Like A Good Neighbor: Childhood Neighbors Influence Occupation Choice*

Michael Andrews† Ryan Hill‡
Joseph Price§ Riley Wilson¶

August 1, 2025

Abstract

We explore the role of immediate next door neighbors in affecting children's later life occupation choice. Using linked historical census records for over 6 million boys and 4 million girls, we reconstruct neighborhood microgeography to estimate how growing up next door to someone in a particular occupation affects a child's probability of working in that occupation as an adult, relative to other children who grew up farther away on the same street. Living next door to someone as a child increases the probability of having the same occupation as them 30 years later by about 10 percent. As an additional source of exogenous variation in exposure to next door neighbors, we exploit untimely neighbor deaths and find smaller and insignificant exposure effects for children who grew up next to a neighbor with an untimely death. We find larger exposure effects when intensity of exposure is expected to be higher, and document larger occupational transmission in more connected neighborhoods and when next door neighbors are the same race or ethnicity or have children of similar ages. Childhood exposure to next door neighbors has real economic consequences: children who grow up next to neighbors in high income or education occupations see significant gains in adult income and education, even relative to other children living on the same street, suggesting that neighborhood networks significantly contribute to economic mobility.

Keywords: Neighborhood networks, peer effects, occupational transmission
JEL Codes: J24, J62, N32

---

*Thanks to Christian Ames, Thomas Barden, Carver Coleman, Maxwell Dunbar, Zach Flynn, Jimena Kiser, Josh Nicholls, Abbie Sanders, Rebekah Todd, Margaret Truman, and Conner Wilkinson for outstanding research assistance, as well as audiences at Clemson; Northwestern; NYU; UC Merced; Towson; Yale; Mannheim; University of Southern Denmark; Rutgers; Oklahoma; Economic History Association Meetings; the NBER Mobility Conference; CEPR Applied Micro-Economic History Workshop; Association for the Study of Economics, Religion, and Culture Conference; and the Cliometric Conference for helpful comments.

†University of Maryland Baltimore County, Email: mandrews@umbc.edu
‡Northwestern University, Email: ryan.hill@kellogg.northwestern.edu
§Brigham Young University, NBER, Email: joseph_price@byu.edu
¶Brigham Young University, CESifo, IZA, Email: riley_wilson@byu.edu.

1 Introduction

Growing evidence suggests that the neighborhood in which a child grows up influences their future earnings and economic mobility (Chetty et al., 2018, 2014), and by exploiting exogenous movement across neighborhoods, several studies concludes that these effects are causal (Bergman et al., 2024; Chetty and Hendren, 2018a,b; Chetty et al., 2016; Chyn, 2018; Chyn et al., 2025; Haltiwanger et al., 2020; Katz et al., 2001; Kawano et al., 2024; Kling et al., 2007; Rosenbaum, 1995). But why do neighborhoods matter? Previous work has focused on the role that neighborhoods play in providing access to education, health, and safety amenities and infrastructure that augment investments in children (Billings et al., 2014; Duncan et al., eds, 2018; Laliberté, 2021; Sharkey and Torrats-Espinosa, 2017).

In this paper, we focus on another potentially important mechanism: the role of neighbors. In addition to amenities, neighborhoods foster person-to-person interactions that potentially generate human capital spillovers from peers, mentors, and role models. We exploit door-to-door census enumeration, a feature of historical population censuses that allows us to reconstruct neighborhood microgeography and identify the influences of immediate next door adult neighbors on the future occupation choices of children, a key decision impacting lifetime earnings. While much of the previous literature on causal effects of neighborhoods focuses on children who grew up in poor neighborhoods, our use of complete census data allows us to estimate the influence on next door neighbors for children growing up across the entire U.S., and to explore various dimensions of neighborhood, individual, and occupational heterogeneity.

Many of the studies that examine social influences on children's later career outcomes investigate the role of parents (Bell et al., 2019; Corak and Piraino, 2011; Fairlie and Robb, 2007; Hvide and Oyer, 2018). But other adult connections in the neighborhood beyond parents also likely play an important role in shaping children's futures. As Chetty and Hendren (2018b) and Chetty et al. (2022) show, the demographic makeup of neighborhoods and the level of connection between high-SES and low-SES individuals are some of the strongest predictors of neighborhood intergenerational mobility effects. Many neighborhood effects are highly localized, and local, neighborhood employment networks are strong.[1] The ability to easily interact face-to-face, as is possible between nearby neighbors, facilitates the spread of ideas.[2] Other studies find that geographically proximate peers influence children's schooling choices (Avdeev et al., 2023; Barrios-Fernández, 2022; Matta and Orellana, 2022). These strands of literature all point to nearby neighbors as a potentially important channel through which neighborhood effects operate.

To identify the impact of next door adult neighbors on occupation choice, we compare individuals who live next door to adults in a particular occupation to other children who live on the same part of a street but farther away from the adult. Our empirical strategy can be illustrated with a simple thought experiment. Suppose that Max lives next door to Dr. Smith. Carl lives five doors away from Dr. Smith on the same street. Is Max more likely to become a doctor than Carl? We scale this thought experiment to estimate exposure effects for all children who lived next door to doctors (and many other occupations) in the 1910 U.S. population census. Importantly, we use narrow geographic fixed effects to focus our analysis on the comparison of children living on the same census manuscript sheet (which is typically a subset of a single street, about 8 households on average), but are exposed to next door neighbors with different occupations. Our analysis relies on the identification assumption that, while selection into neighborhoods may not be random, selection of immediate next door neighbors (sorting of households within a particular subset of a street) is as-good-as-randomly assigned. Given the unpredictability of housing markets and the difficulty of observing occupations of specific neighbors before moving into a residence, we believe this

[1] See Aliprantis (2017); Billings et al. (2024); Grinblatt et al. (2008); McCartney and Shah (2022); Redding and Sturm (2024); Shah and McCartney (2023) for several studies of neighborhood effects that exert effects at highly localized levels, often within a city block or less than about a tenth of a mile. Other studies specifically examine the importance of employment networks among individuals living on the same street (Bayer et al., 2008; Tan, 2022) or census tract (Hellerstein et al., 2011).

[2] See, e.g., Andrews (2019); Andrews and Lensing (2024); Arzaghi and Henderson (2008); Atkin et al. (2022); Catalini (2018); Moretti (2021).

to be plausible. In a complementary design, we exploit unexpected deaths of next door neighbors to identify exogenous variation in children's exposure to adult occupations.

We take advantage of two key features of the de-anonymized historical census data and related data sets. First, we use modern machine learning methods and user-contributed linkages that allow us to track over 10 million children across censuses, from their childhood neighborhood into their adult careers. These approaches offer a large improvement over alternative linking methods, especially for girls. Second, we exploit the fact that, prior to 1970, historical U.S. censuses were collected by enumerators going door-to-door. By examining the ordering of households on census manuscript pages we can reconstruct the microgeography of a neighborhood.

To begin, we illustrate our approach by studying the effects of growing up next door to one exemplary occupation: doctors.[3] We find that boys who live next door to doctors in 1910 are 41% more likely to be doctors as adults in 1940 than are other boys residing on the same census manuscript sheet but farther away from the doctor. To put this magnitude into perspective, having a doctor as a next door neighbor is about one-thirtieth as predictive that a child will become a doctor as is having a doctor in the child's own household; having a doctor in the same household makes a child 12.4 times more likely to become a doctor than other children on the same sheet.

Next, we extend this approach to examine the top 50 largest non-farm occupations for men in the 1910 census and the top 25 largest occupations for women.[4] We first estimate each occupation separately, similar to our doctor analysis. While there is substantial heterogeneity across occupations, the point estimate on next door neighbors is positive for all but four of the male occupations, is statistically significantly greater than zero in 26 out of the 50 occupations, and is never statistically significantly negative. Because fewer women were participating in the labor force in the early 1900s, we have less power to estimate precise

[3] Several studies use doctors as a case study of intergenerational transmission of occupations from parents to children (Lentz and Laband, 1989; Polyakova et al., 2020; Ventura, 2024).

[4] Because women's employment was highly concentrated in both 1910 and 1940 we restrict our analysis to fewer occupations.

transmission effects for women, but we similarly find positive (albeit lower in magnitude) effects for girls. We then combine information on all 50 occupations into a single "stacked regression" in which each child appears as an observation in the regression multiple times for each occupation that appears on their census sheet. A boy is about 10% more likely to enter into the average occupation when they live next door to an individual in that occupation than are other children on the same census sheet. A girl is about 6% more likely to enter their neighbor's occupation compared to other girls on the street. We show that these findings are robust to several alternative specifications and samples of the data that probe at endogeneity and measurement error.

As an additional source of exogenous variation in exposure to next door neighbors, we use data on untimely deaths of adult neighbors. Using data from Price (2024), we identify adult neighbors who were younger than certain age thresholds in the 1910 census but died before 1915. We find smaller and statistically insignificant exposure effects for children who grew up next to adult neighbors that die before 1915, while exposure effects for children who grew up next to adult neighbors who did not die by 1915 see larger and statistically significant effects comparable in magnitude to our baseline effects.

We next show that exposure effects are larger when the intensity of exposure is expected to be greater. We find that exposure effects decline monotonically for children located more doors away from the adult neighbor on the same sheet.

To further understand why childhood neighbors matter, we document how heterogeneity across neighborhood and person-specific characteristics alters the magnitude of exposure effects. Consistent with neighborhood connectedness, boys are more likely to adopt their next door neighbors' occupation in rural areas relative to urban areas, and in places where a smaller share of residents were born in other places. We also find strong patterns of homophily between neighbors. Boys are more likely to adopt their neighbor's occupation when they share a birthplace, are of the same race, are of the same gender, or when the families have children of similar ages. Overall, these results suggest that children are influenced in their career choices by their interactions with people in their social networks.

Importantly, the occupation composition of childhood neighbors has real economic implications. Boys are more likely to adopt the occupation of their next door neighbor when it is a high income occupation. Additionally, boys growing up next door to a high-income, high-education worker, such as a doctor or lawyer, have significantly higher education and earnings as adults relative to other boys on the same census sheet. Growing up next door to a lower income occupation such as a porter, truck driver, or laborer actually results in lower income for the children 30 years later. We find evidence that, in addition to making neighbors more likely to enter their own occupation, neighbors also convey more general human capital to neighboring children. For example, growing up next to a teacher increases the probability of working in other high-education occupations like doctors and lawyers, many of which have higher average incomes than teachers. Results on economic outcomes are particularly striking for girls. The growth of professionalized occupations, such as teachers, nurses, and stenographers, meant that exposure to these new and promising career paths was an important influence on girls educational attainment and future income, which could have intergenerational effects.

Our paper builds on the work of economic and social historians that have exploited the special structure of the historical censuses. The use of census sheet order to measure residential racial segregation was pioneered by Agresti (1980) and Logan and Parman (2017a).[5] Beyond measuring residential segregation, Tan (2022) and Quincy (2022) and compare individuals on the same census sheet to people living on nearby sheets to study industry composition and the effect of income shocks on homeownership, respectively. In an influential study using modern data, Bayer et al. (2008) compare people living on the same city block in Boston to those on neighboring blocks to study local hiring networks. Relative to these studies, we can focus on extremely local geographic differences in exposure, comparing

[5] These measures have subsequently been used to study the relationship between residential racial segregation and lynching (Cook et al., 2018), homeownership (Logan and Parman, 2017b), mortality (Logan and Parman, 2018), and present-day neighborhood-level economic mobility (Andrews et al., 2017). Similar techniques have been used to measure residential segregation of immigrants (Eriksson and Ward, 2019).

next door neighbors to those living on the same census sheet, further minimizing concerns about unobserved sorting within neighborhoods. Perhaps the most similar research design is Bayer et al. (2022), which uses modern mortgage and housing transaction data to compare move rates for homeowners who receive an immediate next door neighbor of a different race to the move rate of others farther away on the same block.[6] Because of historical concerns about residential racial discrimination, the mortgage data include information about borrowers' race; these data are therefore well-suited to study racial change, but do not include the rich occupation and demographic information included in the population censuses. Additionally, the mortgage data only include information on homeowners, whereas the census data we use include anybody residing in a given household. We use the richness and comprehensiveness of the census data to investigate outcome variables, dimensions of heterogeneity, and alternative identification strategies not possible with more restrictive data. Most importantly, relative to all of these studies, we combine the highly detailed microgeographic data on children to adult outcomes, allowing us to study long-run consequences of childhood exposure.

This work adds to our understanding of neighborhood effects, intergenerational mobility, and occupation choice. Although connectedness is a strong predictor of intergenerational mobility (Chetty et al., 2022), there is a dearth of evidence to causally identify this mechanism. We are able to provide causal estimates of how adult neighbors affect children's long-run trajectories. Most studies of adults' influence typically investigate the correlation between parents and children's later career outcomes (Bell et al., 2019; Corak and Piraino, 2011; Fairlie and Robb, 2007; Hvide and Oyer, 2018). Studies of parents' influences, however, are typically unable to disentangle the effects of environment and informational spillovers from genetic endowments and intra-family transfers such as inheriting a family business.[7]

[6] Also see McCartney and Shah (2022); Shah and McCartney (2023), which use these same data and also compare next door neighbors.

[7] Some exceptions are Dal Bó et al. (2009) and Greenberg et al. (2024) which use regression discontinuities in vote shares and AFQT scores to isolate plausibly exogenous variation in the parent's probability of being a politician or serving in the military, respectively.

By estimating the effect of next door neighbors, we remove these last two channels and investigate the importance of information and exposure spillovers. This paper also joins a recent literature focused on the role of information, exposure, and mentoring in children's (especially girls') decisions to enter high-income and high-education careers (Andrews and Zhao, 2025; Breda et al., 2023; Mertz et al., 2024; Olivetti et al., 2020; Porter and Serra, 2020).

The paper proceeds as follows. Section 2 discusses the census data, in particular how we construct links from children in 1910 to adults in 1940 and how we construct measure of microgeographic proximity. Section 3 describes our regression framework and discusses our identification strategy. Section 4 presents our baseline results for doctors, for the 50 largest male occupations separately, the 25 largest female occupations, and for the stacked regressions. Section 5 explores mechanisms of transmission, focusing the intensity of exposure and on homophily within neighborhoods and between neighbors. Section 6 disentangles how exposure to next door neighbors affects economic outcomes of children when they are adults. Finally, Section 7 briefly concludes.

2 Data

Existing work has relied on occupation case studies (e.g., congressional legislatures (Dal Bó et al., 2009), the military (Greenberg et al., 2024)) or administrative employment records, like the Longitudinal Employer Household Dynamics linked to census records (Staiger, 2023), to document intergenerational transmission of occupation. Outside of the US, Norwegian registry data has been used to explore the propensity of sons to follow the occupation choice of fathers (Hvide and Oyer, 2018). However, even these datasets are not suited to answer the question at hand: does a child's exposure to adult neighbors affect their eventual career choice? To answer this question, we must not only be able to link a child to their parents' occupations, but must simultaneously be able to observe the occupations of all their geographically proximate neighbors. Long-running surveys like the Panel Study of Income Dynamics do not contain this information on neighbors and even recent innovations in modern census record linking through the Personal Identification Key (PIK) either restrict analysis to survey samples or the PIKed full-count censuses do not span enough time to observe the children as adults.

For this reason, we exploit individual-linked full count census data for 1910 to 1940. The US Census Bureau releases personally identifiable information from the decennial census after 72 years. Digitized versions of the original census sheets, filled out by hand by Census enumerators, are publicly available and include individuals' names, sex, birth year, birth place, occupation, and address.[8] Each individual's information is recorded on a census line with a family identifier, thus allowing us to connect families within a given point in time.

We use three main data sources: the 1910 full count census, the 1940 full count census, and the Census Tree database of individual links, initially developed in Price et al. (2021) and described in more detail in Buckles et al. (2023).

2.1 Full Count Census Data

We obtain full count census micro data for 1910 through 1940 from Ancestry. This includes all of the digitized information on a census sheet including state, city, enumeration district, address (when recorded), household id, census sheet number, census line number, name, sex, relation to head-of-household, marital status, year of birth, place of birth, employment status, and occupation. We merge the de-identified full count census data to IPUMS full count census records using historical IDs (HISTID) to work with cleaned, categorized measures of occupations. Using the 1940 census, we can observe each child's own occupational choice as an adult, 30 years later. In some robustness analyses, we also use 1920 and 1930 full count census micro data, to explore the dynamics of occupation transmission over time.

[8] As discussed in detail below, address was not recorded for every individual.

2.2 Census Tree Links

The Census Tree is a dataset that provides over 700 million links for individuals across historical US census records. This dataset was originally created by Price et al. (2021) and further expanded and refined in Buckles et al. (2023). The Census Tree builds on family tree data from one of the world's largest internet genealogy platforms, FamilySearch.org. FamilySearch users add and link historic records to the public profiles of their own relatives. These profiles are connected together through family relationships into a large interconnected network of profiles called the Family Tree. FamilySearch users often have private information that allows them to link records that would not be possible for a trained research assistant or machine learning algorithm. Since the the Family Tree is an open-edit wiki-style platform, mistakes made by one user can be edited by other users. The Family Tree can be used to create links between census records by looking for pairs of census records that are attached to the same profile. The census-to-census links for men alone number more than 158 million, greater than the number of conservative links in the Census Linking Project (Abramitzky et al., 2022).

The Census Tree uses the hand-linked records from the Family Tree as training data to develop a new machine learning algorithm to identify additional linkages. The Census Tree then combines these new machine learning links with links from the Census Linking Project (CLP) and the Multigenerational Longitudinal Panel (MLP), as well as the Family Tree links and a set of machine learning links created by FamilySearch. One of the key innovations of the Census Tree is to combine together links from multiple methods and then use a set of decision rules to adjudicate disagreements across the different linking methods. The final result is a dataset with 391 million links for men and 314 million links for women across the 1850 to 1940 censuses.

To construct our analysis sample we start with the universe of children between the ages of 5 and 18 in the 1910 full count census. The children observed in 1910 are then linked to outcomes in the 1940 census, when they are between 35 and 48 years old. In the 1940 census we observe the individual's occupation, along with other outcomes, such as employment, educational attainment, and wage income. To link individuals from 1910 to 1940 we use all available Census Tree links. This yields a sample of approximately 10.5 million, including 6,346,719 men and 4,182,461 women. As seen in Table 1, this represents 40.2 percent of all children observed in 1910.[9] Table 1 shows that our linked sample comes from household settings that are broadly representative of the full population, although our linked sample is more white than the full population, which is perhaps unsurprising given the difficulty linking Black individuals during this time period.

Because women historically change their last name upon marriage, traditional linking methods perform poorly when linking women. For this reason researchers often exclude women or focus on subsamples of women that are easier to follow over time. Our ability to evaluate outcomes for girls is therefore an innovation relative to the prior literature. The patterns and opportunities for employment were quite different for boys and girls in 1910, however. In 1910, only 29.3 percent of women 18-54 were in the labor force, compared to 95.5 percent of men. These gendered patterns were largely unchanged by 1940, when the labor force participation rate of women 18-54 was only 31.7 percent (90.7 percent for men). Given these patterns, we view the occupational choice effect of boys and girls as both interesting, but distinct, questions. As such, the set of occupations we consider and the estimated effects will be sex-specific to correctly interpret patterns in the historical context.

2.3 Measuring Neighbors in the Full Count Census

Our empirical approach relies on identifying geographically proximate households using the 1910 census records. Importantly for us, census enumerators were given explicit instructions on the order in which they were to collect and record household information. "It is your duty *personally* to visit every family and farm within your district... Canvassing of blocks should go in order around the block, not switching back and forth across the street" (U.S. Bureau

[9]This likely understates the true linking rate as some children in 1910 have likely died or emigrated by 1940.

of the Census, 1910, emphasis in original). As specified in the 1910 enumerator instruction guide, the intent was to record families on a census sheet *neighbor-by-neighbor*. Because the enumerator records households in the order that one would walk down the street, the order of households in the census sheet likely captures the kind of proximity most relevant for frequent social interactions. We use information on the enumeration district, sheet (i.e., page) number and line number to identify likely neighbors with fine geographic precision.

