The Good Place: How Networks, Preferences, and Public Policy Determine the Value of Where We Live

Allison Green

A Dissertation
Presented to the Faculty
of Princeton University
in Candidacy for the Degree
of Doctor of Philosophy

Recommended for Acceptance by
the Department of
Economics
Advisers: Leah Boustan and Ilyana Kuziemko

May 2025

Abstract

I study how networks, social preferences, and public policy shape residential choice and its economic consequences. My dissertation uses novel data sources and quasi-experimental research designs to establish causal relationships between social factors and geographic mobility.

In Chapter 1, I examine how networks formed during World War II Navy service impacted labor market choices over a person's lifetime. Using a newly constructed dataset of over 1.4 million sailors and exploiting conditional random assignment to ships, I identify the causal impact of shipmates' geographic diversity on post-war migration. Results reveal that exposure to out-of-state shipmates increased out-migration by 4-5% by 1950, with migration primarily directed toward high-growth regions where these men experienced higher earnings.

In Chapter 2, with Victor Couture, Jonathan Dingel, and Jessie Handbury, I investigate whether demographic preferences explain income segregation in shared spaces. Using smartphone movement data linked to building-level residential demographics, we measure exposure to high-income co-patrons across demographic groups. We find that racial homophily and preferences for high-income individuals are economically significant and explain substantial cross-group variation in exposure to high-income individuals in shared commercial spaces.

In Chapter 3, with Kaan Cankat, I examine how municipal boundary expansions affected local public finance and potentially contributed to the rise of the Sun Belt. Using newly digitized annexation data from 1948 to 1968, we find the average expansion increased municipal population by 24%. Our difference-in-differences analysis reveals persistent declines in per capita expenditures and employment, suggesting cities successfully leveraged economies of scale in high fixed-cost services.

Acknowledgements

I first want to thank my advisors Leah Boustan and Ilyana Kuziemko for the invaluable support and advice you gave me throughout the PhD. You are both models of the economist I aspire to become—brilliant in your work, generous with your time, and outward-looking with your perspective. I also want to thank Steve Redding for his great advice that expanded the scope of my research. I also want to thank the rest of the Industrial Relations section for your feedback over the years. In particular, I want to thank Hank Farber, Ellora Derenoncourt, and Owen Zidar. I also thank the section and all of its staff for generous financial and administrative support.

I want to thank my collaborators Victor Couture, Jonathan Dingel, Jessie Handbury, and Kaan Cankat. Victor, Jonathan, and Jessie, while working with you I have grown substantially as a researcher and I treasure the valuable lessons on what it is to do good research from each of you. Kaan, in you I have found the intellectual and creative partnership that I have always wanted, and I look forward to years of future collaboration (and friendship).

Thank you to my friends in Princeton. My time at Princeton has been the happiest time of my life, largely due to the people I met along the way. I want to thank all of my communities who made this time so joyful—my cohort for making Princeton fun from the start, the EEB department for being my second home, and DBar for being a source of fun and creativity. For making my time special: Chris for late night heart to hearts, Henry for mutual growth, Sebastian for countless movie nights, Carol for always being down, and Sameera for years of fun and adventure. I also want to give a special thanks to my friends from outside Princeton—Deborah, Mariah, Kanako, Dylan, Emilyanne, Lauren, and Dan—I have cherished that our friendship stayed strong throughout even at a distance.

I want to thank my family. I love you dearly and without your encouragement and belief I never would have thought that I could get a PhD. I want to thank my parents for always telling me I could do whatever I set my mind to. I want to thank my siblings Dave, Erin, and Molly, who always gave me someone to look up to. I want to thank my cousins Lauren and Sara who have been the coolest people I have known since birth.

Finally, I want to thank my family I built in Princeton. To my cats, Margot and Rita, I love you the most. And finally to Christian, I feel infinitely lucky to have met you, and I am so excited to do the rest of life with you.

To my parents

Contents

List of Tables

List of Figures

Chapter 1

Networks and Geographic Mobility: Evidence from World War II Navy Ships*

1.1 Introduction

A growing body of research has established a link between economic opportunity and where people live (23). Yet, geographic mobility in the United States is low: 80% of adults under the age of 26 live within 100 miles of where they grew up (48). One potential explanation for limited geographic mobility, despite evidence of significant economic returns from migration, is the prohibitive cost of moving without an established network (61; 13). Most Americans lack out-of-state connections, with geographically broad networks typically concentrated among higher-income and more educated individuals (24).

Using one of the largest natural experiments in network formation in US history–Navy service during World War II, this paper assesses whether expanding the geographic scope of networks can increase migration and expand economic opportunity. World War II provides an attractive empirical setting for exploring the role of networks in shaping migration patterns. The war exposed a substantial portion of young American men to individuals from outside their immediate communities, often for the first time. Nearly 80% of all white men born between 1920 and 1926 served

*Materials from this chapter were presented at the Spring 2025 Urban Economics Association conference in Berlin and at the Summer 2024 NBER Summer Institute Development of the American Economy.

in the war in some capacity.[1] Given this scale, the networks formed through wartime service may have had profound effects on mid-20th-century America, fostering exposure to new ideas, diverse backgrounds, and geographically distant locations.

In this paper, I measure the impact of wartime networks on the migration decisions of rank-and-file sailors serving on Navy ships during World War II. This population was young, usually native-born, and without post-secondary education.[2] In part due to low migration during the Great Depression, the vast majority were living close to where they were born prior to entering naval service.[3] These characteristics make rank-and-file Navy sailors well-suited for studying the potential benefits of expanded geographic diversity in personal networks, as pre-existing research has shown migration gains are typically largest for young movers (22; 62).

To estimate the impact of ship networks on migration, I use quasi-random assignment to Navy ships. According to contemporaneous institutional records, assignments to ships were largely random. However, due to operational constraints in transporting personnel, men were more likely to serve on ships with others from similar geographic origins. The empirical strategy addresses this source of endogeneity by comparing migration outcomes for individuals from the same state who, due to the random chance of ship assignment, were exposed to different sets of peers during their service. This natural experiment, replicated across a large sample of sailors, enables analysis of how newly-formed connections shaped post-war geographic mobility.

I construct a novel dataset from archival records on the near-universe of ships active during the war. By processing 6.5 million scans of Navy personnel documents, I create a dataset that includes detailed information on over 1.4 million sailors across more than 5,000 ships. For each sailor, I document their precise dates of service and pre-war state of residence. This information allows me to identify not only which sailors served together on specific ships at the same time, but also the geographic diversity of shipmates that each individual was exposed to during service. I then

[1] The military was racially segregated throughout World War II. In particular, the Navy was segregated by unit and occupation. This meant that except for a few small cases, Black men and other men of color in the Navy were not serving in combat roles on ships and are thus excluded from the scope of this paper.

[2] Men with college degrees and men enrolled in college were almost always assigned to officer positions.

[3] Figure 1.10 shows that less than 15% of men born in the 1920s (largest World War II cohort) moved across state lines during childhood.

link individuals from Navy ships to migration outcomes five years after the war (1950 Full Count Census) and at the time of their death (Numident Social Security).

Using this dataset and empirical framework, I estimate the impact of these wartime networks on two key dimensions of migration: out-migration and directed migration. Out-migration measures movement away from one's own pre-war residence, capturing overall geographic mobility. Directed migration examines movement toward the pre-war residences of fellow shipmates, reflecting the influence of specific network ties. Together, these two margins reveal both the extent to which direct ties formed during service influenced specific migration destinations and the aggregate network effect of shipmates on general mobility.

I find that shipmates significantly influence both whether a Navy sailor moves (out-migration) and where they move (directed migration). By 1950, a one standard deviation increase in exposure to out-of-state shipmates raises the likelihood of moving out-of-state by 4.2%. The effect is larger for long-distance moves: a one standard deviation increase in exposure to out-of-region shipmates increases the likelihood of moving out-of-region by 6.5%. The impact of shipmates on directed migration is even more pronounced, though heterogeneous across destinations. A one standard deviation increase in exposure to shipmates from fast-growing states in the Pacific increased migration to those areas by over 15% by 1950. Notably, Navy networks had little impact on migration to slower-growing regions like the Midwest, suggesting these networks primarily facilitated moves to areas of high economic opportunity.

To fully leverage variation in ship network composition, I extend the analysis by developing a discrete choice migration model that incorporates network ties formed through Navy service. By embedding network effects into a discrete choice model, I evaluate Navy ties to a particular state within the broader context of all available network connections on a ship. The model includes heterogeneous network effects between states, capturing various channels through which networks influence migration. Notably, the model allows the value of additional network ties to vary based on the pre-existing density of networks in each destination. Using random variation from Navy assignments to estimate the model, I compute counterfactuals that decompose the role of Navy networks in explaining observed migration patterns.

The counterfactual exercises quantify the heterogeneous effects of networks formed during World War II on post-war interstate migration patterns. On average, a 10 percentage point increase in exposure to shipmates from California, the fastest growing state in the 1940s, increases migration to California by 23%. However, effects by origin state range from 7%-37%, with the largest increase observed for people from the East coast who are unlikely to have pre-existing ties to Western states. Examining total migration effects, individuals from the Midwest were proportionally most responsive to these wartime networks, with their moves primarily directed toward high-growth states like Florida, Texas, and California. A back-of-the-envelope calculation suggests that over 14% of migration to California between 1945 and 1950 can be attributed to exposure through wartime networks. Collectively, these counterfactual exercises indicate that networks formed during military service were a major force in driving migration to high-growth areas.

I then explore how network formation on ships, particularly through shared ethnicity between shipmates, influences post-war migration. Shared characteristics between sailors may strengthen social bonds and therefore amplify the impact of shipmate connections on migration decisions. By 1950, a higher share of co-ethnic shipmates increases both the overall likelihood of migrating and the probability of moving to specific states where these co-ethnics are from, with increased exposure to co-ethnic shipmates being about 2.5 times more influential than shipmates of a different ethnicity. However, the influence of co-ethnic largely disappears in the long-run, suggesting that the salience of ethnic connections diminishes over time.

The final section of the paper examines whether moves facilitated by Navy networks led to greater economic opportunity. Using random variation in networks and the estimated discrete choice model, I construct instruments for the probability a Navy sailor moves out-of-state, out-of-region, and to a state in the Pacific Census division. I then estimate the returns to earnings for men induced to move due to their wartime networks using a two-stage least squares estimation procedure. These estimates can be interpreted as the local average treatment effect for compliers–those whose migration decision was marginal to the influence of Navy networks. I find economically large impacts of network-induced moves on lifetime earnings: individuals induced to move out-of-state by their Navy networks reside in zip codes with 59% higher income by the time of their death. While substantial, these estimates are credible within the context of the literature on returns to migration. The findings suggest that networked migration facilitated by Navy service allowed this population of men to move to areas of greater economic opportunity.

This paper makes several contributions to existing literature. First, this work advances understanding of networks' role in shaping migration decisions. Theoretical models predict that networks increase migration rates by lowering migration costs (21) and providing destination-specific information (64; 63; 42). Empirical evidence confirms that origin-destination links increase international migration rates (50; 33; 55; 60). In American settings, (71) employ birth town as a proxy for migration networks in 20th-century migration events, and (30) demonstrate that Civil War veterans from the same companies are more likely to live in close proximity decades after the war. Recent papers have improved measurement of social networks using phone data and Facebook data to show links between where people's connections live and their own location choices (52; 13; 67; 19). For example, (13) utilize phone data in Rwanda to show that individuals are more likely to migrate to areas where they have stronger network ties, while (52) measure networks using Facebook data and employs differences in the timing of friends' moves to show that recent college graduates are more likely to move to Commuting Zones where network ties are located.

Although these studies establish the importance of networks in migration decisions, they typically take individuals' personal networks as fixed and study variation in where network members are located, rather than examining changes in network composition itself. By contrast, I leverage quasi-random assignment to Navy ships during World War II to examine how forming new network ties to people from other places impacts migration decisions. By studying the impact of random exposure to new people through wartime connections, I provide unique evidence on a policy-relevant margin: does expanding access to broader geographic networks increase geographic mobility?

Second, this paper contributes to the literature on returns to migration by introducing a novel approach to estimating these returns. Previous studies have primarily relied on three methods. The first uses displacement events as natural experiments, with several studies finding positive earnings effects from forced relocations - whether from natural disasters (62; 31), wartime displacement (68), or government internment (8). The second approach controls for family background by comparing migrant and non-migrant siblings. This method has been particularly influential in studying the Great Migration, where researchers identified substantial returns for both Black and white men who left the South (26; 15). Related work using sibling comparisons has found moderate positive effects of internal migration on early career earnings more broadly (73). This paper extends this literature by employing random variation in network exposure as an instrument for migration, offering a novel method to estimate the causal effects of migration on economic outcomes. Consistent with studies that find the largest returns to migration for younger movers (62; 22), the analysis focuses on rank-and-file Navy personnel who were typically in their early twenties during service.

The last approach to studying returns to migration uses evidence from policies designed to subsidize migration. The most prominent example is the Moving to Opportunity (MTO) experiment, which subsidized moves from high- to low-poverty neighborhoods (22; 53; 25; 51). While (22) find positive effects for children who moved young, the program showed no earnings gains for the average beneficiary. (10) find housing lottery winners in India experienced little economic gains and greater social isolation after moving. These moving subsidy programs consistently face low take-up rates, which (12) attribute partly to missing destination networks. These papers suggest network barriers may limit the effectiveness of subsidized migration programs by reducing both take-up rates and participants' ability to access economic opportunities in new locations. The analysis of Navy networks facilitating moves to high-opportunity areas provides new evidence on mechanisms that shape the impact of geographic mobility programs.

Third, my paper extends the literature on peer effects and networks, as summarized by (16), with a particular focus on network formation in early adulthood through randomized peer interactions. Research in college and military settings is especially relevant. In college contexts, (66) examines the effects of randomly assigned dorm roommates, and (56) analyzes the impact of dorm assignments at Harvard in the 1920s, and (69) studies networks formed among Harvard MBA students. In military environments, (34) demonstrates that Finnish soldiers in military dorms earn higher incomes when exposed to higher-income peers. Conversely, (20) find negative outcomes from an optimal peer mixing experiment at the Air Force Academy, attributed to endogenous peer formation. (29) observe higher desertion rates in more heterogeneous Civil War battalions, and (44) show that increased exposure to West Point cadets from Northern states makes a person more likely to join the Union over the Confederacy. These studies generally yield mixed results on the benefits of peer mixing, with effects often attenuated when cross-group differences are prominent. This paper advances this literature by examining the impact of randomized peer exposure during World War II and investigating how factors such as ethnicity and geographic origin influence network formation and mediate treatment effects in this unique historical context. This paper is also one of few papers that focuses on the role of randomized peer exposure on geographic mobility.

Finally, this paper contributes to the literature on World War II's impact on the U.S. post-war economy. This research primarily encompasses two main strands: (1) the effects of war mobilization and demobilization on aggregate economic activity (38; 39; 58; 41), and (2) the war's influence on human capital accumulation and wages (14; 27; 6; 11; 4; 5). This study distinguishes itself by being among the first to utilize ship-level variation in studying World War II, building on Suandi (72) use of submarine promotion data. By leveraging granular ship-level data to measure the effect of wartime networks on subsequent economic outcomes, I provide novel insights into the long-term consequences of wartime social connections.

The paper proceeds as follows: Section 1.2 describes aggregate trends in geographic mobility and military service in relation to World War II. Section 1.3 describes the construction of the core World War II Navy dataset. Section 1.4 describes the empirical strategy and presents results on the causal impact of Navy networks on migration. Section 1.5 presents the discrete choice model and conducts counterfactual exercises to quantify the impact of wartime connections on aggregate migration patterns. Section 1.6 explains the role of ethnic ties on network formation aboard Navy ships. Section 1.7 discusses the construction of a migration instrument based on random variation in Navy networks, and uses the instrument to estimate the returns to network-facilitated migration. Section 1.8 concludes.

1.2 Historical Background

1.2.1 Relationship between migration and military service

Geographic mobility for young, white American men peaked in the mid-20th century, coinciding with high rates of military participation during World War II.[4]

Figure 1.1 illustrates this trend, showing the share of men in their 30s living outside their state of birth, controlling for childhood moves.[5] The figure reveals a sharp rise in young-adult migration (ages 20-40) in the first half of the 20th century, peaking for those born in the 1920s—the cohort with the highest level of World War II service. Cross-state prime-age geographic mobility has slowly declined since. Then, overlaying military participation rates for 20th-century birth cohorts shows a striking correspondence between geographic mobility and military service rates.

Military participation remained high between World War II and Vietnam due to the draft (1940-1973). Since the end of conscription, the share of white men joining the military has decreased steadily, with modern enlistment rates below 10% for young white men. The tight relationship between military participation and aggregate geographic mobility is supported by research showing that veterans are more geographically mobile than non-veterans, even when controlling for selection into service (9).

This peak in mobility for the World War II cohort aligns with broader trends in U.S. internal migration throughout the 20th century. (65) document that interstate migration rates reached their highest levels in the mid-20th century, up from their lowest point around the turn of the 20th century. The elevated level of mobility persisted for several decades after the war but has been

[4]See (65) for a discussion of aggregate cross-state migration trends. Historical trends in long distance geographic mobility differ significantly for Black men compared to white men (45).

[5]The figure shows the change in probability of living outside state of birth from ages 10-19 to 30-39.

declining over the past 50 years (57). Research by (48) on contemporary migration trends indicates that 65% of young Americans now live in the same commuting zone where they grew up, and 80% live within 100 miles of their childhood home.

World War II catalyzed large-scale migration flows throughout the United States, extending beyond those directly involved in military service (46). War production centers attracted significant numbers of workers deemed ineligible for military service, including women and men with occupational or physical exemptions (49). The South, particularly agricultural areas, experienced substantial out-migration as individuals relocated to western states with high concentrations of defense production. By spring 1943, the influx of migrant workers to California's defense industry was so substantial that the San Francisco Chronicle dubbed it "The Second Gold Rush" (49). However, many of these wartime migrants ultimately returned to their origin communities after the war ended (46).

Contemporaneous reports indicate that World War II veterans' were highly geographically mobile after the war. According to a 1947 Census report, only 30% of World War II veterans maintained their 1940 residential address by 1947 (46). Former farm residents showed the highest propensity to relocate between 1940 and 1947 though most of these moves occurred within their original county (46).

Several mechanisms may explain the link between military service and increased geographic mobility. First, World War II veterans were entitled to the benefits of the Service Readjustment Act of 1944, more commonly known as the GI Bill. The GI Bill provided veterans access to low-rate mortgages and protection against default, helping accelerate homeownership rates throughout the mid-20th century (37). These benefits coincided with a severe housing shortage in the late 1940s, driven largely by low housing construction during the Depression and war years (47). To address these shortages, particularly for World War II veterans, there were large public and private investments in new housing construction, much of it centered in newly developing suburbs.

Anecdotal evidence suggests social connections formed during military service also played a role in post-war migration patterns. For instance, Major Richard Winters, a member of the Band of Brothers, notes in his memoir how he received a job through one of his commanding officers: "Within two weeks of returning home [from World War II], I accepted Lewis Nixon's invitation to travel to New York City and meet his parents. His father offered me a job and in January 1946, I became personnel manager for the Nixon Nitration Works in Nixon, New Jersey."

Similarly, Walter W. Schumacher, who served aboard the light cruiser USS Omaha, describes in the Veteran History Project how a friendship formed during basic training influenced his decision to move from New York to Toledo, Ohio: "I met [my wife's] brother in the North Atlantic Great Lakes Service School, so I come to Toledo a few times with him [after the war]. And so [me and my wife] got together...in 1946 we got married in New York. She didn't want to leave Toledo, so it didn't make no difference to me [so we moved to Toledo]." These examples demonstrate two distinct mechanisms through which military service shaped migration: network-based job referrals that provided economic opportunities in new locations and social ties that generated non-pecuniary reasons to move through marriage and family connections

1.2.2 US Navy and World War II

The US Navy underwent a dramatic transformation during World War II, evolving from a peacetime force of around 100,000 volunteers in the 1930s to a massive wartime fleet of over 3.5 million personnel.[6] This rapid expansion was set in motion even before the United States officially entered the war. In July 1940, responding to growing global instability, Congress passed the Two-Ocean Navy Act, authorizing a significant increase in naval personnel and funding an extensive shipbuilding program (59).

The buildup of naval forces was closely tied to the implementation of conscription. The first peacetime draft registration in September 1940 required men between the ages of 21 and 45 to register for potential military service. By the time of the Pearl Harbor attack on December 7, 1941,

[6]Source:https://www.history.navy.mil/research/library/online-reading-room/title-list-alphabetically/h/history-of-the-us-navy/personnel-strength-1794-1990.html. Last Accessed: 10/11/24

which precipitated U.S. entry into the war, the Navy had already grown to 325,000 personnel. Over the next four years, this number would increase more than tenfold.

The composition of Navy recruits changed over the course of the war. Initially, the Navy's ranks grew through a combination of the draft and voluntary enlistment. However, in December 1942, voluntary enlistment was suspended.[7] This decision was made to maintain a sufficient agricultural workforce and to alleviate tensions between the Army and Navy, as the Navy's stronger recruiting apparatus had been attracting a disproportionate share of volunteers.

By 1945, the U.S. Navy had reached its largest size in American history with 3.4 million active personnel, of whom approximately 1.5 million were enlisted men serving on ships.[8] who served aboard Navy ships.[9] While Navy personnel were deployed across various theaters, the Pacific was the primary focus, with over 80% of overseas personnel stationed there by August 1945.[10] Despite the rapid expansion and the dangers of wartime service, casualty rates among enlisted personnel were relatively low: out of over 3.5 million who served, 32,925 were killed and 34,478 wounded.

The Navy's fleet expansion mirrored its personnel growth, growing from 478 vessels in 1940 to 6,768 by August 15, 1945 (Victory in Japan day). Navy ships varied significantly in size: 56% were small (under 100 people), 37% medium (150-400 people), and 7% large (over 1000 people). Despite size differences, ships shared a common organizational structure characterized by a vertical hierarchy of officers and enlisted personnel, with rank-and-file making up 70-80% of the crew. Horizontally, ships were organized into functional units like deck, engine room, and mess.

[7]Ended via Executive Order 9279, which closed voluntary enlistment for men aged 18-37. The Navy still maintained an active recruiting arm for enlisting 17-year-olds into voluntary service.

[8]Within Navy terminology, they use the term "enlisted" to refer to all non-officer personnel regardless of whether they volunteer or were drafted.

[9]Other jobs outside of active duty ship service include, domestic shipbuilding and administrative work, overseas shore service, aerial service, etc.

[10]As of August 31, 1945, the Pacific theater (both ashore and afloat) accounted for approximately 82.5% of peak Navy personnel strength across major theaters, with 1,366,716 personnel. The North Atlantic theater had 9.1% (150,046 personnel as of June 30, 1944), the Mediterranean 5.4% (90,175 on August 31, 1944), and continental Europe 3.0% (49,801 on November 30, 1944).Source: https://www.history.navy.mil/research/library/online-reading-room/title-list-alphabetically/u/us-navy-personnel-in-world-war-ii-service-and-casualty-statistics.html. Last Accessed: 10/11/24

The end of the war brought about a massive demobilization of Navy personnel. At the end of the war, most of the rank-and-file personnel were released back into civilian status. Men were returned to a major port or Navy station, and given transportation back to their pre-war residence. Using a sample of around 10,000 individual separation documents, I find that over 90% of all enlisted sailors report intending to return to the same address they were living at prior to the war upon immediately exiting the Navy.[11]

Despite the large-scale demobilization, many veterans maintained ties with the Navy. A substantial portion of men who served in the Navy, even those who were initially drafted, chose to enter the Navy reserves and approximately 25% of men who served in the Korean War also served during World War II.[12]

1.3 Data

To study the long-run impact of network ties formed in the U.S. Navy during World War II, I assemble a dataset that combines detailed information on individuals' service histories with measures of their economic and geographic outcomes after the war.

The primary data source is a newly constructed collection of World War II Navy Muster Rolls, which contain quarterly snapshots of the full roster of enlisted personnel for the universe of Navy ships active during the war. Using these records, I identify the specific ships and time periods of each sailor's service, enabling the construction of comprehensive individual-level shipmate networks. I then link these service records to data from the 1940 and 1950 U.S. Censuses and death records to track geographic mobility patterns over a person's lifetime.

The resulting dataset contains complete World War II Naval service histories and pre- and post-war outcomes for approximately 300,000 individuals. This section details the data construction

[11]Each sailor filled out a separation document (NAVPERS-563) upon exiting the Navy. This document contained address prior to the war, and the address for which the person registered with the selective service board, and the address the person intends to seek employment. I collected over 10,000 separation documents at National Archives Site in St. Louis

[12]Source: https://www.va.gov/vetdata/docs/specialreports/kw2000.pdf

process, including the digitization and cleaning of the Muster Rolls, the record linkage methodology, key variables used in the analysis, and summary statistics of the final sample.

1.3.1 World War II Navy Muster Rolls

The primary data source is a collection of World War II Navy muster rolls, reports submitted quarterly by each U.S. Navy ship to the Bureau of Navy Personnel between 1939 and 1949.[13] For each muster roll, I focus on information from two document types: (1) quarterly censuses of every enlisted sailor on the ship (5-10% of scans), and (2) monthly reports of any personnel changes, such as sailors boarding or leaving the ship, or changes in their rating (20-25% of scans).[14]

I digitize 6.5 million scanned images of Muster Rolls from the National Archives Catalog using optical character recognition (OCR) and LayoutParser.[15] Figure 1.12 Panel A presents examples of the quarterly census and monthly personnel change forms.[16] Since the original documents vary in quality, the digitized data contains encoding errors and missing fields. To address these issues, I use sailors appearing across multiple scans over their service period. By combining information across all records pertaining to a given individual, I am able construct accurate service histories for a much larger sample of individuals than if I relied on individual entries alone. Figure 1.12 Panel B illustrates this cleaning procedure by showing all records pertaining to a particular sailor "Hugh Berry", including OCR errors. The final row displays the corrected service record after combining information across observations. On average, each sailor's service number–a unique seven-digit identifier assigned by the Navy–appears on 8.5 distinct scans, providing ample opportunity to correct errors. Further details on the data cleaning and construction are discussed in Appendix 1.11.1.

[13] All formal "activities" were required to submit these quarterly reports. An activity is defined as "a unit, organization, or installation performing a specific mission or function and established under a commanding officer, officer in charge, etc.; e.g., naval air station, naval shipyard, naval station, a specific air squadron, ship, etc."

[14] The most common personnel changes are individuals boarding the ship ("Received"), exiting the ship ("Transferred"), and changes of rating ("Change of Rating"). Less common types of personnel changes include individuals going AWOL, hospitalizations, and short leaves of absence.

[15] It is important to note that not all 6.5 million images contain usable data; some are cover pages, blank forms, or otherwise irrelevant to my analysis and are thus disregarded.

[16] Each quarterly census scan reports up to 35 sailors, while each monthly report of changes includes up to 15 individuals.

Figure 1.12 Panel C provides summary statistics for the cleaned sample. I identify over 1.4 million unique rank-and-file sailors who served across over 5000 ships between 1941 and 1945. By historical account, around 1.5 million rank-and-file personnel served on Navy ships during the war, so I have fairly complete coverage of the universe of the population of interest. The modal sailor served on only one ship over the course of his service, and over 95% of sailors served on two or fewer vessels. The average person is in my sample for a period of 19 months, with the average stint on a given ship lasting 15 months.

