the number of households away is 9 households. The largest distance threshold used (3.38 miles away) gives a kin propinquity rate of 8.3%, with the average distance being 0.3 miles away.

[3] I calculate the composition of surnames by state and race.

[4] I use 3.38 miles as a potential threshold to investigate because many rural townships were measured using the Public Land Survey System (JCC, 1904). This system divided new lands in the United States into townships and ranges. With the exception of natural and county boundaries, many townships/ranges were 6 miles square (Area=lw=6*6=36). To measure the number of households within a circular area around the household of orientation, 3.38 miles away would be equivalent to miles 6 square ($A=\pi r^2=\pi*3.38^2=36$).

The median distance is 0.1 miles away and 73 households. If these results are representative of the other 38 cities, national kin propinquity rates could increase as much as 2 percentage points for 1880.

Figure 3: Kin Propinquity Rates and Distance Metrics by Method in Minneapolis, 1880

Method *(Sample Size)*	Distance Threshold	% Kin Propinquity	Average Link Distance		Median Link Distance	
			Miles	Households	Miles	Households
Sequential Matching *(N=41,585)*	3 Households	2.8%		1.3		1
	10 Households	3.2%		2.0		1
	25 Households	3.6%		3.5		1
	50 Households	3.7%		4.5		1
Geocoded Matching *(N=38,165)*	0.25 mile	5.7%	0.06	125	0.03	9
	0.5 mile	6.5%	0.10	222	0.06	18
	1 mile	7.6%	0.19	449	0.08	45
	3.38 miles	8.3%	0.30	694	0.10	73

These preliminary results fit with my hypothesis of urban kin propinquity rates. The remaining 38 cities still need to be processed, and further diagnostics need to be run to confirm whether the surname links made should be considered kinship associations, or if the analysis of urban kin propinquity needs to be further revised. Diagnostics include the ages of linked households (assuming a family life cycle pattern, we should see younger households linking with similar age households and households approximately 20-40 years older than them), and older households linking with younger generations approximately 20-40 years younger. One limitation of the geocoded method that I want to explore is the bias of outer-ring residents of a city. For example, someone living near the center of the city has a greater probability of being removed from analysis because of common surnames appearing randomly than someone living near the edge of the city. Because of this, I hypothesize that residents living further away from the city

center are more likely to not be removed from analysis and exhibit higher kin propinquity rates than those living closer to the city center.

Sources Cited

Ditz, Toby L., *Property and Kinship: Inheritance in Early Connecticut, 1750-1820,* (Princeton: Princeton University Press, 1986).

Gjerde, Jon, *The Minds of the West: Ethnocultural Evolution in the Rural Middle West, 1830-1917,* (Chapel Hill: University of North Carolina Press, 1997).

Grigoryeva, Angelina, & Ruef, Martin, "The Historical Demography of Racial Segregation," *American Sociological Review*, 80:4 (2015), 814-842.

Hacker, J. David, & Roberts, Evan, "The Impact of Kin Availability, Parental Religiosity, and Nativity on Fertility Differentials in the Late 19th-Century United States," *Demographic Research*, 37:34 (2017), 1049-1080.

Journals of the Continental Congress, 1774-1789, ed. Worthington C. Ford et al. (Washington, D.C.: U.S. Government Printing Office, 1904-37), 28:375-381.

Litwak, Eugene, "Geographic Mobility and Extended Family Cohesion," *American Sociological Review*, 25:3 (1960), 385-394.

Litwak, Eugene "Extended Kin Relations in an Industrial Democratic Society," in *Social Structure and the Family: Generational Relations,* ed. Ethel Shanas, & Gordon F. Streib, (Englewood Cliffs: Prentice-Hall, Inc., 1965), 290-325.

Logan, John R., Jindrich, Jason, Shin, Hyoungjin, & Zhange, Weiwei, "Mapping America in 1880: The Urban Transition Historical GIS Project," *Historical Methods* 44:1 (2011), 49-60.

Logan, John R., & Bellman, Benjamin, "Before *The Philadelphia Negro:* Residential Segregation in a Nineteenth-Century Northern City," *Social Science History* 40:4 (2016), 683-706.

Logan, Trevon, & Parman, John M., "The National Rise in Residential Segregation," *Journal of Economic History*, 77:1 (2017), 127–170.

Nelson, Matt, *Relieved of These Little Chores: Agricultural Neighbor Labor, Family Labor, and Kinship in the United States, 1790-1940,* (Dissertation: University of Minnesota, 2018).

Ruggles, Steven, "The Decline of Intergenerational Coresidence in the United States, 1850 to 2000," *American Sociological Review,* 72:6 (2007), 964-989.

Ruggles, Steven, "Patriarchy, Power, and Pay: The Transformation of American Families, 1800-2015," *Demography*, 52:6 (2015), 1797-1823.

Ruggles, Steven, Genadek, Katie, Goeken, Ron, Grover, Joe, & Sobek, Matthew, *Integrated Public Use Microdata Series: Version 7.0* [dataset], Minneapolis: University of Minnesota, 2017, https://doi.org/10.18128/D010.V7.0.

Shanas, Ethel, "Living Arrangements of Older People in the United States," *The Gerontologist*, 1:1 (1961), 27-29.

Shanas, Ethel, "Help Patterns and Social Class in Three Countries," *Journal of Marriage and Family*, 29:2 (1967), 257-266.

Shanas, Ethel, "Family-Kin Networks and Aging in Cross-Cultural Perspective," *Journal of Marriage and Family*, 35:3 (1973), 505-511.

Smith, Daniel Scott, "'All in Some Degree Related to Each Other': A Demographic and Comparative Resolution of the Anomaly of New England Kinship," *The American Historical Review*, 94:1 (1989), 44-79.