We estimate occupational transmission for many occupations, but for clarity we describe our data creation process just looking at doctors. First, we flag all households across all 1910 census sheets that include a child age 5-18 and all households across all 1910 census sheets that include an individual with the target occupation: physicians and surgeons (1950 occupation code 075). We then collapse the data to include one observation per household, preserving the geographic measures (city, state, enumeration district, census sheet number) and the within-sheet ordering of households. If household members spread across multiple sheets we reassign them to the census sheet of the household head. For each household we then construct a set of binary indicators. The first equals one if there is a doctor in one's own household. The second equals one if there is a doctor on the same sheet and one household above or one household below one's own household ("next door"). Thus when we limit the sample to focal children we can observe if they have a doctor in their own household or one next door. For our baseline analysis we do not restrict the gender of the next door neighbor in the target occupation; we explore gender homophily in Section 5. We repeat this process for each target occupation separately. We focus on the 50 largest occupations in terms of 1910 male workforce for boys and the 25 largest occupations in terms of 1910 female workforce for girls. In both cases we exclude farmers and farm laborers.[10]

[10] We exclude farmers from this set of occupations for four reasons. First, farmers often live in communities where most households have the same occupation; it is likely that farmers who live on a street with no other farmers, for instance a lone farmer in an urban neighborhood, are very different from the average farmer. Second, the nature of farming changed substantially between 1910 and 1940. Third, because farmers often lived on large farms, the sheet ordering is less tied to geographic proximity. Finally, farmers often report no wage income; in the analyses in Section 6, we consider heterogeneity across occupations in terms of income, among other dimensions, making it difficult to know how to classify farmers. We show that our baseline results are robust to including farmers in Appendix Table C1.

Appendix A provides much more information on using the census sheet order to reconstruct neighborhood microgeography. Given census enumeration instructions, sheet ordering should allow us to accurately identify geographically proximate neighbors. In a subset of sheets, street addresses have been recorded and transcribed; we use these, in conjunction with historical Sanborn Fire Insurance maps(Sanborn Map Company, Various Years), to validate the accuracy on the census sheet order in correctly identifying next door neighbors.[11] To the best of our knowledge, we are the first to attempt to assess the extent to which census sheet order successfully measures the proximity of neighbors. We find that the sheet-order based next door neighbor measure is correct 91 percent of the time. Appendix A additionally details potential sources of measurement error when using sheet order to determine proximity, as well as showing various approaches to address this potential measurement error and showing that our results are robust. We discuss alternative methods to reconstruct neighborhood geography using the street addresses, and using the street addresses to instrument for measurement error when using the sheet order, in Section 4.5 and Appendix D.

3 Empirical Approach

To fix ideas, we introduce our empirical approach with an illustrative example of a single occupation: doctors. We construct an OLS regression that tests whether a boy living next door to a doctor in 1910 is more likely to become a doctor by 1940 relative to other boys on the same census sheet. With this simple framework in hand, we can then broaden the analysis to study the transmission of an arbitrary occupation for both girls and boys. We will also introduce a stacked OLS regression framework that allows us to test average transmission effects for the full set (or subsets) of occupations in a unified approach.

[11] About 35% of households in the 1910 census do not contain any street address information. In contrast, we can use the census sheet order for all households in the census.

Our baseline estimating equation is as follows

$$\text{Doctor1940}_{is} = \alpha + \beta_0 \text{OwnDoc1910}_i + \beta_1 \text{NextDoorDoc1910}_{is} + \text{Age}_i + \gamma_s + \epsilon_{is}, \quad (1)$$

where OwnDoc1910_i is an indicator equal to one if there is an adult doctor living in the same household as i and $\text{NextDoorDoc1910}_{is}$ is an indicator equal to one if one of the households nextdoor to i has an adult who is a doctor. In other words, if $\text{NeighborDoc1910}_{is} = 1$, then the focal child has at least one doctor that lives one house away in either direction on their side of the street as represented by the census sheet ordering. Age_i is a fixed effect for the focal child's age in the 1910 census. Finally, γ_s is a census sheet fixed effect, and ϵ_{is} is an idiosyncratic error term that includes all unobservable local amenities and individual- and household-level characteristics. We cluster standard errors at the household level to account for common treatment across siblings. The inclusion of census sheet fixed effects (γ_s) is key to our identification strategy, allowing us to compare households in the same narrow geographic location that have different next-door neighbor occupation exposure. We discuss our identification assumptions further in the next section.

Generalizing this doctor regression to any arbitrary occupation, we can replace the doctor next door and own household indicators with an indicator for any particular occupation living next door or in the same household in 1910. We then replace the outcome with an indicator for the focal child reporting that occupation in 1940. For example, we can test whether a girl living next door to a teacher in 1910 becomes a teacher by 1940.

Lastly, we introduce a stacked regression approach to estimate average transmission effects within a group of occupations. This allows us to systematically explore patterns of heterogeneity and causal mechanisms later in the paper. To do this, we take each occupation-specific regression sample described above and stack them, so that all 50 occupations for boys appear in one regression sample and all 25 occupations for girls appear in another regression sample. We can also restrict to specific subsets of occupations, such as all high-income occupations, to explore dimensions of heterogeneity. The unit of observation in the stacked regressions is a child-occupation pair.

We estimate

$$\begin{aligned}\text{Occ1940}_{ijs} =& \alpha + \beta_0 \text{OwnOcc1910}_{ij} + \beta_1 \text{NextDoorOcc1910}_{ij} + \text{Occ}_j \times \text{Age}_i \\ & + \text{Occ}_j \times \gamma_s + \epsilon_{ijs},\end{aligned} \tag{2}$$

where j indexes each occupation. In this specification, each child i is included in the regression multiple times, one for each occupation that occurs on the child's street.[12] Since the same household is used multiple times, in this specification we two-way cluster standard errors at the household level and at the individual i level. Our fixed effects for age and census sheet are interacted with occupation to make the same within-occupation comparisons that we made in the single occupation specification.

3.1 Identification

In order to identify a causal occupation transmission effect, we assume that while households might endogenously select into a particular neighborhood, the assignment of their specific next-door neighbors is as-good-as-randomly assigned.

Given the unpredictability of housing markets and imperfect information, it is difficult for individuals to observe the occupations of potential neighbors before moving into a specific residence. Moreover, when deciding where to live, households can only move into available housing, limiting their ability to choose immediate neighbors' occupations. In other words, conditional on living in the same portion of a street, a household's proximity to a neighbor in a particular occupation is as good as random.

[12] Because we include sheet fixed effects, the estimated effects are identified off of sheets where there is variation in OwnOcc1910 and NextDoorOcc1910. As such, we can eliminate child observations on sheets without people in the target occupation without affecting the identifying variation. This reduces the overall size of the stacked sample from over 316 million possible individual-occupation pairs for the 50 male occupations, to 26.6 million pairs.

To see this idea more formally, consider a simple linear model of children occupational choice:

$$\text{Occ1940}_{is} = \alpha \text{Amenities}_s + \beta \text{NextDoorOcc1910}_i + \nu_i, \tag{3}$$

where Occ1940_{is} is an indicator equal to one if a child in household i that we observe living on census manuscript sheet s in 1910 has a particular occupation in 1940. Amenities_s capture any amenities or other local characteristics that may change the probability that a child enters a particular occupation as an adult. NextDoorOcc1910_i is equal to one if the adult living next door to household i in 1910 is in that occupation. ν_i capture any individual- or household-level idiosyncratic characteristics that affect the probability that a child enters a particular occupation.

We are interested in estimating the importance of exposure to adult neighbors in a particular occupation, given by β. The challenge is that we lack data on ν_i, and likely also lack data on relevant local amenities, and both are likely to be correlated with neighbors in a particular occupation. For an extreme example, children growing up in coal mining towns are more likely to live next to coal miners than other children observed in the full count census because that is the primary occupation in their town; they are also more likely to become coal miners, but this is likely driven as much by proximity to coal mines as to exposure to adult coal miners living next door. As another example, if wealthy households, where many of the adults hold occupations like doctors or lawyers, sort into the same high prestige neighborhoods and hence are more likely to live next door to one another, then children of high income families are more likely to go into the occupations of adults living in their neighborhoods than the average child, even if true causal exposure effects were zero, since high income parents have the resources and familial social capital to encourage their children to enter high income occupations. This is an example of the well-known reflection problem articulated by Manski (1993).

We argue that the ability to compare individuals over extremely local geographies resolves the reflection problem, since individuals that reside on the same census manuscript sheet (on average comparing across only 7.6 households) are exposed to the same local amenities, within that microgeographic area, cannot select the occupations of their immediate next door neighbors. Therefore, conditional on being on the same manuscript sheet, all households are equally likely to be immediately next door to a neighbor in a given occupation. Our identification assumption is therefore that

$$\text{Cov}(\text{NextDoorOcc1910}_i, \text{Amenities}_s|s) = \text{Cov}(\text{NextDoorOcc1910}_i, \nu_i|s) = 0, \tag{4}$$

where we condition on the census manuscript sheet s. Similar identification has been used in housing transaction data in a modern context (Bayer et al., 2022).

3.2 Plausibility of the Identification Assumption

Our identification strategy would fail if households sort *within census manuscript sheets* in ways that are correlated with both their neighbor's occupation and their child's proclivity to work in that occupation. In the data, this would likely result in individuals in the same occupation living next to each other or individuals in similar occupations (along dimensions such as industry, skill-level, or income) living next to each other.

Sorting specifically by occupation within sheets is empirically rare in our data. We find little evidence that adults having the same occupation are able to sort into adjacent houses. For the 50 largest, non-farm occupations in the 1910 census, among the sheets that contain at least one individual with each occupation, a large fraction of sheets contain only one occurrence of that occupation. With the exception of resource-based occupations that are geographically constrained like mining, most people are the only person in their occupation on their census sheet. In the results below, we show that our results are robust to focusing on only the sheets that have one of each occupation, in which case it is by construction impossible for households with the same occupation to sort to be adjacent to one another.[13] Occupation-specific measures of segregation, analogous to measures of racial segregation in Logan and Parman (2017a), show similar results. Across occupations, occupation-specific segregation is close to zero (consistent with as good as random placement) and substantially lower than racial segregation or even ethnic segregation for nearly all occupations.[14] See Appendix B for a detailed analysis of neighborhood sorting.

To further bolster the identification assumptions, we use a natural experiment that quasi-randomly shift exposure to particular occupations through neighbors' untimely deaths. If our estimated effects were driven by match-specific sorting we would expect similar transmission, even if the next door neighbor to the adult in the focal occupation died before the child could be "exposed" to the career. As we show in the results section, the effects for these children are near zero and statistically insignificant.

A second approach is to condition directly on features that may predict within-sheet sorting to rule out that effects are driven by households with similar occupations (and potentially similar aspirations for their children's future careers) sorting next door to each other. Results are similar when we control for a suite of household-level characteristics or include fixed effects for the occupation of the household head in each child's household. This specification allows us to, for example, account for the elevated rate at which sons of lawyers become doctors and determine if proximity to a doctor has an additional effect on the child becoming a doctor relative to the baseline rate at which children of lawyers become doctors. To minimize concerns about within-sheet variation in access to amenities, we can include fixed effects for the position of each child's household on the census sheet and compare outcomes for even smaller subsets of census sheets. We discuss these results in more detail in the robustness section below.

[13]Consistent with this evidence, if we regress an indicator that equals one if the household head is in a particular occupation on an indicator that equals one if someone in the neighboring household is in that particular occupation with sheet fixed effects, the estimated coefficient is large, significant, and *negative*. This is because children from the neighboring household are now in the counterfactual comparison group and people in the same occupation are unlikely to sort next door to each other.

[14]The exception is once again for resources based occupations like farming, mining, and lumbermen.

We conduct our analysis for a broad range of occupations that vary across skill-level, income-level, and prestige, and find similar patterns of transmission. Once we move from particular occupations (such as doctor) to the full set of large occupations, the kind of sorting necessary to invalidate our identifying assumption becomes more nuanced. Remaining threats to identification must be match-specific sorting where families sort into homes next to people in particular occupations because their child has an idiosyncratic proclivity for that particular occupation in a way that is unique and not driven by general characteristics like occupational income or prestige.

4 Results: Occupation Transmission

We present our main results in three steps: First we focus on transmission of a single occupation: doctors. Second, we estimate transmission for the largest occupations for girls and boys individually. Third, we estimate the average transmission effects across all occupations in our stacked OLS regression framework. Following the main results, we analyze the unexpected deaths natural experiment, and present further robustness tests.

4.1 Doctors

Table 2 shows results for boys in the doctor regression described in equation (1). To help assess our identification assumptions, we present different levels of fixed effects in each column. Column 1 does not include any location fixed effects and includes all households with a child in the 1910 census. Column 2 also does not include any location fixed effects, but to include a consistent sample in which there is identifying variation in more restrictive specifications, we restrict attention to census sheets with a child in the 1910 census and for which there is at least one household with a doctor. Column 3 adds broad geographic fixed effects for state-city-enumeration district, and Column 4 adds our preferred, more narrow sheet-level fixed effects. Across all four specifications, we see a very consistent positive effect of own household transmission. Boys that grow up with a doctor in the home are 9.54-9.83

percentage points more likely to become a doctor than other boys on the street. We report the mean probability of becoming a doctor for the counterfactual group – boys on the same sheets, with no doctor in the home or next door. Those with doctors in the home are at least 10 times more likely to become a doctor than their peers on the same sheet, a result that matches other estimates of the intra-household transmission rate in other contexts.

Our key coefficient of interest is the effect of living next door to a doctor in 1910. For this coefficient, the level of fixed effects is relevant for our identification assumption. The estimate with no fixed effects includes not only the causal effect of exposure, but also a selection effect where kids that live in the same neighborhood as other doctors may have been more likely to become doctors even absent exposure. We see that the coefficient in column 1 of 0.0072 falls to 0.0036 in Column 2 when considering only sheets with doctors, and falls slightly farther to 0.0033 in Column 3 with enumeration district fixed effects. In Column 4, with sheet-level fixed effects, the estimate is 0.0032. This estimate removes the street and neighborhood-level sorting, and we interpret the remaining coefficient as the causal exposure effect. Boys that live next door to a doctor are 0.32 percentage points more likely to become a doctor, or 41% more likely than the boys who live further down the street.[15]

4.2 All Occupations Separately

Now that we have established a significant next door neighbor transmission effect in the case of doctors, we can consider a broader set of occupations, including occupations available to women in the time period.

For each occupation we replace the indicators Doctor1940_{is}, OwnDoc1910_{i}, and $\text{NextDoorDoc1910}_{is}$ with analogous indicators for that particular occupation. To simplify the analysis and to ensure that we have a sufficient number of individuals in each occupation, we focus on the 50 largest non-farmer occupations in terms of male employment in 1910, and the 25 largest

[15]In Appendix Table C2, we repeat this exercise but link children to the 1920, 1930, and 1940 censuses. For both own-household and next door coefficients, effect sizes after scaling by the untreated mean are similar across all three decades, suggesting that we do not lose out on addition information by focusing on 1940 outcomes. We find similar results for the full stack for girls and boys in Appendix Table C3.

non-farm occupations for women.

These coefficients are provided for boys and girls for each occupation separately in Figure 1. Both within household and next door neighbor effects are presented. In both cases, we scale the coefficients by the mean among kids with no individuals in the occupation in their household or next door so that the plotted coefficients can be interpreted as percent changes relative to the untreated mean.[16] Both are separately sorted by effect size and solid filled markers indicate that the coefficient was statistically significant at the 5% level. We find that, for all 50 of the largest occupations for men, having an individual in the same household significantly predicts that the child is more likely to enter that occupation. There is, however, substantial heterogeneity in effect size across occupations. For own-household transmission for boys, doctors are indeed ranked highly, with the fifth largest estimate in percentage terms. But other occupations of various type, such as brickmasons, bakers, meat cutters, blacksmiths, clergymen, barbers, and lawyers also predict that kids are more than ten times more likely to go into the occupation if an adult in their own household is in that occupation. Own household transmission for girls is positive and statistically significant for most occupations, but the effect sizes are much lower than for boys. Girls are likely to choose the occupation of a parent one to four times as often as an untreated peer. The muted transmission effect for girls may be due to lower overall employment among women (in both 1910 and 1940) or expanding occupational opportunities that were available to women in 1940 but not 1910.

Focusing on the next door coefficients, we see that most occupations show positive transmission effects from next door neighbors. In contrast to the own-household coefficients, however, not all are statistically different from zero and three (stationary firemen, doorkeepers, and shipping and receiving clerks) are negative, although not statistically significantly so. The most predictive occupation is for brickmason, the same as for the own-household coefficients, although the next several most predictive occupations differ between the two

[16] We plot the un-transformed coefficients in Appendix Figure C1.

lists. Doctors once again rank highly, in eighth place. Girls show more mixed evidence on neighbor transmission. While the majority of the top-25 occupations have positive transmission, only six are statistically significant. So although the effects for girls are suggestive, we lack the same level of statistical precision for women.

4.3 All Occupations Combined

In Table 3, we present results from the stacked regressions. Panel A shows average effects across the 50 largest occupations for boys, and Panel B shows average effects across the 25 largest occupations for girls. Columns use the same location fixed effects strategy as in Table 2. In Columns 1 and 2, we have no location fixed effects (although we still include a fixed effect for each occupation j). Recall that Column 1 includes all census manuscript sheets, while Column 2 includes only sheets that have at least one individual of each occupation j.

In contrast to the results when examining only doctors, in the stacked regression coefficients for both own-household effects and next door neighbor effects become substantially smaller as we include geographically smaller location fixed effects. This is consistent with substantial locational sorting for the average occupation.

Our preferred specification in Column 4 includes a fixed effect for each census manuscript sheet interacted with a fixed effect for each occupation. In this specification, growing up with an adult in the same household in an average occupation increases the probability that the boy enters that occupation by 3.81 percentage points, or about 115%. Growing up next door to an individual with an average occupation increases the probability that the boy enters that occupation by 0.34 percentage points, or about 10.3%. For both the own-household and next door coefficients, the scaled estimates in Table 3 are smaller than those for doctors in Table 2, which is consistent with our findings in Figure 1 that exposure effects for doctors are larger than for an average occupation. For girls, shown in Panel B, the effects are also positive and significant, but smaller in percentage point change. Girls are 0.85 percentage points, or about 54%, more likely to choose their own parent's occupation, and 0.1 percentage points, or about 6.4%, more likely to choose their neighbor's occupation. The relative increase is about half that for boys off of a much lower base. This smaller effect size likely reflects the changing nature of female employment during this time period, and the lower overall propensity to participate in the labor force at all compared to men.[17]

4.4 Untimely Neighbor Deaths

To interpret these transmission effects causally, we assume that the placement of occupations relative to children is as-good-as-random. However, it is possible that households have some ability to endogenously sort to live next door to particular neighbors along dimensions of unobserved household characteristics in a way that biases our estimates. In particular, if families can sort toward occupations that their children are more likely to choose (e.g. more educated families living together on the part of the street with the best houses), this could bias our estimates upward. In this section, we use the untimely deaths of neighbors as a natural experiment that quasi-randomly shifts occupation exposure, reducing the potential bias from sorting.

To describe the analysis, we again use doctors as an example. Using the Database of Human Lifespan (Price, 2024), we tag doctors with documented deaths between 1910 and 1915, the five-year window following the 1910 census in which we measure occupation exposure. We limit the sample to sheets with doctors who were under the age of 60 in 1910, to capture unexpected deaths. We then compare outcomes for boys that live next to a doctor in 1910 who dies before 1915 (6% of the sample of boys) to boys that live next to a doctor who survived past 1915.[18] This quasi-random shift in the duration of exposure to a doctor neighbor can help solidify the causal link between exposure and outcomes. In both cases, the boy's family similarly selected into homes next to doctors, but we can hold this selection fixed and determine if boys with more limited exposure experience similar treatment effects.

[17]In Appendix Table C4 we show that results are robust to logit specifications, which account for concerns about sparse outcomes when using linear probability models.

[18]Nearly 31% of the sample live next to a doctor with a missing death date, and are therefore excluded from the analysis.

Table 4 shows regression results for the specification in Equation 1, but we separate out the regressors to indicate if child lived in the same household or next door to a doctor that died by 1915 or one that remained alive. We find that doctor death sharply decreases the probability that a child becomes a doctor both for the doctor's own children as well as the next door neighbor. Children of surviving doctors have a 10.8 percentage point higher probability of becoming a doctor compared to the untreated mean compared to only 4.8 percentage points for children of doctors that died prematurely. The effect of next door neighbors similarly falls from 0.3 percentage points down to -0.002 percentage points, a coefficient that is not statistically distinguishable from zero if the neighboring doctor died prematurely. We further restrict to doctors that died before the age of 50 and 40 to better capture unexpected death and find similar results. We find similar neighbor-to-neighbor transmission when the doctor survives, but countering the selection story, we do not find transmission effects if the doctor dies prematurely.

Since there are few doctors that died prematurely in our data, the zero effects are estimated with some noise. We therefore perform this analysis for all 50 occupations in the stack regression framework for boys and report results in Columns 5-8 of Table 4. We likewise find that nextdoor neighbor effects are smaller and statistically insignificant when the next door neighbor dies, with further coefficient decreases with younger deaths. Interestingly, we also see falling own household transmission for adults in the household that die younger, suggesting that duration of exposure matters both within and across households. Together, these findings support the causal nature of the influence of exposure on occupation choice and helps to rule out the possibility that effects are driven by selection or sorting within a sheet.