Supplementary Military Data

I supplement muster roll data with additional military records to explore how varying peer interactions and service circumstances might impact outcomes. The *Dictionary of American Naval Fighting Ships* (DANFS), published by the Navy from 1959 to 1991 and later digitized by volunteers, offers detailed historical accounts for each vessel. These accounts include ship specifications, operational timelines, major engagements, casualty reports, and commendations. The basic details on operational dates and ship dimensions serve to validate the cleaned muster roll data. Moreover, the battle and conflict records enable analysis of cross-ship variation in combat intensity, potentially shedding light on how these experiences influenced the strength of bonds formed among crew members. I also use a wartime report from the Naval Health Research Center to measure the share of wounded, missing in action, and killed for each of the 150 distinct categories of ships.

To gain additional insights into the types of ties that a person might form within a ship, I use historical Navy Occupation Ratings data. This data provides information on each rating code (e.g., ”EM3C”), full-title occupation (e.g., ”Electrician’s Mate 3rd Class”), pay grade (e.g., ”III”), and rating branch (e.g., ”Artificer’s Branch”). By using this information, I can explore how the nature of relationships might vary based on occupational proximity and hierarchy. For instance, two people of the same rating are more likely to work together and thus may be more likely to form close ties. Similarly, the relative positions of individuals within the rating hierarchy could have different impacts on the types of ties formed.

Finally, I use archival internal documentation from the Bureau of Navy Personnel on how service numbers were allocated across enlistment centers to identify the place each person enlisted by their service number.

1.3.2 Measuring Characteristics of Shipmates

To analyze the relationship between shipmate characteristics and long-run migration outcomes, I construct networks of shipmates for each individual in the data.

Individual Characteristics

I first construct individual characteristics using data available in the Muster Rolls. These characteristics serve as proxies for geographic, economic, and demographic attributes, which I then use to construct network measures of each sailor's exposure on their ship. By utilizing information only available in the Muster Rolls, I can create network measures for the entire ship network, and not just the subset of individuals I subsequently link to other datasets.

I use place of enlistment as a proxy for pre-war geographic residence. During World War II, over 100 receiving stations were assigned specific service number ranges. For example, sailors enlisting in Buffalo, New York, received numbers between 8052000 and 8066999. I use these ranges to determine enlistment locations for the entire sample.[17]

To impute ethnicity/ancestry, I employ the method developed by (2) using ethnic differentiation among names.[18] I compute indices from white men in the 1940 census, and construct these indices for the sixteen largest ethnicities/nationalities. I supplement this with an index for Jewish names

[17] Appendix Figure 1.11 maps the enlistment locations, demonstrating their wide geographic distribution and supporting the use of place of enlistment as a proxy for pre-war residence. At least one enlistment center was active in each of the 48 contiguous states.

[18] This method calculates the likelihood of a name belonging to one of eight ethnic groups, assigning ethnicity when the score exceeds 0.7. The formula for Italian ethnicity, for example, is:

$$F_{nh} = \frac{\frac{\text{\# of People with last name n and Italian}}{\text{\# of people Italian}}}{\frac{\text{\# of People with last name n and Italian Father}}{\text{\# of people with Italian}} + \frac{\text{\# of People with name n not Italian}}{\text{\# of people not Italian}}}$$

I construct name indices by full name, first name, and last name.

index used in (1). To impute pre-war education, income, and occupation, I use complete names and state of enlistment.[19] For all three economic characteristics, I impute the value belonging to the head of household of an individual with that name. I do so because a substantial portion of Navy sailors during World War II were still living at home in 1940, and it avoids confounding life-cycle effects with differences in economic status.

Panel C of Figure 1.12 summarizes these imputation measures. I successfully identify the state of enlistment for over 97% of the sample, ethnicity for 96%, and occupational score for 89%. Appendix 1.11.1 discusses each imputation method in more detailing and provides additional validation.

Computing Network Measures

Using these imputed characteristics, I then construct individual-level measures of their exposure network to other sailors on Navy ships. The primary focus is on the geographic network of exposure, which captures the extent to which a sailor was exposed to shipmates from different states. I supplement this with additional characteristics of shipmates including their occupational role on the ship and imputed ethnicity and income.

To measure geographic exposure to shipmates, I create a vector of 49 elements for each sailor. The first 48 elements correspond to exposure to people from the 48 contiguous states, while the 49th element represents an “other” category that includes territories and non-American enlistments.[20] Each element in this vector represents the duration-weighted share of shipmates from that particular state or category.[21] For sailors who served on multiple ships, I focus solely on

[19]For instance, a John Smith from Ohio in my sample is assigned the average 1940 occupational score of all John Smiths in Ohio. For men whose complete full name and state do not match to the 1940 Census, I then impute characteristics based on surname and state of enlistment.

[20]Because Alaska and Hawaii were territories until 1959, enlistments in both of these places are grouped into the “’other” category.

[21]Weighted and unweighted measures of ship characteristics have a correlation of 0.83, reflecting that the set of shipmates a person was serving with was relatively stable across time.

their first ship assignment to construct these network measures. This approach reflects that while assignment to the first ship is conditionally random, assignment to subsequent ships is less so.[22]

To complement the geographic exposure measure, I incorporate data on shipmates' boarding times and occupational roles within the ship. The timing of when shipmates boarded relative to each other allows for exploration of how interaction duration might influence network formation and group dynamics. Using Navy rating codes, I identify each sailor's ship occupation, enabling the examination of hierarchical dynamics within the ship. Sailors with similar occupations likely had greater exposure to one another, potentially forming stronger network ties. This occupational data also facilitates analysis of how interactions between sailors in different hierarchical positions might vary. Finally, I use imputed characteristics on ethnicity, income, and occupation based on sailors' names and enlistment locations. By incorporating these imputed characteristics into the network measures, I can analyze how exposure to diverse backgrounds and socioeconomic statuses within the ship's network might influence network formation and post-war migration decisions.

1.3.3 Record Linkage

Linking individuals from Navy service records to three key periods in their lives – pre-war (1940 Census), prime-age (1950 Census), and death (Numident Social Security Records, Veteran Affairs, and FindAGrave) – forms a crucial part of this analysis. The Muster Rolls provide limited characteristics for linkage, necessitating a multi-step record linkage procedure that maximizes link rates while minimizing false positive matches.[23]

Step 1: Restrict Data. I restrict the sample to men born between 1905 and 1928, covering World War II draft eligibility and active combat service requirements. Men known to have served in the Army are excluded.[24]

[22] Assignment to ships is discussed in full in Section 1.4. Future robustness checks will relax this assumption and pool exposure across ships.

[23] The Muster Rolls provide name, date of enlistment, and place of enlistment. Date of enlistment provides an upper bound on year of birth and facilitates linking to social security records. Place of enlistment aids in linking to 1940 Census residence.

[24] Over 99% of Navy enlistees during this period were born within this range, according to the Veterans Affairs index. Branch switching after formal enlistment was exceedingly rare.

Step 2: Bilateral Links. I construct bilateral links between datasets using both deterministic and probabilistic approaches. The deterministic method, following (3), utilizes all common variables with non-missing values. For instance, when linking the Muster Rolls to the 1940 Census, I use first and last names, allowing for both exact and fuzzy matching, and verify uniqueness across datasets.[25]

Complementing this, I employ probabilistic record linkage using the Python package `splink`, which implements (36). This approach is well-suited for cases with missing observations, measurement error, and continuous variables such as distance (35). For example, when linking the Muster Rolls to the 1940 Census, I use first and last names, birth year (upper bound in the Muster Rolls), middle initial, and the geographic distance between 1940 residence and enlistment city. To prevent false positives, I require all non-missing variables to match within a specified bandwidth and verify uniqueness across both datasets.[26]

Step 3: Harmonize Deterministic and Probabilistic Links. I harmonize links established through deterministic and probabilistic methods. Discrepancies between the two procedures, occurring in less than 0.5% of cases, are dropped from the analysis.

Step 4: Establish a Chain of Links. In this final step, I use the network of bilateral links to connect individuals across multiple data sources. This approach enables indirect links between datasets that lack a direct match, maximizing the use of available information. For example, I may link a sailor from the Muster Rolls to the 1940 Census via the Social Security Death records, using name and date of enlistment for the first link, and name and place of birth for the second.

The final linked sample comprises 478,000 sailors observed in both the muster rolls and the 1940 Census. Using the IPUMS 1940-1950 MLP Crosswalk, 266,000 of these individuals are linked to the 1950 Census. Additionally, 578,000 individuals are linked to at least one death loca-

[25]To account for common OCR encoding errors and the limited set of linking variables, names must be unique over a fuzziness bandwidth. For example, a name that is unique on exact match must also be unique within a Jaro-Winkler distance of 0.05.

[26]For instance, when linking between the 1940 Census and Social Security records, I use the first names of the person's mother and father when available. To avoid false positives between household members, I restrict links to those where the Jaro-Winkler distance between first names is less than 0.3 and birth years fall within a 3-year bandwidth.

tion measure (either Social Security records or FindAGrave). The achieved match rates of 35%, 20%, and 42% respectively are at the higher end of match rates typically found in the literature. Appendix 1.11.3 provides detailed information on each step of the linking process. Match rates for bilateral linkages and overall linkage across all datasets are reported in Appendix Table 1.3.

1.3.4 Additional Data Sources

Full Count Censuses

The 1940 Census provides pre-war geographic, demographic, and household characteristics of men who served in the Navy during World War II. With a median age of 16 in 1940 for men in the linked sample of Navy men, the measured household characteristics represent a mix of parental household attributes (for those still living at home) and individual characteristics (for those who were household heads). I measure prime-age outcomes using the 1950 Full Count Census, which contains comprehensive data on residence, occupation, and household characteristics.[27] For measuring income in both Censuses, I use occupational score which reports the average income (in hundreds of 1950 dollars) for each 1950 occupational code.[28]

Location and Characteristics at Death

To measure individuals' outcomes over their lifetimes, I use mortality records from the Social Security Administration's Numident file as the primary data source. These records cover a high proportion of individuals who died between 1985 and 2007, totaling approximately 50 million records. The Numident file provides the zip code of residence at the time of death, which serves as the main measure for end-of-life location and economic outcomes.

To supplement this data and extend the sample outside of deaths covered by Numident, I collect data from the website "FindAGrave.com", the world's largest online database of gravestones. This

[27] The 1950 Census asked fewer questions of all individuals than the 1940 Census. Questions on education and income were only asked for the sample-line which covered 20% of the population.

[28] Household income is reported only on the sample line in 1950, and there are current irregularities in the field in the public release of the 1950 Full Count Census. I use occupational score in both periods for consistency.

publicly searchable database contains burial locations, demographic information including places of birth and death, and military service details including branch and conflict. After restricting to veterans born between 1900 and 1928 and excluding those identified as Army servicemen, the database yields 1.8 million individuals eligible for linking.

Finally, I use data acquired from a Freedom of Information Act (FOIA) request to the Veterans Affairs Bureau.[29] These data contain records on over 1.8 million veterans who served in the Navy, enlisted between 1941 and 1945, and died prior to 2013. Each observation is identified by an individual's Social Security number, but name and location at death are not included. However, the date of enlistment into the Navy (a field also reported in muster rolls) is included, making this data particularly useful for linking in tandem with other data sources.

Further information on each data source, reliability, and variables contained is reported in Appendix 1.11.1.

1.3.5 Sample for Main Analysis

Subsequent analysis focuses on a specific subset of individuals: white men born between 1905 and 1928 who can be linked to Navy records and at least one post-war migration period. I include only those whose state of enlistment matches their pre-war state of residence.[30] This restriction eliminates ambiguity about whether post-war locations represent moves relative to pre-war or wartime residences. The resulting sample comprises 380,000 individuals linkable to either their 1950 location or place of death.

Two key considerations arise when assessing how men in the sample may differ from the general population: (1) how well the linked sample represents the population of white men in rank-and-file Navy ship positions, and (2) how this Navy population compares to the general cohort of white men born between 1905 and 1928. Figure 1.15 addresses these issues by comparing the

[29] FOIA 23-09430

[30] Pre-war state of residence is determined from either the 1940 Census or the state of Social Security number issuance for those issued between 1937 and enlistment. Around 15% of men in my sample have a different state of enlistment than pre-war state of residence.

linked sample to all men in the Muster Rolls data (Panel A) and to same-age men in the 1940 Census (Panel B). Each panel in Figure 1.15 compares three linked samples: those linked to 1940 (All Links), those linked to both 1940 and 1950 (1950 Links), and those linked to both 1940 and death records (Death Links). Coefficients represent the standardized difference between the linked sample and the full population.

The linked sample appears largely representative, with all coefficients within 0.2 standard deviations of the general population. Notable selection patterns include a slight under-representation of individuals from the South and over-representation from the Midwest in both comparisons. When compared to the 1940 Census, the sample shows a lower share of individuals from farm households, likely reflecting the agricultural exemption for Navy service. Additionally, the sample exhibits slight positive selection relative to the general U.S. population, with higher education, occupation scores, and geographic mobility, consistent with typical patterns in linked samples.

1.4 Impact of WW2 Ship Networks on Migration

This section estimates the impact of geographic networks formed on Navy ships on post-war migration decisions. I begin by describing the assignment process of sailors to ships and provide evidence for conditional random assignment. I then show how variation in the geographic mix of shipmates impacted state of residence in 1950 and by time of death.

1.4.1 Empirical Strategy

Estimating the causal impact of exposure to shipmates from different states on migration patterns requires ruling out systematic assignment of sailors to ships based on their likelihood of moving or preference for specific locations (54). It also necessitates separating out the role of shipmates from other impacts of the Navy that might impact migration such as dislocation, exposure to specific locations through service, and networks formed in other aspects of service.

To disentangle network effects from these other channels that might impact migration, it is therefore crucial to understand the institutional details of how the Navy assigned sailors to ships. This section describes how sailors were assigned to ships and demonstrates that, after accounting for transportation logistics, the geographic composition of shipmates was effectively random. This section concludes by discussing how other impacts of Navy service may interact with shipmate networks.

Assignment to Ships

Historical accounts indicate that newly trained rank-and-file personnel were highly substitutable during World War II.[31] These men, who comprised approximately 75% of the Navy force throughout the war, were typically young and minimally trained. Advancements in Navy bureaucracy and early computing technology facilitated a systematic assignment process for these sailors relative to previous wars. Upon completion of basic training, men were assigned to ships primarily based on immediate operational needs rather than individual characteristics or preferences.

The Navy's assignment process, while broadly random, was shaped by three logistical and operational considerations. First, men were transported in groups to minimize transportation costs,increasing the likelihood of serving with others who enlisted at similar times and from the same region.[32] Second, changing Navy policies, particularly lowering the minimum draft age to 18 in July 1942 and ending volunteer enlistment in late 1942, altered the composition of new entrants

[31]From *Administration of Navy Department in World War II* (40): "The distribution of enlisted personnel was in some respects more difficult than the distribution of officers because the numbers were greater, but the clerical work in the Bureau was less because the movement of enlisted men was usually accomplished in drafts and did not require the writing of orders for each individual. Eventually the distribution of such personnel was further simplified by the Navy Classification Code system. Under this system a code number was assigned to every enlisted man in the Navy which stood for his background of vital statistics and education, rating, and skill; this code number, in effect, described the man and his qualifications. Each billet in the Navy also had a code number describing the skills and special qualifications needed by an individual to fill it satisfactorily. Bringing the code number of the individual and of the billet description together resulted in locating the man qualified to fill the billet. This work was accomplished in the Bureau by the use of punched cards and tabulating machines, and removed the process from the realm of excessive detailed clerical work."

[32]The transportation of enlisted personnel was usually made in groups known as drafts. Orders were issued to individuals on the rare occasions when some highly qualified specialist was needed in an emergency to fill a specific billet. The drafts were normally distributed to ships through the organizations of the service commands or through receiving stations. At times a sudden requisition for a large draft made heavy demands on the distribution service of Bureau of Personnel (BuPers).

over time.[33] Third, different ship types required varying skill mixes, with submarines, for example, having a higher proportion of skilled technical positions than destroyers.[34]

Given these factors, I consider assignment to Navy ships to be random when conditioning on three variables: state of enlistment, first quarter received on a ship, and ship type. Thus, what distinguishes two men from Pennsylvania who enlisted in early 1943 and served on destroyers is the chance assignment to a specific ship based on immediate needs at the time they completed basic training. Consequently, any difference in outcomes can be attributed to experiences in the Navy and not underlying differences in preferences or migration propensities.

It is likely that other experiences of Naval service impacted a person's migration likelihood beyond networks formed with shipmates. Certain forces were broadly shared across all servicemen, including dislocation, access to GI benefits, and the opportunity cost of service years. However, other factors that influence migration may be particular to a person's service experience: networks formed during training, experiences in other places along the ship's route, the impact of combat and trauma, promotions, etc. While these forces likely influence migration, I abstract away from them, by assuming ship networks are orthogonal to these other forces.[35]

Balance Tests

The identification strategy relies on the assumption that assignment to Navy ships was random, conditional on three factors: state of enlistment, first quarter on ship, and ship type. To validate this assumption, I test whether individual characteristics from the 1940 Census predict the composition of shipmates they encounter, after controlling for these three factors.

[33]Executive Order 9279 reduced the inflow of men from agricultural states and eliminated the option to choose Navy over Army service.

[34]Some ships also had specific physical requirements or relied on volunteers, as with the inherently dangerous submarine service.

[35]One factor that may be of particular concern would be if individuals who are systematically exposed to people from a particular state also spend more time in that state during the course of service. Current work in progress is to establish groupings of ships that had similar geographic paths over the course of World War II.

I estimate the following balance test to show the relationship between a sailor's pre-war characteristics and the ethnic, geographic, or economic composition of his later shipmates.

$$y_{ik} = \beta \underbrace{X_i}_{\text{Baseline Characteristics}} + \underbrace{\gamma_{g(i),h(i)}}_{\text{Ship type, quarter, and state FE}} + \varepsilon_{ik} \tag{1.1}$$

Here, y_{ik} represents various characteristics of the ship k on which individual i served, X_i are baseline characteristics of individual i, and $\gamma_{g(i),h(i)}$ includes fixed effects for enlistment state, first quarter on ship, and ship type.[36] Coefficients are scaled to represent 1/100 of a standard deviation of y_{ik}.

Figure 1.2 presents the results of these balance tests, with each panel showing balance on a different ship characteristic y_{ik}. Each plot displays three specifications: no controls, state fixed effects, and fixed effects for state, ship type, and first quarter on ship.[37]

Without controls, baseline characteristics and ship characteristics are correlated: individuals living in a different state than their birth state have, on average, a 2 p.p. (20% of a SD) greater share of shipmates from western states. This relationship is not surprising given the coordinated distribution of people to training and ships, and reflects that people from western states in 1940 are 31 p.p. more likely to be living in a different state than birth state relative to the national average.

Adding state controls almost fully eliminates any relationship between baseline and ship characteristics. For instance, conditional on being from the same state, a person living in a different state than their birth state is on average on ships with only a 0.1 p.p. (1% of a SD) greater share of shipmates from Western states. For almost all other ship characteristics and baseline characteristics, the coefficient is not statistically different from zero. Finally, adding time and ship type controls largely does not change coefficients. However, it importantly attenuates the coefficient on age such that it is not statistically different from zero, reflecting that people who enlist at similar times are more likely to be of similar ages.

[36] Throughout the paper, $h(i)$ will represent a person's pre-war residence, while $g(i)$ represent his "type" classified by when he served and what type of ship he served upon.

[37] Panel C which shows the share of co-ethnic shipmates also include controls for own ethnicity in all three specification.

Two main factors explain the residual correlation between baseline characteristics and ship assignment. First, state of enlistment imperfectly proxies the Navy's spatial distribution constraints, as enlistees from different parts of large states may systematically end up on different ships. Second, the bias towards similar shipmates scales with the number of same-day boarders, and individuals boarding with larger groups tend to be somewhat differently selected.[38] An alternative specification with county controls and restricting the sample to those with less than 10% same-day boarders fails to reject the joint nullity in a Wald test, supporting these explanations. In Section 1.4.3, I re-run core estimates using this more conservative specification and all results of the paper remain unchanged.

These results provide strong support for the assumption that, conditional on the key variables, assignment to ships was essentially random with respect to characteristics that might influence later migration decisions.

1.4.2 Main Empirical Specifications

I estimate the impact of ship networks on migration along two margins: out-migration and directed migration. Out-migration refers to the likelihood of leaving one's geographic origin due to exposure to shipmates from other areas. Directed migration captures how exposure to people from a specific location influences the probability of migrating to that location.

The analysis examines these effects at three levels of geographic granularity (state, Census division, and Census region) and two points in time (1950 and time of death).

The following equation represents the causal impact of shipmates on out-migration:

$$\underbrace{y_{ikmt}}_{\text{Moves out of own } m} = \beta_{mt} \underbrace{X_{ikm}}_{\text{Share of ship not from own } m} + \underbrace{\gamma_{h(i),g(i)}}_{\text{Ship type, quarter, and state FE}} + \varepsilon_{ikmt} \tag{1.2}$$

[38] Looking across all sailors in the sample, shipmates who boarded on the same day account for 5% of total shipmate exposure at the median and 25% of exposure at the 75th percentile.

where y_{ikmt} indicates whether individual i from boat k has moved out of their original location m by time t. X_{ikm} is the share of shipmates not from location m. β_{mt} measures the effect of increased exposure to out-of-area shipmates on the probability of out-migration by period t.

For directed migration, I estimate:

$$\underbrace{y_{ikjt}}_{\text{Lives in } j} = \beta_{jt} \underbrace{X_{ikj}}_{\text{Share of Ship from } j} + \underbrace{\gamma_{h(i),g(i)}}_{\text{Ship type, quarter, and state FE}} + \varepsilon_{ikjt} \tag{1.3}$$

Here, y_{ikjt} indicates whether individual i lives in location j at time t. X_{ikj} represents the share of shipmates on boat k from location j. β_{jt} captures the effect of increased exposure to shipmates from j on the probability of living in j at time t. $\gamma_{h(i),g(i)}$ includes fixed effects for enlistment state, enlistment time, and ship type.

These two migration margins provide complementary insights into how exposure to shipmates from other geographic locations influences an individual's migration decision. Networks affect migration decisions by altering the option value of moving to specific locations through the addition of network ties. The directed migration specification tests this mechanism directly by measuring how random exposure to shipmates from specific geographic locations changes the likelihood of moving to that location. The out-migration specification then considers the aggregation of these direct migration effects by measuring the impact of total exposure to individuals from other places on the decision to migrate at all.

While informative, this analytical framework does have limitations: each specification isolates only one component of the ship network and measures its impact on a single margin of the migration decision. To address these limitations and provide a more comprehensive analysis, Section 1.5 employs a discrete choice model that considers the role of the entire ship network on the full migration decision.

1.4.3 Results

This section presents empirical findings on how Navy ship networks influenced post-war migration patterns. I first examine the impact of exposure to shipmates from other places on out-migration. Then I analyze the role of exposure to shipmates in predicting whether people systematically move to where their shipmates live prior to the war (directed migration). Finally, I explore heterogeneity in these effects and run a series of robustness tests to validate the findings.

Impact of Navy Networks on Out-Migration

Figure 1.3 quantifies the relationship between exposure to shipmates from different geographic locations and post-war out-migration. Panel A demonstrates how out-migration rates vary with the share of shipmates from one's own location across states, Census divisions, and Census regions. Each sub-panel presents a bin scatter of out-migration by 1950 and over a person's lifetime, using within-home state quintiles to group observations. The plots reveal that men with a higher share of shipmates from their own state were substantially less likely to migrate - men in the highest quintile of own-state shipmates were 1.1 percentage points less likely to leave their state by 1950 compared to those in the lowest quintile, with this gap widening to 2 percentage points by time of death.

Panel B presents regression results from equation (1.2), which formalizes this relationship between shipmates and out-migration by providing causal estimates of how exposure to shipmates from different locations influences both short-term (by 1950) and long-term (by time of death) out-migration. Exposure to shipmates from different geographic areas significantly increases the likelihood of subsequent migration. A one standard deviation (6 p.p.) increase in exposure to out-of-state shipmates raises the likelihood of out-of-state migration by 0.5 percentage points in 1950 and 1.1 percentage points by time of death, corresponding to a 4.2 percent increase by 1950 and 2.8 percent increase by time of death. While the coefficients in Panel B of Figure 1.3 are smaller for out-of-division and out-of-region migration, this difference in effect size reflects that the rate of out-migration over longer distances is lower. Proportionally, a one standard deviation increase in exposure to out-of-division and out-of-region shipmates increases the likelihood of migration by 4.5% and 6.5% in 1950 and 4.4% and 4.1% by time of death, respectively.

The effect of exposure to shipmates from other places on out-migration exhibits striking proportionality across different geographic levels and time periods, despite substantial variation in baseline migration rates. For instance, out-of-state migration increases from 12% in 1950 to 42% by time of death, and out-of-region migration increases from 5%-26%. While the proportional network effect size on out-migration is somewhat weaker over a person's lifetime than by 1950, particularly for out-of-state migration, these relative magnitudes are more stable than those observed for directed migration. The relatively consistent proportional effect of networks on out-migration over time aligns with a scenario where direct ties drive initial moves, followed by second-order mobility effects stemming from network-induced migration.

Impact of Navy Networks on Directed Migration

This section examines how random exposure to shipmates on Navy ships affects the likelihood of a person migrating to the places where those shipmates lived prior to the war. I analyze the impacts of shipmate exposure on directed migration across three levels of geographic granularity: Census regions, Census divisions, and states.[39] Figure 1.4 presents the main results at the Census division level. Panel A reports the causal estimates from equation (1.3), while Panel B shows these estimates normalized as percent increases in directed migration from a one standard deviation increase in exposure. Appendix Figure 1.18 illustrates the underlying relationship between exposure and migration rates for each Census division, while Appendix Figures 1.17 and 1.16 present complementary results at the region and state levels.

Navy networks most strongly influence migration to areas rapidly growing in the post-war period. For instance, a one standard deviation (7.2 p.p.) increase in exposure to shipmates from Pacific states is associated with a 18 percent (0.4 p.p.) increase in migration to those states by 1950. As discussed in Section 1.2 the West was by far the fastest growing area of the country in the 1940s.

[39] There are nine Census divisions nested within the four Census regions (West, Midwest, Northeast, and South).

Similar patterns extend to other high-growth areas, with significant network effects observed for migration to the coastal South (10% increase) and Mid-Atlantic Census divisions (10% increase).

While the magnitude of the effect of shipmate exposure on migration is largely proportional to the underlying migration rate to specific areas, the Midwest presents a notable exception to this pattern. Despite being the most common migration destination in 1950 for men in my sample and experiencing substantial in-migration from the South, Navy networks had little influence on migration to these states in the Midwest. This divergence likely stems from pre-existing migration patterns between the South and Midwest for white individuals; the established flow of people between these regions meant that potential migrants already had access to information and support networks, making new Navy connections less critical for these moves.

Looking across both time horizons, while the impact of increased shipmate exposure is more pronounced in absolute terms by time of death, the proportional effect is roughly two times as large in 1950. For instance, while a one standard deviation increase in exposure to shipmates from Pacific states increases the likelihood of moving by 0.4 p.p. in 1950 and 0.9 p.p by time of death, the proportional impact by time of death is only 55 percent of that in 1950. This pattern suggests that these direct ties were most salient in the short run, but continued to be influential over a person's lifetime.

Examining the results across levels of geographic granularity reveals that treatment effects of directed migration are unevenly distributed within regions and divisions. For example, a one standard deviation increase in exposure to people from Western states increases the likelihood of moving to the West by 17%. In contrast, a one standard deviation increase in exposure to Californians raises the likelihood of moving to California by 23%. State-level heterogeneity in treatment will be further explored in the discrete choice analysis in Section 1.5.