4.5 Robustness

We provide two further sets of robustness checks for our main results, which we briefly summarize here and provide further analysis in the Appendix. First, we probe our main results using more restrictive regression specifications and alternative sample definitions. Second, we consider an alternative method for constructing neighbor proximity using street addresses.

To further assess robustness to endogenous sorting, in Table 5 we present results using additional sets of household controls and fixed effects in the doctor regression. First, we control directly for household head's age, education, race, and nativity. Then we include fixed effects for the household head's occupation, allowing for the comparison of children on the same street with same parental occupations but different neighbors. Finally, we restrict the sample and the scope of the fixed effects to only include children that live at most three houses from a doctor. In all cases, we find similar results: having a doctor in one's own household increases the likelihood that a boy becomes a doctor in 1940 by about ten percentage points, and living next door to a doctor also significantly increases the probability that a boy becomes a doctor in 1940 by between .25 and .4 percentage points.

In Appendix Tables C5 and C6, we implement additional strategies to minimize concerns that results are driven by unobservable within-sheet sorting. First, we restrict our sample to census sheets on which only a single doctor resides, which both simplifies the source of treatment variation and ensures that doctors are not sorting to live next to one another. Second, we drop any individuals living in group quarters (dormitories, barracks, prisons, etc.). Another concern is that access to unobservable amenities may vary even within census sheets. For instance, houses close to the end of a block may be larger or have easier access to resources on a cross street, and so doctors and those with a predilection to become doctors may be more likely to sort to live close to the ends of a block. To minimize these concerns, we include a fixed effect for each household's position on the census sheet (e.g., a control for households at the top of the page, one house down from the top of the page, etc.).

As described in Section 2, while the census sheet order usually captures the microgeography of neighborhoods accurately, in some cases it may be measured with error. In Appendix D, we consider an alternative measure of next door neighbors using street addresses recorded in the census.[19] We consider a household to be next door to a doctor if it has the closest house number (either larger or smaller) on the same street as the doctor. In Appendix Tables D1 and D2 we use this street address-based measure for next door neighbors and find results that are similar to our baseline estimates. We further use the street address-based measure of proximity to correct for non-classical measurement error in the census sheet-order analysis as suggested by Pischke (2007) and find that, if anything, our next door neighbor transmission results are slightly attenuated.

5 Why Neighbors Matter: Social Ties and Homophily

Given the strong patterns of occupation transmission that we have documented, we next explore the mechanisms that drive the effects. We do this in three parts. First, we consider the intensity of treatment by looking at how transmission decays with distance on the street and duration of exposure. Second, we consider heterogeneity by neighborhood type. Third, we look at the nature of the neighbors themselves, testing whether homophily in individual characteristics increases transmission.

5.1 Intensity of Treatment

Occupation transmission appears to occur through very localized networks. One reason that children next door to a doctor might steer toward that occupation is that they have frequent interactions with their doctor neighbor, increasing the amount of information, mentoring, apprenticeships, or other tangible or human capital that can be passed on. Transmission likely increases by the proximity to the doctor on the street as well as the duration of exposure.

To test how proximity affects transmission, we extend our baseline stack regression to include indicators for additional households further down the block from a particular occu-

[19] Akbar et al. (2019) and Quincy (2022) also use street addresses in historical censuses to determine the proximity of houses.

pation j:

$$\text{Occ1940}_{ijs} = \alpha + \beta_0 \text{OwnOcc1910}_{ij} + \sum_{d=1}^{5} \beta_d \text{NeighborOcc1910}_{ijd} + \text{Occ}_j \times \text{Age}_i + \text{Occ}_j \times \gamma_s + \epsilon_{ijs}, \tag{5}$$

where the indicators $\text{NeighborOcc1910}_{ijd}$ are equal to one if at least one of the households d steps away from the focal child in the 1910 census sheet has an adult that lists occupation j. For example, if $\text{NeighborOcc1910}_{ij3} = 1$ in the doctor section of the stack, there is a doctor exactly three doors away from the focal child. Children who appear on the same census sheet but are more than five houses away are the omitted group.

Figure 2 plots each β_d (the coefficient for $\text{NeighborOcc1910}_{ijd}$). We see a monotonic and steep slope of the coefficients as we move from immediate next door neighbor towards five houses down the street. This suggests that the strength of the local ties on the street diminish with each subsequent house, lessening the likelihood of occupation transmission.[20] In our baseline results, children 2-5 households away from the adult neighbor are part of the set of counterfactual children; while these children have smaller exposure effects than the children living next door, the effects are still positive, and so our baseline results should be seen as a lower bound on the magnitude of exposure effects.

5.2 Neighborhood Characteristics

The interconnectedness of neighborhoods might affect the magnitude of exposure effects. We examine three measures that potentially capture dimensions of neighborhood interaction and

[20] To estimate Equation 5, we require census sheets with at least six households (the household with the doctor and five neighbors). Sheets with several larger households may have fewer than six households on the sheet, and so looking at the effects on children farther away from a doctor mechanically reduces the sample size. Also, if we examine effects up to five doors down, households at the top and bottom of the census sheet are treated differently than households in the middle of the census sheet, since a household at the edge of a sheet only has neighbors in one direction. This is another reason to focus our main analysis only on next door neighbors.

connectedness: whether the census sheet is in an urban or rural county, the immigrant share, and the out-of-state share. Neighborhoods that are less transient and have more permanent residents are likely to have stronger social ties. For this reason, we explore heterogeneity by the share of the population that was not born in that state, or the out-of-state share. If network strength matters, we expect to see neighbor exposure effects to be stronger in counties with fewer out-of-state residents. For the urban or rural counties, the theoretical implications are more ambiguous. In urban areas, residences are likely to be geographically close together, and so individuals may have more frequent interactions with their neighbors. At the same time, because of this proximity, children may be able to interact with more distant neighbors more easily in urban areas. It is also possible that urban life is more anonymous, and so interactions are more formative in rural areas (Dunkelman, 2017; Wirth, 1938). This is similar for immigrant shares. Households may interact more in immigrant enclaves (Damm, 2009), but immigrant communities could also provide alternative networks, making next door neighbors relatively less important for the transmission of information about occupations.[21]

We present the scaled next door neighbor coefficients split by county-level urban or rural status, immigrant share, and the out-of-state share in Figure 3. We find that exposure effects are significantly larger in neighborhoods that are in rural counties, have lower immigrant shares, and that have more residents who were born in the state (low out-of-state share). These patterns are consistent with more homogenous places with stronger connectedness facilitating occupational transmission.[22]

[21] Living in an area with a large share of immigrants might be particularly important if the focal individual themselves comes from an immigrant family or shares an ethnicity with their neighbors. We explore the importance of the similarity between two neighbors in the next section.

[22] In Appendix Table C7, we explore heterogeneity among other neighborhood characteristics, including the share of non-Whites, whether or not there is an institution of higher education in the same county, and region of the U.S. In all specifications, we find positive average transmission effects, suggesting that most types of neighborhoods in this time period were facilitating occupational spillovers.

5.3 Individual Characteristics

Given the heterogeneity across neighborhoods, it is plausible that there is heterogeneity across characteristics of the neighbors themselves. In particular, we test whether homophily among neighbors makes exposure effects stronger. To do this, we modify equation (2) but allow the effects to vary if the focal child and neighbor have the same characteristics, as follows

$$\begin{aligned}\text{Occ1940}_{ijs} =& \alpha + \beta_0 \text{OwnOcc1910}_{ij} + \beta_1 \text{NextDoorOcc1910}_{ij} \times \text{SameCharacteristic}_{ij} \\ &+ \beta_2 \text{NextDoorOcc1910}_{ij} \times \text{DifferentCharacteristic}_{ij} + \text{Occ}_j \times \text{Age}_i \\ &+ \text{Occ}_j \times \gamma_s + \epsilon_{ijs}, \end{aligned} \tag{6}$$

where $\text{SameCharacteristic}_{ij}$ is set equal to one if i has a next door neighbor in occupation j who is the same as i along various demographic and economic characteristics. Similarly, $\text{DifferentCharacteristic}_{ij}$ equals one if i has a next door neighbor in occupation j who is different along a characteristic. For example, when considering racial homophily, a Black boy living next to a White doctor would have $\text{SameCharacteristic}_{ij} = 0$ and $\text{DifferentCharacteristic}_{ij} = 1$.

In Table 6 Column 2, we estimate exposure effects for next door neighbors who are and are not from the same birthplace, using the state of birth variable in the census. Boys are more likely to enter the occupation of their next door neighbor when the neighbor is born in the same state, 0.47 percentage points more likely for neighbors from the same state versus 0.23 percentage points more likely for neighbors from different states. We present F-test statistics showing that these two coefficients are strongly significantly different from one another. We see similar patterns in Column 3 when we examine differences by birth country (with all native-born grouped together).

In Column 4, we examine whether the race of the next door neighbor affects the magnitude of the exposure effect. Children with same-race next door neighbors are 0.38 percentage points more likely to enter into the occupation of that neighbor. Children are 0.83 percentage points *less* likely to go into the occupation of their different-race neighbors. Importantly, this negative effect must be interpreted relative to the appropriate counterfactual group. These coefficients compare the occupational choice of a different-race child relative to the choices of children who live farther away, but still on the same street. Racial segregation at the time is high, so these comparison children are more likely to share the same race as the person in the target occupation. For example, a non-White child living next to a White doctor is less likely to become a doctor than other (mostly White) children who grew up on the same street.[23] In Appendix Table C8 we also show differences by household income, household education, and last name homophily.[24]

In Column 5 we test whether the presence of a child with the same age in the next door household affects the magnitude of the exposure effect. If children are more likely to interact with other children of the same age on their streets, then they are also more likely to be exposed to the occupations of the parents of these similar-age children. Recent studies indicate that the parents of children's peers are important in determining educational outcomes in modern settings (Chung, 2020; Fruehwirth and Gagete-Miranda, 2019). We do find that next door neighbors are more predictive of children's future occupations when they have a child of the same age in the household relative to next door neighbors without a same age child, although the differences in magnitude are modest (0.35 percentage points versus 0.31 percentage points) and not statistically significant.

We also explore dimensions of gender homophily. To this point we have estimated the effect of a neighbor's occupation regardless of their gender. We now re-estimate a version of equation (2), but in place of the next door indicator include two mutually exclusive

[23]In unreported analysis, we restrict to sub-samples of children of the same race, and find that non-White children have statistically zero next-door effects when living next to a White adult of the target occupation. Because of strong racial segregation in 1910, there are very few cases of mixed race neighborhoods, making it difficult to estimate cross-race effects with much statistical power.

[24]One concern is that extended families co-locate, which would potentially bias our next door neighbor estimates. However, we find that neighbor transmission effects are nearly identical even when the neighbors do not have the same last name.

indicators that equal one if the neighbor is in the target occupation and either male or female. We provide these results for our sample of boys and girls in Table 7. For both boys and girls we find significant exposure effects from both male and female neighbors. However, the exposure effects for boys are three times as large if the neighbor is male relative to female, while the exposure effects for girls are twice as large if the neighbor is female, relative to male. Consistent with evidence of gender homophily between teachers and female students affecting school performance and major choice (?), we find that exposure effects are larger when genders match. Taken together, the heterogeneity by neighborhood and neighbor characteristics are consistent with exposure effects being larger when social ties are potentially stronger.

6 Exposure Effects on Economic Outcomes

Growing up next to someone in a particular occupation changes a child's likelihood of working in that occupation, but does this change the child's long-run outcomes? Some high-income occupations, such as doctors and lawyers, have higher exposure effects (see Figure 1) but several low-income occupations, such as brickmasons, tailors, and waiters, are also highly transmissible. It is possible that the type of neighbor a child grows up next to could affect their earnings and educational attainment as an adult. Alternatively, growing up next door to a neighbor in a particular occupation may have little effect on a child's economic outcomes if, absent exposure to the neighbor's occupation, the child would have gone into a different occupation that had similar economic characteristics. For instance, growing up next door to a doctor may induce children who would otherwise become lawyers to become doctors instead, leading to only modest changes to the child's income and educational attainment. We consider income and education effects for both boys and girls.

6.1 What Occupation Characteristics Matter?

We first show that occupation transmission varies by the average level of income and education in that occupation. We re-estimate the stacked regression from Equation 2 but restrict the sample to only include stacked panels for occupations that have particular characteristics. Because of the more limited employment opportunities and lower labor force participation for women in the time period, we focus this section on our sample of boys, but provide analogous estimates for girls in Appendix Table C9. We focus on variation in occupational income levels and education requirements, as captured by occupation income scores and education scores. We plot the next door neighbor effects, scaled by the untreated mean, in Figure 4.[25] Exposure effects for high income occupations are slightly larger on average, around 12 percent, while effects for high education occupations are slightly lower, at 9 percent. However, there are stark difference across the four mutually exclusive groups when interacting income and education level. Exposure effects are larger for high income occupations that have either high and low education requirements, although the point estimate is largest for high income occupations with low educational attainment. Effects are also larger for low-income, low education occupations. However the exposure effects for low income jobs that require high educational attainment (like teachers or clergymen) are significantly lower, at less than 5 percent.[26]

6.2 Effects on Adult Income

Given the varying size of the transmission effect by occupation, we now turn our attention to the overall effects of occupation exposure on future income. Using the same approach as in Equation 1 but for each of the top 50 male occupations, we estimate the effect of

[25] Unscaled estimates, along with estimates for having someone of a particular occupation in the child's own household are provided in Appendix Table C10.

[26] We explore other occupational characteristics, such as heterogeneity by occupational self-employment rates, occupations with apprenticeship structures, or whether occupations are growing or in decline but find few meaningful differences. Our main neighbor spillover finding is likely not driven by direct transmission of businesses, neighbor-provided apprentices, or structural labor market changes (see Table C11.)

living next door to someone of a particular occupation in 1910 on wage income in 1940.[27] As seen in Figure 5, many neighbors' occupations have a significant impact on boy's adult income and there is substantial heterogeneity. Living next to a lawyer or doctor during childhood is associated with a $45-50 (1940$) increase in annual earnings in 1940, relative to other children on the census sheet. Relative to average annual income of $1,071 in 1940 among similarly-aged (30-50) working men, this represents a 4.2-4.7 percent increase in annual income. Importantly, this is relative to other children who were living in the same neighborhood in 1910, and thus accounts for socioeconomic neighborhood sorting.

There are other high-income occupations such as managers and locomotive engineers that also lead to increases in the neighbor child's adult income, but there are also low- or middle-income occupations such as teachers, clergymen, and clerical workers that lead to significant increases in neighbor children's adult income. Once again, since these estimates are relative to other children living the same neighborhood, this likely speaks to the occupation choice counterfactual. Teachers in general did not live in the most affluent neighborhoods, but living next door to a teacher exposed a child to opportunities that were relatively better than their peers.

There are also occupations that significantly *reduced* next door children's wage income in 1940 relative to other neighbor children. Living next to a porter, a truck driver, or a laborer—all of which are low-income occupations—led to annual income reductions at least half as large as the gain from living next to a lawyer or doctor. Why would exposure to neighbors in some occupations lead to downward economic mobility relative to other kids growing up on the same sheet? We conjecture that this once again reflects the counterfactual. Kids that grow up next to adults in lower income occupations than the average on their sheet miss out on the exposure to neighbors in an average occupation. Hence, incomes fall not because low income neighbors directly encourage kids to pursue low income occupations, but

[27] The 1940 census only collected information on wage income. Many workers in high skill occupations such as doctors, lawyers, managers, and real estate agents report no wage income, but report that they had non-wage income (but not the amount). If anything, this under-reporting attenuates the results for high skill occupations.

because they prevent exposure to higher income occupations.

Girls also have a few professions whose influence increases income. Living next to a milliner, bookkeeper, or stenographer during childhood increases annual earnings by $15-20, a 7-9 percent increase relative to average annual wage income ($204) and a 2-3 percent increase relative to average annual wage income among employed women. These results likely differ in magnitude from men because of differences between men and women in the extensive margin decision to participate in the labor force during this time period, which is influenced by neighbors' occupations, as seen in Figure 1b. Similar to boys, living next to laborers or private laundresses actually led to lower annual income relative to other girls on the same sheet.

Although it is clear that living next to someone in a particular occupation as a child has an impact on future earnings, the channels through which this operates is not clear. It could all be driven by an occupation match effect if, for example, growing up next to a doctor increased a child's likelihood of becoming a doctor, but had no other effects. However, it is also possible that having a doctor for a neighbor increases exposure to information about high-paying jobs in general, or the human capital requirements necessary to qualify for a high-paying job. As noted above, there are some low income, high education occupations, such as teacher or clergyman, that are not very transmittable to neighbor children (see Figure 1) but do lead to income gains (see Figure 5). For this reason, we next explore how childhood neighbors' occupations affect children's educational attainment.

6.3 Effects on Educational Attainment

In Figure 6, we find that living next door to someone of a particular occupation has a similarly heterogeneous impact on boy's educational attainment. Living next to a doctor or lawyer in 1910 increased average years of schooling by nearly 0.4 relative to other children on the census sheet. But even many of the low income, high education occupations that led to income gains lead to increases in educational attainment as well. The occupations that have the largest effects on neighbor children's educational attainment are occupations that require formal education, while living next to a neighbor in some trade occupations, such as lumberman, miner, teamster, or laborer, actually reduced children's educational attainment. Figure 6b shows similarly large effects for many professional occupations with educational requirements for girls. Girls living next to nurses, bookkeepers, and teachers attain about 0.2 more years of education than the other girls on the street, but many other occupations also lead to educational gains for girls, which could have persistent, intergenerational effects. Living next to some occupations decreases educational attainment, including laundresses, laborers, and kindred workers. These effects speak to the network and mentoring channel of occupation exposure.

Since neighborhood exposure to many occupations changes a child's ultimate educational attainment, it seems plausible that growing up next door to a doctor changes more than just the child's probability of becoming a doctor. We next estimate the occupation-by-occupation spillover between each of the 50 largest occupations. To do this, we estimate equation (1) but iterate through each of the top 50 occupations as both the outcome and the independent variables. The scaled effect of growing up next to a neighbor in any given occupation on the probability of being in each of the top 50 occupations is presented in Figure 7. The occupations are sorted by education score, with the most educated occupations at the left and bottom.

Several clear patterns arise. First, the largest, positive effects are concentrated along the diagonal, suggesting high levels of occupation-specific transmission. Second, it is clear that there are positive spillovers to occupations with similar education requirements and negative spillovers to occupations where the educational requirements differ substantially. The positive, significant spillovers are concentrated in the bottom left (high education occupations) and top right (low education occupations) quadrants. In the off diagonal quadrants the spillover are mostly negative. For example, living next door to a doctor as a child significantly increased the probability that the child entered a high-paying, high-skilled occupation that requires educational training, such as lawyer, teacher, insurance agent, and manager, but reduces the likelihood that the boy becomes a laborer, kindred worker, miner, or porter (low education occupations). Growing up next to a doctor also does not increase the likelihood of being in a high-paying, low-education manual or trade occupations such as foreman or compositor/typesetter. The patterns by educational requirements appear to be an important dimension. The benefit of growing up next to someone in a high-income occupation, like doctor, seems both general, directing children to occupations that require more education and training, but also specific, having the largest impact on children's decisions to become doctors.

The implications are similar when looking at teachers, a low income, high education occupation. Growing up next to a teacher increases the probability of being in a high-paying, high-education occupation such as doctor, lawyer, and insurance agent by almost as much as the effect on being a teacher. This general effect can help explain why low income, high education occupations such as teacher had unexpectedly large effects on future income.

7 Conclusion

Where an individual spends their childhood has important implications for their future economic success (Chetty et al., 2018). Using neighborhood microgeography reconstructed from historic, door-to-door census enumeration, we show that part of this can be explained by the composition of neighbors a child grew up next to. Among boys in 1910, living next door to someone in a particular occupation, increased the likelihood that they worked in that occupation in 1940 by 10 percent, relative to other boys who were living on the same street. Girls similarly had positive transmission of occupations, albeit concentrated among different occupations and at slightly lower rates.

This neighbor-to-neighbor transmission of occupation varies across occupation, with larger exposure effect from neighbors in high income, high education occupations. The occupations of childhood neighbors also have long-term economic impacts on children. If a person's childhood neighbors were in particular occupations (such as doctor, lawyer, teacher, or clerical worker), the child experienced higher income and educational attainment as an adult, relative to other children that grew up on the same street. Part of this comes through the direct effect of occupation matching, but growing-up next door to neighbors in more educated occupations also increases educational attainment and the likelihood of working in a high-education occupation. At the same time, some occupations such as truck driver and laborer actually led to reductions in income and educational attainment for neighboring children. Positive transmission effects on education are striking for girls, who in this time period were beginning to enter professional occupations at higher rates and perhaps benefited uniquely by networking with older professional women in their neighborhood.