These results highlight two potential mechanisms through which Navy networks influenced migration: (1) by facilitating moves to high-growth, high-opportunity areas, and (2) by providing connections to areas where individuals were less likely to have pre-existing ties.

Heterogeneity and Robustness

Heterogeneity

This section examines whether Navy networks had differential effects on migration across subgroups of sailors. In particular, I explore whether individuals with geographically isolated networks are disproportionately influenced by exposure to new connections. Theoretically, individuals from rural areas, lower-income backgrounds, or those without prior migration experience might have fewer ties outside their home region and thus benefit more from access to geographically expanded networks.

Figure 1.20 tests for these differential effects by estimating equation (1.2) separately for subgroups defined by baseline characteristics from the 1940 Census. Panel A presents the direct coefficients from these estimations, and Panel B converts these coefficients into standardized effects by showing the percent increase in migration probability from a one standard deviation increase in out-of-state exposure, facilitating comparison across groups with different baseline migration rates.

The results show remarkable consistency in the impact of Navy networks across subgroups. A one standard deviation increase in exposure to out-of-state shipmates raises migration probabilities by 4%-5% regardless of pre-war characteristics. This uniformity likely reflects the relatively homogeneous nature of the sample–predominantly young, white, non-college educated men.

While the overall effects are similar, point estimates are slightly larger for men from rural and lower-income counties, who likely had more limited pre-war networks. However, these differences are not statistically distinguishable from zero. The evidence thus provides only modest support for the idea that geographically isolated individuals benefit more from new network connections.

Robustness

In this section, I address a main threat to the causal interpretation of the results: systematic assignment of individuals to ships based on migration propensity. To address potential assignment bias, I conduct two exercises. First, I predict the likelihood of moving to Pacific states by 1950 using 1940 baseline characteristics. I construct this prediction for both the linked Navy sample and the full population of white men born between 1905 and 1928. Figure 1.22 shows that predicted migration is uncorrelated with ship characteristics for both samples, indicating that baseline characteristics do not explain the observed increase in migration to Pacific states.

Second, I reconstruct my main estimates using a more restricted sample and robust controls. Figure 1.23 presents results using a sample where individuals' exposure to those they boarded with is less than 10% of their total shipmate exposure. This restriction minimizes potential selection effects from boarding groups. I also include more granular geographic controls (county) and fully interacted fixed effects. The estimated treatment effects remain consistent across all sample and fixed effect combinations, supporting the robustness of the main results.

These exercises provide strong evidence that the observed migration patterns are not driven by systematic assignment bias or confounding factors, but rather reflect the causal impact of Navy-formed networks on migration decisions.

1.5 Discrete Choice Model of Migration with Networks

In this section, I introduce a discrete choice model of cross-state migration that incorporates networks from Navy ships. I document that the expansion of network ties increased the likelihood of migration, largely driven by direct connections over long distances. Then, I use model estimates to quantify the contribution of ship networks to aggregate migration trends.

1.5.1 Preference over Location

The following model describes a World War II veteran's migration decision in the post-war period. Following (13), this model embeds place-specific networks into a discrete choice framework.

The utility an individual i receives from moving to state d in period t:

$$U_{idt} = \underbrace{\mathbb{1}\left(d \neq h(i)\right)\left[\beta^{dest}_{h(i)dt} X_{k(i)d} + \pi_{h(i)dg(i)t}\right]}_{\text{Utility from not home state}} + \underbrace{\mathbb{1}(d = h(i))[\beta^{home}_{dt} X_{k(i)d} + \gamma_{dg(i)t}]}_{\text{Utility from home state}} + \underbrace{\varepsilon_{idt}}_{\text{T1EV Logit shock}} \tag{1.4}$$

Here, $k(i)$ indexes the Navy ship served on during the war, $g(i)$ represents a person's type (defined by the category of ship and first quarter served), and $h(i)$ denotes pre-war state of residence. Individuals choose the state that maximizes their utility. Assuming ε follows a type-I extreme-value distribution yields conditional logit preferences over states.

This model captures three key components of the migration decision: First, type-specific push-pull factors between states ($\pi_{h(i)dg(i)t}$ and $\gamma_{dg(i)t}$) such as distance, wages, and amenities. Second, network-specific factors ($\beta^{dest}_{h(i)dt} X_{k(i)d}$ and $\beta^{home}_{dt} X_{k(i)d}$) capture the impact of ship ties, with β^{home} and β^{dest} allowing for differential effects between the home state and other states. The $\beta^{dest}_{h(i)dt}$ term further allows the value of networks in potential destinations to vary by both origin and destination state, reflecting that ties may have different values depending on the specific state pair involved in a potential move. Finally, all other factors are captured by the individual idiosyncratic term ε.

The discrete choice model is similar in specification to the OLS models discussed in Section 1.4. Both measure the role of the geographic mix of shipmates on the migration decision, but the functional form of the discrete choice models yields some significant advantages. First, the discrete choice model controls for multilateral resistance in ship networks– when measuring the impact of an additional shipmate from California on moving to California, the OLS estimation (Equation (1.3)) is agnostic to the characteristics of other shipmates, while the discrete choice model controls for the relative attractiveness of other network ties. Second, while the OLS estimates captures the average impact of additional ties to a location on subsequent migration to that place, the discrete choice model allows destination locations to vary in attractiveness based on pre-war residence. This heterogeneity in the discrete choice model occurs through two channels: the mean attractiveness of destination states varies by origin state and type (captured by $\pi_{h(i)dg(i)t}$), and the value of network ties to that destination state vary by origin state (captured by $\beta^{dest}_{h(i)dt}$).

Parameterizing Network Effects

The value of additional network ties may vary by home and destination state, reflecting the complex ways in which exposure to people from different places can influence migration decisions. These network effects encompass several mechanisms: information sharing, reduction of migration costs, job referrals, and the amenity value of friendship. These various channels suggest that the value of additional network ties vary both based on existing migration patterns and the relative desirability of destinations. I will not distinguish between these mechanisms but instead allow for enough flexibility to capture the general effect of networks on migration decisions.

Both information sharing and reduction of migration costs suggest that network ties are more valuable in places with sparse existing migration networks. New connections to shipmates from locations with fewer established ties could therefore be particularly influential in shaping migration decisions.

The relative desirability of destinations also plays a role in determining network value. If expectations about wages and amenities are anchored to personal experiences or national averages, network connections would have varying impacts based on destination desirability. Information about places with higher wages or better amenities than one's current location would increase the perceived value of moving. Conversely, learning about lower wages or fewer amenities elsewhere could decrease migration likelihood.

In addition to reducing migration costs and providing information, networks may also be valued as a place-based amenity. If people value having friends in a location, this would contribute equally to the appeal of all destinations, regardless of their economic characteristics.

Ideally, to capture these varied effects, one would estimate network effects for every origin and destination state combination. However, this estimation procedure would require a prohibitively large number of parameters given the available data. To capture key variations in network value while maintaining model parsimony, I parameterize network effects along two dimensions: destination region and pre-existing migration patterns:

$$\beta^{\text{dest}}_{h(i)dt} = \sum_{r \in \text{Reg}} \beta^{r,dest}_t \cdot \mathbb{1}(d \in r) + \beta^{\text{dist}}_t \log \text{dist}_{h(i)d} \tag{1.5}$$

Here "dist" is the distance in kilometers between the centroids of state h and d. The term $\beta^{r,dest}_t$ allows the value of new ties to vary across regions, capturing regional differences in information content and network density. The β^{dist}_t term proxies for existing networks between states, reflecting that the value of new networks may scale with the density of pre-existing ones.[40]

The home network effects capture ways in which an increase in the share of people from one's home state changes the utility of remaining in that state, encompassing both intra-state moves and improving the option value of local labor market opportunities. To account for regional variation in the value of expansion to one's home network, I parameterize home networks effects as:

$$\beta^{\text{home}}_{dt} = \sum_{r \in \text{Reg}} \beta^{r,home}_t \cdot \mathbb{1}(h \in r)$$

This formulation allows the value of home-state ties to vary across regions, capturing potential regional differences in the importance of local networks.

Estimation and Identification

The conditional choice probability that individual i chooses to live in state d by time t is:

$$P_{d|it} = \frac{\exp(V_{idt})}{\sum_{d'} (\exp(V_{id't}))} \tag{1.6}$$

[40]An alternative specification uses existing migration flows instead of distance: $\beta^{\text{dest}}_{h(i)dt} = \sum_{r \in \text{Reg}} \beta^{r,dest}_t \cdot \mathbb{1}(d \in r) + \beta^{\text{mig}}_t \text{cond mig}_{h(i)d,35-40}$, where cond $\text{mig}_{h(i)d,35-40}$ represents migration flows between states from the 1935-1940 census, capturing established migration patterns.

where V_{ikdt} is the common component of the indirect utility function.[41] The log-likelihood associated with this choice probability and parameters Γ_t can be written as

$$LL(\Gamma_t) = \sum_i \sum_d I_{idt} P_{d|it} \tag{1.7}$$

where I_{idt} is an indicator for if individual i lives in state d in time t. I estimate this model separately for both periods using Poisson Pseudo Maximum Likelihood (PPML). As described in (43) in the general case, and (70) for its use in gravity models, PPML is a tractable and numerically equivalent alternative to estimating discrete choice models via maximum likelihood estimation.[42]

The parameters of the model (β_{dt}^{home} and $\beta_{h(i)dt}^{dest}$) are identified if the idiosyncratic component of utility is independent of the characteristics of an individual's ship network $X_{k(i)d}$. As discussed in Section 1.4, assignment to ships is random conditional on type $g(i)$ and pre-war residence $h(i)$. Therefore with the inclusion of type-specific home and destination effects ($\pi_{h(i)dg(i)t}$ and $\gamma_{dg(i)t}$), this assumption is satisfied.

However, using a fully flexible parameterization $\pi_{h(i)dg(i)t}$ would result in over one hundred thousand distinct estimates. I therefore assume that type-specific (g) utility from living in state d is independent of state h.

$$\pi_{h(i)dg(i)t} = \bar{\pi}_{h(i)dt} + \delta_{dg(i)t}$$

This additional assumption allows for a tractable parameterization of push-pull factors between states.

---

[41] $U_{ikdt} = V_{ikdt} + \varepsilon_{idt}$ where

$$V_{idt} = \underbrace{\mathbb{1}\left(d \neq h(i)\right)[\beta_{h(i)dt}^{dest} X_{k(i)d} + \pi_{h(i)dg(i)t}]}_{\text{Utility from not home state}} + \underbrace{\mathbb{1}(d = h(i))[\beta_{dt}^{home} X_{k(i)d} + \gamma_{dg(i)t}]}_{\text{Utility from home state}}$$

[42] For other recent uses see (13) and Dingel and Tintelnot (32). I implement the procedure using the Stata package `ppmlhdfe` (28)

1.5.2 Results

Figure 1.5 presents the results from estimating equation (1.7). Panel A reports estimates for migration by 1950 and by time of death, with destination network effects parameterized relative to distance as shown in equation (1.5). Panel B visualizes the implied pairwise coefficients for $\beta^{dest}_{h(i)dt}$ and β^{home}_{dt} for both periods. Since I control for origin-destination push-pull effects through $\pi_{h(i)dg(i)t}$ and $\gamma_{dg(i)t}$, these network effects can be interpreted as the causal impact of Navy networks on the migration decision.

The results indicate that network ties to states outside one's home state positively influence migration decisions for most origin-destination state pairs. Specifically, $\beta^{dest}_{h(i)d}$ is positive for 78% of home-destination pairs in 1950, increasing to 92% when examining migration by time of death.

The positive value of β^{dist} suggests that additional network connections are more influential in migration decisions when a new connection is from a distant state, aligning with the interpretation that Navy networks are most valuable where pre-existing networks were sparse. An alternative specification supports this interpretation (Appendix Figure 1.24), as using pre-existing migration patterns yields similar relative utility estimates, with correlations of 0.92 in 1950 and 0.83 by time of death.

The value of network ties varies significantly across destination regions. Notably, ties to the Midwest are least valuable, with all instances of negative utility from destination networks occurring in this region in 1950. This finding is consistent with results from Section 1.4.3, which showed that new ties to the Midwest did not increase migration to those states. Ties within one's origin state have minimal impact, suggesting that the out-migration results in Section 1.4.3 are driven by shifts in ties to other locations rather than differential effects of home networks.

1.5.3 Counterfactual Exercises

Using the estimated discrete choice model, I conduct several counterfactual exercises. First, to compare the predictions of the conditional logit model with the OLS results from Section 1.4.3, I estimate the predicted change in migration to California with a ten percentage point increase in exposure to Californian shipmates. Second, to extend the analysis from Section 1.4.3, relative to examining how exposure to shipmates from a single location impacts migration, I measure how the overall geographic mix of shipmates affects levels of out-migration and directed migration. Third, I conduct a back-of-the-envelope calculation of the share of population growth in California between 1935 and 1940 that can be explained by World War II networks.

Counterfactual: Moving to California

Using California as a case study, I compare migration predictions from the discrete choice model against OLS estimates. California provides an appealing test case as the fastest growing state between 1940 and 1950 in population. For both models, I estimate the impact of a ten percentage point increase in the share of Californian shipmates on the probability of moving to California.

For the discrete choice model, I construct state-specific predictions by comparing two counterfactual ships for each origin state h: a low-exposure ship $\underline{k}_h$ and a high-exposure ship $\bar{k}_h$. These ships differ in their share of Californian shipmates by ten percentage points, with all other state shares held proportionally constant at their empirically observed levels.

$$\Delta_{10} P_{\text{CA}|ht} = \frac{1}{\sum\limits_{i \in h}} \sum_{i \in h} \left(P_{\text{CA}|i\bar{k}_{h(i)}t} - P_{\text{CA}|i\underline{k}_{h(i)}t} \right) \tag{1.8}$$

where $\Delta_{10} P_{\text{CA}|ht}$ represents the change in probability of moving to California by period t for individuals from state h, averaged across all individuals i, given a ten percentage point increase in Californian shipmates. To get a comparable estimate using OLS, I estimate equation (1.3) for directed migration to California, which yields a uniform effect across all origin states.

Figure 1.6 illustrates how Navy networks differentially impact migration to California across origin states. Panel A maps overall migration flows to California for men in my sample, Panel B shows the percentage point increase in migration from a ten percentage point rise in Californian shipmates, and Panel C presents these effects as percent change relative to average flows.[43] Figure 1.25 replicates this analysis for Texas as the destination state.

While the OLS and discrete choice model predict similar average effects, the discrete choice model reveals substantial geographic heterogeneity in migration responses to Californian shipmates. A ten percentage point increase in Californian shipmates raises average migration to California by 0.5 percentage points in 1950 and one percentage point by death in both models. However, the discrete choice model shows that effects vary from 0.2 to 2.0 percentage points in 1950 and 0.3 to 1.9 percentage points by death, with the largest impacts concentrated in states neighboring California. These heterogeneous effects mirror overall migration patterns, indicating network effects amplify existing migration propensities–consistent with results from Section 1.4.3 showing the impact of migration networks is largely proportional to average migration rates.

Navy networks most strongly influenced migration to California for individuals from states with limited pre-existing migration ties to California. Converting these effects into percent changes reveals that exposure to Californian shipmates increased migration from Eastern states by 30-38% compared to just 7-20% for Western states in 1950, with similar but attenuated patterns by death (18-22% for Eastern states versus 7-11% for Western states). These differential effects by state of origin demonstrate how Navy networks attenuated gravity patterns in migration flows, with network ties offsetting distance frictions where average migration flows are weakest.

Counterfactual: Out-Migration and Directed Migration

To better understand how Navy networks shaped migration decisions, I decompose migration into two components: the probability of moving out-of-state and the probability of choosing a specific destination state conditional on moving. This analysis extends Section 1.4 by considering how a ship's entire geographic composition, rather than individual components, affected migration patterns. It also allows me to explore heterogeneity in how the impact of ship networks varied across different origin and destination states.

[43]Percentage changes in Panel C are computed by dividing the values in Panel B divided by Panel A, and then multiplying by 100 for each state.

I conduct two counterfactual exercises that compare observed migration patterns to those predicted if individuals had been assigned to ships with network composition matching the average for others from their home state. The first exercise quantifies how variation in shipmate characteristics affects the likelihood of leaving one's home state. The second analyzes how networks influence the choice of destination state, conditional on moving.

Out-Migration Counterfactual

For each home state h, I measure out-migration relative to the average ship.[44] The counterfactual exercise is expressed as:

$$\Delta P_{d\neq h(i)|it} = \left(P_{(d\neq h(i))|k(i)it} - P_{(d\neq h(i))|\tilde{k}_{h(i)}it}\right) \tag{1.9}$$

where k_h is the average ship for home state h, and $\Delta P_{d\neq h|ht}$ captures the change in out-migration probability compared to placement on the average ship. This exercise extends Section 1.4.3 by incorporating the full ship composition rather than just the share of out-of-state shipmates. The counterfactual accounts for both the presence of home-state connections and the relative attractiveness of connections to different destinations. For each origin state, I present results in terms of a standard deviation of $\Delta P_{d\neq h|ht}$ across individuals. This represents how a one standard deviation change in effective variation of shipmate characteristics affects the probability of moving out-of-state, measured in percentage points.

Ship networks substantially influence an individuals' propensity to leave their home state, with effects varying greatly across states of origin. Panel B of Figure 1.7 shows a one standard deviation change in ship characteristics raises out-migration probabilities by 0.2 to 1.1 percentage points in 1950 and 0.1 to 1.6 percentage points by time of death. When scaled by average migration rates (Panel C), these changes represent increases of 1.3% to 6.9% in 1950 and 0.8% to 3.9% by death, aligning with the average effect of share of shipmates from different states on out-migration of 4.2% (1950) and 2.8% (death) found in Section 1.4.3. The impact on longer-distance

[44] The average ship for state h is computed as the geometric mean of the vector of ship shares across all individuals from state h.

moves is slightly larger: Appendix Figures 1.26 and 1.27 show that similar network variation increases cross-division and cross-region migration by between 1.3% and 10.7% by 1950. The large heterogeneity in the impact of Navy networks on out-migration is only weakly correlated with the average out-of-state migration rates shown in Panel A (correlation of 0.58 in 1950 and 0.02 by time of death).

Geographic patterns in where networks were mostly influential in driving migration shifts over time. In 1950, networks generated the largest increases in state-level out-migration from the Sun Belt, Upper Midwest, and the Northeast, while by time of death the strongest effects were concentrated in the Northeast and Pacific. For longer-distance moves, however, the pattern was more stable: Figure 1.27 shows that networks consistently had the largest impact on cross-region migration out of the Midwest and Northeast across both time horizons.

Conditional Directed Migration Counterfactual

This counterfactual examines how networks influenced destination choice conditional on moving out-of-state, and is computed as:

$$\Delta \text{Dir Mig}_{idt} = \frac{P_{(d|ik(i)t)}}{1 - P_{(h(i)|ik(i)t)}} - \frac{P_{(d|i\tilde{k}_{h(i)}t)}}{1 - P_{(h(i)|i\tilde{k}_{h(i)}t)}} \tag{1.10}$$

where k_h is the average ship composition for individuals from home state h. One standard deviation of variation in ΔDir Mig$_{idt}$ across individuals from state h measures the effective impact (in percentage points) of shipmate characteristics on migrating to state d, conditional on leaving one's own state.

Figure 1.8 Panels A and B illustrate how increased exposure to shipmates from specific states shapes bilateral migration flows. A one standard deviation increase in exposure raises conditional directed migration between certain state pairs by up to 4 percentage points in 1950 and 1 percentage point by time of death, with the largest effect observed for migration from Colorado to California. These treatment effects (SD$_h\Delta$Dir Mig$_{idt}$) are strongly correlated with average bilateral migration flows between states (correlation of 0.75 in both periods). Converting the treatment effects into percent changes, as shown in Panel B, reveals that variation in shipmate exposure increases conditional migration between state pairs by 5-15% in most cases, with some pairs experiencing increases in conditional migration of up to 50%.

While network effects on migration flows are largely proportional to average conditional migration patterns, these effects are particularly strong for moves to distant, higher-income states. Comparing average conditional migration flows and bilateral network effects with differences in state income levels in 1940 reveals this systematic pattern: average conditional migration flows show little correlation with income differences between states, while migration driven by ship network variation exhibits a strong positive correlation with state income differences, indicating these networks particularly facilitated moves to higher-opportunity areas. Networks effects are also stronger for geographically distant state pairs, suggesting these connections helped overcome barriers to long-distance migration.

Figure 1.8 Panel C aggregates these bilateral effects to measure how variation in shipmate exposure shapes migration flows to each destination state. For each destination state, the impact of a one standard deviation increase in exposure is calculated by first computing the percent change in migration from each origin state, then taking a weighted average across origin states, with weights proportional to migration flows. This normalized measure captures how increased Navy exposure affects the likelihood of choosing a particular destination state over alternatives.

The strongest effects emerge for migration to rapidly growing states in the West and South. California shows the largest response, with a 17% increase in migration following a one standard deviation increase in exposure to Californian shipmates. Similarly large effects appear for Texas (11%) and Florida (12%). These states also received the highest share of overall conditional migration flows - Florida and California each account for 15% of conditional migration, followed by Texas at 5%. This pattern suggests that Navy networks amplified existing migration trends toward high-growth areas, making already-attractive destinations even more likely to receive migration inflows.

Counterfactual: Share of population growth attributable to war networks

To contextualize the magnitude of these migration effects and their potential role in explaining the rise in overall migration during this period, I conduct a counterfactual exercise focusing on California's population growth between 1945 and 1950. California, the fastest-growing state during this period, experienced a population increase of 1.4 million (15% growth) over these five years.[45]

Using predictions from the discrete choice model, I estimate migration flows to California under two scenarios: one where everyone had average exposure to Californians (8%) and another with no exposure. I then compute the predicted increase in flows from each state of origin. To scale these predictions to the entire population of World War II veterans, I use the state-of-enlistment shares from my Navy data and Army enlistment data, combined with general population estimates for each branch (9.5 million Army, 3.5 million Navy).

This approach predicts 47,500 additional veterans moving to California due to World War II networks. Assuming these veterans moved with their households, and using the average household size in the sample of 4.1 (in 1950), this back-of-the-envelope calculation implies approximately 195,000 people moved to California due to wartime networks. This figure represents roughly 14% of California's total population growth during this period.

The back-of-the-envelope calculation likely underestimates the total effect for two reasons. First, it only considers migration spillovers at the household level, ignoring potential broader network effects once links between places are established. Second, this estimate is based solely on the differential cross-Navy ship effect, potentially missing a higher "base level" effect from other aspects of war experience or general Navy service. I note, however, this exercise does not account for potential crowd-out–in the absence of these war network-induced moves, alternative migration flows to California might have occurred.

I replicate this exercise for two other fast-growing states during this period: Texas and Florida. Texas experienced population growth of 930,000 (14%) between 1945 and 1950, while Florida's population increased by 330,000 (14%). Using the same methodology, I estimate that World War

[45] Population data sourced from https://fred.stlouisfed.org/series/CAPOP (18)

II networks account for 3.9% of Texas's growth and 2.8% of Florida's growth between 1945 and 1950. These findings suggest that wartime networks potentially played a significant role in shaping migration patterns, particularly for migration to the West.

1.6 Shared Ethnicity and Network Formation

This section investigates whether shared ethnicity among shipmates influences post-war migration patterns. The motivation for this analysis stems from the nature of the treatment in this study: exposure to individuals from diverse geographic backgrounds. While general exposure may influence migration choices, it is likely that closer ties formed during service would have a more pronounced effect on post-war decisions.

I focus primarily on shared ethnicity as a key factor influencing network formation and subsequent migration decisions. The salience of ethnic differences among white individuals by country of origin in the 1940s suggests that co-ethnic interactions may have been particularly influential in the context of racially homogeneous but ethnically diverse Navy ships.

This approach builds on previous research by Carrell et al. (20), which demonstrated that patterns of endogenous network formation with peer randomization can significantly influence outcomes. By analyzing the differential impact of shared ethnicity on later migration, I identify one characteristic through which men randomly placed together on a ship, non-randomly formed closer ties.

Appendix 1.12.1 extends this analysis to three additional factors that may influence the strength of ties formed on ships: (1) occupational proximity on ships, (2) socioeconomic background, and (3) the level of combat the ship experienced during the war. These factors provide further insight into the nuanced ways in which shipboard experiences might have shaped post-war migration patterns.

1.6.1 Empirical Framework

To estimate the impact of the ethnic mix of shipmates on an individual's migration decision, I expand upon the empirical frameworks in both Section 1.4 and Section 1.5.

First, I augment equation (1.2) to examine how ethnic composition influences out-migration by 1950 or by death. The analysis includes three specifications. I begin by estimating the impact of the overall share of co-ethnic shipmates. I then decompose out-of-state shipmates into those who share the individual's ethnicity and those who do not. Finally, I incorporate the share of co-ethnic shipmates from one's home state to examine whether ethnic ties influence the decision to stay. For all specifications, I include state by ethnicity fixed effects in addition to the usual time and ship type fixed effects.

Next, I modify the utility specification within the discrete choice framework (Equations (1.4) and (1.5)) by decomposing the share of shipmates from each state who share the individual's ethnicity and those who do not.

$$\begin{aligned}
&\beta_{hd}^{\text{dest}} X_{k(i)d} = \beta_{\text{dest}}^{\text{co-eth}} X_{k(i)d}^{\text{co-eth}} + \beta_{\log \text{dist}}^{\text{co-eth}} \log \text{dist}_{h(i)dt} X_{k(i)d}^{\text{co-eth}} + \beta_{\text{dest}}^{\neg\text{co-eth}} X_{k(i)d}^{\neg\text{co-eth}} + \beta_{\log \text{dist}}^{\neg\text{co-eth}} \log \text{dist}_{h(i)dt} X_{k(i)d}^{\neg\text{co-eth}}, \\
&\beta_{d}^{\text{home}} X_{k(i)d} = \beta_{\text{home}}^{\text{co-eth}} X_{k(i)d}^{\text{co-eth}} + \beta_{\text{home}}^{\neg\text{co-eth}} X_{k(i)d}^{\neg\text{co-eth}}.
\end{aligned} \tag{1.11}$$

Here, $X_{ikdt}^{\text{co-eth}}$ represents the share of shipmates from state d who are also co-ethnics, while $X_{ikdt}^{\neg\text{co-eth}}$ denotes the share of shipmates from state d who are not co-ethnics. By decomposing effects into the contribution of co-ethnic shipmates and shipmates of a different ethnicity, I capture how the presence of co-ethnic shipmates differentially influences migration and whether the differential impact of co-ethnic shipmates varies with pre-existing network density (proxied by $\log$ dist). This specification also includes ethnic-specific destination pull factors in the discrete-choice model to account for potential differences in spatial returns to migration across ethnic groups.

1.6.2 Results

The analysis shows that shared ethnicity between shipmates significantly influences the likelihood of migration, suggesting stronger social ties formed between co-ethnic shipmates. Figure 1.9

presents both OLS estimates (Panel A) and discrete choice results (Panel B). The OLS estimates reveal that a higher overall share of co-ethnic shipmates increases out-of-state migration by 1950, though this effect dissipates when examining lifetime migration patterns. This increased mobility in 1950 operates through two channels: first, co-ethnic shipmates from other states have a stronger impact on migration compared to non-co-ethnic shipmates; second, a higher share of co-ethnic shipmates from one's home state also increases overall geographic mobility.[46]

The discrete choice model corroborates these findings, showing that a higher share of co-ethnic shipmates from any state increases the likelihood of out-of-state migration. This broader increase in mobility from co-ethnic exposure may reflect enhanced network centrality within ships–due to the presence of more co-ethnic shipmates, men form closer ties with shipmates of other ethnicites from different states.