Overall, our results suggest that childhood neighbors matter, and who you grew up with can help explain some of the effect of place on children's long-run economic mobility. We find that neighbor exposure effects are larger in neighborhoods that are plausibly more connected and among neighbors who share common characteristics such as similar-aged children, place of origin, and race. These patterns emphasize the role of social ties and connectedness and are consistent with both information and exposure channels. Interacting with someone in a high-income or high-education occupation can make the returns of that occupation salient or remove information barriers that keep people from being eligible to work in those occupations. However, these effects could also be simply driven by exposure. Children might not know that a particular occupation exists in their choice set unless they know someone in that occupation. Although social connectedness between physical neighbors might be weaker now than in the past, these patterns suggest that the composition of socially connected adults in the neighborhood and community influence children's long-term outcomes.

References

Abramitzky, Ran, Leah Boustan, Katherine Eriksson, Myera Rashid, and Santiago Pérez, "Census Linking Project: 1900-1910 Crosswalk," April 2022.

Agresti, Barbara F., "Measuring Residential Segregation: In Nineteenth-Century American Cities," *Sociological Methods & Research*, May 1980, *8* (4), 389–399. Publisher: SAGE Publications Inc.

Aigner, Dennis J., "Regression with a binary independent variable subject to errors of observation," *Journal of Econometrics*, March 1973, *1* (1), 49–59.

Akbar, Prottoy A., Sijie Li, Allison Shertzer, and Randall P. Walsh, "Racial Segregation in Housing Markets and the Erosion of Black Wealth," May 2019.

Aliprantis, Dionissi, "Human capital in the inner city," *Empirical Economics*, November 2017, *53* (3), 1125–1169.

Andrews, Michael J., "Bar talk: Informal social interactions, alcohol prohibition, and invention," *Alcohol Prohibition, and Invention (November 18, 2019)*, 2019.

— **and Chelsea Lensing**, "Cup of Joe and Knowledge Flow: Coffee Shops and Invention," 2024.

— **and Yiling Zhao**, "Home economics and women's gateway to science," 2025.

Andrews, Rodney, Marcus Casey, Bradley L. Hardy, and Trevon D. Logan, "Location matters: Historical racial segregation and intergenerational mobility," *Economics Letters*, September 2017, *158*, 67–72.

Arzaghi, Mohammad and J. Vernon Henderson, "Networking off Madison Avenue," *The Review of Economic Studies*, 2008, *75* (4), 1011–1038. Publisher: [Oxford University Press, Review of Economic Studies, Ltd.].

Atkin, David, M. Keith Chen, and Anton Popov, "The Returns to Face-to-Face Interactions: Knowledge Spillovers in Silicon Valley," June 2022.

Avdeev, Stanislav, Nadine Ketel, Hessel Oosterbeek, and Bas van der Klaauw, "Spillovers in Fields of Study: Siblings, Cousins, and Neighbors," 2023.

Barrios-Fernández, Andrés, "Neighbors' Effects on University Enrollment," *American Economic Journal: Applied Economics*, July 2022, *14* (3), 30–60.

Bayer, Patrick, Marcus D Casey, W. Ben McCartney, John Orellana-Li, and Calvin S Zhang, "Distinguishing Causes of Neighborhood Racial Change: A Nearest Neighbor Design," Working Paper 30487, National Bureau of Economic Research September 2022.

_, Stephen L. Ross, and Giorgio Topa, "Place of Work and Place of Residence: Informal Hiring Networks and Labor Market Outcomes," *Journal of Political Economy*, 2008, *116* (6), 1150–1196. Publisher: The University of Chicago Press.

Bell, Alex, Raj Chetty, Xavier Jaravel, Neviana Petkova, and John Van Reenen, "Who becomes an inventor in America? The importance of exposure to innovation," *The Quarterly Journal of Economics*, 2019, *134* (2), 647–713. Publisher: Oxford University Press.

Bergman, Peter, Raj Chetty, Stefanie DeLuca, Nathaniel Hendren, Lawrence F. Katz, and Christopher Palmer, "Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice," *American Economic Review*, May 2024, *114* (5), 1281–1337.

Billings, Stephen B., David J. Deming, and Jonah Rockoff, "School Segregation, Educational Attainment, and Crime: Evidence from the End of Busing in Charlotte-Mecklenburg *," *The Quarterly Journal of Economics*, February 2014, *129* (1), 435–476.

_, Mark Hoekstra, and Gabriel Pons Rotger, "The scale and nature of neighborhood effects on children," *Journal of Public Economics*, December 2024, *240*, 105260.

Breda, Thomas, Julien Grenet, Marion Monnet, and Clémentine Van Effenterre, "How effective are female role models in steering girls towards STEM? Evidence from French high schools," *The Economic Journal*, 2023, *133* (653), 1773–1809.

Buckles, Kasey, Adrian Haws, Joseph Price, and Haley E.B. Wilbert, "Breakthroughs in Historical Record Linking Using Genealogy Data: The Census Tree Project," September 2023.

Bó, Ernesto Dal, Pedro Dal Bó, and Jason Snyder, "Political Dynasties," *The Review of Economic Studies*, January 2009, *76* (1), 115–142.

Catalini, Christian, "Microgeography and the Direction of Inventive Activity," *Management Science*, September 2018, *64* (9), 4348–4364. Publisher: INFORMS.

Chetty, Raj and Nathaniel Hendren, "The impacts of neighborhoods on intergenerational mobility I: Childhood exposure effects," *The Quarterly Journal of Economics*, 2018, *133* (3), 1107–1162. Publisher: Oxford University Press.

_ **and** _ , "The impacts of neighborhoods on intergenerational mobility II: County-level estimates," *The Quarterly Journal of Economics*, 2018, *133* (3), 1163–1228. Publisher: Oxford University Press.

_ **, John N. Friedman, Nathaniel Hendren, Maggie R. Jones, and Sonya R. Porter**, "The opportunity atlas: Mapping the childhood roots of social mobility," Technical Report, National Bureau of Economic Research 2018.

_ **, Matthew O. Jackson, Theresa Kuchler, Johannes Stroebel, Nathaniel Hendren, Robert B. Fluegge, Sara Gong, Federico Gonzalez, Matthew Jacob, Drew Johnston, Martin Koenen, Eduardo Laguna-Muggenburg, Florian Mudekereza, Tom Rutter, Nicolaj Thor, Wilbur Townsend, Ruby Zhang, Mike Bailey, Pablo Barbera, Monica Bhole, and Nils Wernerfelt**, "Social capital I: measurement and associations with economic mobility," *Nature*, 2022, *608*, 108–121.

_ **, Nathaniel Hendren, and Lawrence F. Katz**, "The effects of exposure to better neighborhoods on children: New evidence from the moving to opportunity experiment," *American Economic Review*, 2016, *106* (4), 855–902. Publisher: American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

_ **,** _ **, Patrick Kline, and Emmanuel Saez**, "Where is the land of Opportunity? The Geography of Intergenerational Mobility in the United States *," *The Quarterly Journal of Economics*, November 2014, *129* (4), 1553–1623.

Chung, Bobby W., "Peers' parents and educational attainment: The exposure effect," *Labour Economics*, June 2020, *64*, 101812.

Chyn, Eric, "Moved to Opportunity: The Long-Run Effects of Public Housing Demolition on Children," *American Economic Review*, October 2018, *108* (10), 3028–3056.

_ , Robert Collinson, and Danielle H Sandler, "The Long-Run Effects of America's Largest Residential Racial Desegregation Program: Gautreaux*," *The Quarterly Journal of Economics*, August 2025, *140* (3), 2213–2267.

Cook, Lisa D., Trevon D. Logan, and John M. Parman, "Racial Segregation and Southern Lynching," *Social Science History*, 2018, *42* (4), 635–675. Publisher: Cambridge University Press.

Corak, Miles and Patrizio Piraino, "The Intergenerational Transmission of Employers," *Journal of Labor Economics*, 2011, *29* (1), 37–68. Publisher: [The University of Chicago Press, Society of Labor Economists, NORC at the University of Chicago].

Damm, Anna Piil, "Ethnic Enclaves and Immigrant Labor Market Outcomes: Quasi-Experimental Evidence," *Journal of Labor Economics*, April 2009, *27* (2), 281–314. Publisher: The University of Chicago Press.

Duncan, Dustin T., Ichiro Kawachi, and Foreword by Ana V. Diez Roux, eds, *Neighborhoods and Health*, second edition, new to this edition:, second edition, new to this edition: ed., Oxford, New York: Oxford University Press, April 2018.

Dunkelman, Marc J., "Next-Door Strangers: The Crisis of Urban Anonymity," *The Hedgehog Review*, 2017, *19* (2).

Eriksson, Katherine and Zachary Ward, "The Residential Segregation of Immigrants in the United States from 1850 to 1940," *The Journal of Economic History*, December 2019, *79* (4), 989–1026.

Fairlie, Robert W. and Alicia Robb, "Families, Human Capital, and Small Business: Evidence from the Characteristics of Business Owners Survey," *ILR Review*, January 2007, *60* (2), 225–245. Publisher: SAGE Publications Inc.

Fruehwirth, Jane Cooley and Jessica Gagete-Miranda, "Your peers' parents: Spillovers from parental education," *Economics of Education Review*, December 2019, *73*, 101910.

Greenberg, Kyle, Matthew Gudgeon, Adam Isen, Corbin L. Miller, and Richard W. Patterson, "Intergenerational Transmission of Occupation: Lessons from the United States Army," Technical Report 33009, National Bureau of Economic Research September, revised November 2024. NBER Working Paper Series.

Grinblatt, Mark, Matti Keloharju, and Seppo Ikheimo, "Social Influence and Consumption: Evidence from the Automobile Purchases of Neighbors," *The Review of Economics and Statistics*, 2008, *90* (4), 735–753. Publisher: The MIT Press.

Haltiwanger, John C., Mark J. Kutzbach, Giordano E. Palloni, Henry Pollakowski, Matthew Staiger, and Daniel Weinberg, "The Children of HOPE VI Demolitions: National Evidence on Labor Market Outcomes," November 2020.

Hellerstein, Judith K., Melissa McInerney, and David Neumark, "Neighbors and Coworkers: The Importance of Residential Labor Market Networks," *Journal of Labor Economics*, 2011, *29* (4), 659–695.

Hvide, Hans K. and Paul Oyer, "Dinner Table Human Capital and Entrepreneurship," January 2018.

Katz, Lawrence F., Jeffrey R. Kling, and Jeffrey B. Liebman, "Moving to opportunity in Boston: Early results of a randomized mobility experiment," *The quarterly journal of economics*, 2001, *116* (2), 607–654. Publisher: MIT Press.

Kawano, Laura, Bruce Sacerdote, William L. Skimmyhorn, and Michael Stevens, "On the Determinants of Young Adult Outcomes: Impacts of Randomly Assigned Neighborhoods For Children in Military Families," July 2024.

Kling, Jeffrey R, Jeffrey B Liebman, and Lawrence F Katz, "Experimental Analysis of Neighborhood Effects," *Econometrica*, January 2007, *75* (1), 83–119.

Laliberté, Jean-William, "Long-Term Contextual Effects in Education: Schools and Neighborhoods," *American Economic Journal: Economic Policy*, May 2021, *13* (2), 336–377.

Lentz, Bernard F. and David N. Laband, "Why So Many Children of Doctors Become Doctors: Nepotism vs. Human Capital Transfers," *The Journal of Human Resources*, 1989, *24* (3), 396–413. Publisher: [University of Wisconsin Press, Board of Regents of the University of Wisconsin System].

Library of Congress, U.S., "Fire insurance maps at the Library of Congress: a resource guide," n.d. accessed July 12, 2025.

Logan, Trevon D. and John M. Parman, "The national rise in residential segregation," *The Journal of Economic History*, 2017, *77* (1), 127–170. Publisher: Cambridge University Press.

_ **and** _ , "Segregation and Homeownership in the Early Twentieth Century," *American Economic Review*, May 2017, *107* (5), 410–414.

_ **and** _ , "Segregation and mortality over time and space," *Social Science & Medicine (1982)*, February 2018, *199*, 77–86.

Manski, Charles F., "Identification of Endogenous Social Effects: The Reflection Problem," *The Review of Economic Studies*, 1993, *60* (3), 531–542. Publisher: [Oxford University Press, Review of Economic Studies, Ltd.].

Matta, Juan and Alexis Orellana, "Neighbors' Spillovers on High School Choice," 2022.

McCartney, W. Ben and Avni M. Shah, "Household mortgage refinancing decisions are neighbor influenced, especially along racial lines," *Journal of Urban Economics*, March 2022, *128*, 103409.

Mertz, Mikkel, Maddalena Ronchi, and Viola Salvestrini, "Female representation and talent allocation in entrepreneurship: The role of early exposure to entrepreneurs," *Available at SSRN 4920176*, 2024.

Moretti, Enrico, "The Effect of High-Tech Clusters on the Productivity of Top Inventors," *American Economic Review*, October 2021, *111* (10), 3328–3375.

Olivetti, Claudia, Eleonora Patacchini, and Yves Zenou, "Mothers, peers, and gender-role identity," *Journal of the European Economic Association*, 2020, *18* (1), 266–301.

Pischke, Steve, "Lecture notes on measurement error," 2007.

Polyakova, Maria, Petra Persson, Katja Hofmann, and Anupam B. Jena, "Does medicine run in the family—evidence from three generations of physicians in Sweden: retrospective observational study," *BMJ*, December 2020, *371*, m4453. Publisher: British Medical Journal Publishing Group Section: Research.

Porter, Catherine and Danila Serra, "Gender differences in the choice of major: The importance of female role models," *American Economic Journal: Applied Economics*, 2020, *12* (3), 226–254.

Price, Joseph, "Database of Human Lifespan," 2024.

_ , Kasey Buckles, Jacob Van Leeuwen, and Isaac Riley, "Combining family history and machine learning to link historical records: The Census Tree data set," *Explorations in Economic History*, 2021, *80*, 101391. Publisher: Elsevier.

Quincy, Sarah, "Income shocks and housing spillovers: Evidence from the World War I Veterans' Bonus," *Journal of Urban Economics*, November 2022, *132*, 103494.

Redding, Stephen J and Daniel M Sturm, "Neighborhood Effects: Evidence from Wartime Destruction in London," April 2024, (32333).

Rosenbaum, James E., "Changing the geography of opportunity by expanding residential choice: Lessons from the Gautreaux program," *Housing Policy Debate*, January 1995, *6* (1), 231–269. Publisher: Routledge _eprint: https://doi.org/10.1080/10511482.1995.9521186.

Sanborn Map Company, "Collection: Sanborn maps," Various Years. Library of Congress Digital Collections.

Shah, Avni and W. Ben McCartney, ""I'll Have What She's Having": Neighborhood Social Interactions Lead to Policy Spillovers," *Journal of the Association for Consumer Research*, October 2023, *8* (4), 403–415. Publisher: The University of Chicago Press.

Sharkey, Patrick and Gerard Torrats-Espinosa, "The effect of violent crime on economic mobility," *Journal of Urban Economics*, November 2017, *102*, 22–33.

Staiger, Matthew, "The Intergenerational Transmission of Employers and the Earnings of Young Workers," October 2023.

Tan, Hui Ren, "How Widespread Are Social Network Effects? Evidence from the Early Twentieth-Century United States," *Journal of Labor Economics*, January 2022, *40* (1), 187–237. Publisher: The University of Chicago Press.

U.S. Bureau of the Census, Department of Commerce and Labor, "Thirteenth Census of the United States, April 15, 1910, Instructions to Enumerators," 1910.

Ventura, Maria, "Following in the family footsteps: Incidence and returns of occupational persistence," 2024.

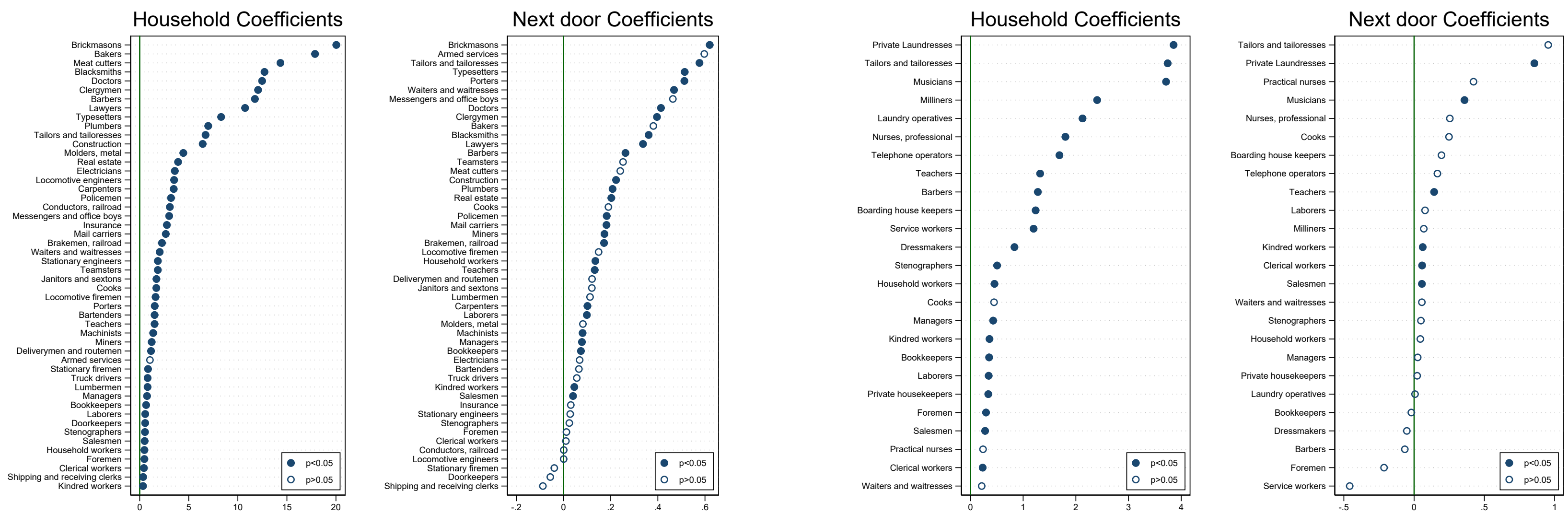
Wirth, Louis, "Urbanism as a Way of Life," *American Journal of Sociology*, 1938, *44* (1), 1–24. Publisher: The University of Chicago Press.

Tables and Figures

Figure 1: Household and Next Door Coefficients, Individual Occupations

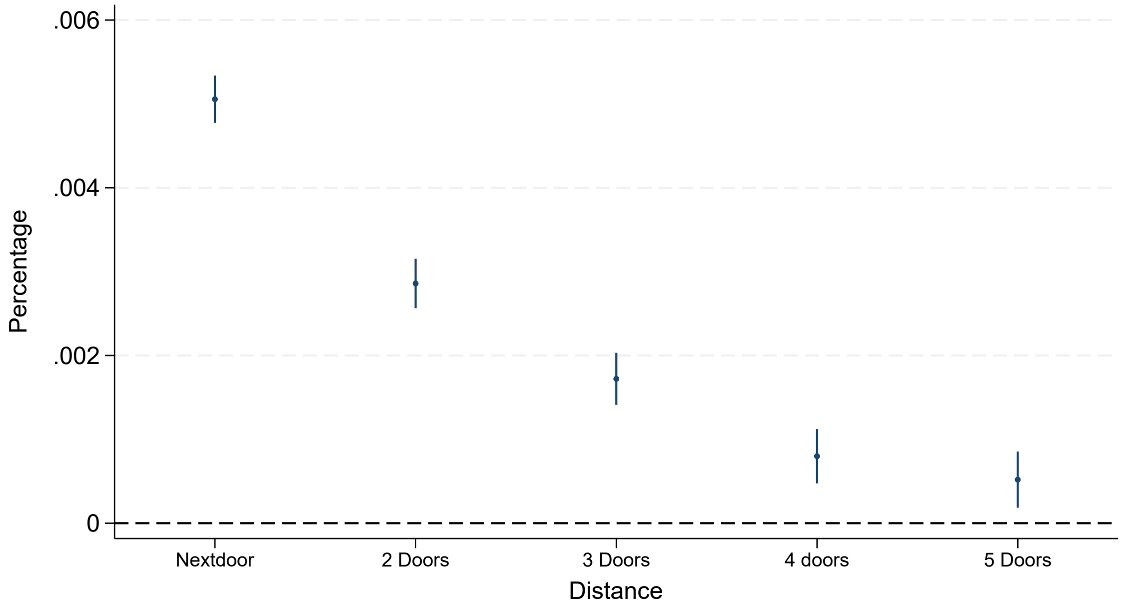
(a) 50 Largest Occupations for Men

(b) 25 Largest Occupations for Women



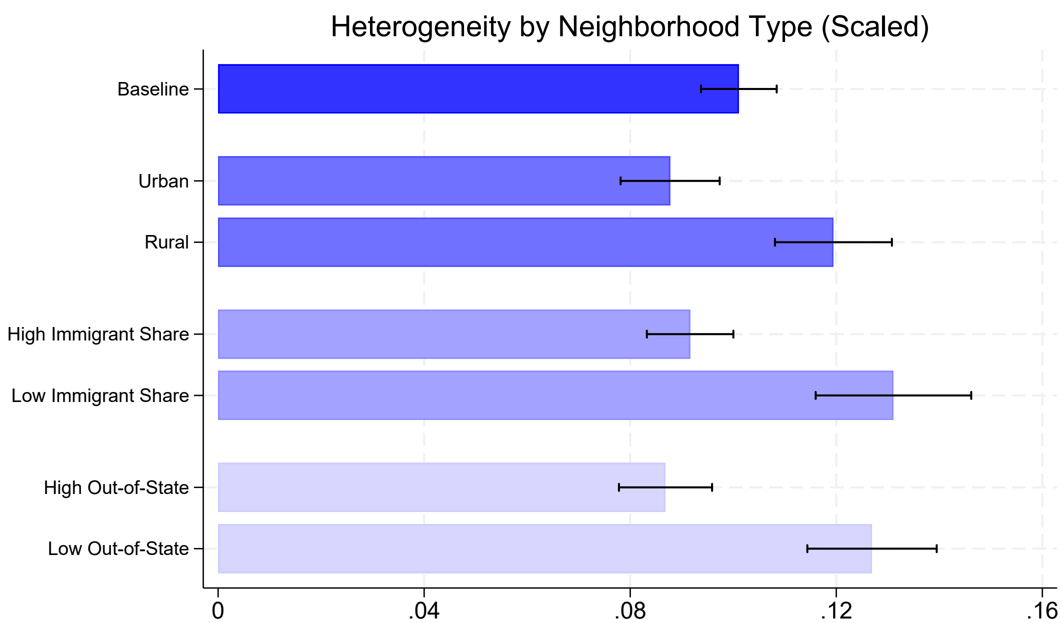
Notes: These figures report the results of next door regressions for each occupation run separately. The outcome is a binary measure that equals one if the child reports working in that occupation in 1940. The independent variable of interest is an indicator for living next door to someone with that occupation in 1910. Each marker represents the coefficient scaled by the untreated mean of the outcome. Markers are filled if the coefficient is statistically significant at the 5% level. All regressions also include age fixed effects.