Panel B allows for direct comparison between co-ethnic and different-ethnicity shipmate effects on migration probability. Appendix Figure 1.28 shows the predicted coefficient for β_{hd}^{dest} for when shipmates are co-ethnics or of a different ethnicity correlated with the distance between states. The results demonstrate that co-ethnic shipmates had substantially stronger effects on migration. In 1950, a co-ethnic tie was on average 2.5 times more influential than a non-co-ethnic tie. While this effect moderated over time, co-ethnic ties remained on average 1.5 times more influential by time of death. As shown in Figure 1.28, the relative importance of ethnic ties varies with distance. In 1950, co-ethnic ties had a largely uniform impact on the likelihood of moving across destinations regardless of distance, while the impact of non-co-ethnic ties increases with distance between states, suggesting the relative value of co-ethnic ties was highest for close moves.

These results indicate that shared ethnicity significantly shaped network formation aboard ships and influenced subsequent migration patterns. The presence of more co-ethnic shipmates not only strengthened specific geographic ties but also increased overall mobility, highlighting how social connections formed during service affected post-war migration decisions.

[46]These two channels can be seen in columns 2 and 3 when the share of shipmates from out-of-state is decomposed into the share of shipmates who are from out-of-state and co-ethnic and the share of shipmates who are from out-of-state and of a different ethnicity.

In Appendix 1.12.1, I present results on the impact of on-ship occupational proximity and pre-war income of shipmates on migration propensity. I find little evidence that either factor meaningfully changes the likelihood of migration, though both factors present substantial measurement challenges. Finally, in Figure 1.21, I test for heterogeneity in treatment effects across different ship sizes and levels of combat exposure. I find no measurable impact of these factors on the magnitude of the treatment effect.

1.7 Returns to Networked Migration

This section leverages random variation in Navy ship networks to estimate returns to migration. The previous analysis established that networks formed during Navy service influenced both whether individuals moved and their choice of destination after the war. An important question remains: did individuals induced to move through these Navy networks experience greater economic opportunity?

Using an instrument for the probability of moving derived from the discrete choice model in Section 1.5, I estimate both prime-age (1950) and lifetime (death) returns to migration. This approach provides a unique quasi-experimental setting for studying the impact of network-facilitated migration on later earnings.

1.7.1 Navy Network Migration Instrument

To estimate how migration causally affects earnings, I construct instruments predicting the probability of individuals moving out-of-state, out-of-region, and to the Pacific by 1950 and by time of death. These predicted probabilities are derived from conditional choice probabilities (equation (1.6)) using estimated parameters from the discrete choice model in Section 1.5.

The instrument for the predicted probability of moving depends on two components. First, observable fixed characteristics that inform ship assignment – an individual's pre-war residence $h(i)$ and their type $g(i)$ (defined by ship type and first quarter on ship) – determine common push-pull factors between states. Second random variation in geographic composition of shipmates determine the role of networks in driving migration. Therefore, this approach distills multi-dimensional variation in shipmate characteristics into a single instrument, where conditional on $h(i)$ and $g(i)$, all variation in the instrument is random. The instrument, therefore, captures how random chance in ship assignment influences a person's likelihood of later migration.

Using this instrument, I estimate the returns to migration using two-stage least squares regression.

$$\begin{aligned} \underbrace{y_{ikt}}_{\text{Lives in } y} &= \alpha \underbrace{P_{y|it}}_{\text{Probability of moving to } y} + \underbrace{\gamma_{h(i),g(i)}}_{\text{Ship type, year, geography FE}} + \epsilon_{ikt} \\ \log \text{Inc}_{ikt} &= \beta \widehat{y_{ikt}} + \gamma_{h(i),g(i)} + \nu_{ikt} \end{aligned} \tag{1.12}$$

where y_{ikt} represents whether individual i in time t is living in y, where y represents living in a different state, a different region, or a Pacific state.[47] In 1950, I proxy individual income using occupational score.[48] By time of death, I proxy income using the median household income in the zipcode a person is last known to reside.[49] To extend this exercise, I also measure the impact of migration on non-pecuniary outcomes such as family formation and mortality (Appendix Table 1.6).

In this setting, compliers are individuals who are marginal movers whose migration decision is sensitive to the characteristics of shipmates. An individual's classification as a complier, always-taker, or never-taker is governed by the push-pull factors between states as well as their idiosyncratic logit draws as described in Section 1.5. The idiosyncratic logit draw captures all residual factors besides ship networks and common state push-pull factors that influence a person's migration decision.[50]

[47] For Pacific state moves, individuals originally from Pacific states are excluded.

[48] This is for two reasons. Income is sample-line in the 1950 Census and only available for 20% of individuals. Additionally, known irregularities in the current public release of the 1950 Census make occupational score a more reliable measure of income

[49] Median houshold income is reported in the year 2000 in nominal levels.

[50] A non-exhaustive list of residual factors that are captured in the idiosyncratic logit draw are: pre-existing network connections to other states, idiosyncratic preferences over amenities, idiosyncratic labor components that change spatial return to different states.

Across states, the population of compliers, always-takers, and never-takers differs. For a high-growth state like California where pull factors to California often outweigh idiosyncratic reasons for moving, there is a high share of never takers–people who will always stay in California. Conversely, for a low-growth state like West Virginia where push factors are large, there is a greater share of always takers–people who move regardless of their ship experience. The instrument accounts for these differences: states with high shares of always takers and never takers will exhibit smaller variability in the instrument compared to states with a higher share of compliers. Therefore, for the IV estimate to be interpreted as the local average treatment effect (LATE) for compliers, the key assumption is that the variance in idiosyncratic error terms is constant across states (7; 62).[51]

Beyond these considerations, the standard instrumental variable assumptions apply. The relevance of the instrument holds if random variation in ship networks induces out-migration, which is demonstrated in the first stage results shown in Appendix Table 1.30. The exogeneity of the instrument to baseline characteristics is shown in Appendix Table 1.29. Monotonicity in the instrument holds if there are no defiers. In this context, defiers are individuals who behave contrary to the average effect. While the instrument allows for exposure to shipmates from a particular state to increase or decrease the probability of moving, for monotoncicity the directional effect of exposure to people from a destination state d must be stable for all people from the same origin state h.[52] Given this construction, I realistically assume the population of defiers to be negligible. Finally, the exclusion restriction states the geographic mix of shipmates must only impact incomes through migration. While this assumption is not directly testable, the main potential confounder would be if other peer characteristics correlated with geography drive increases in income through other channels.

[51]For the LATE interpretation to hold under strictest assumptions, I need the first stage to be common across strata, particular pre-war residence. By controlling for push-pull factors directly in the instrument, only variation in the idiosyncratic logit term remains.

[52]i.e. It does not violate monotonicity if exposure to people from West Virginia decreases the likelihood of someone from Oregon moving out-of-state. It does violate monotonicity if for some people from Oregon exposure to people from West Virginia increases the likelihood of moving out-of-state and for others it decreases the likelihood of moving out-of-state

1.7.2 Results

Network-facilitated migration led to large increases in lifetime earnings, as shown in Table 1.1, which reports estimates from equation (1.12). Coefficients for results in 1950 are reported in Appendix Table 1.5. Columns 1-3 report the OLS estimates, while columns 4-6 report the IV estimates. The IV estimates show out-of-state movers experience 46 log points higher income, out-of-region movers see a 53 log point increase, and movers to the Pacific experience 69 log points increase, translating to 59%, 70%, and 99% higher earnings, respectively. These estimates indicate that men induced to move due to experiences on their Navy ship experienced much higher earnings over the course of their lifetimes. In 1950, I note small but significant returns to migration (1% for out-of-state moves, 3% for out-of-region moves, and 4% for moves to Pacific states).

The large pecuniary returns to migration raise the question of why migration rates were not higher in the absence of these network connections. As discussed in (62), high returns to migration are consistent with substantial migration costs and young individuals being mismatched to their birthplaces. Networks formed during Navy service likely reduced these costs, facilitating moves to high-opportunity areas. Furthermore, if networks act as a place-based amenity, the Navy-formed connections may have lessened the trade-off between economic opportunity and network-based amenities.[53] Both of these factors may explain why individuals were willing to move when provided with new network connections, despite previously forgoing apparently lucrative migration opportunities.

The IV estimates being larger than the OLS estimates is perhaps surprising, especially as more geographically mobile people tend to be positively selected. However, several mechanisms could explain this pattern. First, compliers – those whose migration decisions are influenced by Navy networks – may experience larger returns than always-takers. If compliers live in areas with more geographically isolated networks, and these are also lower opportunity, then might reasonably be expected to have a higher return to migration than someone who is an always taker. Second, the

[53] Networks may act as a place-based amenity if individuals value having friends and community near where they live.

IV specifically identifies returns to network-facilitated migration. While the OLS sample includes all Navy veterans, those who leverage networks to migrate may achieve better outcomes, perhaps through accessing higher-paying jobs through their connections. Third, given findings about the relevance of networks for long-distance moves in Section 1.5, the IV might be in part capturing substitution from some otherwise shorter distance cross-state moves to higher opportunity long-distance moves.

Appendix Figure 1.33 illustrates the heterogeneity in returns to networked migration across various subgroups. The results reveal a distinct pattern: in the short run, higher-income, urban individuals from manufacturing households in 1940 experience the greatest benefits from migration. However, this pattern reverses in the long run. Individuals from lower-income, rural areas, and farm households in 1940 exhibit the highest lifetime returns to migration. This shift suggests that those who ultimately benefited most from migration took longer to realize these gains.

Table 1.6 presents estimates for marriage, household size, and mortality, which allow me to examine the impact of migration on non-pecuniary outcomes. Migration shows no significant effect on marriage status, household size, or age at death. However, there is a strong positive impact on the likelihood of having a spouse born in a different state: moving out-of-state increases the probability of marrying someone born in a different state by 29 percentage points, conditional on being married.

While, these findings suggest that Navy networks were highly beneficial for those induced to move, I do note a few caveats. Median household income of zip codes is in nominal levels, so these returns may be in part offset by cost-of-living differences across places, and this effect might be particularly salient for the returns to moving to Pacific states. In addition, I am using zip code of residence at time of death as a proxy for income. Additionally, if there were additional peer effects from shipmates from high opportunity states, then this treatment might be a mixture of the returns to migration and the returns to other benefits such as human capital accumulation. The bundled treatment effect might in part explain the gap between the OLS and IV estimates.

The large returns to migration I am estimating for Navy sailors are largely in line with similar estimates from the literature.[54] For lifetime estimates, (62) find that young individuals displaced by a volcanic eruption had 82% higher earnings 35 years later. For a similar time period and looking to internal migration within the US, both (15) and (26) find large returns for migration from the South to the Midwest during the Great Migration, with returns for Black men exceeding 80%; (15) finds returns for white men of around 60%. (73) examines brothers in early 20th century census records and finds that those who migrate out-of-state or out-of-region have 15% and 18% higher occupational scores by their 30s compared to their non-migrant siblings.

This analysis demonstrates that network-facilitated migration, as induced by Navy ship assignments during World War II, led to substantial increases in lifetime earnings. These findings highlight the importance of networks in facilitating beneficial migration and suggest that reducing migration costs through network formation can have significant long-term economic impacts.

1.8 Conclusion

This paper leverages the unique historical context of World War II Navy service to examine how the expansion of geographic networks influences migration patterns and economic outcomes. By constructing a novel dataset of Navy sailors during World War II and exploiting conditional random assignment to ships, the study provides causal evidence on the impact of newly formed social connections on geographic mobility. The findings demonstrate that exposure to individuals from diverse geographic backgrounds significantly increases both the likelihood of moving anywhere and in particular the choice of where to migrate.

The analysis reveals that war-formed networks may have played a crucial role in shaping postwar migration trends across the United States. Navy connections were particularly influential in driving migration from slower-growing areas to rapidly expanding regions particularly in the West, accelerating broader demographic shifts. These findings highlight how large-scale events such as

[54] A comparison of estimates from the literature is shown in Appendix Figure 1.32

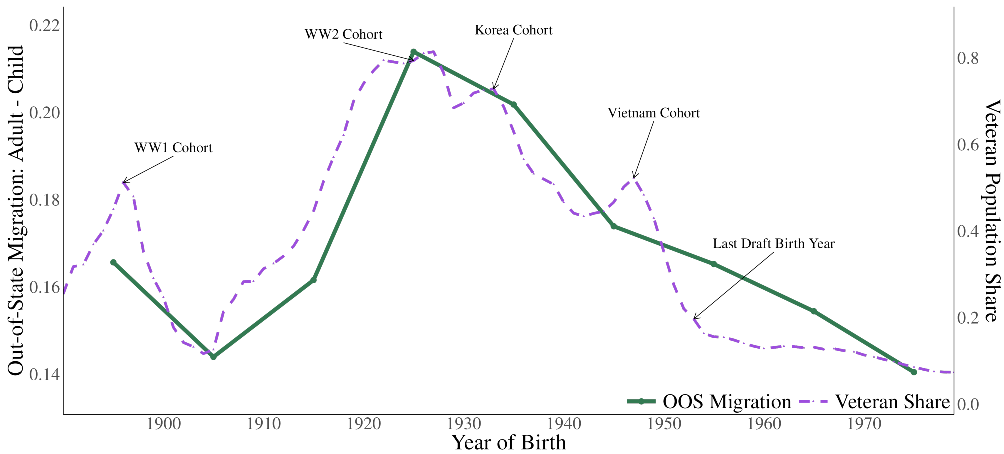
wartime mobilization, can have lasting impacts on the geographic distribution of population and economic activity.

Importantly, the paper bolsters knowledge on the substantial economic benefits from migration. Individuals induced to move due to their expanded geographic networks relocated to areas where they achieved higher lifetime earnings. This finding underscores the potential for social connections to enhance economic opportunity and reduce spatial inequality in opportunity by facilitating access to more productive labor markets.

The results have broader implications for understanding social connectivity and economic mobility in the United States. The dramatic expansion of geographic networks during World War II likely contributed to higher levels of interstate connectivity persisting for years if not decades after the war. Given the important role of networks in facilitating migration and the potential for positive spillovers once disparate locations are connections, policies that foster connections between Americans from different areas may have long-lasting impacts on economic opportunity and spatial economic disparities.

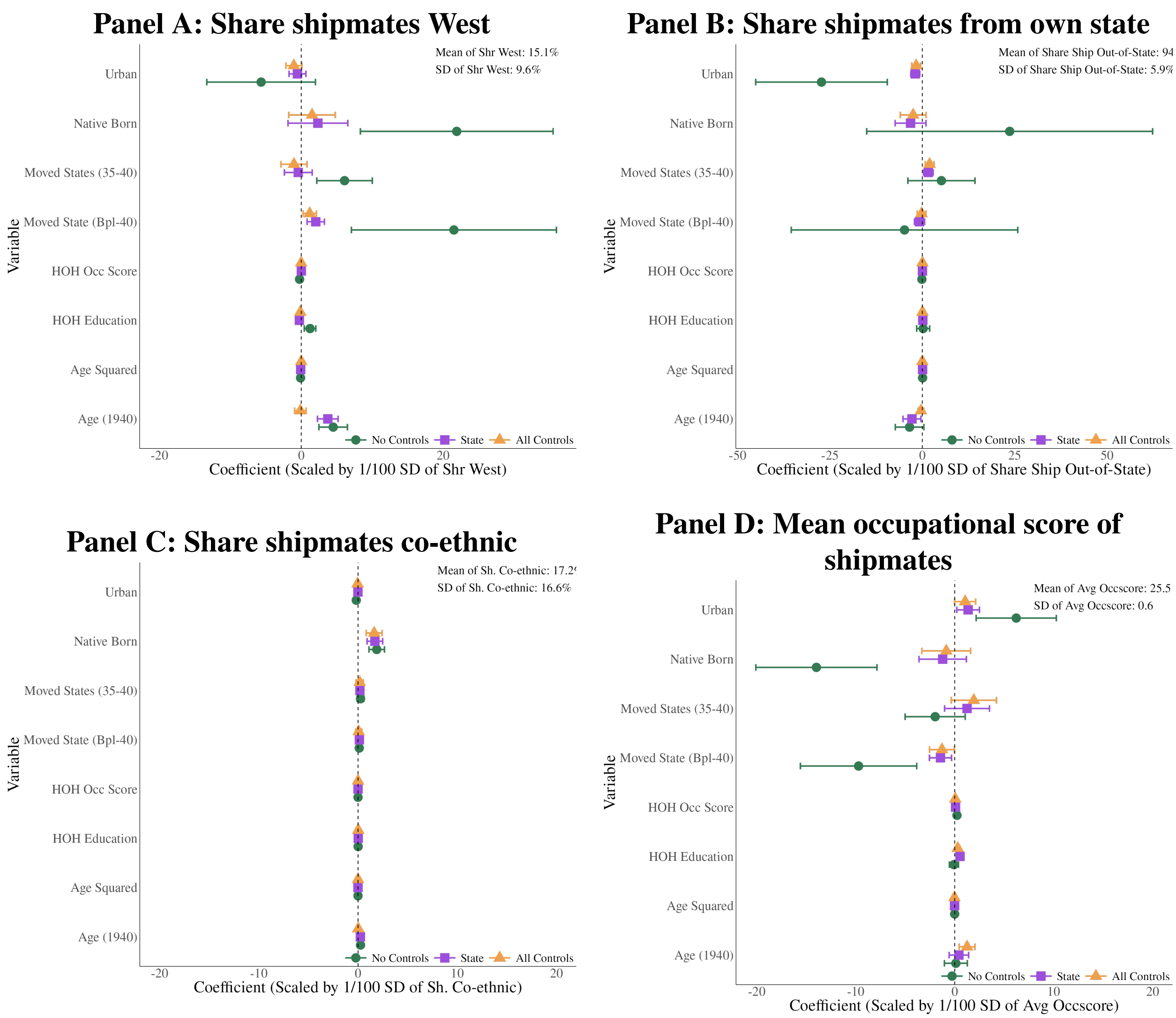
1.9 Exhibits

Figure 1.1: Change in probability of living outside state of birth from ages 10-19 to 30-39



NOTES: This figure displays interstate mobility and military participation rates for white men born between 1890 and 1979. The green solid line shows the difference in probability of living outside one's birth state between ages 30-39 and 10-19, by birth decade. Levels are shown in Appendix Figure 1.10. The dashed purple line represents the share of white men over the age of 35 who report being veterans, by birth year. Data source: Decennial census public use samples (1900-2010) and American Community Survey (2005-2024).

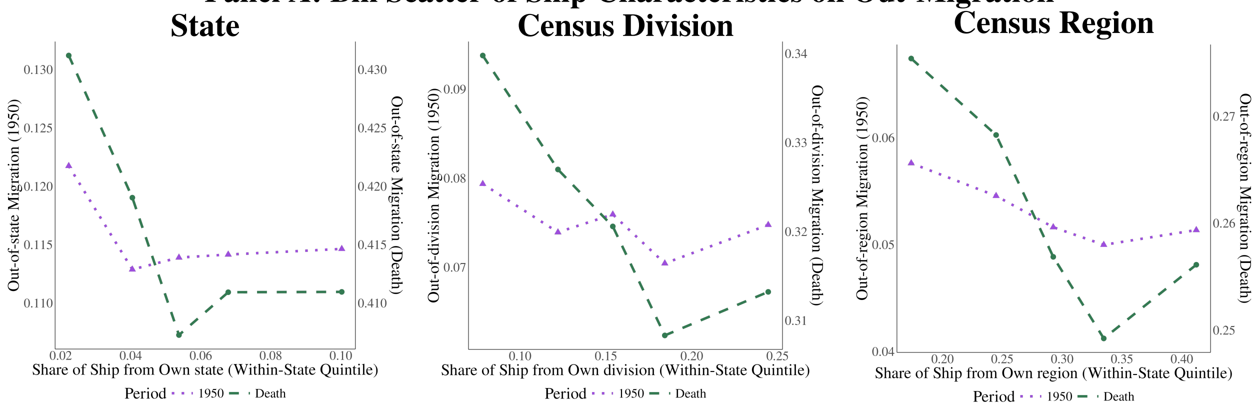
Figure 1.2: Balance Coefficient Plots



NOTES: This figure presents tests of whether baseline individual characteristics predict shipmate composition, controlling for state of enlistment, enlistment timing, and ship type. Each panel shows results from estimating equation (1.1) with a different shipmate characteristic as the dependent variable: Panel A shows the share of shipmates from Western states, Panel B shows the share from one's own state, Panel C shows the share of co-ethnic shipmates, and Panel D shows mean occupational score of shipmates. For each panel, coefficients from three specifications are displayed: no controls (green circles), state fixed effects only (purple squares), and the full set of fixed effects for state, ship type, and enlistment timing (yellow triangles). Panel C additionally controls for own ethnicity across all specifications. All coefficients are scaled to represent 1/100 of a standard deviation of the dependent variable. Horizontal bars represent 95% confidence intervals. Standard errors are clustered at the ship level.

Figure 1.3: Impact of Navy Networks on Out-Migration

Panel A: Bin Scatter of Ship Characteristics on Out-Migration

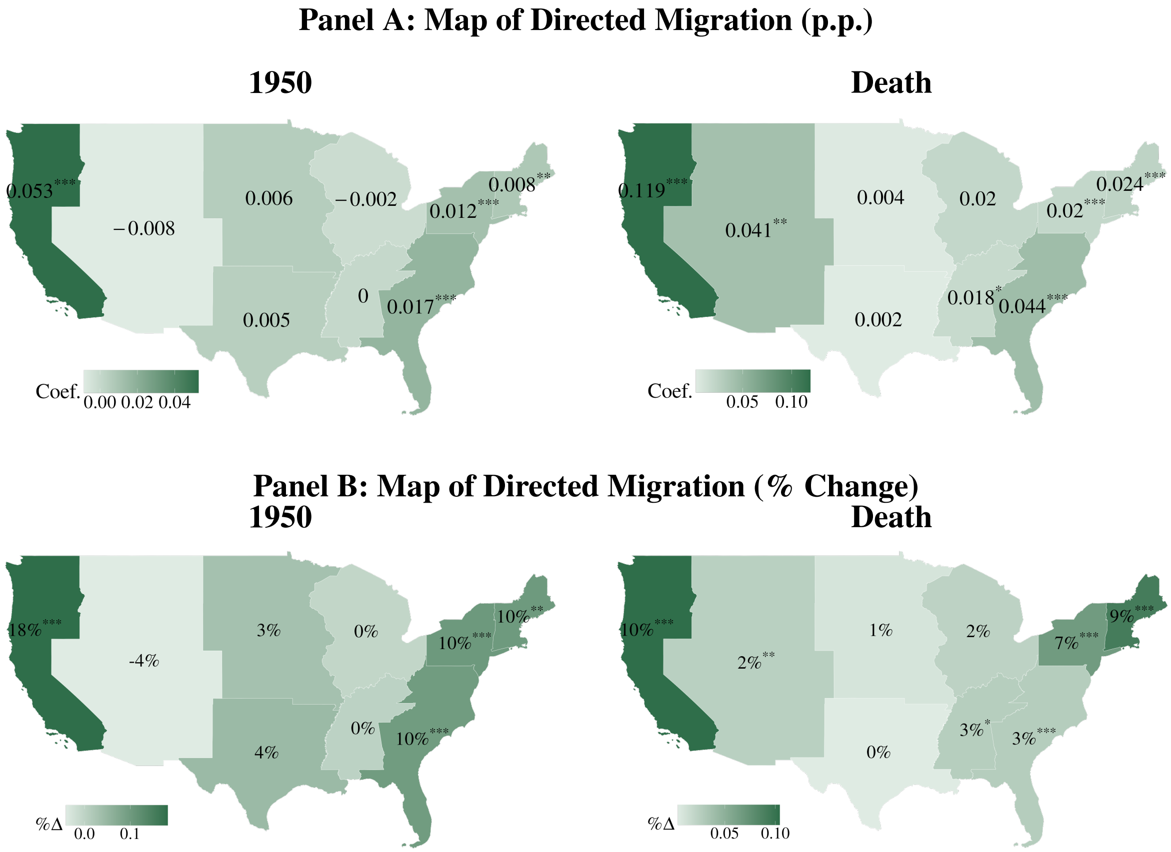


Panel B: Causal Impact of Shipmates on Out-Migration

	1950			Death		
	State (1)	Division (2)	Region (3)	State (4)	Division (5)	Region (6)
Sh. Ship Out-of-State	0.08***			0.19***		
	(0.02)			(0.03)		
Sh. Ship Out-of-Division		0.04***			0.17***	
		(0.01)			(0.01)	
Sh. Ship Out-of-Region			0.03***			0.11***
			(0.007)			(0.01)
Dep Var Mean	0.115	0.075	0.053	0.416	0.322	0.261
Observations	175,276	175,276	175,276	243,108	243,062	243,062
R^2	0.0288	0.0210	0.0143	0.0333	0.0356	0.0466
Within R^2	0.00013	0.00011	0.00020	0.00029	0.00062	0.00054
Ship type FE	✓	✓	✓	✓	✓	✓
State enlist FE	✓	✓	✓	✓	✓	✓
First quarter FE	✓	✓	✓	✓	✓	✓

NOTES: Panel A shows the relationship between out-migration at three geographic levels (state, Census division, Census region) and shipmate characteristics. Each geographic level is represented by a separate plot, displaying migration by 1950 (purple, left axis) and by death year (green, right axis). Both axes span an equal range but are re-leveled to accommodate different baseline migration rates in each period. For each state of origin, five within-state quintiles are created based on the share of shipmates from the same state, division, or region. The average out-migration rate is calculated for each quintile, with a sample population-weighted average across home states reported on the plot. Panel B presents results from Equation (1.2) for the three geographic levels. Columns 1-3 report migration estimates from the 1950 Full Count census. Columns 4-6 report migration estimates over lifetime. All specifications include fixed effects for type of ship served on, state of enlistment, and the first quarter a person served on the ship. Standard errors are clustered at the ship-level.

Figure 1.4: Impact of Navy Networks on Directed Migration (Census Divisions)



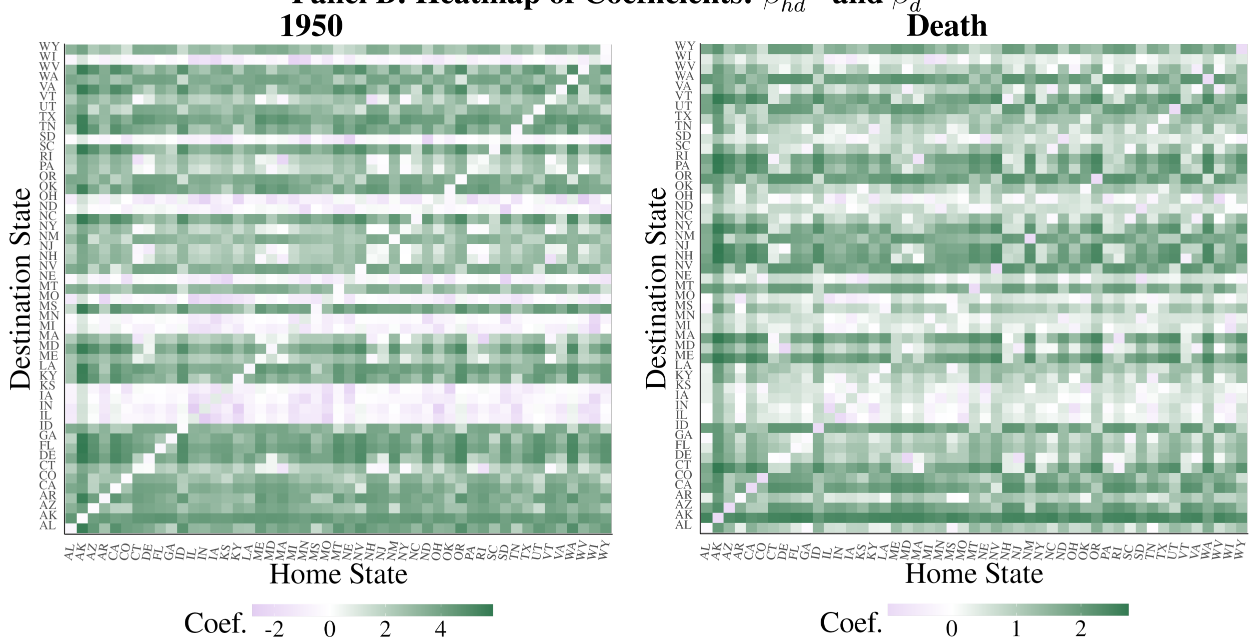
NOTES: This figure shows the impact of Navy ship networks on directed migration across Census divisions, by showing estimates from equation (1.3). Panel A presents β_{jt} showing the effect of increased exposure to shipmates from each Census division on migration to that division, measured both in 1950 and at time of death. The coefficients are depicted geographically on maps of the United States to highlight spatial patterns. Panel B shows these effects normalized as percentage increases in directed migration resulting from a one standard deviation increase in exposure to shipmates from each destination division. Standard errors are clustered at the ship-level.

Figure 1.5: Discrete Choice Model Coefficients

Panel A: Discrete Choice Estimates

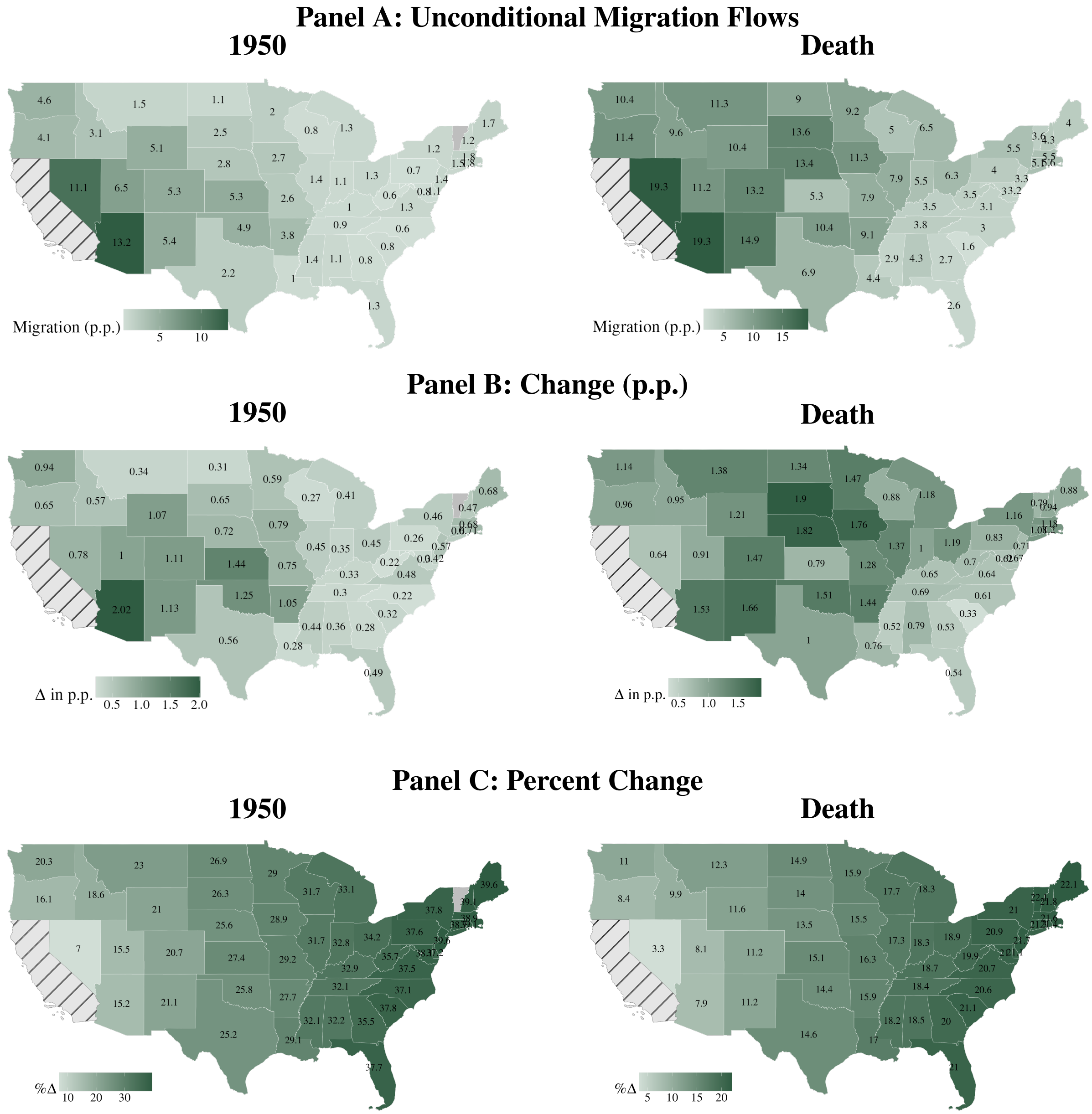
	Coefficient Estimates								
	β^{dist}	$\beta^{dest}_{Midwest}$	β^{dest}_{North}	β^{dest}_{South}	β^{dest}_{West}	$\beta^{home}_{Midwest}$	β^{home}_{North}	β^{home}_{South}	β^{home}_{West}
1950	1.35	-10.20	-7.85	-5.92	-7.05	0.79	0.18	-0.46	-0.23
	(0.34)	(2.30)	(2.21)	(2.38)	(2.59)	(0.44)	(0.30)	(0.54)	(0.41)
Death	0.77	-5.17	-3.90	-4.57	-4.10	0.59	0.69	-0.12	-0.91
	(0.18)	(1.25)	(1.16)	(1.28)	(1.37)	(0.23)	(0.13)	(0.26)	(0.20)
R-squared	0.65 (1950), 0.40 (Death)								
N	7,124,979 (1950), 11,276,968 (Death)								

Panel B: Heatmap of Coefficients: β^{dest}_{hd} and β^{home}_{d}



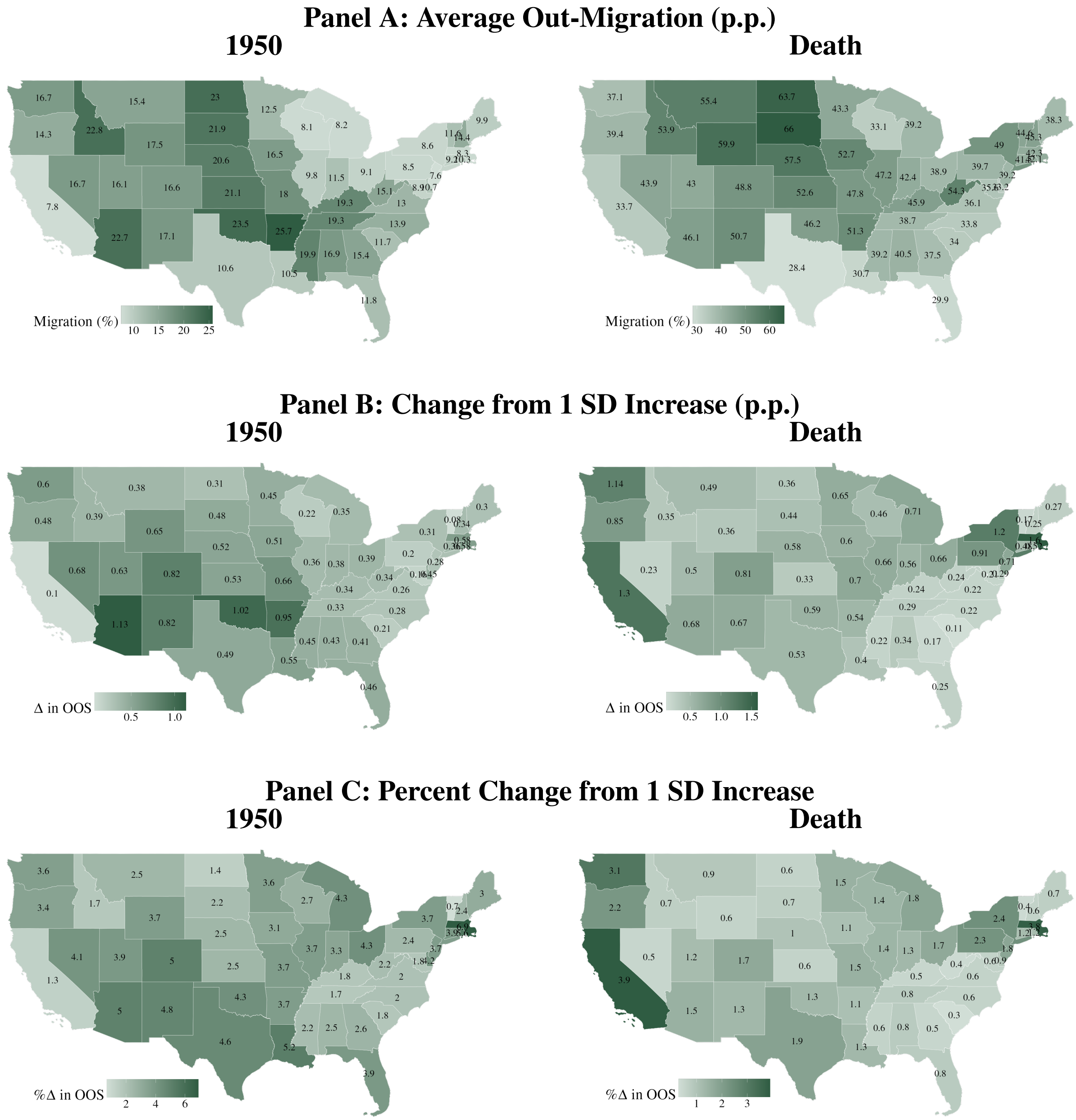
NOTES: Panel A reports discrete choice estimates from Equation (1.7). Row 1 reports estimates for migration by 1950, and Row 2 reports estimates for migrate by time of death. The pseudo-R^2 and number of observations is reported for each specification at the bottom. Coefficients are estimated by Poisson-Pseudo Maximum Likelihood (PPML) estimation. Panel B reports the implied home-destination coefficients of β^{dest}_{hd} and β^{home}_{d} from the coefficients estimated in Panel A and the parameterization described in Equation (1.5). Each cell is a home-destination state pair. Distance is computed as kilometers between state centroids. Logged distance between contiguous states vary between 4.1 (RI to MA) to 8.4 (ME to CA). The median origin-destination pair has a logged distance of 7.3 (MA to WI).

Figure 1.6: Impact of a 10 p.p. increase in Californian ship share on migration, by state of origin



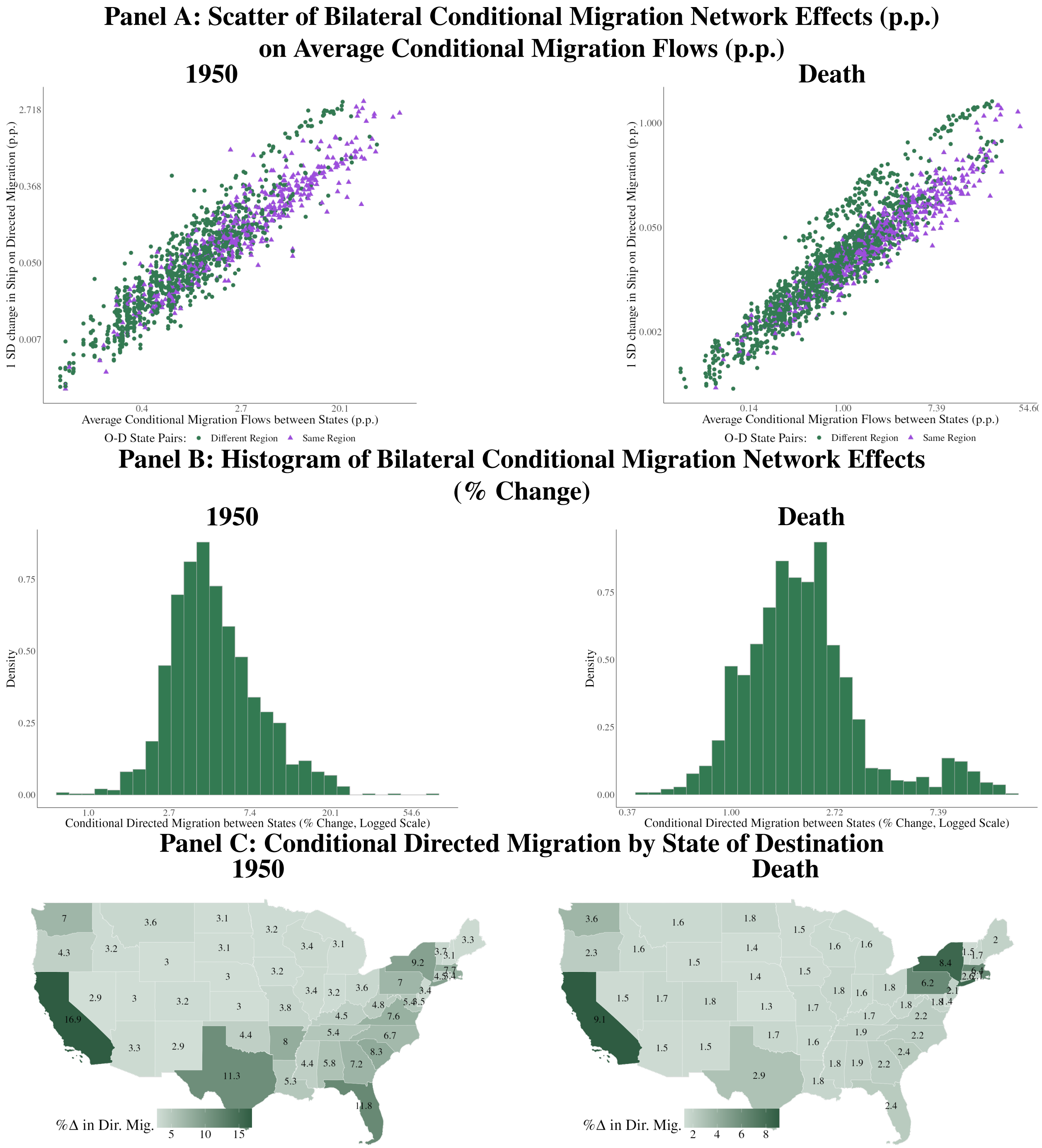
NOTES: This figure presents results from the counterfactual exercising showing the impact of increased exposure to Californian shipmates on migration to California from each origin state. The counterfactual compares the predicted migration response from serving on a high-exposure ship (10 percentage points higher) to a low-exposure ship as described in Equation (1.8) using discrete choice estimates reported in Figure 1.5. Panel A reports average unconditional migration flows to California by origin state by 1950 and by a person's death. Panel B shows the increase in migration probability of serving on a high-exposure ship relative to a low-exposure ship in percentage points. Panel C displays the percent increase in migration probability relative to the average unconditional migration rate. The left map in each panel reports results in 1950, while the right map in each panel reports results by death.

Figure 1.7: Impact of one SD increase in ship exposure on cross-state migration, by state of origin



NOTES: This figure shows how variation in shipmate composition affects cross-state migration probabilities. Following Equation (1.9) and using estimates from Figure 1.5, the counterfactual, ($\Delta P_{d \neq h(i)|it}$), computes the difference in probability of moving out-of-state when assigned to one's actual ship versus a ship with average composition for their pre-war residence $h(i)$. Panel A displays average out-of-state migration rates in 1950 and by time of death for men in the linked sample. Panel B reports one standard deviation in $\Delta P_{d \neq h(i)|it}$ across individuals from state h, representing the change in migration probability from a one standard deviation increase in effective shipmate variation. Panel C shows this effect as a percentage change relative to average out-migration rates from each state (Panel B/Panel A). Each panel presents results for both 1950 (left) and time of death (right).

Figure 1.8: Impact of Navy shipmates on conditional migration



NOTES: This figure shows how variation in shipmate composition affects directed conditional migration between states. Following equation (1.10), I compute for each origin-destination state pair (h, d) the change in probability of moving to state d conditional on moving, comparing assignment to one's actual ship versus a ship with average composition for home state h (ΔDir Mig). Panel A presents scatter plots relating one standard deviation of ΔDir Mig (representing the impact of a one standard deviation increase in effective shipmate variation) to average conditional migration flows between state pairs. Panel B shows the distribution of these one standard deviation effects when normalized as percent increases over average conditional flows. Panel C aggregates these bilateral effects to the destination state level by computing weighted averages of the percent changes across origin states, with weights proportional to migration flows. Each panel compares effects in 1950 (left) to those measured at time of death (right).

Figure 1.9: Network Formation: Role of Co-ethnic Exposure in Migration

Panel A: OLS Estimates

	State Mover					
		1950			Death	
	(1)	(2)	(3)	(4)	(5)	(6)
Share Ship Out-of-State	0.08***			0.18***		
	(0.02)			(0.02)		
Sh. Co-ethnic	0.03***			-0.03		
	(0.01)			(0.03)		
Share Co-ethnic + Not State		0.10***	0.11***		0.15***	0.15***
		(0.02)	(0.03)		(0.008)	(0.04)
Share Not Co-eth + Not State		0.07***	0.08**		0.18***	0.19***
		(0.02)	(0.03)		(0.02)	(0.03)
Share Co-ethnic + Own State			0.05			0.03
			(0.05)			(0.12)
Observations	165,145	165,145	165,145	229,026	229,026	229,026
R^2	0.03703	0.03703	0.03703	0.04032	0.04032	0.04032
Within R^2	0.00013	0.00012	0.00013	0.00025	0.00025	0.00025
State by Ethnicity FE	✓	✓	✓	✓	✓	✓
Ship type FE	✓	✓	✓	✓	✓	✓
First Quarter FE	✓	✓	✓	✓	✓	✓

Panel B: Discrete Choice Estimates

	$\beta^{\text{co-eth}}_{\text{dest}}$	$\beta^{\text{co-eth}}_{\text{log dist}}$	$\beta^{\neg\text{co-eth}}_{\text{dest}}$	$\beta^{\neg\text{co-eth}}_{\text{log dist}}$	$\beta^{\text{co-eth}}_{\text{home}}$	$\beta^{\neg\text{co-eth}}_{\text{home}}$	R^2	N
1950	4.82	-0.11	-13.25	2.11	-1.94	0.62	0.65	5,243,710
	(4.64)	(0.67)	(2.42)	(0.34)	(0.59)	(0.29)		
Death	-5.28	1.01	-4.26	0.77	-1.34	0.67	0.41	8,017,913
	(2.67)	(0.37)	(1.21)	(0.17)	(0.31)	(0.13)		

NOTES: This figure presents estimates of how co-ethnic ties formed during Navy service influenced post-war migration patterns. Panel A reports OLS estimates for whether an individual moves out-of-state by 1950 (Columns 1-3) or by time of their death (Columns 4-6). All specifications include fixed effects for state of enlistment by ethnicity, first quarter on ship, and ship type, and standard errors are clustered at the ship level. Columns 1 and 4 include regressors for the share of shipmates from out-of-state and share of co-ethnic shipmates. Columns 2 and 5 include regressors for share of shipmates who are from out-of-state and co-ethnic, and those who are from out-of-state and not co-ethnic. Columns 3 and 6 add an additional regressor for the share of shipmates who are co-ethnic and from own's own state. Panel B reports discrete choice estimates from equation (1.11). Row 1 reports estimates for migration by 1950, and Row 2 reports estimates for migrate by time of death, and the pseudo-R^2 and number of observations is reported for each period t. Coefficients are estimated by Poisson-Pseudo Maximum Likelihood (PPML) estimation.

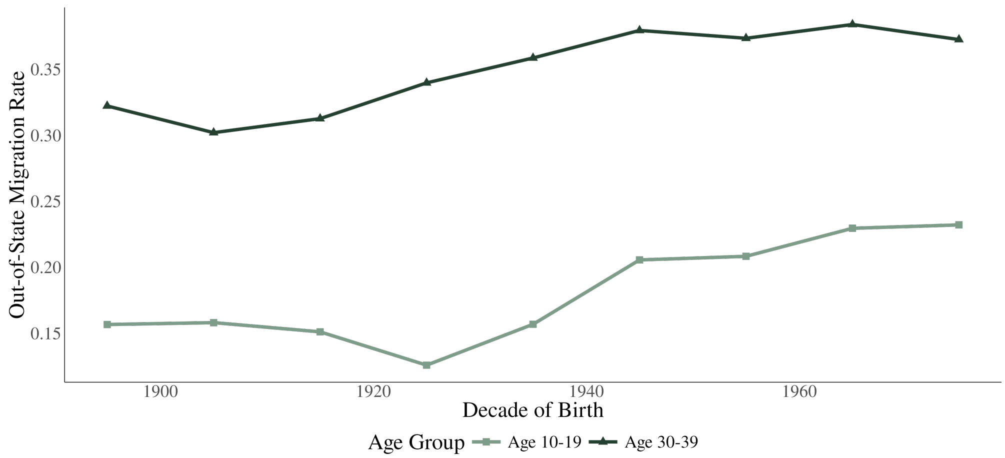
Table 1.1: Lifetimes Returns to Networked Migration

	Zipcode Inc Death (logged)					
		OLS			IV	
	State	Region	Pacific	State	Region	Pacific
	(1)	(2)	(3)	(4)	(5)	(6)
State Mover	0.05***			0.46***		
	(0.003)			(0.08)		
Region Mover		0.08***			0.53***	
		(0.004)			(0.09)	
Pacific Mover			0.61***			0.69***
			(0.006)			(0.13)
Observations	145,723	145,692	128,688	144,203	144,171	127,270
F-stat				228.67	201.38	103.16
1940 State by County FE	✓	✓	✓	✓	✓	✓
Ship type FE	✓	✓	✓	✓	✓	✓
First Quarter FE	✓	✓	✓	✓	✓	✓

NOTES: This table reports coefficients from Equation (1.12) showing the returns to migration by time of death (lifetime). Columns 1-3 report OLS estimates, while Columns 4-6 report IV estimates. In all specifications, the outcome is logged median household income of the last zip code the person lived in prior to their death. Income is reported for the year 2000 in nominal dollars. State mover is an indicator for if a person moved across state lines. Region mover is an indicator if a person moved between Census regions. Pacific mover is an indicator if someone who was not previously living in the Pacific Census division moved to a state in that division (excluding Alaska and Hawaii). In columns 4-6, the instrument is constructed as the predicted probability that a person will move out-of-state, out-of-region, or to a Pacific state using estimates from the discrete choice model described in Section 1.5. Standard errors are clustered at the ship-level.

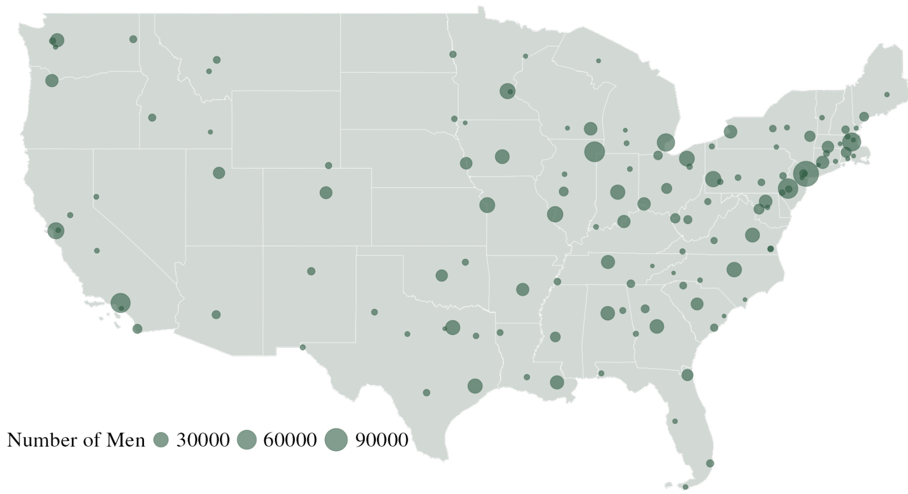
1.10 Additional Figures

Figure 1.10: Probability of living in different state than birthplace for white men



NOTES: This figure displays interstate geographic mobility for white men born 1890-1979. The dark green line shows the share of men living outside their state of birth for men ages 30-39, by birth decade. The lighter green line shows the share of men living outside their state of birth for men ages 10-19, by birth decade. The difference between these two lines is shown in Figure 1.1. Data source: Decennial census public use samples (1900-2010) and American Community Survey (2005-2024).

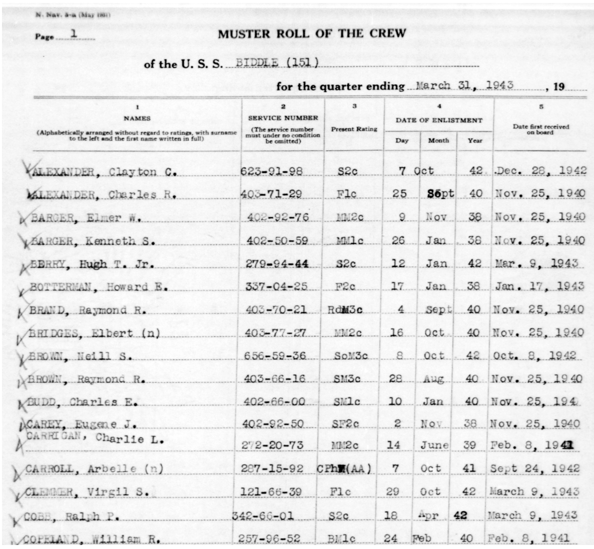
Figure 1.11: Map of Cities of Enlistment



NOTES: This figure depicts the 133 cities where at least 100 sailors enlisted in the muster rolls. The size of each circle is scaled to the number of sailors who reported enlisting into the city. Data source: Place of enlistment is recorded either through direct report on the report of changes or through service number. Ranges of service numbers were assigned to each receiving station in the US. See 1.11.2 for more details on data preparation.