Figure 2: Door Distance Regressions (50 Largest Male Occupations)



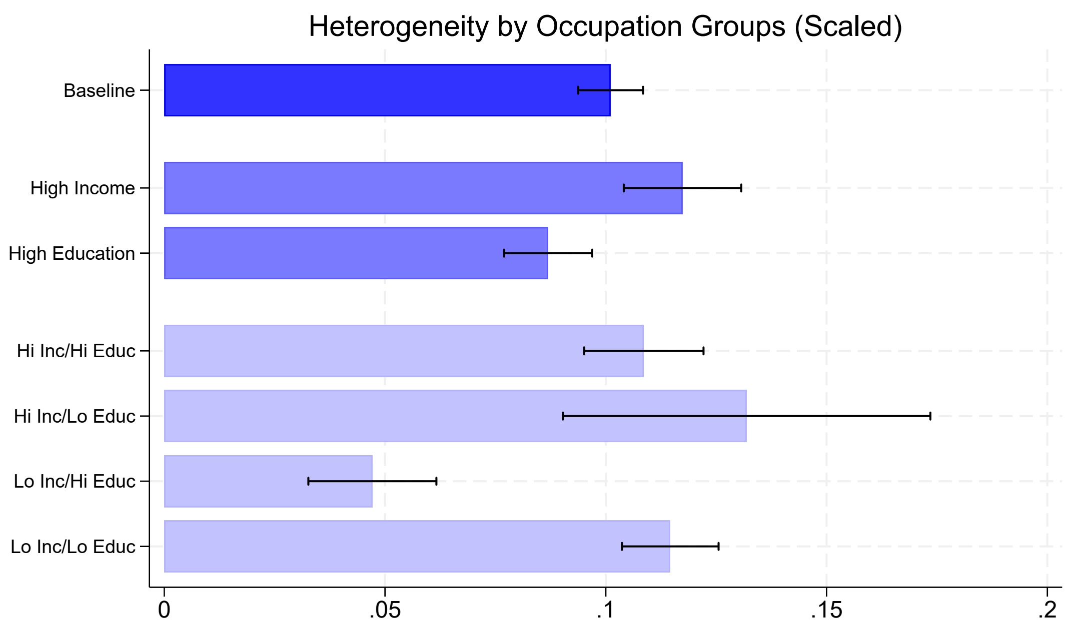
Notes: Coefficients from the stacked regression specification for the 50 largest occupations for boys. Sample restricted to boys between the ages of 5 and 18 in 1910 that can be linked to the 1940 census. The outcome is a binary measure that equals one if the boy chooses the exposed occupation in 1940. This is regressed on binary measures that equal one if the boy's family lived next door, two doors away, etc. from a particular occupation in 1910, as well as a binary measure that equals one if the child had that occupation in their own household (not plotted). Regression includes census sheet fixed effects. 95% confidence intervals included with standard errors clustered at the individual by occupation and household level.

Figure 3: Neighborhood Characteristics: Connectedness



Notes: Each bar reports the next door coefficient scaled by the untreated mean from the stacked regression specification described in equation 2. Each regression sample is restricted to groups of sheets that fit the description on the y-axis. Sample of interest is boys 5-18 years old in 1910 that are linked to the 1940 census. Error bars represent the 95% confidence interval (scaled using the the delta method).

Figure 4: Heterogeneity by Occupation Type

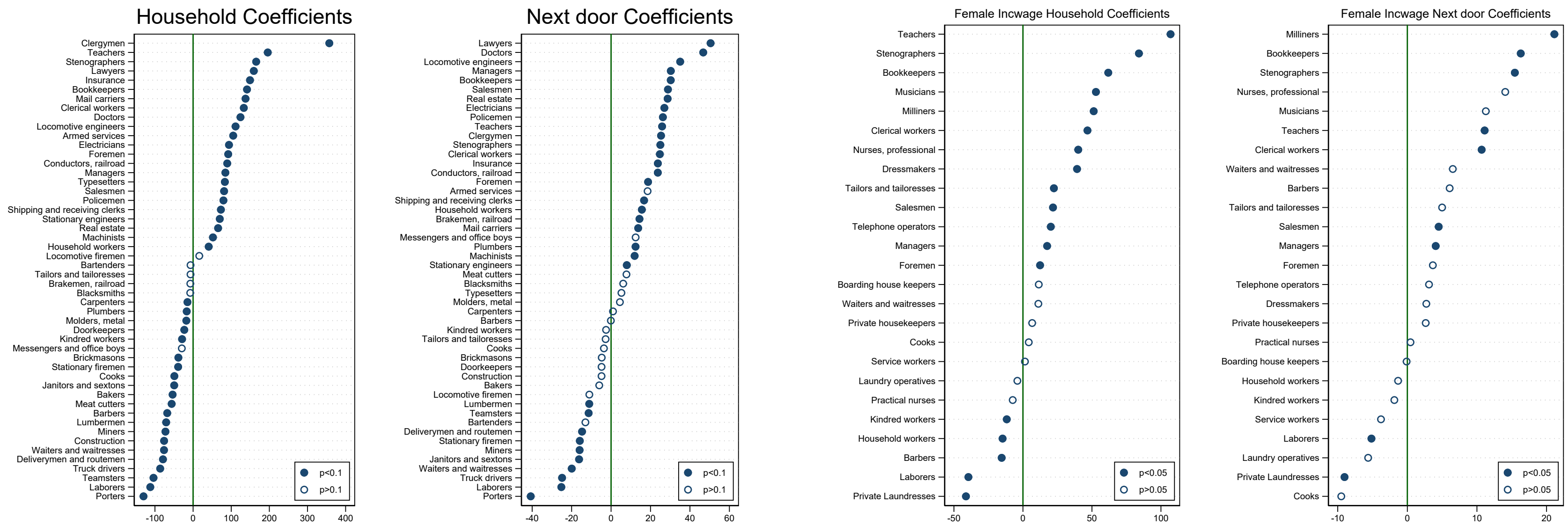


Notes: Each bar reports the next door coefficient scaled by the untreated mean from the stacked regression specification described in equation 2. Each regression sample is restricted to the subset of occupations in the stack denoted on the y-axis. Sample of interest is boys 5-18 years old in 1910 that are linked to the 1940 census. Error bars represent the 95% confidence interval (scaled using the the delta method). The occupation groups are divided using the median income and education occupation scores for each occupation from IPUMS.

Figure 5: Household and Next Door Coefficients on Income

(a) 50 Largest Occupations for Men

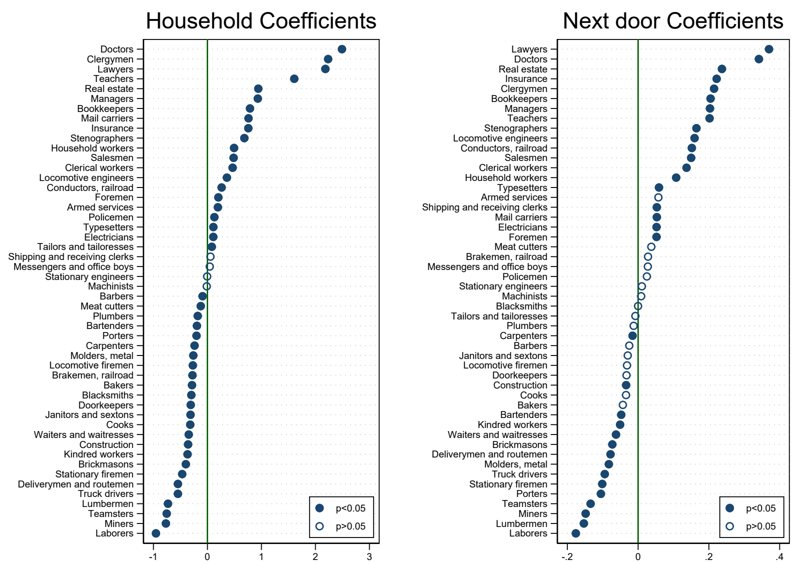
(b) 25 Largest Occupations for Women



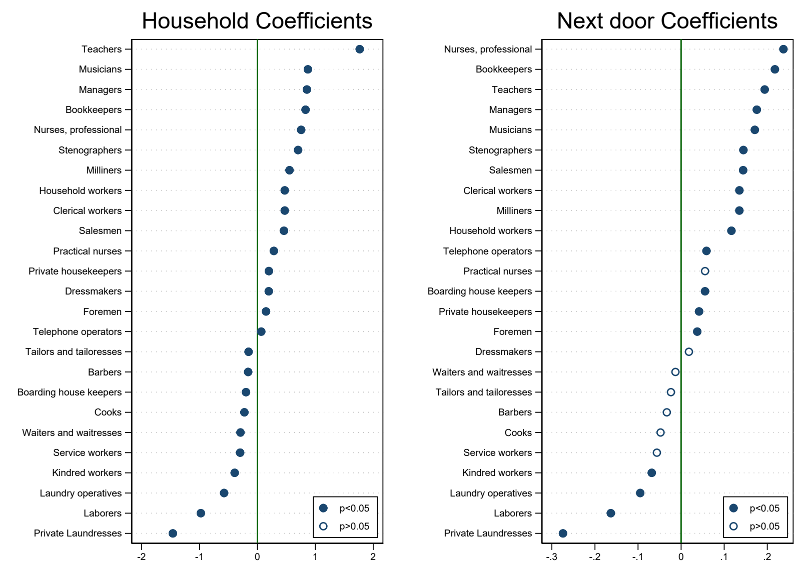
Notes: These figures report the results of income on next door regressions run for each occupation separately. The outcome is reported income in the 1940 census. The independent variable of interest is an indicator for living next door to someone with that occupation in 1910. Each marker represents the next door coefficient scaled by the untreated mean of the outcome. Markers are filled if the coefficient is statistically significant at the 5% level. All regressions also include own household indicators (not reported) and age fixed effects.

Figure 6: Household and Next Door Coefficients on Education

(a) 50 Largest Occupations for Men

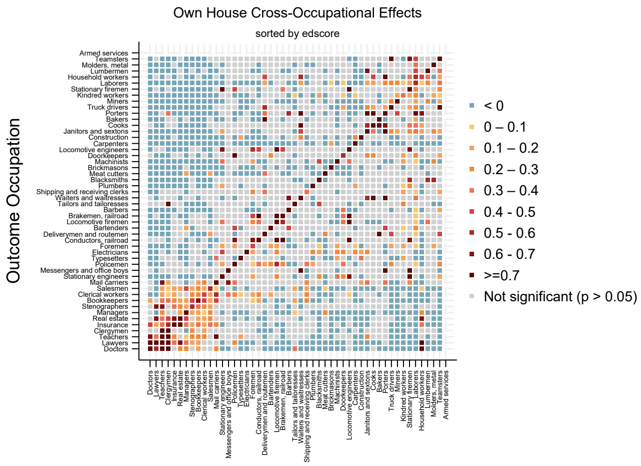


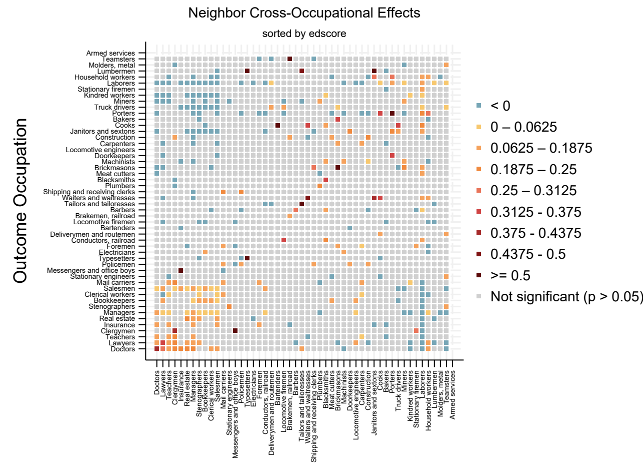
(b) 25 Largest Occupations for Women



Notes: These figures report the results of education on next door regressions run for each occupation separately. The outcome is years of educational attainment reported in the 1940 census. The independent variable of interest is an indicator for living next door to someone with that occupation in 1910. Each marker represents the next door coefficient scaled by the untreated mean of the outcome. Markers are filled if the coefficient is statistically significant at the 5% level. All regressions also include own household indicators (not reported) and age fixed effects.

Figure 7: Occupation Spillover Matrices





Notes: These Figures report the coefficients for occupation spillovers for own household (left panel) and nextdoor neighbor (right panel) transmission. The x-axis corresponds to the occupation that the child is exposed to in 1910 and the y-axis corresponds to the potential occupation that a child could report in 1940. Each square in the matrix is colored to represent the sign, magnitude, and significance level of the coefficient from a regression of the y-axis outcome on the x-axis exposure as described in the text. For example, living next door to a doctor has a large positive effect on becoming a doctor, an insignificant effect on becoming a plumber, and a significant negative effect on becoming a laborer.

Table 1: Summary Statistics

	All 1910 children	1910 - 1940 Linked children	1910 Boys	1910 - 1940 Linked boys	1910 Girls	1910 - 1940 Linked girls
	(1)	(2)	(3)	(4)	(5)	(6)
Age in 1910	11.380	11.244	11.366	11.228	11.395	11.268
Male	0.503	0.603	1.000	1.000	0.000	0.000
Nonwhite	0.130	0.047	0.128	0.060	0.132	0.027
Rural	0.603	0.646	0.611	0.634	0.595	0.666
Northeast	0.252	0.226	0.250	0.240	0.253	0.204
Midwest	0.315	0.378	0.316	0.368	0.314	0.395
South	0.371	0.324	0.371	0.323	0.371	0.325
West	0.063	0.072	0.063	0.070	0.062	0.076
Head of household						
Married	0.884	0.914	0.883	0.907	0.885	0.923
Foreign born	0.311	0.277	0.311	0.292	0.310	0.255
Income score	21.975	22.107	21.899	22.170	22.051	22.011
Education score	12.342	12.867	12.124	12.712	12.563	13.103
Total	26,161,014	10,529,180	13,146,449	6,346,719	13,014,565	4,182,461

Notes: This table shows summary statistics for the sample data. Column 1 describes all children age 5-18 in the 1910 census. Column 2 describes the children linked from the 1910 census to the 1940 census. Columns 3-4 further restricts the sample to boys only, while columns 5-6 restrict to girls only.

Table 2: Baseline Doctor Results

Dependent variable: Doctor occupation in 1940	All sheets	At least one doctor occupation per sheet: No geographic FE	City - enumeration district FE	Sheet FE
	(1)	(2)	(3)	(4)
Own household	0.0983***	0.0954***	0.0963***	0.0970***
	(0.0018)	(0.0019)	(0.0018)	(0.0018)
Next door neighbor	0.0072***	0.0036***	0.0033***	0.0032***
	(0.0005)	(0.0006)	(0.0006)	(0.0007)
R-squared	0.011	0.046	0.138	0.327
Untreated mean	0.0038	0.0078	0.0078	0.0078
Observations	6,335,660	305,059	305,059	305,059

Notes: Column 1 includes all sheets in 1910 census, while columns 2-4 include only sheets with at least one adult in the target occupation. All columns include age fixed effects.
**p*<0.1, ***p*<0.05, ****p*<0.01.

Table 3: Baseline Stacked Regression Results

Dependent variable: Target occupation in 1940	All sheets	At least one person in target occupation per sheet: No geographic FE	City - enumeration district FE	Sheet FE
	(1)	(2)	(3)	(4)
Panel A: Men				
Own household	0.0548***	0.0467***	0.0424***	0.0381***
	(0.0001)	(0.0002)	(0.0002)	(0.0002)
Next door neighbor	0.0198***	0.0112***	0.0070***	0.0034***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
R-squared	0.038	0.048	0.087	0.323
Untreated mean	0.0100	0.0331	0.0331	0.0331
Observations	316,783,008	26,649,908	26,649,908	26,649,908
Panel B: Women				
Own household	0.0158***	0.0116***	0.0095***	0.0085***
	(0.0001)	(0.0001)	(0.0001)	(0.0002)
Next door neighbor	0.0087***	0.0041***	0.0020***	0.0010***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
R-squared	0.011	0.018	0.083	0.356
Untreated mean	0.0072	0.0156	0.0156	0.0156
Observations	104,389,272	11,135,893	11,135,893	11,135,893

Notes: This table reports estimates from the stacked regression specification, with Panel A showing results for the 50 largest male occupations, and Panel B for the 25 largest female occupations. Column 1 includes all sheets in 1910 census, while columns 2-4 include only sheets with at least one adult in the target occupation. All columns include age fixed effects.
**p*<0.1, ***p*<0.05, ****p*<0.01.

Table 4: Untimely Deaths Analysis

Dependent variable: Target occupation in 1940	Doctors Only: All	Age < 60	Age < 50	Age < 40	50 Largest Male Occupations (Stack): All	Age < 60	Age < 50	Age < 40
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Own household: died	0.0439***	0.0479***	0.0619***	0.0508***	0.0283***	0.0292***	0.0293***	0.0261***
	(0.0058)	(0.0070)	(0.0101)	(0.0179)	(0.0013)	(0.0018)	(0.0021)	(0.0028)
Next door neighbor: died	-0.0004	-0.0020	0.0011	0.0017	0.0026***	0.0015	0.0006	0.0007
	(0.0024)	(0.0033)	(0.0045)	(0.0050)	(0.0009)	(0.0014)	(0.0017)	(0.0022)
Own household: *did not* die	0.1047***	0.1076***	0.1150***	0.1026***	0.0423***	0.0436***	0.0439***	0.0387***
	(0.0023)	(0.0024)	(0.0028)	(0.0040)	(0.0003)	(0.0003)	(0.0003)	(0.0004)
Next door neighbor: *did not* die	0.0031***	0.0032***	0.0040***	0.0043***	0.0024***	0.0026***	0.0026***	0.0026***
	(0.0009)	(0.0010)	(0.0011)	(0.0013)	(0.0002)	(0.0002)	(0.0003)	(0.0003)
Share of children with a neighbor that died	0.0787	0.0558	0.0423	0.0315	0.0478	0.0277	0.0225	0.0192
R-squared	0.322	0.324	0.330	0.318	0.275	0.273	0.272	0.268
Untreated mean	0.0071	0.0072	0.0070	0.0066	0.0305	0.0310	0.0312	0.0313
Observations	123,833	111,446	88,938	49,707	5,317,478	4,224,135	3,731,294	2,732,185

Notes: Regressions use the baseline doctor or stack sample restricting to only sheets with exactly one doctor (or each target occupation in the stack). We exclude individuals that have a missing death date. Columns 2-4 and 6-8 further restrict to sheets with doctors (or target occupation) that are below a particular age.
**p*<0.1, ***p*<0.05, ****p*<0.01.