Figure 1.12: Navy Muster Rolls

Panel A: Examples of Muster Roll and Quarterly Report of Changes for USS Biddle



N. Nav. 5-a (May 1931)

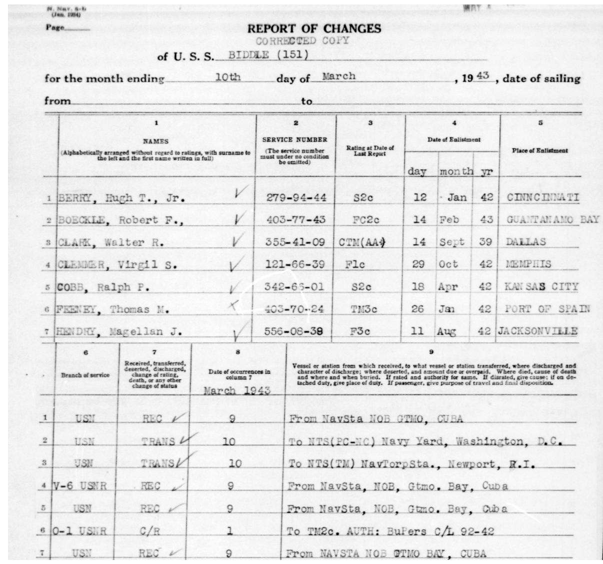
Page 1

MUSTER ROLL OF THE CREW

of the U. S. S. BIDDLE (151)

for the quarter ending March 31, 1943, 19

1 NAMES (Alphabetically arranged without regard to ratings, with surname to the left and the first name written in full)	2 SERVICE NUMBER (The service number must under no condition be omitted)	3 Present Rating	4 DATE OF ENLISTMENT Day	Month	Year	5 Date first received on board
ALEXANDER, Clayton C.	623-91-98	S2c	7	Oct	42	Dec. 28, 1942
ALEXANDER, Charles R.	403-71-29	F1c	25	Sept	40	Nov. 25, 1940
BARGER, Elmer W.	402-92-76	MM2c	9	Nov	38	Nov. 25, 1940
BARGER, Kenneth S.	402-50-59	MM1c	26	Jan	38	Nov. 25, 1940
BERRY, Hugh T. Jr.	279-94-44	S2c	12	Jan	42	Mar. 9, 1943
BOTTERMAN, Howard E.	337-04-25	F2c	17	Jan	38	Jan. 17, 1943
BRAND, Raymond R.	403-70-21	RdM3c	4	Sept	40	Nov. 25, 1940
BRIDGES, Elbert (n)	403-77-27	MM2c	16	Oct	40	Nov. 25, 1940
BROWN, Neill S.	656-59-36	SoM3c	8	Oct	42	Oct. 8, 1942
BROWN, Raymond R.	403-66-16	SM3c	28	Aug	40	Nov. 25, 1940
BUDD, Charles E.	402-66-00	SM1c	10	Jan	40	Nov. 25, 194
CAREY, Eugene J.	402-92-50	SF2c	2	Nov	38	Nov. 25, 1940
CARRIGAN, Charlie L.	272-20-73	MM2c	14	June	39	Feb. 8, 1941
CARROLL, Arbelle (n)	287-15-92	CPhM(AA)	7	Oct	41	Sept 24, 1942
CLEMMER, Virgil S.	121-66-39	F1c	29	Oct	42	March 9, 1943
COBB, Ralph P.	342-66-01	S2c	18	Apr	42	March 9, 1943
COPELAND, William R.	257-96-52	BM1c	24	Feb	40	Feb. 8, 1941



N. Nav. 5-b (Jan. 1934)

Page

REPORT OF CHANGES

CORRECTED COPY

of U. S. S. BIDDLE (151)

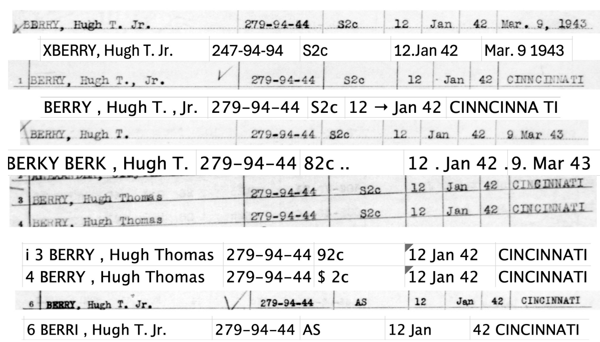
for the month ending 10th day of March, 1943, date of sailing

from to

	1 NAMES (Alphabetically arranged without regard to ratings, with surname to the left and the first name written in full)	2 SERVICE NUMBER (The service number must under no condition be omitted)	3 Rating at Date of Last Report	4 Date of Enlistment day	month	yr	5 Place of Enlistment
1	BERRY, Hugh T., Jr.	279-94-44	S2c	12	Jan	42	CINNCINNATI
2	BOECKLE, Robert F.,	403-77-43	FC2c	14	Feb	43	GUANTANAMO BAY
3	CLARK, Walter R.	355-41-09	CTM(AA)	14	Sept	39	DALLAS
4	CLEMMER, Virgil S.	121-66-39	F1c	29	Oct	42	MEMPHIS
5	COBB, Ralph P.	342-66-01	S2c	18	Apr	42	KANSAS CITY
6	FEENEY, Thomas M.	403-70-24	TM3c	26	Jan	42	PORT OF SPAIN
7	HENDRY, Magellan J.	556-08-38	F3c	11	Aug	42	JACKSONVILLE

	6 Branch of service	7 Received, transferred, deserted, discharged, change of rating, death, or any other change of status	8 Date of occurrences in column 7 March 1943	9 Vessel or station from which received, to what vessel or station transferred, where discharged and character of discharge; where deserted, and amount due or overpaid. Where died, cause of death and where and when buried. If rated and authority for same. If disrated, give cause; if on detached duty, give place of duty. If passenger, give purpose of travel and final disposition.
1	USN	REC	9	From NavSta NOB GTMO, CUBA
2	USN	TRANS	10	To NTS(PC-NC) Navy Yard, Washington, D.C.
3	USN	TRANS	10	To NTS(TM) NavTorpSta., Newport, R.I.
4	V-6 USNR	REC	9	From NavSta, NOB, Gtmo. Bay, Cuba
5	USN	REC	9	From NavSta, NOB, Gtmo. Bay, Cuba
6	O-1 USNR	C/R	1	To TM2c. AUTH: BuPers C/L 92-42
7	USN	REC	9	From NAVSTA NOB GTMO BAY, CUBA

Panel B: Example of Identifying "Hugh Berry" from Muster Rolls



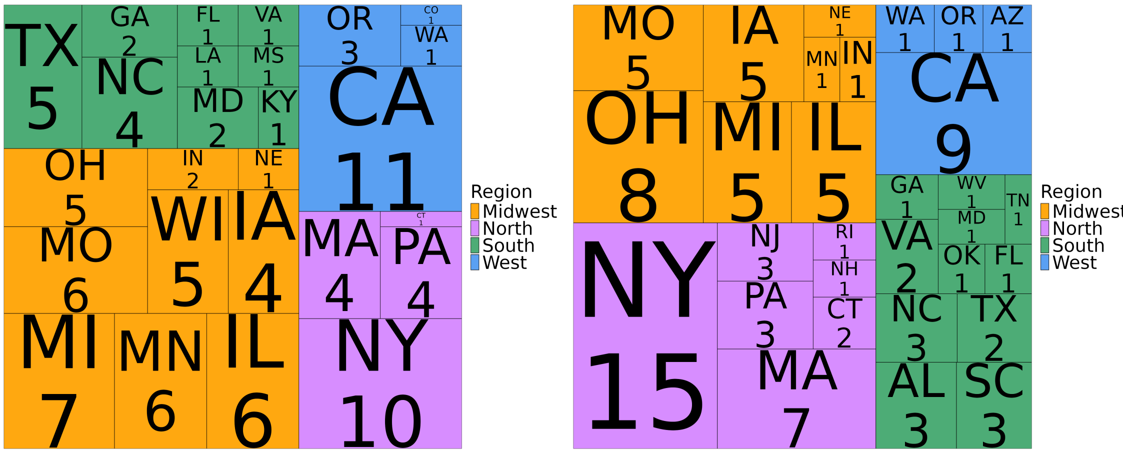
servno	lname	fname	mname	# scans	place	date e	date b
2799444	BERRY	HUGH	THOMAS	8	CINCINNATI OH	1-12-1942	3-9-1943

Panel C: Summary of Cleaned Muster Data

Metric	Value
Total number of people	1,456,484
Mean number of scans per person	9.6
Share on 1 boat	65%
Non-missing name	99%
Non-missing date enlist	61%
Non-missing place	96%
Non-missing ethnicity	93%
Non-missing HOH Occ. Score	89%

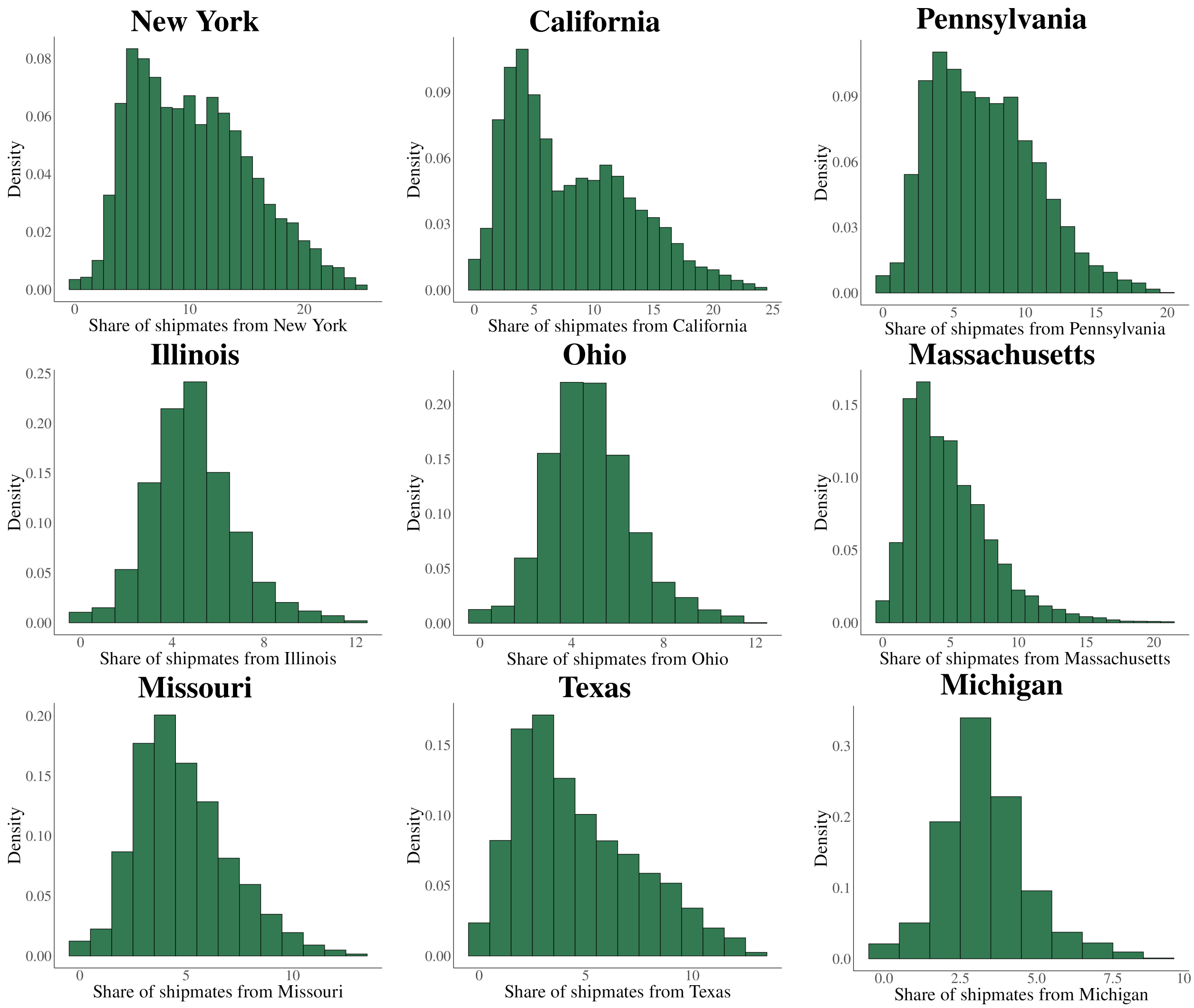
NOTES: Panel A displays example scans from muster rolls from the USS Biddle (DE-151). The left image is a quarterly census for for the quarter ending March 31, 1943, while the right image is a monthly report of changes from March 10, 1943. Panel B tracks all scanned entries and OCR transcriptions of Navy sailor Hugh Berry in the muster rolls, with the final row showing his cleaned characteristics post-processing. Panel C presents summary statistics for the cleaned data.

Figure 1.13: Example Ship Exposure: USS Mero and USS Seacat



NOTES: This plot shows the geographic mix of shipmates for two different individuals on two different ships: the submarine USS Mero and the submarine USS Seacat. For each destination state, the number of distinct individuals each person is exposed to is represented on the plot.

Figure 1.14: Shipmate Distribution by State



NOTES: This figure displays histograms of shipmate composition for the nine states with the largest representation in the Navy muster roll data. Each panel shows the distribution across individuals of the share of their shipmates who originated from that state. For example, the New York panel shows what fraction of each sailor's shipmates were from New York.

Table 1.2: Summary Table of Ship Characteristics

	All Ships	Minesweepers	Destroyers
Number of Ships	5,019	618	381
Median Num. Ppl/Ship/Quarter	84	46	345
Median Num Quarters/Ship	7	9	10
Characteristics of Shipmates: Median [IQR]			
Num. of States	26.6 [15.50]	20.9 [10.0]	37.5 [5.9]
Share West (%)	11 [15]	12 [19]	11 [16]
Share North (%)	28 [20]	28 [23]	30 [18]
Share Midwest (%)	27 [10]	28 [23]	26 [18]
Share South (%)	27 [10]	26 [9]	27 [9]
# Ethnic Groups	15.6 [5.60]	13.1 [4.1]	18.0 [0.3]
Mean Occ Score	25.6 [1.03]	25.5 [1.2]	25.5 [0.6]
Occ in Farming (%)	18 [5]	19 [6]	19 [4]

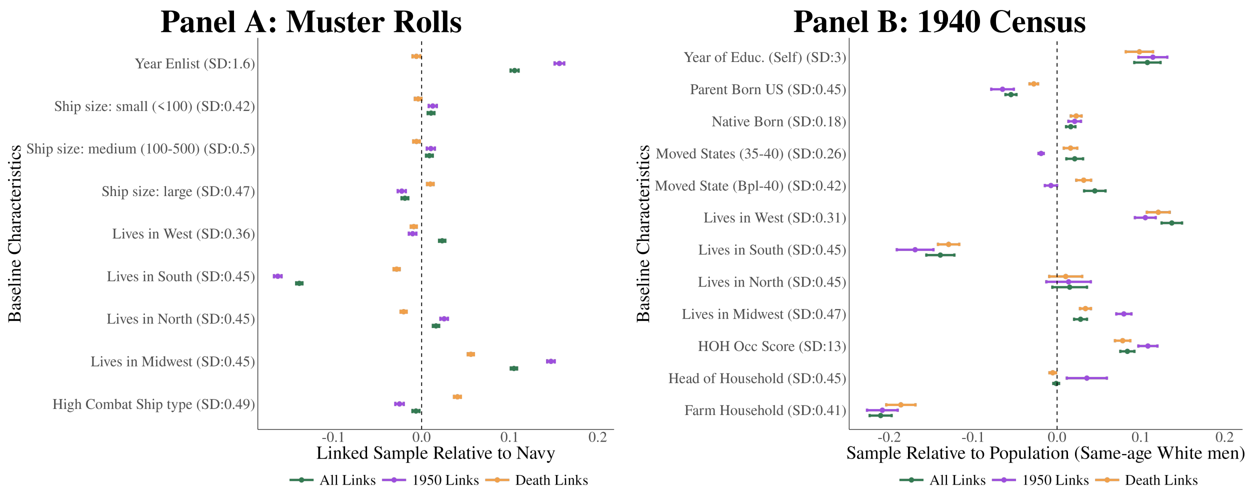
NOTES: This table presents summary statistics for Navy ships active between 1942 and 1945. Column 1 reports statistics for all ships in the sample, while columns 2 and 3 focus on two specific ship categories: minesweepers and destroyers. The first panel reports ship-level statistics. The second panel presents the distribution of shipmate characteristics, reporting the median and interquartile range [in brackets] across ships. Geographic shares reflect the distribution of sailors' pre-war state of residence across four Census regions. Ethnic groups are identified using name-based ethnicity classification following Abramitzky et al. (2). Occupational score and share in farming are constructed using characteristics of heads of household with the same name and state of residence in the 1940 Census. Ships with fewer than 10 sailors in a quarter are excluded from the sample.

Table 1.3: Summary of Linked Sample

	Navy	Census 1940	Numident	FindAGrave
Pre-link restricted sample	1,355,514	26,465,390	14,734,046	1,836,012
Bilateral Links:				
Census 1940	343,330	–	–	–
Numident	428,213	4,420,699	–	–
FindAGrave	202,436	473,798	–	–
Number of Links from Muster Rolls:				
	1940 Census	1950 Census	Death (Any)	
Number of Links	478,000	266,000	578,000	

NOTES: Panel A compares three linked samples to the general population of rank-and-file personnel aboard Navy ships. Comparisons are made for three separate samples: All Links (Navy men linked to 1940 Census), 1950 Links (Navy men linked to 1940 and 1950 Census), and Death Links (Navy men linked to 1940 and death). Each row reports the estimates from a separate estimation; for each variable the characteristic is regressed on being in the linked sample. The coefficient for this is reported scaled to a standard deviation of the outcome variable. Panel B replicates this procedure but comparing linked individuals to same-age white men in the 1940 Census.

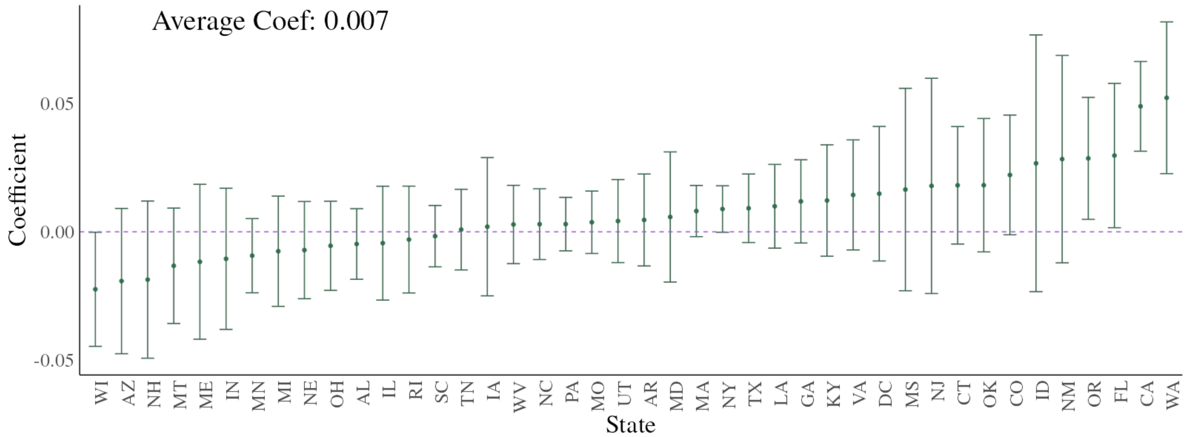
Figure 1.15: Benchmarking Linked Sample



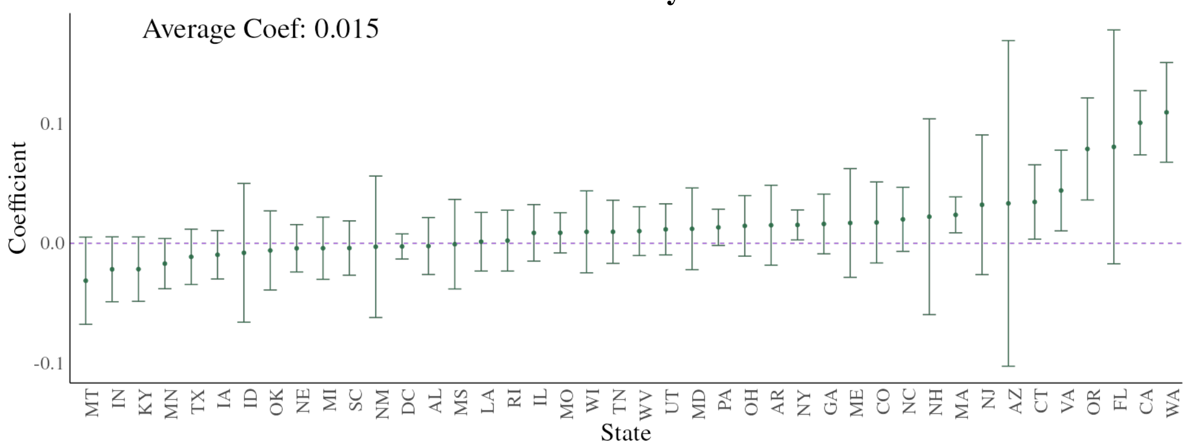
NOTES: This figure compares three linked samples to the general population of rank-and-file personnel aboard Navy ships. Panel A compares characteristics between the linked sample and all men in the muster rolls data. Panel B compares characteristics between the linked sample and white men aged 16-40 in the 1940 Census. All Links refers to Navy men linked to the 1940 Census, 1950 Links refers to Navy men linked to both 1940 and 1950 Census records, and Death Links refers to Navy men linked to both 1940 Census and death records. Each row reports the estimates from a separate regression where the characteristic is regressed on being in the linked sample. The coefficient is reported scaled to a standard deviation of the outcome variable.

Figure 1.16: Directed Migration to States

Panel A: Results in 1950



Panel B: Results by Death



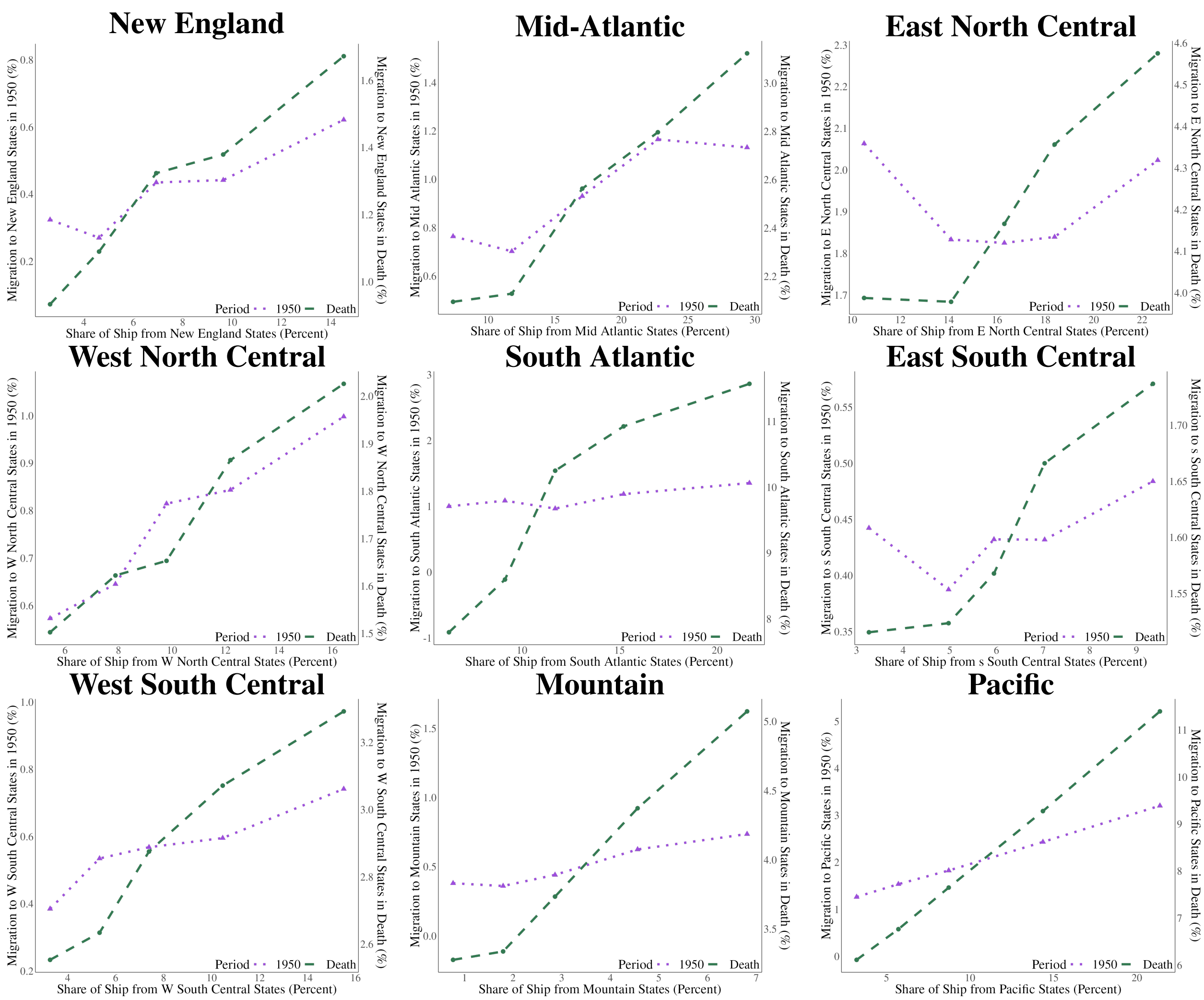
NOTES: These plots show results from Equation (1.3) at the state-level. For each state depicted on the x-axis, the coefficient represents regressing the likelihood someone moves to state on the share of shipmates from that state, excluding individuals originally from that state. Along with the point-estimate, the 95% confidence interval is reported for each state. Standard errors are clustered at the ship-level. Panel A reports estimates for all 48 contiguous state in 1950, while Panel B reports estimates by death. The average coefficient is reported at the top of each plot.

Figure 1.17: Directed Migration to Census Regions: Coefficients

	1950				Death			
	Midwest	North	South	West	Midwest	North	South	West
Variable	**(1)**	**(2)**	**(3)**	**(4)**	**(5)**	**(6)**	**(7)**	**(8)**
Share of Ship	0.002	0.012***	0.018**	0.010***	0.010	0.026***	0.029*	0.046***
from Region	(0.007)	(0.003)	(0.008)	(0.003)	(0.011)	(0.005)	(0.015)	(0.005)
Sample	Non-Midwest	Non-North	Non-South	Non-West	Non-Midwest	Non-North	Non-South	Non-West
Dep Var Mean	0.022	0.012	0.018	0.007	0.053	0.033	0.140	0.040
Observations	117,493	129,518	142,318	169,242	172,089	181,944	188,154	233,820
R^2	0.02	0.01	0.02	0.01	0.02	0.01	0.02	0.03
Within R^2	0.00	0.00	0.00	0.00	0.00	0.00	0.00	0.00
State FE	✓	✓	✓	✓	✓	✓	✓	✓
Category FE	✓	✓	✓	✓	✓	✓	✓	✓
Time FE	✓	✓	✓	✓	✓	✓	✓	✓

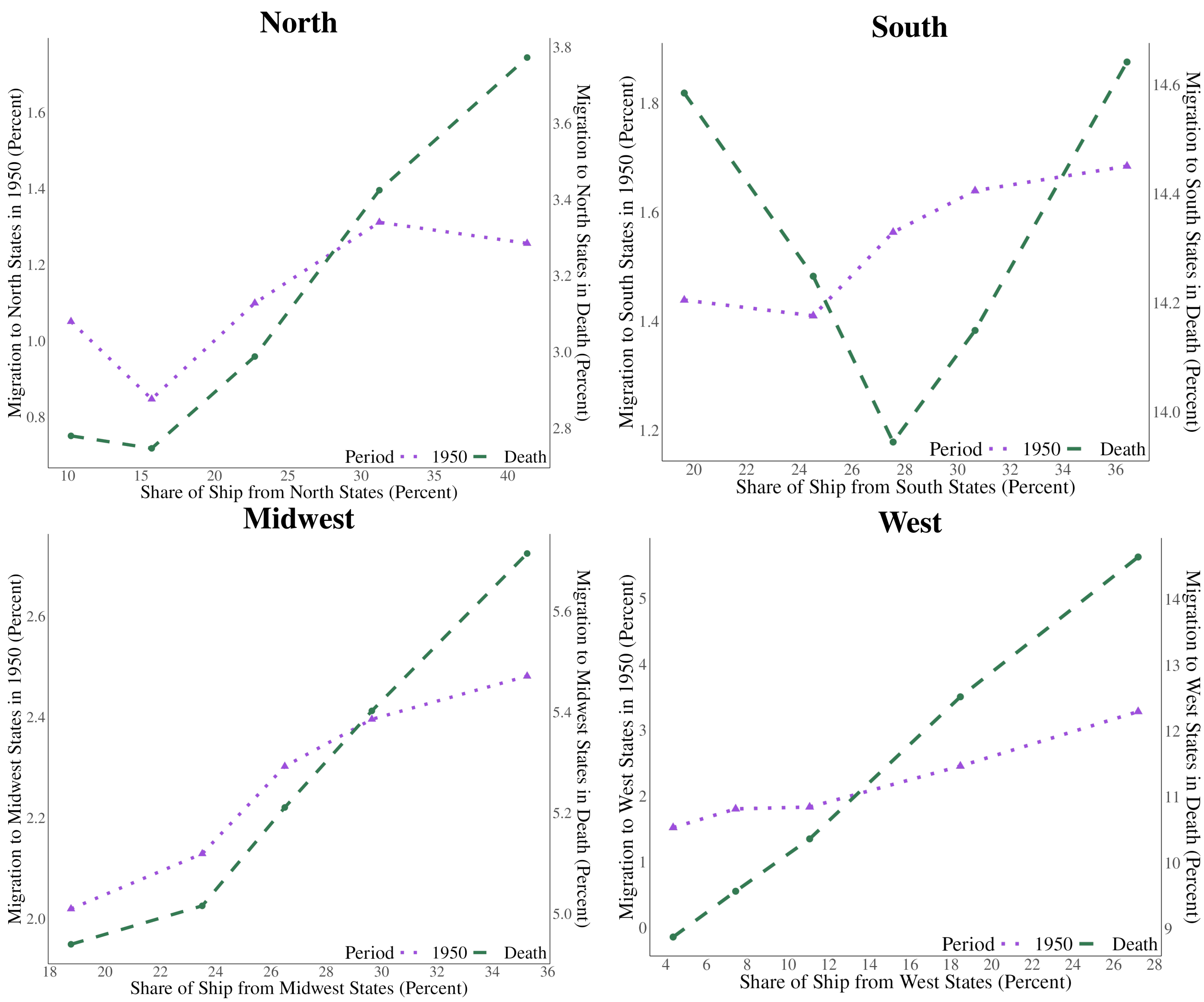
NOTES: This plot is analogous to Figure 1.4 except it shows directed migration at the regional level. This figure shows the impact of Navy ship networks on directed migration across Census region, by showing estimates from equation (1.3). The table presents $\beta^{j,t}$ showing the effect of increased exposure to shipmates from each Census region on migration to that region, measured both in 1950 and at time of death. Standard errors are clustered at the ship-level.

Figure 1.18: Directed Migration (Census Divisions)



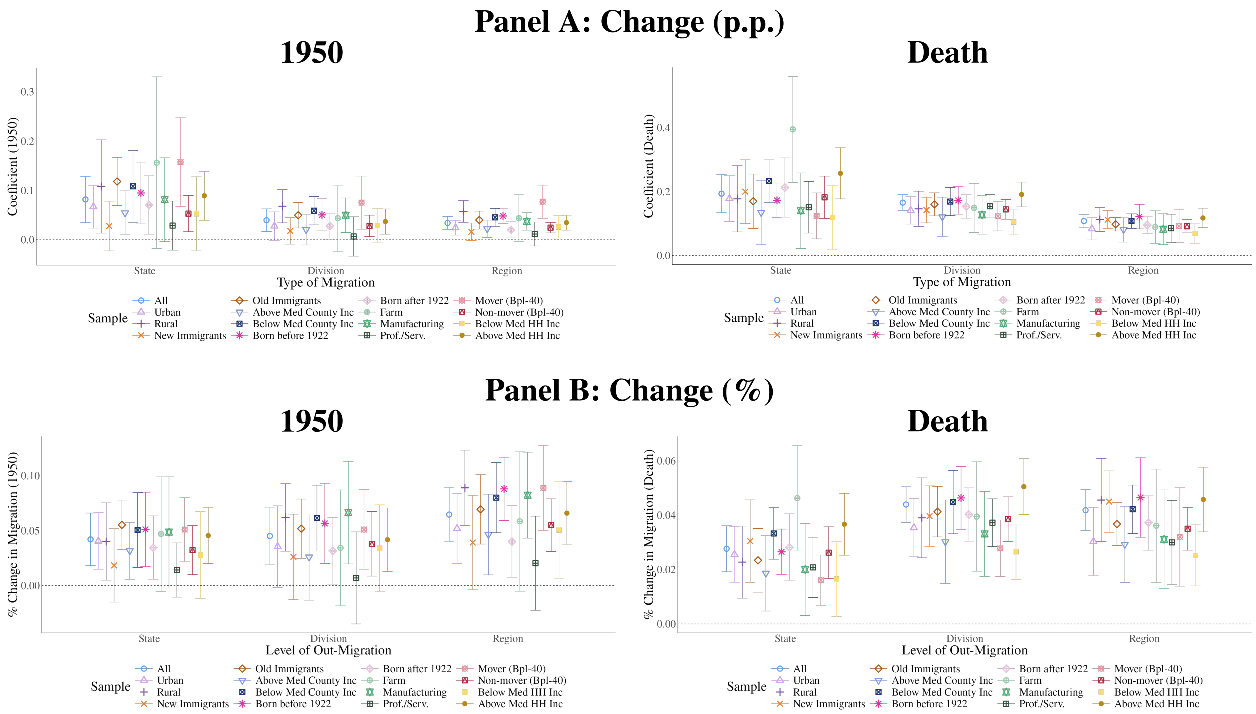
NOTES: This plot is a continuation of 1.4. Each plot shows the relationship between unconditional migration to Census divisions and shipmate characteristics. Each Census divisions is represented in a separate plot displaying unconditional migration by 1950 (purple, left axis) and by death year (green, right axis). Both axes span an equal range but are re-leveled to accommodate different baseline migration rates in each period. The average unconditional migration rate is calculated for each quintile for each home state, with a sample population-weighted average across home states reported on the plot.

Figure 1.19: Directed Migration to Census Regions



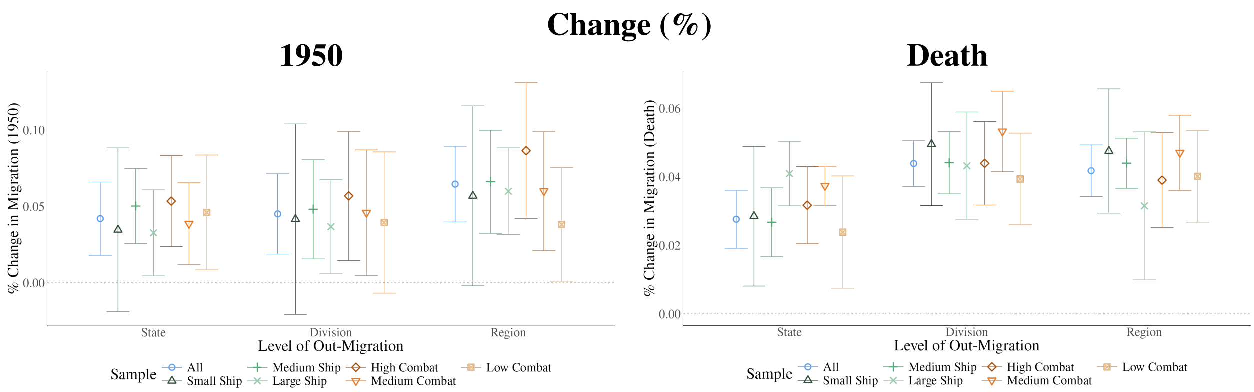
NOTES: This plot is a continuation of 1.17. This plot is analogous to Figure 1.18 except it shows directed migration at the regional level. Each plot shows the relationship between unconditional migration to Census regions and shipmate characteristics. Each Census region is represented in a separate plot displaying unconditional migration by 1950 (purple, left axis) and by death year (green, right axis). Both axes span an equal range but are re-leveled to accommodate different baseline migration rates in each period. The average unconditional migration rate is calculated for each quintile for each home state, with a sample population-weighted average across home states reported on the plot.

Figure 1.20: Heterogeneity in Out-migration by Baseline Characteristics



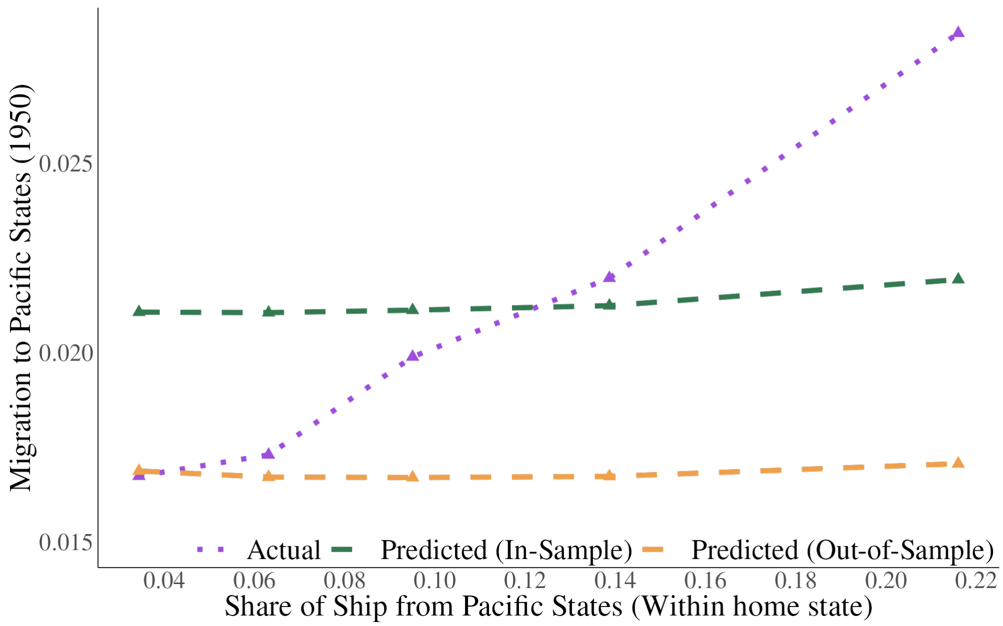
NOTES: This plot depicts results from Equation (1.2) for different sub-samples of the population. Each coefficient depicted shows the result of a single regression restricted to a specific sub-population. On each plot, coefficients are reported for 16 different samples and three different levels of geography. Panel A, reports the raw coefficient, and in Panel B the coefficient is scaled to report percent change from a 1 SD increase in exposure. The 95% confidence interval is also shown for each coefficient. Standard errors are clustered at the ship level.

Figure 1.21: Heterogeneity in Out-migration by Ship Type



NOTES: This plot depicts results from Equation 1.2 for different sub-samples of the population. Each coefficient depicted shows the result of a single regression restricted to a specific sub-population. On each plot, coefficients are reported for 16 different samples and three different levels of geography. Panel A, reports the raw coefficient, and in Panel B the coefficient is scaled to report percent change from a 1 SD increase in exposure. The 95% confidence interval is also shown for each coefficient. Standard errors are clustered at the ship level.

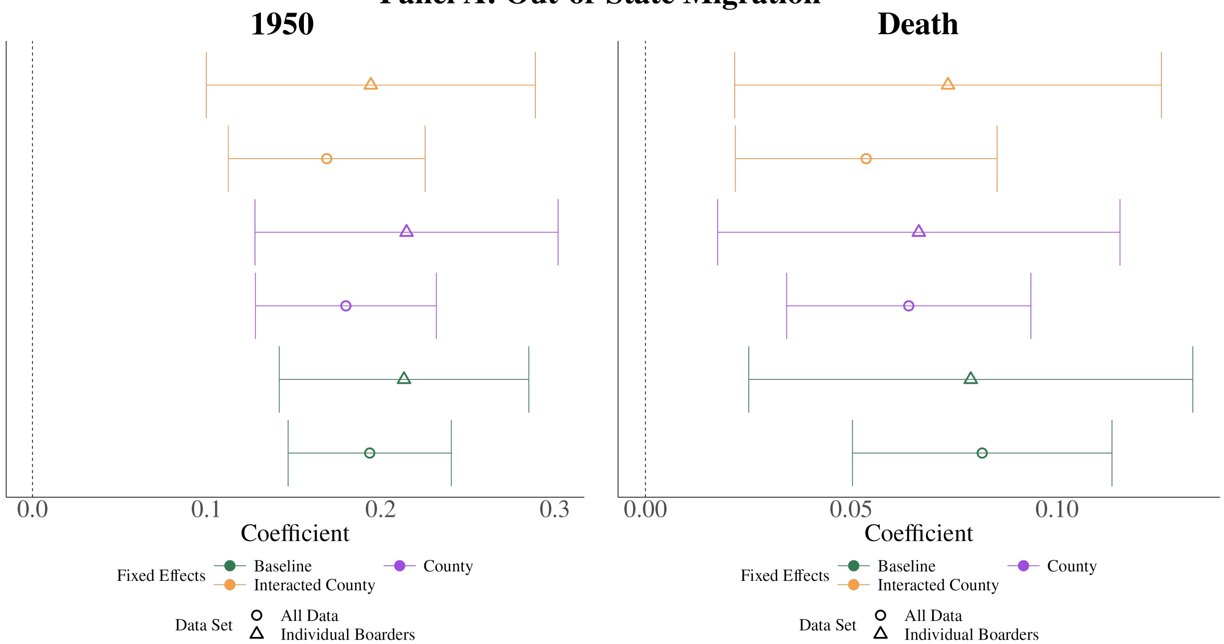
Figure 1.22: Robustness: Predicted vs. Actual Migration to Pacific States



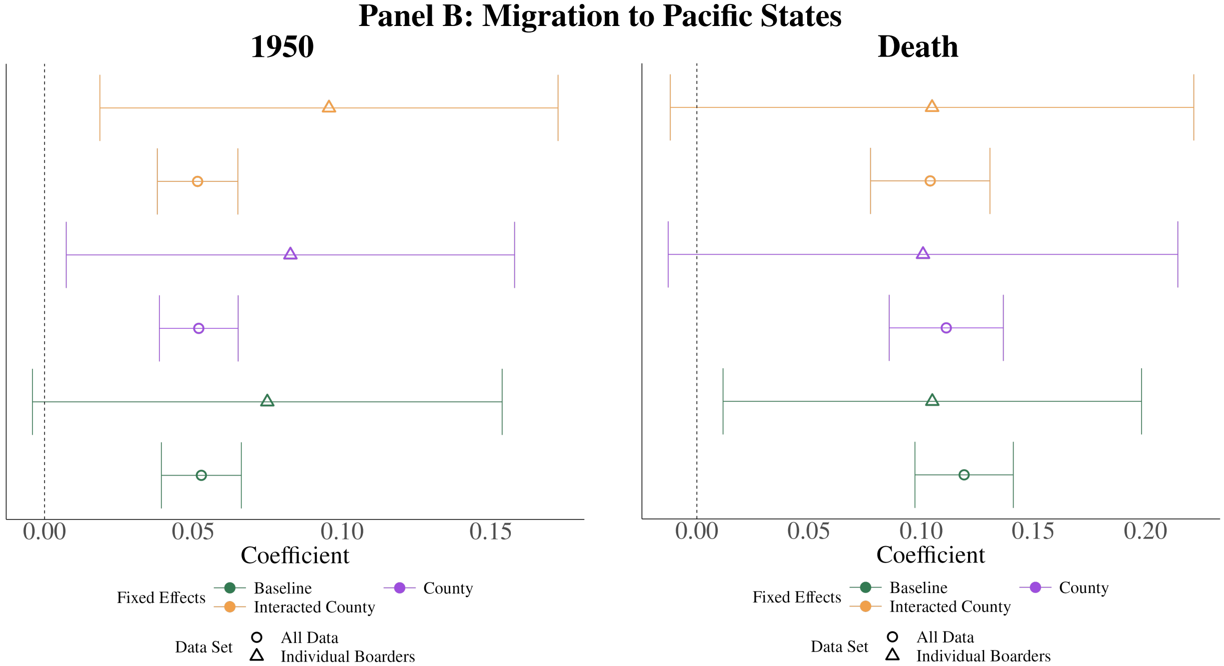
NOTES: This figure demonstrates the relationship between predicted and actual migration to Pacific states in 1950. Individuals from the same home state h in the linked Navy sample are are separated into five equal sized bins by the share of their shipmates from Pacific states. For each person, migration to Pacific states is predicted using pre-war characteristics from the 1940 Census (age, education, household size, home ownership, farm status, occupational score, and state and county of residence). This predicted migration is generated from two samples: individuals in the Navy sample (in-sample,green) and white men born between 1905 and 1928 (out-of-sample,yellow). The purple dotted line plots actual 1950 migration to Pacific states against these predicted values and is constructed analogously to plots in Figure 1.18. For each of the three series–predicted (in-sample,green), predicted (out-of-sample,yellow), and actual (purple)–the average unconditional migration rate is calculated for each quintile for each home state, with a sample population-weighted average across home states reported on the plot.

Figure 1.23: Robustness Exercises

Panel A: Out-of-State Migration

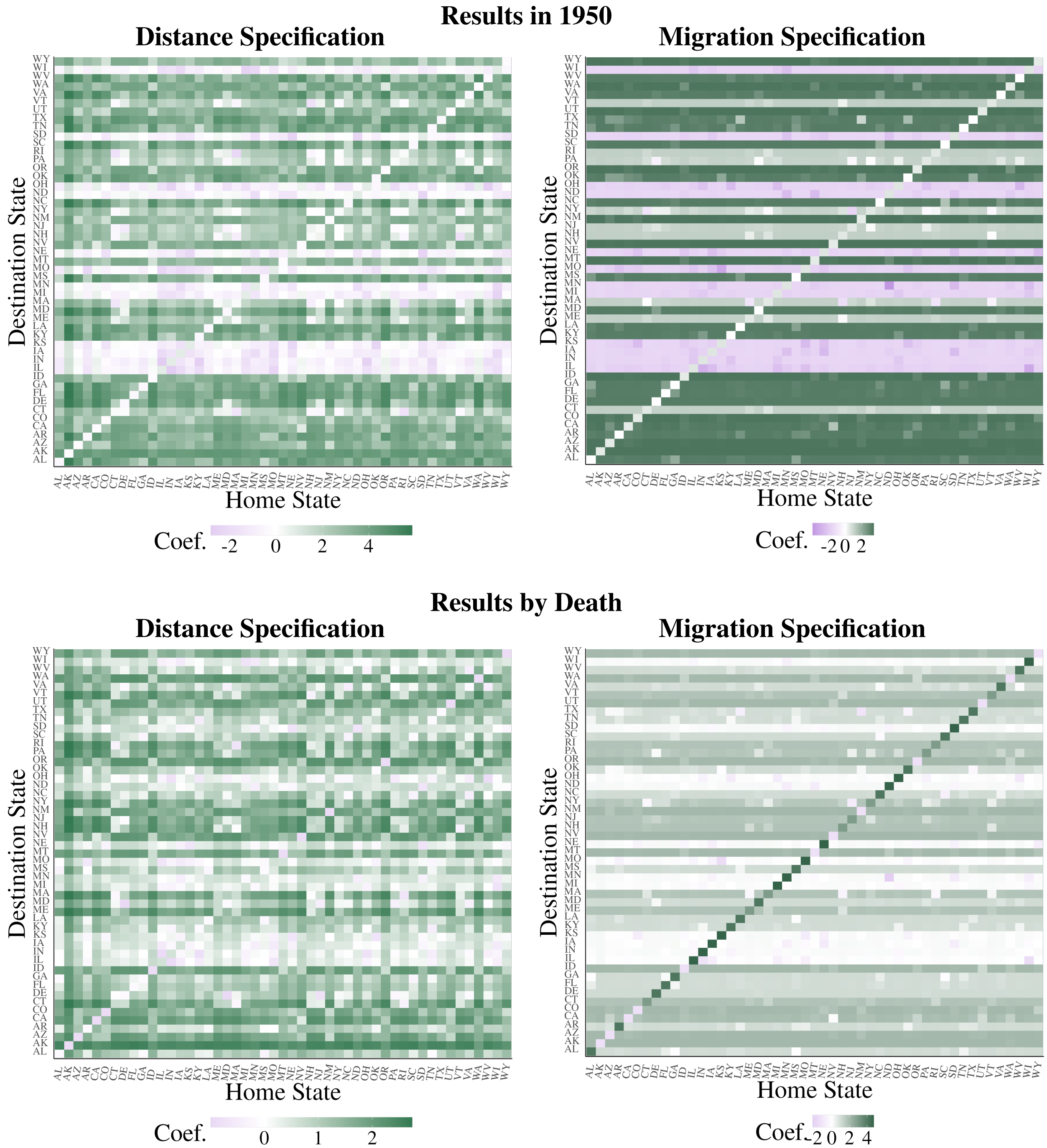


Panel B: Migration to Pacific States



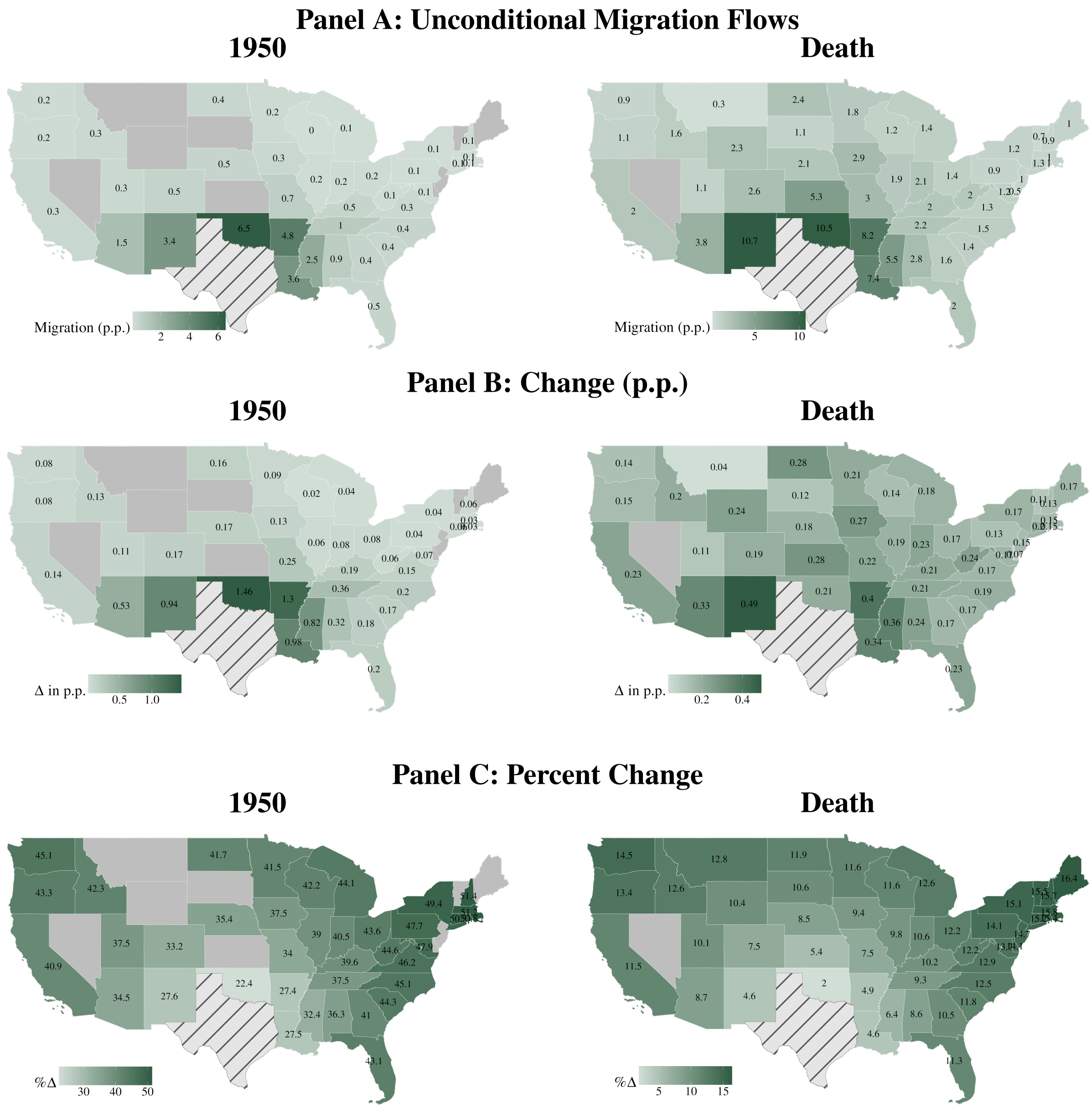
NOTES: This figure presents robustness checks for the main migration estimates. Panel A shows estimates from Equation (1.2) for out-of-state migration, while Panel B shows estimates from Equation (1.3) migration to Pacific states. Each panel reports estimates for both 1950 outcomes (left) and outcomes at time of death (right). Each point represents a separate regression estimate under different specifications and sample restrictions. The baseline specification includes pre-war state, quarter-of-enlistment, and ship type fixed effects. Additional specifications add: (i) county-level fixed effects and (ii) fully interacted fixed effects (with county). Sample restrictions include: (i) limiting to individuals whose exposure to same-day boarders is less than 10%. Horizontal bars represent 95% confidence intervals with standard errors clustered at the ship level.