Table 5: Doctor Regression – Alternative Control Strategies

Dependent variable: Doctor occupation in 1940	Baseline	Head of Household controls	Head Occupation FE	Only kids within 3 doors of doctor	Within 3 doors + Head Occupation FE
	(1)	(2)	(3)	(4)	(5)
Own household	0.0970***	0.0709***	0.0284***	0.0984***	0.0300***
	(0.0018)	(0.0021)	(0.0026)	(0.0022)	(0.0029)
Next door neighbor	0.0032***	0.0025***	0.0027***	0.0040***	0.0031***
	(0.0007)	(0.0007)	(0.0006)	(0.0011)	(0.0007)
Head occscore		0.0001**			
		(0.00004)			
Head ed score		0.0003***			
		(0.00003)			
Head age		-0.0002***	-0.0002***		-0.0003***
		(0.00003)	(0.00002)		(0.00004)
White		-0.0039*	-0.0001		-0.0016
		(0.0021)	(0.0011)		(0.0022)
US Native		-0.0028***	-0.0047***		-0.0067***
		(0.0010)	(0.0007)		(0.0012)
Education flag		0.0100***			
		(0.0013)			
R-squared	0.327	0.329	0.055	0.406	0.062
Untreated mean	0.008	0.008	0.008	0.008	0.008
Observations	305,059	305,047	305,036	143,141	143,133

Notes: Column 1 corresponds to Column 4 of the baseline doctor table. Column 2 includes sheet fixed effects and adds controls for observable characteristics of the head of household. Column 3 inludes head of household occupation fixed effects as well as the head of household controls. Columns 4 and 5 restrict the sample to only kids who live within three doors of a doctor. All columns include sheet and age fixed effects.

$^*p<0.1$, $^{**}p<0.05$, $^{***}p<0.01$.

Table 6: Heterogeneity by Neighbor Similarity

Dependent variable: Target occupation in 1940	Baseline	Birthplace	Birth country	Race	Same age
	(1)	(2)	(3)	(4)	(5)
Own household	0.0381***	0.0381***	0.0381***	0.0381***	0.0382***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Next door neighbor	0.0034***				
	(0.0001)				
Treatment 1: Same characteristic		0.0047***	0.0040***	0.0038***	0.0035***
		(0.0002)	(0.0002)	(0.0001)	(0.0003)
Treatment 1 average		*0.0984*	*0.1595*	*0.2284*	*0.0311*
Treatment 2: Different characteristic		0.0023***	0.0019***	-0.0083***	0.0031***
		(0.0002)	(0.0002)	(0.0006)	(0.0001)
Treatment 2 average		*0.1453*	*0.0829*	*0.0086*	*0.1639*
F-test		100.28	66.90	404.93	1.60
Two-tailed p-value		0.0000	0.0000	0.0000	0.2058
R-squared	0.323	0.323	0.323	0.323	0.323
Untreated mean	0.0331	0.0331	0.0331	0.0331	0.0338
Observations	26,649,908	26,649,908	26,649,908	26,649,908	26,649,908

Notes: The sample is all boys in the 50 occupations of the baseline stack regressions. In Columns 2-5, treatments compare head of household of the focal child to the target occupation holder next door. In Column 5, same age is defined as any child in the target occupation house being within one year of the focal child.

*p<0.1, **p<0.05, ***p<0.01.

Table 7: Heterogeneity by Gender Homophily

Dependent variable: Target occupation in 1940	Boys Baseline	Boys Homophily	Girls Baseline	Girls Homophily
	(1)	(2)	(3)	(4)
Own household	0.0381***		0.0085***	
	(0.0002)		(0.0002)	
Next door neighbor	0.0034***		0.0010***	
	(0.0001)		(0.0001)	
Male own household		0.0431***		0.0051***
		(0.0002)		(0.0002)
Male next door neighbor		0.0037***		0.0008***
		(0.0001)		(0.0002)
Female own household		0.0095***		0.0149***
		(0.0003)		(0.0003)
Female next door neighbor		0.0012***		0.0014***
		(0.0003)		(0.0002)
R-squared	0.323	0.323	0.356	0.356
Untreated mean	0.0331	0.0331	0.0156	0.0156
Observations	26,649,908	26,649,908	11,135,893	11,135,893

Notes : The sample in Columns 1 and 2 are the boys in the stack regression of 50 largest male occupations. The sample in Columns 2 and 4 are the girls in the stack regression of 25 largest female occupations. Columns 2 and 4 separate the treatment variables by the gender of the adult in the target occupation. All columns include sheet and age fixed effects.

*$p<0.1$, **$p<0.05$, ***$p<0.01$.

A Reconstructing Neighborhood Microgeography from Historical Censuses

Figure A1 shows an example page from the 1910 census, covering parts of West Division and Grace Streets in Chicago, IL.[28] Our goal is to use these census data to reconstruct the microgeography of a neighborhood. We do this in two ways:

1. Use the order in which households are listed on each sheet to determine (this is what we use for our baseline results).

2. Use reported street addresses (this is what we use for robustness analyses in Appendix D).

In this section, we describe how we implement each method, as well as documenting strengths of each, potential sources of measurement error, and how we address those potential errors.

To implement the sheet order method, we rely on the fact that prior to 1970 censuses were enumerated door-to-door in a predictable manner, and consequently household that are listed immediately above or below one another ar the immediate next households one would visit when walking down the street. Figure A2 provides an illustration from the 1910 census enumerator instructions showing how the enumerators were required to interview households and confirming that enumerators were required to proceed in a way that preserves microgeographic proximity.

To implement the street address method, we use the fact that street names are written vertically on the left hand side of census manuscript sheets. House numbers are then filled into the first column. Figure A1 shows both of these for a neighborhood in Chicago. One limitation of the street addresses is that they are only recorded in urban areas and when available. Additionally, difficulties in transcribing street names and numbers may result in

[28]These sheets contain portions of a neighborhood just east of Goose Island and the North Branch Canal, just south of W. Scott Street (called Vedder Street in the 1910 census). Grace Street no longer exists; the buildings have been demolished and the block is now vacant.

additional missing street addresses. In total, about 65% of households in the 1910 census have a street address. The unavailability of street addresses for a large fraction of households is one reason why we use the census sheet order for our baseline results.

For the subset of census sheets that contain street addresses, we use the reported street addresses in combination with historical maps from the Sanborn Fire Insurance Company (Sanborn Map Company, Various Years) to validate that the sheet order does in fact accurately reflect neighborhood microgeographic proximity. Figure A3 shows an example of a Sanborn map for a portion of northern Chicago between Lincoln Park and the Chicago River. Each colored square represents a page in the map book with a map and house numbers for each portion of a neighborhood. For more information on the Sanborn maps, see Library of Congress (n.d.).

In Figure A4, we zoom in on a portion of the census sheet presented in Figure A1 along with its corresponding Sanborn map. By matching census pages to the maps, we can "follow the enumerator down a street." The blue numbers map a household in the census sheet to the corresponding house in the map. The first household reported on this portion of Grace Street after turning the corner from West Division is house number 1206; the map shows that it is the second physical structure on the street, but 1200 Grace Street (the first structure) is non-residential and so does not contain a household enumerated by the census. The census sheet reveals that 1206 Grace Street is a multi-family home containing four households. Because census enumerators typically only write the house number for the first household in multi-family homes, the last three households in 1206 Grace Street have missing addresses in the transcribed census data; hence, without additional imputation rules for missing addresses, the street address method would not be able to detect the geographic proximity between the four households living at 1206 Grace Street.[29] The census sheet correctly orders the next three households, at 1210, 1212, and 1216 Grace Street; these are

[29]The census sheet order may introduce a different kind of measurement order if the census sheet order does not capture the frequency of interactions when households all live in the same structure. We further address multi-family homes below.

the last four households on the census sheet.[30]

Figure A5 follows the same enumerator a few census sheets later. On the first line of this census sheet, the enumerator finished the last house on a small section of Vedder Street (Vedder intersects Grace Street), a multi-family home at 743 Vedder Street. Following the enumerator instructions, the enumerator then turned a corner and proceeded down North Halsted Street. The street address method, which determines microgeographic proximity by identifying close street numbers on the same street, is unable to identify the close proximity between the last house on Vedder Street and the first house on Halsted Street; the sheet order method, in contrast, preserves this proximity.

Figure A5 highlights the fifth household on this census sheet, 1933 Halsted Street. As can be seen on the census sheet, the written address is difficult to read. In the transcribed census records, this address is mis-recorded as 1223 Halsted. If we used the street address method, we would therefore erroneously place this household further down its street and hence measure nextdoor neighbors with error; the census sheet method again avoids this source of measurement error because it does not rely on transcribed addresses to determine microgeographic proximity.

We next use the street addresses and Sanborn maps to quantify the level of measurement error in the sheet order definitions of neighbors. To do this, we first restrict the sample to focal child on the same census sheet as a doctor. We then choose 1 percent of doctors (on 383 sheets) at random to verify by hand. About one quarter of these census sheets have addresses and can be matched to a Sanborm map. We then hand check each observation and construct a measure of neighbors based on the addresses on census image and the Sanborn maps. For this subsample, we find that the sheet-order based next door neighbor measure is correct 91 percent of the time.

Figure A6 presents additional suggestive evidence, using the full sample with recorded street addresses, that sheet order typically does a good job of capturing microgeographic

[30] 1212 Grace Street is also a multi-family home.

proximity. Panel A shows that next door households are much more likely to have a house number that is within two of one another (e.g., 1210 Grace Street and 1212 Grace Street). Panel B shows that households that are closer to one another on a census sheet are more likely to be on the same street; note that if there were no measurement error in the census sheets, the share of next door households on the same street would still be less than 100% since census enumerators are instructed to turn corners, as shown in Figure A5.

A.1 Supplemental Sheets

A potential source of measurement error with the census sheet method arises if a family is not home at the time the enumerator first canvasses the neighborhood. As the enumerator instructions detail:

> "In case a family is out at the first visit, or in case the only persons at home are...not able to supply the required information about the members of the family, you must enumerate this family at a later visit. But no space should be left blank for this family upon the schedule you are filling at the time of your first visit unless you have positive and reliable information as to the number of persons in the family so that you will know exactly how many lines to leave blank...
>
> Use an extra sheet or sheets of the population schedule for enumerating those families who were out at the time of your first visit or those individuals for whom no spaces were left blank or no names were entered on the sheet of schedule you were filling at that time. At the head of these extra sheets write the word "Supplemental" and number them, finally, as the last sheets used in your work" (U.S. Bureau of the Census, 1910, p. 22-23).

Unfortunately, when the census manuscript pages were transcribed, no record was kept of which pages were supplemental. If supplemental sheets are used, this leads to measurement error in nextdoor neighbors for two reasons. First, households recorded on adjacent rows of a supplemental sheets do not in fact reside nextdoor to one another. Second, when a household's immediate neighbor appears on the supplemental sheet, the next household listed on the manuscript sheet would not actually be the nextdoor neighbor; the household would mistakenly be identified as nextdoor to a neighbor two doors away rather than the true nextdoor neighbor. To the best of our knowledge, the issue of supplemental sheets has not been acknowledged in previous work using census sheets to determine neighborhood proximity. Notice that this issue will not arise when using the street address method, since we can use street addresses recorded on the supplemental sheets to place the supplemental households adjacent to their actual neighbors.

To ensure that the presence of supplemental sheets are not affecting our results, in Tables A1 and A2 we repeat our baseline doctor and stacked regression results, respectively, but drop sheets likely to be supplemental. Since supplemental sheets are included at the end of each enumeration district, we discard the last one, two, and five sheets from each district. All results are similar to our baseline results. We additionally find that no occupations were especially overrepresented on the likely supplemental sheets.

Table A1: Doctor Results while Dropping Possible Supplemental Sheets

Dependent variable: Target occupation in 1940	Baseline	Number of sheets dropped		
		Last sheet	Last two sheets	Last five sheets
	(1)	(2)	(3)	(4)
Own household	0.0970***	0.0969***	0.0965***	0.0956***
	(0.0018)	(0.0018)	(0.0019)	(0.0020)
Next door neighbor	0.0032***	0.0030***	0.0029***	0.0032***
	(0.0007)	(0.0007)	(0.0007)	(0.0008)
R-squared	0.327	0.326	0.325	0.323
Untreated mean	0.0078	0.0078	0.0079	0.0080
Observations	305,059	298,801	288,966	250,270

Notes: Columns 1-4 include only sheets with at least one adult in the target occupation. All columns include sheet and age fixed effects.
**p* <0.1, ***p* <0.05, ****p* <0.01.

Table A2: Stacked Regression Results while Dropping Possible Supplemental Sheets

Dependent variable: Target occupation in 1940	Baseline	Number of sheets dropped		
		Last sheet	Last two sheets	Last five sheets
	(1)	(2)	(3)	(4)
Own household	0.0408***	0.0408***	0.0408***	0.0407***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Next door neighbor	0.0042***	0.0041***	0.0041***	0.0041***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)
R-squared	0.368	0.367	0.367	0.365
Untreated mean	0.0346	0.0345	0.0345	0.0344
Observations	33,117,306	32,372,452	31,235,482	26,389,976

Notes: Columns 1-4 include only sheets with at least one adult in the target occupation. All columns include sheet and age fixed effects.
**p* <0.1, ***p* <0.05, ****p* <0.01.

Figure A1: Example 1910 Census Manuscript Page, West Division and Grace St., Chicago

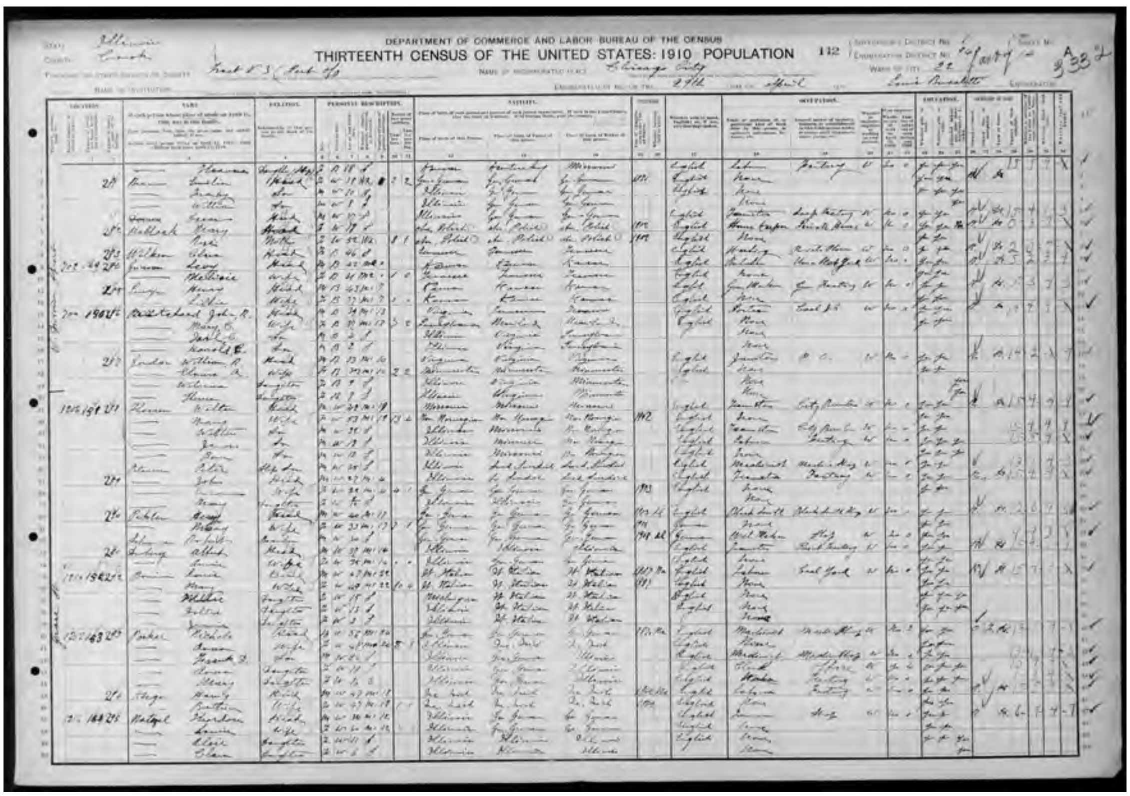
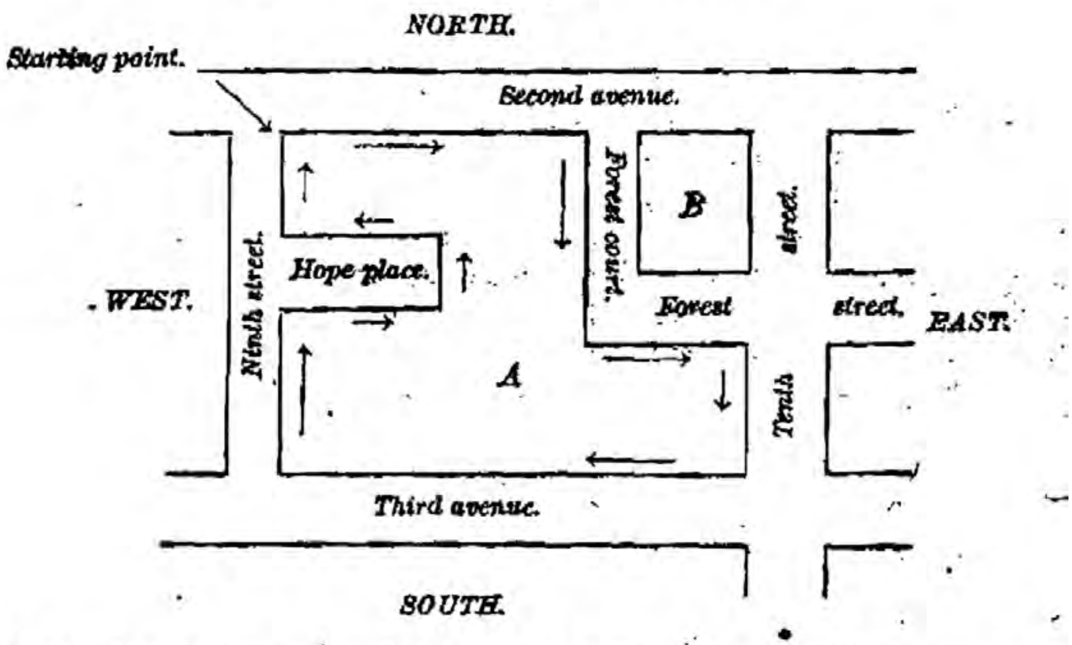


Figure A2: 1910 Census Enumerator Instructions

69. The arrows in the following diagram indicate the manner in which a block containing an interior court or place is to be canvassed:



(Note that block marked A is to be fully canvassed before work is undertaken in block B.)

Figure A3: Example of Sanborn Fire Insurance Map for a Portion of Chicago, IL

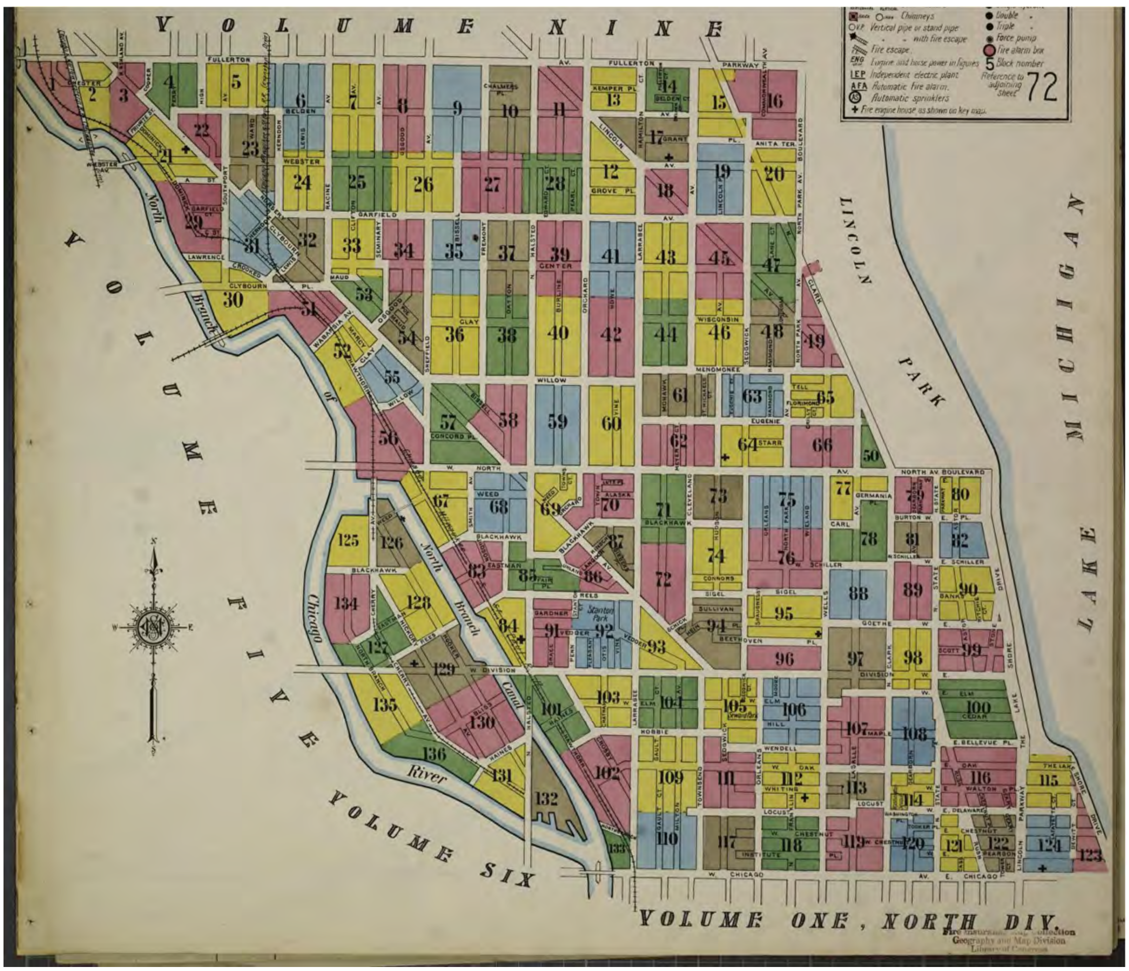
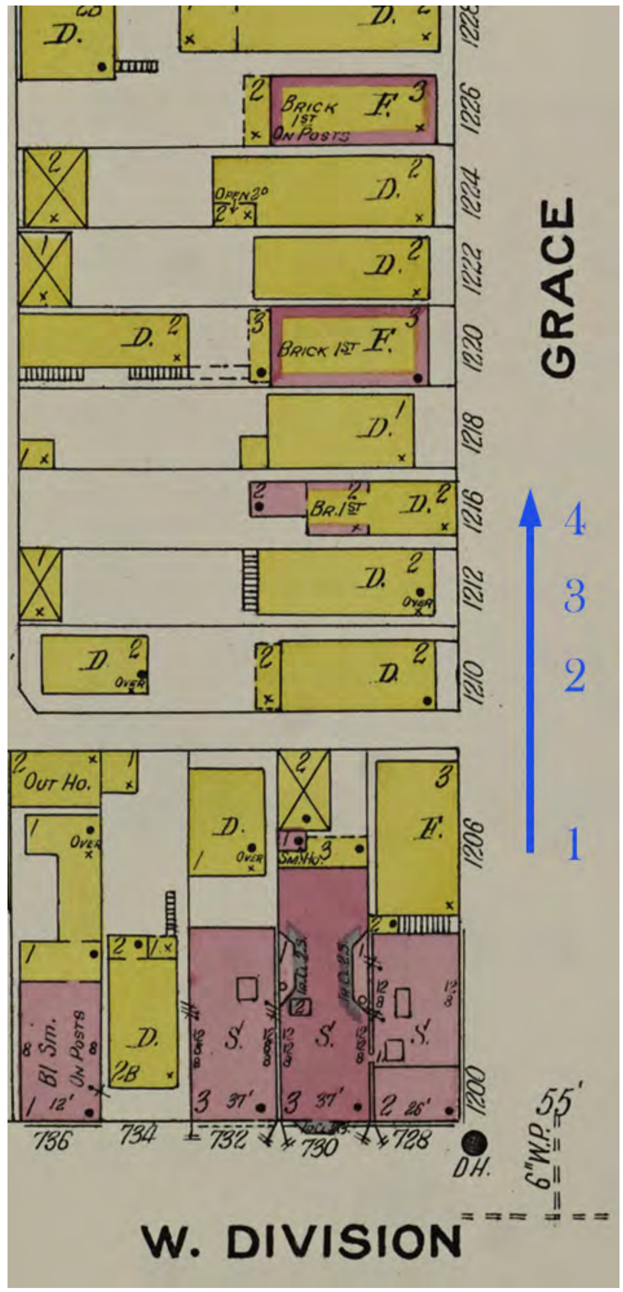
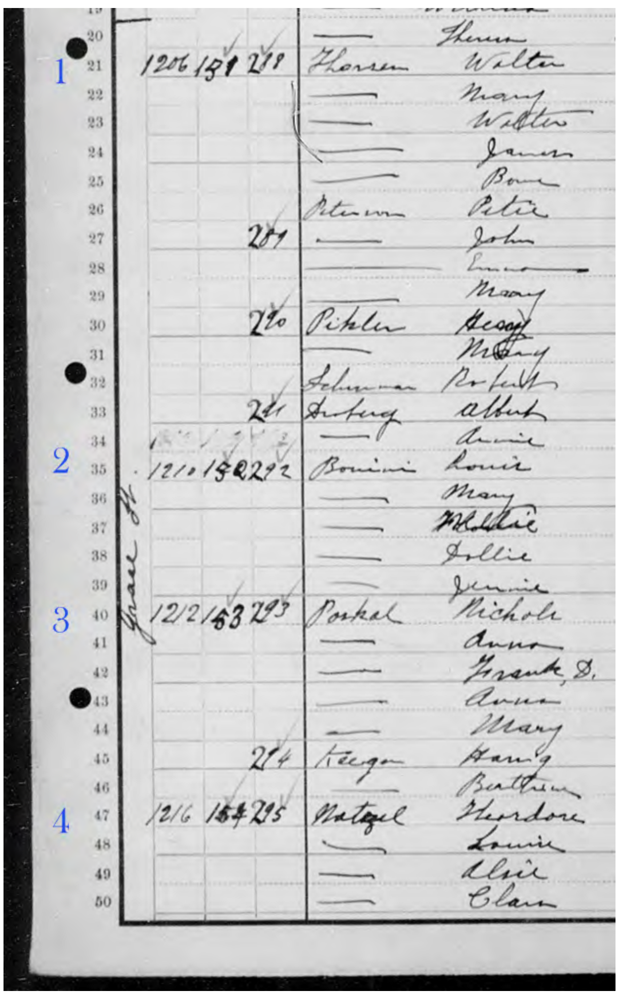


Figure A4: Detailed View of Grace Street, Chicago, IL

(b)



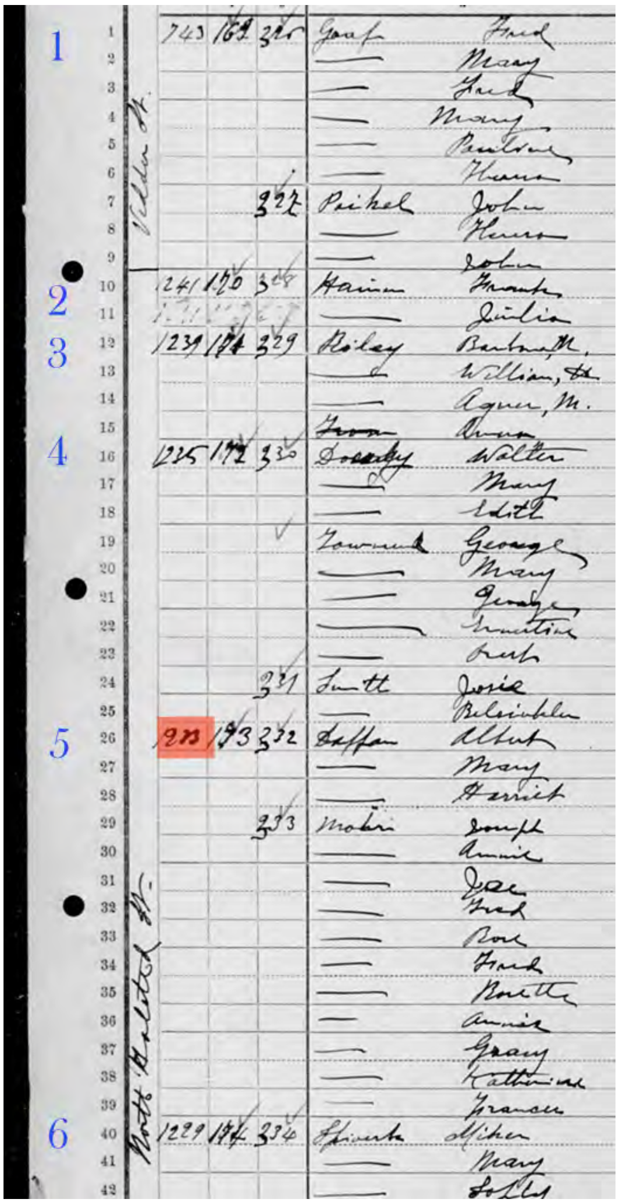
(a)



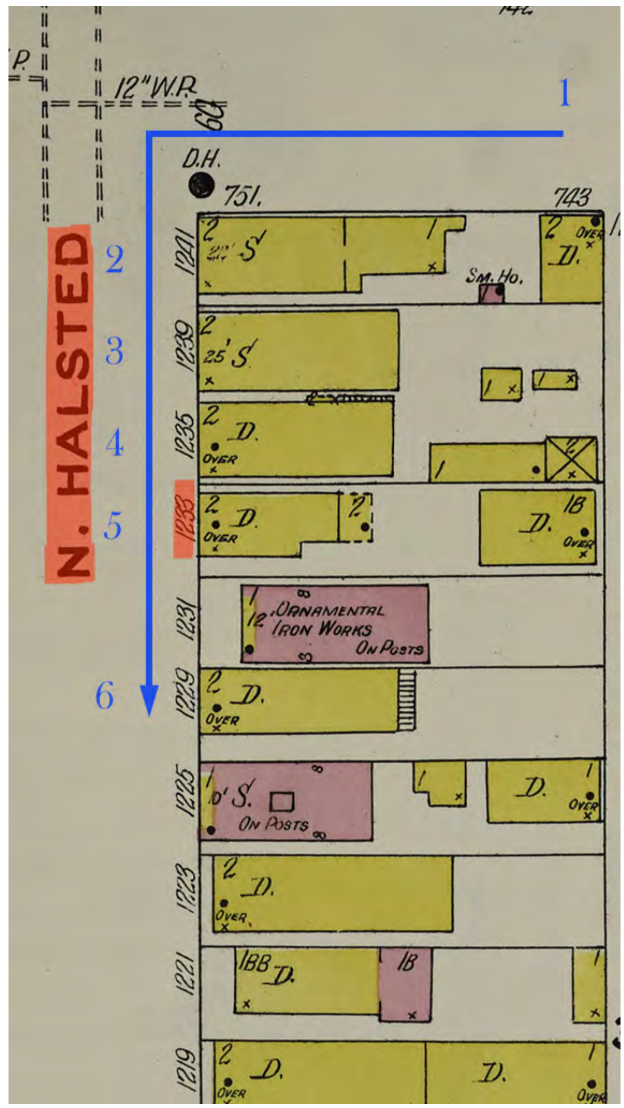
Notes:

Figure A5: Detailed View of North Halsted Street, Chicago, IL

(a)

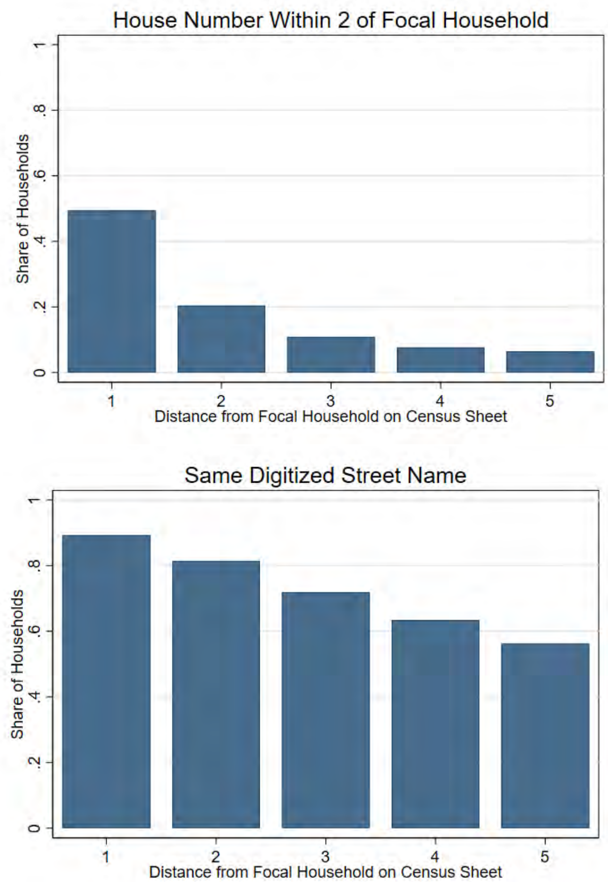


(b)



Notes:

Figure A6: Sheet Proximity and Address Measures



Notes: Observation at the household level from the 1910 full count census. Sample restricted to enumeration districts where 80 percent of households have digitized address information. In the top panel, for each household, the share of households where the house number is within 2 digits of the house number is plotted by the number of households apart based on the sheet ordering definition. In the bottom panel, for each household, the share of households that are on the same street is plotted by the number of households apart based on the sheet ordering definition.

B Within-Neighborhood Sorting in the 1910 Census

Another way to show that adults with the same occupation do not cluster adjacent to one another is to use a measure of within-neighborhood residential segregation, such as that proposed by Logan and Parman (2017a). We adopt their measure, which calculates how likely an adult in one occupation is to live next to an adult in a different occupation relative to what would be expected when randomly allocating occupations across households. For a given occupation j, this is given by:

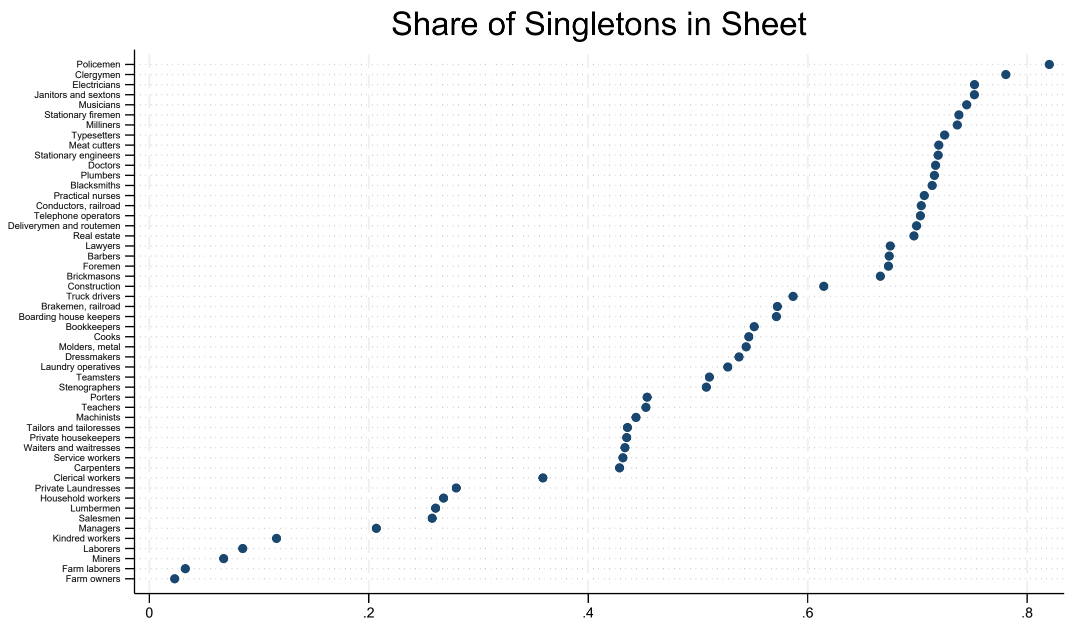
$$\eta_j = \frac{E(\bar{x_j}) - x_j}{E(\bar{x_j}) - E(\underline{x_j})}, \tag{7}$$

where x_j is the observed number of pairs of adjacent households in which one household has occupation j and the other does not, $E(\bar{x_j})$ is the expected x_j if households sorted randomly given the total number of j in the population, and $E(\underline{x_j})$ is the expected x_j if households were perfectly segregated by occupation.[31] $\eta_j = 0$ therefore corresponds to no residential segregation for occupation j, while $\eta_j = 1$ corresponds to perfect segregation. In all cases, η_j is close to zero. To put these segregation measures into perspective, we compare them to segregation measures by race and ethnicity (country of origin) for the 5 largest foreign born groups: German, Italian, Irish, Russian, and Canadian. Every occupation that we study is far less segregated than are race and ethnicity. Comparing across occupations, η_j tends to be larger for occupations related to agriculture or natural resource extraction (farm laborers, miners), where we expect most households on a census sheet to have the same occupation, although even in these cases segregation is much less than by race or ethnicity. In our baseline results, we exclude farmers and farm laborers, in part because of this potential for

[31] In the simplest case of an enumeration district, $E(\underline{x_j}) = 2$, since all households of occupation j would be clustered together on one part of the sheet and hence only one occupation j household in the first row of the sheet containing j occupations and one occupation j household at the last row containing occupation j households would be adjacent to a household with a different occupation. However, if this enumeration district was split into multiple, smaller sheets it is possible that $E(\underline{x_j}) = 0$, if every household has someone in the same occupation j.

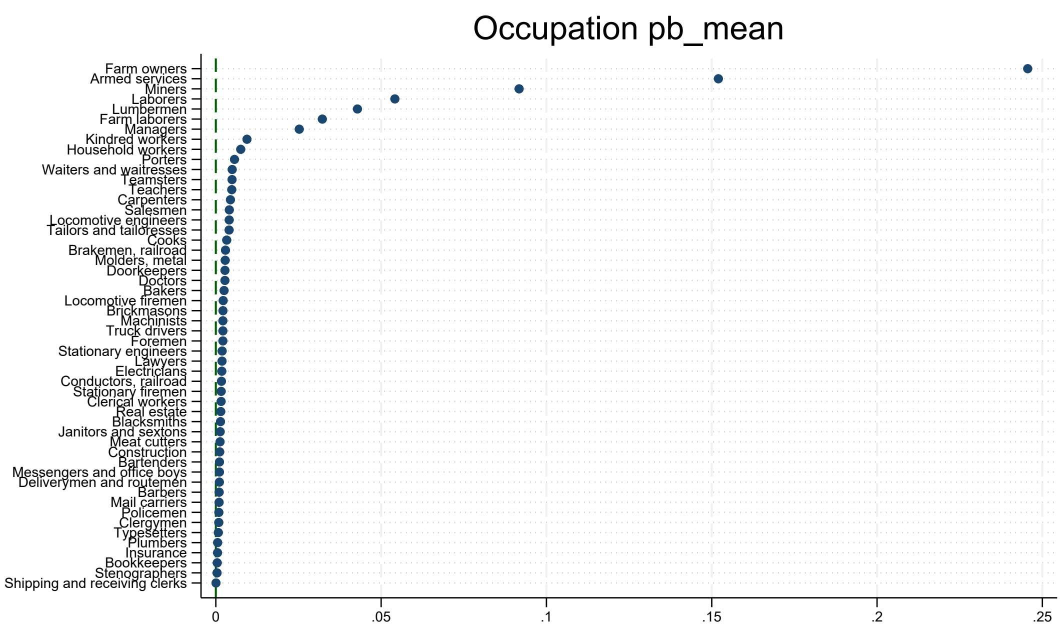
sorting.

Figure B1: Share of Singleton Occupations



Among the census manuscript pages with at least one occurrence of each occupation, the share of pages in which the occupation appears exactly once.

Figure B2: Logan-Parman (2017) Measure of Segregation



C Additional Tables and Figures

Figure C1: Household and Next Door Coefficients for 50 Largest Occupations for Men, Without Scaling by Mean of Untreated Kids

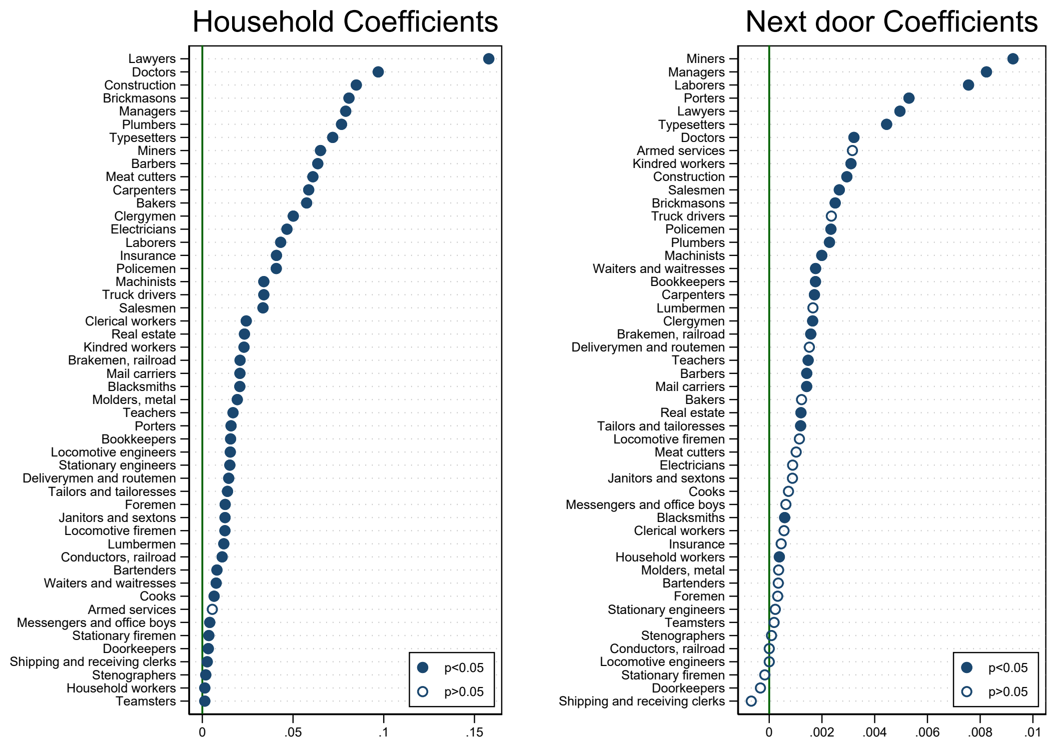


Table C1: Stacked Regression Results while Handling Farmers in Different Ways

Dependent variable: Target occupation in 1940	Baseline occupations (no farm owners or farm laborers)	Baseline occupations plus farm owners & laborers	No farm owners	No farm laborers	Only farm owners	Only farm laborers	Only farm owners & farm laborers
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Own household	0.0381***	0.0408***	0.0341***	0.0454***	0.1028***	0.0105***	0.0498***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0007)	(0.0004)	(0.0004)
Next door neighbor	0.0034***	0.0042***	0.0031***	0.0045***	0.0112***	0.0013***	0.0073***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0008)	(0.0003)	(0.0003)
R-squared	0.323	0.368	0.315	0.376	0.335	0.250	0.376
Untreated mean	0.0331	0.0346	0.0336	0.0342	0.0690	0.0402	0.0484
Observations	26,649,908	33,117,306	29,644,924	30,122,288	3,472,382	2,995,017	6,467,399

Notes: Column 1 excludes farm owners and laborers, Column 3 excludes farm owners, Column 4 excludes laborers, Columns 5-6 uses only farm owner and/or farm laborers.
*p<0.1, **p<0.05, ***p<0.01.

Table C2: Doctor Results in 1920, 1930, and 1940 Censuses

Dependent variable: Doctor in…	1920	1930	1940
	(1)	(2)	(3)
Own household	0.0164***	0.0852***	0.0970***
	(0.0007)	(0.0016)	(0.0018)
Next door neighbor	0.0008***	0.0025***	0.0032***
	(0.0003)	(0.0006)	(0.0007)
R-squared	0.266	0.321	0.327
Untreated mean	0.0014	0.0065	0.0078
Observations	410,385	323,151	305,059

Notes: Columns 1-3 uses the outcome of being a doctor in 1920-1940. Sample excludes sheets with no doctors in 1910.
*p<0.1, **p<0.05, ***p<0.01.

Table C3: Stacked Regression Results - Alternative Outcome Years

Dependent variable: Target occupation in…	1910 kids linked to at least one census			1910 kids linked to all three censuses		
	1920	1930	1940	1920	1930	1940
	(1)	(2)	(3)	(4)	(5)	(6)
Panel A: Boys						
Own household	0.0365***	0.0450***	0.0381***	0.0375***	0.0471***	0.0410***
	(0.0001)	(0.0002)	(0.0002)	0.0002	0.0002	0.0002
Next door neighbor	0.0032***	0.0037***	0.0034***	0.0032***	0.0036***	0.0034***
	(0.0001)	(0.0001)	(0.0001)	0.0002	0.0002	0.0002
R-squared	0.318	0.324	0.323	0.405	0.385	0.377
Untreated mean	0.0262	0.0329	0.0331	0.0257	0.0327	0.0330
Observations	37,437,872	28,531,044	26,649,908	15,473,431	15,473,431	15,473,431
Panel B: Girls						
Own household	0.0168***	0.0107***	0.0085***	0.0140***	0.0090***	0.0080***
	(0.0001)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Next door neighbor	0.0016***	0.0011***	0.0010***	0.0011***	0.0011***	0.0009***
	(0.0001)	(0.0001)	(0.0001)	(0.0002)	(0.0002)	(0.0002)
R-squared	0.338	0.351	0.356	0.426	0.395	0.393
Untreated mean	0.0255	0.0180	0.0157	0.0200	0.0146	0.0149
Observations	23,530,340	13,458,635	11,165,884	6,885,595	6,885,595	6,885,595

Notes: Table reports coefficient from the stack regressions where outcomes are indicators of the child reporting the target occupation in the census year listed at the top of the column. The sample in Columns 1-3 are kids that are linked to the corresponding outcome census, and Columns 4-6 are kids that are linked from 1910 to 1920, 1930, and 1940. Panel A sample is boys exposed to the 50 largest occupations for men in 1910, and Panel B is girls exposed to the biggest 25 occupations for women in 1910. All regressions include sheet and age fixed effects.

*p<0.1, **p<0.05, ***p<0.01.

Table C4: Logit Specification

Dependent variable: Doctor occupation in 1940		At least one doctor occupation per sheet		
	All sheets	No geographic FE	City - enumeration district FE	Sheet FE
	(1)	(2)	(3)	(4)
Own household	1.030***	0.6220***	0.6623***	0.9209***
	(0.0022)	(0.0024)	(0.0028)	(0.0039)
Next door neighbor	0.4332***	0.0486***	0.0360***	0.0851***
	(0.0020)	(0.0023)	(0.0027)	(0.0037)
R-squared	0.16513	0.01645	0.04106	0.09262
Observations	316,783,000	4,410,177	4,410,032	4,410,177

Notes: Column 1 includes all sheets in 1910 census, while columns 2-4 include only sheets with at least one adult in the target occupation. All columns include age fixed effects.
*p<0.1, **p<0.05, ***p<0.01.

Table C5: Doctor Results: Various Household Controls

Dependent variable: Doctor occupation in 1940	Baseline	One doctor per sheet	Drop group quarters	Include sheet position fixed effects
	(1)	(2)	(3)	(4)
Own household	0.0970***	0.0912***	0.0979***	0.0970***
	(0.0018)	(0.0020)	(0.0018)	(0.0018)
Next door neighbor	0.0032***	0.0024***	0.0031***	0.0032***
	(0.0007)	(0.0007)	(0.0007)	(0.0007)
R-squared	0.327	0.319	0.327	0.327
Untreated mean	0.008	0.008	0.008	0.008
Observations	305,059	268,500	300,416	305,059

Notes: Column 1 reports the baseline doctor regression for children with at least one adult doctor on the sheet, and including sheet fixed effects. Column 2 restricts to children on a sheet with a single adult doctor. Column 3 drops children living in "group quarters". Column 4 adds a fixed effect for the line the household appears on. Columns 5-6 sequentially drop the last sheets of the enumeration district which were often labeled as "supplementary" and included households that were missed during the initial enumeration and are therefore out of order. All columns include sheet and age fixed effects.

**p*<0.1, ***p*<0.05, ****p*<0.01.

Table C6: Stacked Regression Results: Robustness

Dependent variable: Target occupation in 1940	Baseline	One target occ. per street	Drop Group Quarters	Sheet Position Fixed Effects	Excluding Occupations Likely to be on the Last Two Sheets
	(1)	(2)	(3)	(4)	(5)
Own household	0.0381***	0.0325***	0.0385***	0.0381***	0.0389***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Next door neighbor	0.0034***	0.0018***	0.0033***	0.0033***	0.0028***
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
R-squared	0.323	0.301	0.323	0.323	0.327
Untreated mean	0.0331	0.0264	0.0332	0.0331	0.0304
Observations	26,649,908	17,838,972	26,328,032	26,649,908	21,829,224

Notes: Column 1 shows baseline stack regression including sheet by occupation fixed effects. Column 2 includes only sheets that have one worker in the target occupation. Column 3 drops children living in group quarters. Column 4 includes position on the sheet fixed effects. Column 5 drops occupations that are most likely to appear on last two sheets, such as laborers.

**p*<0.1, ***p*<0.05, ****p*<0.01.

Table C7: Heterogeneity by Additional Neighborhood Characteristics

Dependent variable: Target occupation in 1940	High non-white share	Low non-white share	University in county	No university in county	Northeast	Midwest	South	West
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Own household	0.0396***	0.0370***	0.0360***	0.0402***	0.0361***	0.0385***	0.0425***	0.0338***
	(0.0003)	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0003)	(0.0004)	(0.0006)
Next door neighbor	0.0038***	0.0030***	0.0029***	0.0038***	0.0027***	0.0032***	0.0049***	0.0025***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0002)	(0.0003)	(0.0004)
R-squared	0.336	0.312	0.328	0.319	0.330	0.308	0.337	0.315
Untreated mean	0.0338	0.0327	0.0324	0.0340	0.0335	0.0318	0.0365	0.0296
Observations	11,676,940	14,972,663	13,604,749	13,044,706	9,198,224	9,743,871	5,464,572	2,243,240

Notes: Each column reports results within different subsets of the census sample based on neighborhood or characteristics or region.

**p*<0.1, ***p*<0.05, ****p*<0.01.

Table C8: Heterogeneity by Neighbor Similarity for Last Names that Are Unique to Each Sheet

Dependent variable: Target occupation in 1940	Baseline	Last name	Income score	Education score
	(1)	(2)	(3)	(4)
Own household	0.0381***	0.0381***	0.0376***	0.0375***
	(0.0002)	(0.0002)	(0.0002)	(0.0002)
Next door neighbor	0.0034***			
	(0.0001)			
Treatment 1: Same characteristic		0.0220***	0.0059***	0.0076***
		(0.0010)	(0.0002)	(0.0002)
Treatment 1 average		*0.0041*	*0.1300*	*0.1334*
Treatment 2: Different characteristic		0.0030***	-0.0012***	-0.0038***
		(0.0001)	(0.0002)	(0.0002)
Treatment 2 average		*0.2328*	*0.0848*	*0.0814*
F-test		325.74	849.42	2223.73
Two-tailed p-value		0.0000	0.0000	0.0000
R-squared	0.323	0.323	0.323	0.323
Untreated mean	0.0331	0.0331	0.0338	0.0338
Observations	26,649,908	26,649,128	26,649,908	26,649,908

Notes: In columns 2-5, treatments compare head of household of the focal child to the target occupation holder next door. In columns 3-4 the treatments evaluate if both head of household of focal kid and target occupation next door are above/below median in occscore and edscore. The number of observations in column 2 is slightly less due to the fact that individuals whose last name is missing and whose neigbhors' last names are also missing are dropped from the regression.
*p<0.1, **p<0.05, ***p<0.01.

Table C9: Heterogeneity by Occupation Income and Education – Girls

Dependent variable: Target occupation in 1940	Baseline	High Income	High Education	High income: High education	High income: Low education	Low income: High education	Low income: Low education
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Own household	0.0085***	0.0045***	0.0114***	0.0045***	0.00000	0.0162***	0.0057***
	0.0002	0.0003	0.0003	0.0003	0.0000	0.0004	0.0002
Next door neighbor	0.0010***	0.0002	0.0012***	0.0002	0.00000	0.0019***	0.0008***
	0.0001	0.0002	0.0002	0.0002	0.0000	0.0003	0.0002
R-squared	0.356	0.316	0.339	0.316	0.000	0.342	0.380
Untreated mean	0.0156	0.0098	0.0204	0.0098	0.0000	0.0250	0.0109
Observations	11,135,893	1,822,587	5,461,292	1,822,587	0	3,638,705	5,674,601

Notes: Each column reports results within groups of occupations based on whether the occupation is above or below the median income and education score for female occupations.
*p<0.1, **p<0.05, ***p<0.01.

Table C10: Heterogeneity by Occupation Income and Education – Boys

Dependent variable: Target occupation in 1940	Baseline	High Income	High Education	High income: High education	High income: Low education	Low income: High education	Low income: Low education
	(1)	(2)	(3)	(4)	(5)	(6)	(7)
Own household	0.0381***	0.0537***	0.0439***	0.0609***	0.0396***	0.0242***	0.0334***
	(0.0002)	(0.0003)	(0.0003)	(0.0004)	(0.0004)	(0.0004)	(0.0002)
Next door neighbor	0.0034***	0.0040***	0.0037***	0.0053***	0.0014***	0.0017***	0.0036***
	(0.0001)	(0.0002)	(0.0002)	(0.0003)	(0.0002)	(0.0003)	(0.0002)
R-squared	0.323	0.337	0.323	0.334	0.290	0.301	0.322
Untreated mean	0.0331	0.0339	0.0421	0.0485	0.0103	0.0358	0.0314
Observations	26,649,908	8,481,845	10,662,704	5,356,209	3,125,636	5,306,495	12,861,567

Notes: Each column reports results within groups of occupations based on whether the occupation is above or below the median income and education score.

**p*<0.1, ***p*<0.05, ****p*<0.01.

Table C11: Additional Heterogeneity by Occupation Type

Dependent variable: Target occupation in 1940	Baseline	High self-employment	High Apprentice Share	Growing Occupations	Shrinking Occupations
	(1)	(2)	(3)	(4)	(5)
Own household	0.0381***	0.0571***	0.0509***	0.0343***	0.0426***
	(0.0002)	(0.0003)	(0.0003)	(0.0002)	(0.0002)
Next door neighbor	0.0034***	0.0036***	0.0018***	0.0031***	0.0036***
	(0.0001)	(0.0002)	(0.0002)	(0.0002)	(0.0002)
R-squared	0.323	0.339	0.286	0.323	0.321
Untreated mean	0.0331	0.0276	0.0140	0.0405	0.0243
Observations	26,649,908	10,839,995	4,852,184	14,477,861	12,172,046

Notes: Column 1 includes all occupations. Columns 2-3 include occupations with high self-employment and high apprenticeship share respectively. Columns 4 and 5 split occupations by the percent change in employment between 1910 and 1940 and categorize them as growing or shrinking.

**p*<0.1, ***p*<0.05, ****p*<0.01.

D Results Using Street Addresses

As discussed in Section 2, our indicator for living next door to a doctor may be subject to misclassification error in some cases, potentially biasing the OLS estimates. Misclassification could be due to skips or other inconsistencies in the enumeration process. Unlike continuous regressors, misclassification of binary regressors is necessarily non-classical. If the dummy is one (zero), measurement error can only be negative (positive), meaning that the error is mechanically negatively correlated with the true dependent variable. However, as Pischke (2007) notes, it is possible to recover the magnitude of bias by using the extent of misclassification.

Following the notation in Pischke (2007), consider the estimated $\hat{\beta}$ coefficient associated with some treatment (d_i) such as living next door to a doctor. If we only observe a version of the treatment ($\tilde{d}_i$) that misclassifies some observations then

$$plim \hat{\beta} = \beta \left[P(d_i = 1 | \tilde{d}_i = 1) - P(d_i = 1 | \tilde{d}_i = 0) \right]$$

We are able to estimate these probabilities using the subsample of doctors that we hand checked against the Sanborn Fire Insurance maps. In this sample, we find that among the people we classify as neighbors of doctors ($\tilde{d}_i = 1$) 59.3 percent actually live next door to doctors based on the address and map information. Among people we classify as living more than one door away from doctors ($\tilde{d}_i = 0$), 2.12 percent actually live next door to doctors. Based on these measures, our estimates of the effect of living next door to a doctor are biased downward by 57.2 percent. The causal effect is about 1.75 times larger than our estimate for doctors.[32]

One approach to dealing with misspecification error is to instrument our indicator for being next door to a doctor on the census sheet with an indicator for being next door using

[32] Across each of the different occupations, as long as $P(d_i = 1|\tilde{d}_i = 1) > P(d_i = 1|\tilde{d}_i = 0)$ our point estimates will be biased downward.

the street addresses. But because our treatment of interest is binary and the associated measurement error is non-classical, this commonly used IV approach to correct classical measurement error does not apply in our setting. The IV estimate will, however, still be useful to compute in a bounding exercise, as recommended by Pischke (2007) and Aigner (1973). For a subset of census sheets, enumerators list street names and numbers. We can use these street addresses to construct an alternative measure of proximity.

Next door indicators defined under this definition are likely also measured with error, but for reasons that we argue are likely to be orthogonal to the error in the sheet-ordered dummies (e.g. transcription errors of house numbers and street names); see the discussion in Appendix A. We exploit this fact in Table D1. In Column 1, we report the baseline estimates from Column 4 in Table 2 and compare them to Column 2, where we restrict to households that also have recorded street addresses. Both the coefficients for having a doctor in the focal child's own household and next door are slightly larger than the baseline estimates. In Column 3, we use the street addresses to measure proximity. We consider a household to be next door to the focal household if it has the closest house number (either larger or smaller) on the same street.[33] Instead of sheet-based fixed effects, we include street name-by-enumeration district fixed effects.

We find results that are qualitatively similar to our results in Column 1, although coefficients for both own household and next door are slightly smaller in magnitude. When using street addresses, growing up next door to a doctor makes a child 0.28 percentage points, or about 46%, more likely to become a doctor. Since the measurement error in both of our measures of household's proximity are believed to be orthogonal, in Column 4 we use the measure of NextDoorDoc1910_i using recorded street addresses as an instrument for NextDoorDoc1910_i using census sheet order. The coefficient on NextDoorDoc1910_i is 2.4

[33] Bayer et al. (2022) use a similar method to identify next door neighbors using modern data on parcels from CoreLogic. One of the reasons we use the sheet ordering for our preferred specifications is because addresses are not recorded for about half of the 1910 census observations. To ensure that we are capturing neighbors with the street address ordering, we restrict the sample to enumeration districts where at least 80 percent of households have address recorded. This allows us to focus on neighborhoods where the enumerator actually recorded addresses and missing values are due to transcription or digitization error.

Table D1: Doctor Results Using Street Addresses

Dependent variable: Doctor in 1940	Sheet ordering: Baseline (1)	Sheet ordering: Street sample (2)	Street ordering (3)	Street order IV (4)
Own household	0.0970***	0.1269***	0.1176***	0.1278***
	0.0018	0.0045	0.0040	0.0045
Next door neighbor	0.0032***	0.0036*	0.0028**	0.0087**
	0.0007	0.0019	0.0014	0.0043
R-squared	0.283	0.369	0.200	0.012
Untreated mean	0.0038	0.0061	0.0061	0.0061
Observations	6,044,153	1,053,150	1,053,150	1,053,150

Notes: Columns 1-2 use the ordering of census sheets to define neighborhood proximity. Columns 3 uses street ordering. Column 4 is an instrumental variables regression with street order variables instrumenting for sheet order variables.

$^{*}p<0.1$, $^{**}p<0.05$, $^{***}p<0.01$.

times larger than in our estimates using only the census sheet order in Column 2 and 3.1 times larger than in our estimates using only the street address in Column 3. The OLS and IV estimates bound our estimated effect of growing up next to a doctor on the probability of becoming a doctor between 0.32 percentage points and 0.87 percentage points. Our measurement error corrections place the estimate squarely between these estimates (0.56 percentage points).

Table D2: Stacked Regression Results Using Street Addresses

	Sheet Ordering			Instrument	
	Baseline	Street sample	Street ordering	First stage	Second stage
Dependent variable: Target occupation in 1940	(1)	(2)	(3)	(4)	(5)
Own household	0.0410***	0.0372***	0.0371***	-0.1207***	0.0379***
	(0.0002)	(0.0003)	(0.0003)	(0.0007)	(0.0003)
Next door neighbor	0.0042***	0.0026***	0.0026***	0.4067***	0.0066***
	(0.0001)	(0.0002)	(0.0002)	(0.0007)	(0.0005)
R-squared	0.332	0.371	0.269	0.663	0.371
Untreated mean	0.0104	0.0110	0.0109	0.0155	0.0121
Observations	314,295,968	56,042,752	56,042,752	56,042,752	56,042,752

Notes: Columns 1-2 use the ordering of census sheets to define neighborhood proximity. Column 3 uses street ordering. Column 4 and 5 report the first and second stage of an instrumental variables regression with street order variables instrumenting for sheet order variables.
*p<0.1, **p<0.05, ***p<0.01.