Figure 1.24: Heatmap of Gravity Coefficients β_{hd}^{dest} and β_{d}^{home}



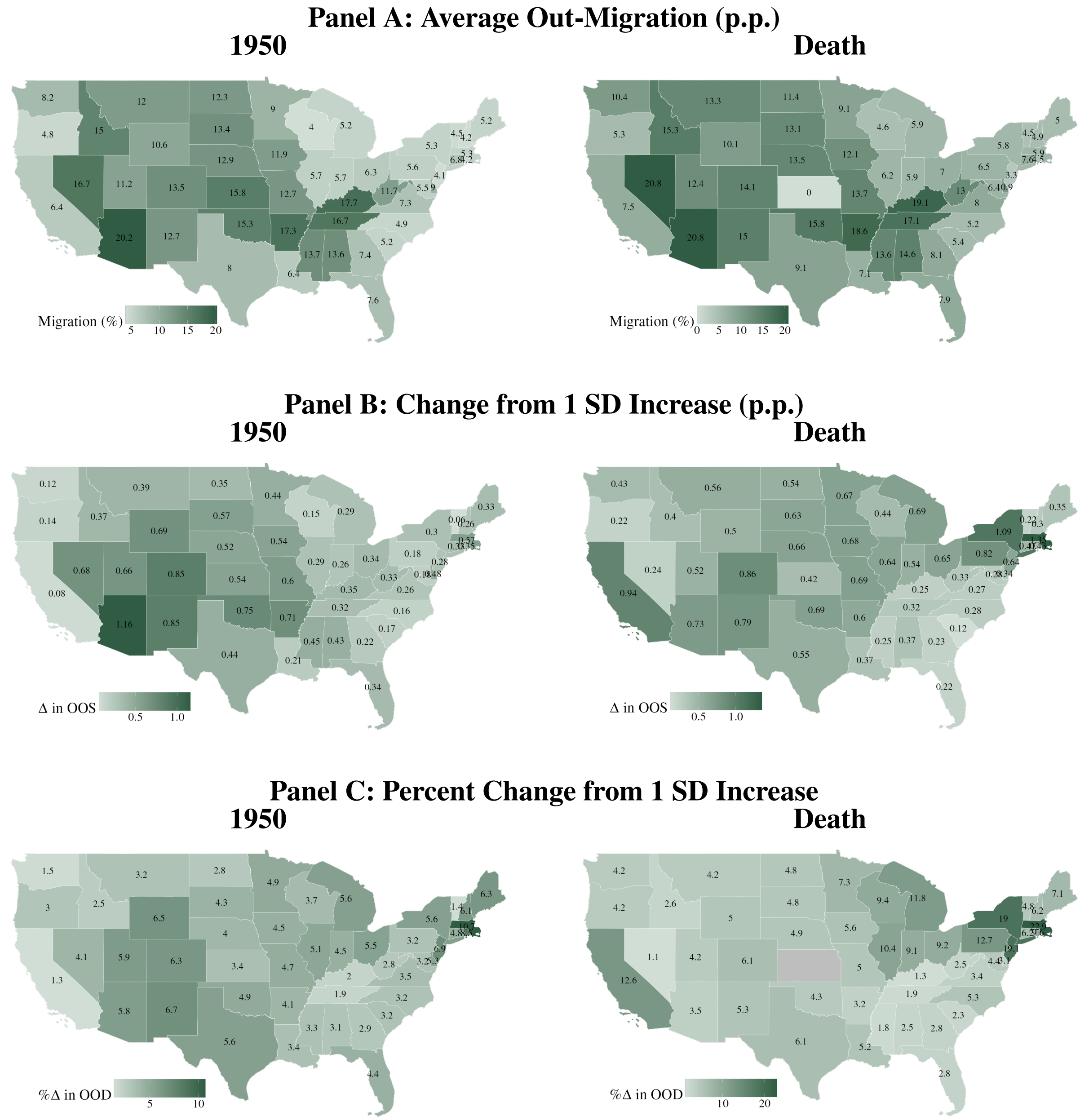
NOTES: This figure displays heatmaps of the gravity model coefficients estimated from equation (1.4). Each cell represents a home-destination state pair, with color intensity indicating the strength of network effects. The top row shows coefficients for 1950 outcomes, while the bottom row shows coefficients for outcomes at time of death. The left panels show coefficients from the distance specification, where network effects are parameterized as a function of log distance between states. The right panels show coefficients from the migration specification, where network effects are parameterized using pre-existing migration flows between 1935 and 1940. Diagonal elements represent β_{d}^{home}, measuring the impact of ties to one's home state. Off-diagonal elements represent β_{hd}^{dest}.

Figure 1.25: Impact of a 10 p.p. Increase in Texas Share on Migration



NOTES: This figure is an analog to Figure 1.6, but for measuring the impact of increased exposure to Texan shipmates on migration to Texas from each origin state. The counterfactual compares the predicted migration response from serving on a high-exposure ship (10 percentage points higher) to a low-exposure ship as described in Equation (1.8) using discrete choice estimates reported in Figure 1.5. Panel A reports average unconditional migration flows to California by origin state by 1950 and by a person's death. Panel B shows the increase in migration probability of serving on a high-exposure ship relative to a low-exposure ship in percentage points. Panel C displays the percent increase in migration probability relative to the average unconditional migration rate (Panel B/Panel A). The left map in each panel reports results in 1950, while the right map in each panel reports results by death.

Figure 1.26: Impact of 1 SD increase in ship exposure on division migration, by state of origin

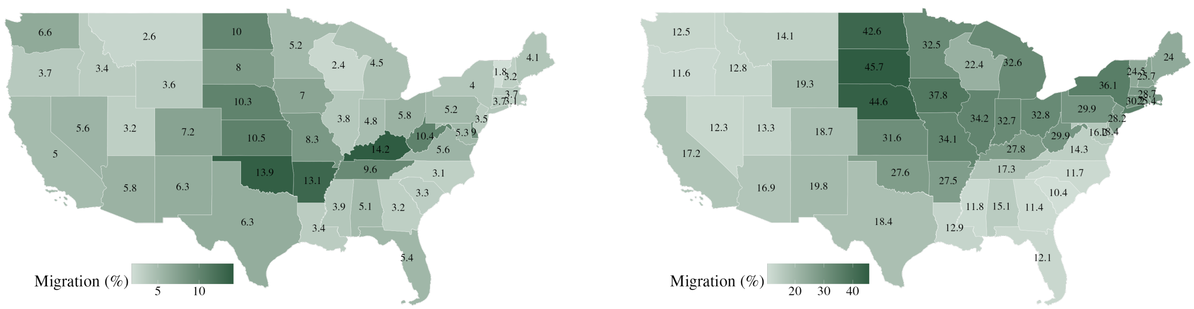


NOTES: This figure is analogous to Figure 1.7, but shows results for migration across Census divisions. Following Equation (1.9) and using estimates from Figure 1.5, the counterfactual, $(\Delta P_{div_d \neq div_{h(i)}|it})$, computes the difference in probability of moving out-of-division when assigned to one's actual ship versus a ship with average composition for their pre-war residence state $h(i)$. Panel A displays average out-of-division migration rates in 1950 and by time of death for men in the linked sample. Panel B reports one standard deviation in $(\Delta P_{div_d \neq div_{h(i)}|it})$ across individuals from state h, representing the change in migration probability from a one standard deviation increase in effective shipmate variation. Panel C shows this effect as a percentage change relative to average out-migration rates from each state (Panel B/Panel A). Each panel presents results for both 1950 (left) and time of death (right).

Figure 1.27: Impact of 1 SD increase in ship exposure on region migration, by state of origin

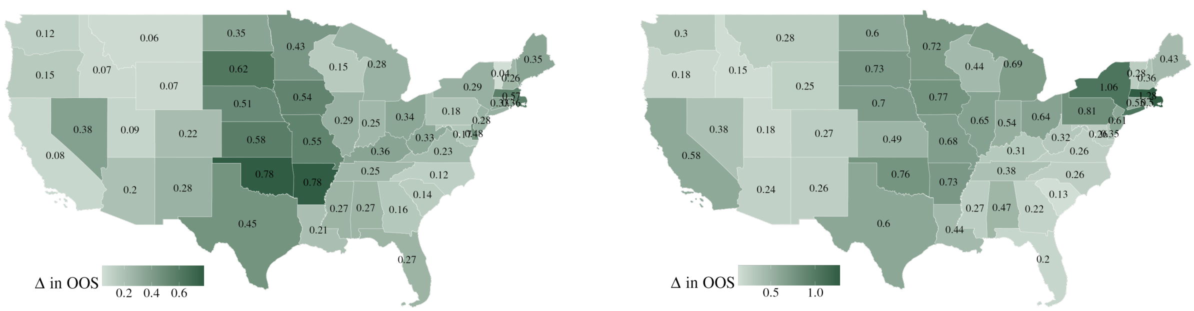
Panel A: Average Out-Migration (p.p.)

1950 **Death**



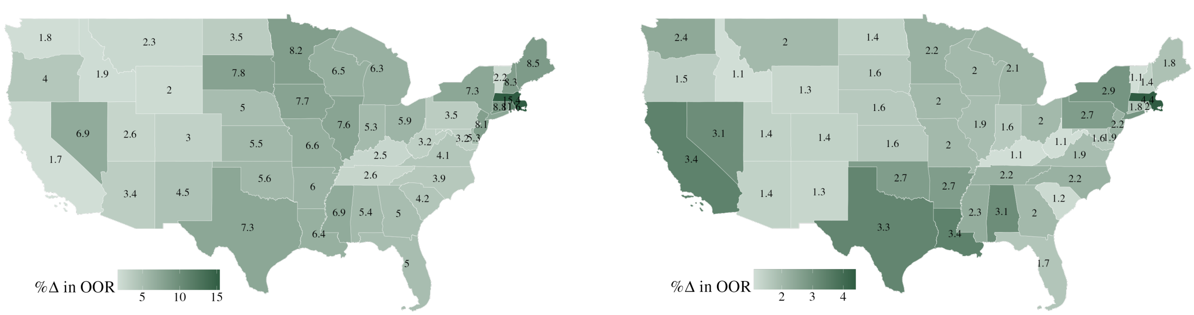
Panel B: Change from 1 SD Increase (p.p.)

1950 **Death**



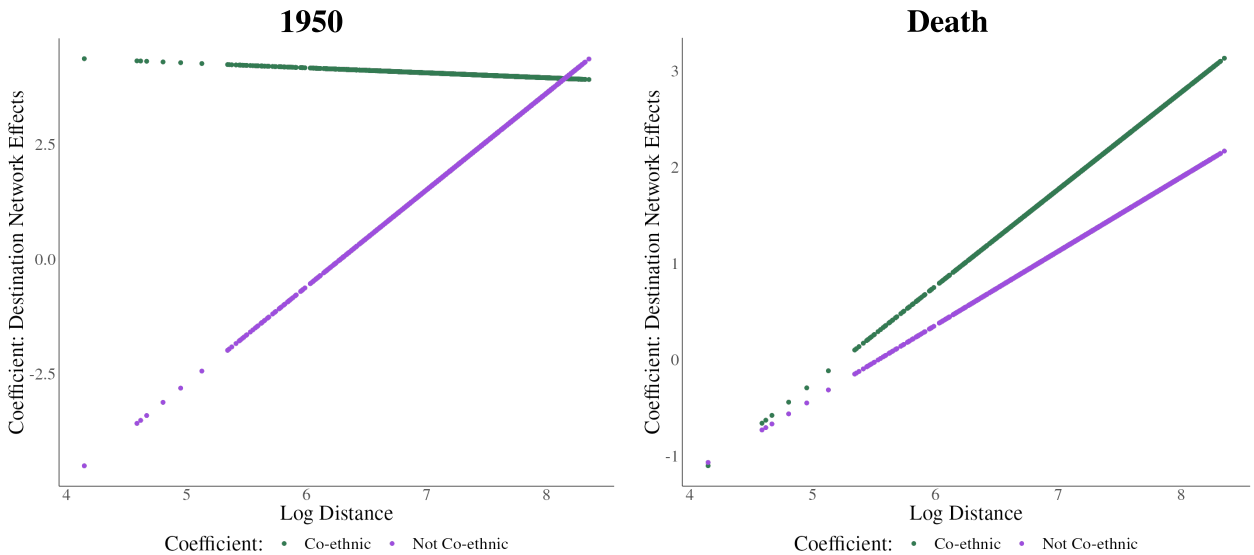
Panel C: Percent Change from 1 SD Increase

1950 **Death**



NOTES: This figure is analogous to Figure 1.7, but shows results for migration across Census regions. Following Equation (1.9) and using estimates from Figure 1.5, the counterfactual, $(\Delta P_{reg_d \neq reg_{h(i)}|it})$, computes the difference in probability of moving out-of-region when assigned to one's actual ship versus a ship with average composition for their pre-war residence state $h(i)$. Panel A displays average out-of-region migration rates in 1950 and by time of death for men in the linked sample. Panel B reports one standard deviation in $(\Delta P_{reg_d \neq reg_{h(i)}|it})$ across individuals from state h, representing the change in migration probability from a one standard deviation increase in effective shipmate variation. Panel C shows this effect as a percentage change relative to average out-migration rates from each state (Panel B/Panel A). Each panel presents results for both 1950 (left) and time of death (right).

Figure 1.28: Co-ethnics and Network Formation



NOTES: This figure plots the predicted coefficient β_{hd}^{dest} from equation (1.11) against logged distance between states, separately for co-ethnic and different-ethnicity shipmates. The left panel shows results for migration by 1950, while the right panel shows results for migration by time of death. Logged distance between states varies from 4.1 (RI to MA) to 8.4 (ME to CA), with the median origin-destination pair having logged distance of 7.3 (MA to WI).

Table 1.4: Network Formation: Occupational Proxmity and Pre-War Occupational Score

	Rating Group		Occupation Score	
	1950	Death	1950	Death
β_{dest}	-8.39***	-4.42***	-9.75***	-4.45***
	(2.18)	(1.14)	(1.96)	(1.04)
$\beta_{logdist}$	1.53***	0.84***	1.70***	0.83***
	(0.31)	(0.16)	(0.28)	(0.14)
β_{rg}	0.21**	0.29***		
	(0.23)	(0.11)		
$\beta_{rg\ logdist}$	-0.03	-0.05**		
	(0.05)	(0.02)		
β_{occ}			-0.00	-0.00
			(0.02)	(0.01)
$\beta_{occ\ logdist}$			0.00	0.00
			(0.00)	(0.00)
β_{home}	-0.14**	0.12***	0.13	0.34***
	(0.25)	(0.12)	(0.22)	(0.10)
$\beta_{home\ rg}$	0.21***	0.21***		
	(0.04)	(0.02)		
$\beta_{home\ occ}$			-0.01	0.00
			(0.00)	(0.00)
R^2	0.65	0.41	0.65	0.41
N	4,348,954	6,756,488	5,536,734	8,430,555

Notes: This table reports discrete choice estimates from Equation (1.13) (Columns 1 and 2) and Equation (1.14) (Columns 3 and 4). Columns 1 and 3 report coefficient estimates for 1950, while columns 2 and 4 report esimates by time of death. Coefficients are estimated by Poisson-Pseudo Maximum Likelihood (PPML) estimation. Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1

Figure 1.29: Returns to Migration IV: Balance

	Migration Instrument (in SD)					
		1950			Death	
	State	Region	Pacific	State	Region	Pacific
	(1)	(2)	(3)	(4)	(5)	(6)
Native Born	0.001	-0.01*	-0.0007	-0.001	-0.004	0.001
	(0.007)	(0.007)	(0.006)	(0.008)	(0.005)	(0.005)
HOH Education	0.0003	0.0002	0.0002	-0.0004	-7.9×10^{-5}	-0.0004
	(0.0004)	(0.0004)	(0.0004)	(0.0004)	(0.0003)	(0.0003)
HOH Occ Score	1.8×10^{-5}	0.0001	-8.7×10^{-5}	0.0001	2.6×10^{-5}	-0.0002*
	(0.0001)	(0.0001)	(0.0001)	(0.0001)	(8.5×10^{-5})	(8.5×10^{-5})
Age	0.001	0.001	9.8×10^{-5}	-8.5×10^{-5}	-3.7×10^{-5}	-0.0006
	(0.002)	(0.002)	(0.001)	(0.001)	(0.0009)	(0.0009)
Age Squared	-2.2×10^{-5}	-2.4×10^{-5}	2.4×10^{-6}	-9.8×10^{-7}	-2.3×10^{-6}	1.6×10^{-5}
	(3.3×10^{-5})	(3.4×10^{-5})	(2.9×10^{-5})	(2.8×10^{-5})	(2×10^{-5})	(2×10^{-5})
Moved States (35-40)	0.008	0.01**	0.01	0.01*	0.01***	0.007*
	(0.006)	(0.006)	(0.007)	(0.005)	(0.004)	(0.004)
Observations	140,278	140,277	124,641	140,877	140,877	124,048
R^2	0.80850	0.83229	0.87108	0.85534	0.92943	0.94543
Within R^2	2.72×10^{-5}	7.45×10^{-5}	5.89×10^{-5}	4.82×10^{-5}	9.96×10^{-5}	0.00011
1940 State by County FE	✓	✓	✓	✓	✓	✓
Ship type FE	✓	✓	✓	✓	✓	✓
First Quarter FE	✓	✓	✓	✓	✓	✓

NOTES: This tables show the underlying variation between the migration instruments defined in Section 1.7 and individual baseline characteristics in 1940. The dependent variable is the migration instrument normalized into standard deviations. The instrument is constructed as the predicted probability that a person will move out-of-state, out-of-region, or to a Pacific state using estimates from the discrete choice model described in Section 1.5. Columns 1-3 show results using migration instruments constructed for 1950 outcomes, while columns 4-6 show results using instruments for lifetime migration. The baseline characteristics are measured using the 1940 Census. Native Born is an indicator for U.S. birth, HOH variables refer to characteristics of the household head in 1940, and Moved States (35-40) indicates migration across state lines between 1935-1940. Standard errors are clustered at the ship-level.

Figure 1.30: Returns to Migration IV: First Stage

	1950			Death		
	State (1)	Region (2)	Pacific (3)	State (4)	Region (5)	Pacific (6)
Migration Instrument	0.76***	0.69***	0.85***	0.70***	0.72***	0.93***
	(0.04)	(0.06)	(0.09)	(0.04)	(0.05)	(0.09)
Observations	173,504	173,503	154,312	177,190	177,158	156,125
R^2	0.13156	0.13000	0.06327	0.05766	0.06743	0.05548
Within R^2	0.00290	0.00118	0.00199	0.00172	0.00166	0.00144
1940 State by County FE	✓	✓	✓	✓	✓	✓
Ship type FE	✓	✓	✓	✓	✓	✓
First Quarter FE	✓	✓	✓	✓	✓	✓

NOTES: This table shows the first stage from Equation (1.12). For each specification the migration instrument is constructed for that period t and migration type y, where instrument is constructed as the predicted probability that a person will move out-of-state, out-of-region, or to a Pacific state by time t using estimates from the discrete choice model described in Section 1.5. The dependent variable is an indicator for whether a person moved out-of-state, out-of-region, or to the Pacific by time period t. Columns 1-3 report estimates for migration in 1950, while columns 4-6 report estimates for lifetime migration. Columns 1 and 4 report estimates for migration out-of-state. Columns 2 and 5 report estimates for migration out-of-region. Columns 3 and 6 report estimates for Pacific migration where individuals original from the Pacific are dropped. Standard errors are clustered at the ship-level.

Figure 1.31: Returns to Migration IV: Reduced Form

	1950			Death		
	Occscore 1950 (logged)			Zipcode Inc Death (logged)		
	(1)	(2)	(3)	(4)	(5)	(6)
Migration Instrument	0.009**	0.03***	0.03***	0.31***	0.38***	0.65***
	(0.004)	(0.009)	(0.01)	(0.05)	(0.06)	(0.13)
Observations	173,504	173,503	154,312	144,203	144,202	127,270
R^2	0.94222	0.94222	0.94418	0.24403	0.24403	0.21607
Within R^2	3.11×10^{-5}	6.78×10^{-5}	3.82×10^{-5}	0.00038	0.00040	0.00026
State-county40 fixed effects	✓	✓	✓	✓	✓	✓
Category fixed effects	✓	✓	✓	✓	✓	✓
Time fixed effects	✓	✓	✓	✓	✓	✓

NOTES: This table shows reduced form estimates from from Equation (1.12). For each specification the migration instrument is constructed fof period t and migration type y, where instrument is the predicted probability a person will move out-of-state, out-of-region, or to a Pacific state by time t using estimates from the discrete choice model described in Section 1.5. The dependent variable is a proxy logged income in period t. In 1950, income is measured by occupational score of the individual. By time of death, income is measured by median household income in the zipcode a person resided in at death. Income at death is reported for the year 2000 in nominal dollars. Columns 1-3 report estimates for migration in 1950, while columns 4-6 report estimates for lifetime migration. Columns 1 and 4 report estimates for migration out-of-state. Columns 2 and 5 report estimates for migration out-of-region. Columns 3 and 6 report estimates for Pacific migration where individuals original from the Pacific are dropped. Standard errors are clustered at the ship-level.

Table 1.5: Returns to Networked Migration (1950)

	Occscore 1950 (logged)					
		OLS			IV	
	(1)	(2)	(3)	(4)	(5)	(6)
State Mover	0.02***			0.01**		
	(0.0006)			(0.005)		
Region Mover		0.02***			0.04***	
		(0.0010)			(0.01)	
Pacific Mover			0.01***			0.03***
			(0.001)			(0.01)
Observations	175,275	175,275	155,995	173,504	173,503	154,312
F-stat				370.76	133.45	87.031
1940 State by County FE	✓	✓	✓	✓	✓	✓
Ship type FE	✓	✓	✓	✓	✓	✓
First Quarter FE	✓	✓	✓	✓	✓	✓

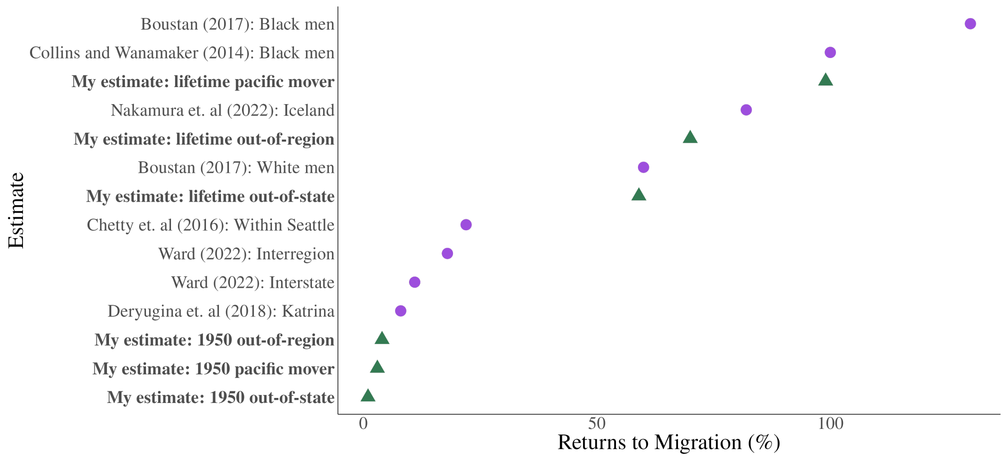
NOTES: This table reports coefficients from Equation (1.12) showing the returns to migration by time of death (lifetime). Columns 1-3 report OLS estimates, while Columns 4-6 report IV estimates. The outcome is logged occupational score of the person in the 1950 Census, where the occupational score as the average income for each occupation in 1950 in hundreds of 1950 dollars. Individuals without recorded occupations are or non-earning occupations are dropped from estimation. Income is reported for the year 2000 in nominal dollars. State mover is an indicator for if a person moved across state lines. Region mover is an indicator if a person moved between Census regions. Pacific mover is an indicator if someone who was not previously living in the Pacific Census division moved to a state in that division (excluding Alaska and Hawaii). In columns 4-6, the instrument is constructed as the predicted probability that a person will move out-of-state, out-of-region, or to a Pacific state using estimates from the discrete choice model described in Section 1.5. Standard errors are clustered at the ship-level.

Table 1.6: Impact of Networked Migration on non-pecuniary outcomes

	HH Size (1950) (1)	Married (1950) (2)	Wife Different Birthplace (3)	Age at Death (4)
State Mover	-0.28	-0.04	0.29***	-0.45
	(0.41)	(0.06)	(0.07)	(1.2)
Observations	173,504	173,504	152,259	176,731
Mean of Dep Var	4.1	0.73	0.39	75
F-stat	370.76	370.76	340.79	296.59
State-county40 fixed effects	✓	✓	✓	✓
Category fixed effects	✓	✓	✓	✓
Time fixed effects	✓	✓	✓	✓

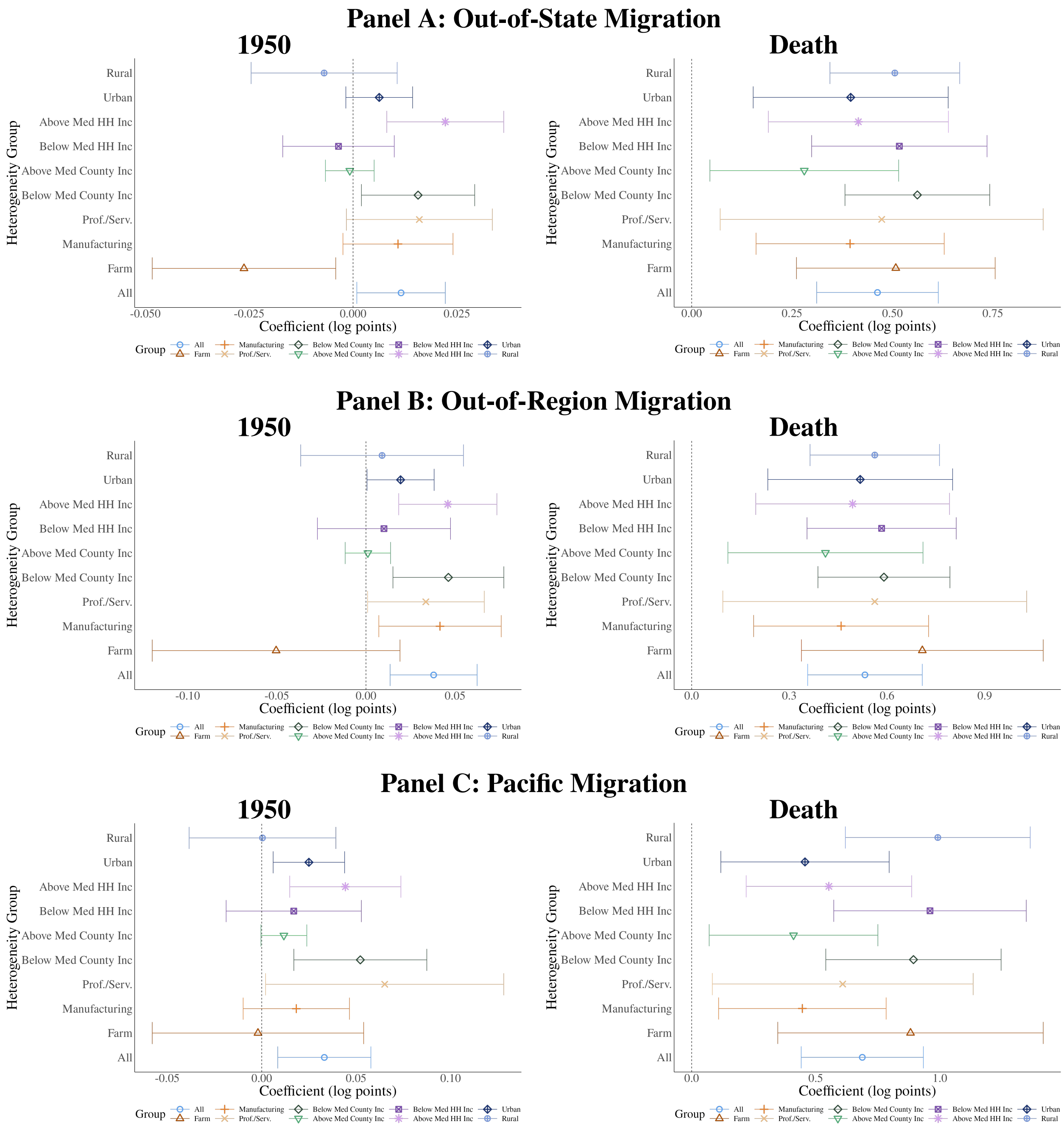
NOTES: This table reports IV estimates analogous to Equation (1.12) for non-pecuniary effects of networked migration. Each specification measures the impact of moving out-of-state, where the instrument is constructed as the predicted probability that a person will move out-of-state using estimates from the discrete choice model described in Section 1.5. Columns 1-3 report estimates for outcomes in 1950 and use the predicted probability of moving out-of-state by 1950, while column reports a lifetime outocme and uses the predicted probability of moving out-of-state by time of death. Column 1 reports the impact on total household size. Column 2 reports the impact for whether someone is married. Column 3 reports whether a married man's wife is born in either a different state or country. Finally, column 4 reports the impact of out-of-state migration on age of mortality. Standard errors are clustered at the ship-level.

Figure 1.32: Benchmarking Returns to Migration against the Literature



NOTES: This figure presents estimates from Tables 1.1 and 1.5, showing the corresponding percent change in income for people induced to migrate due to Navy networks. Additional estimates from various economics papers are included to provide comparisons of the pecuniary returns to migration. All calculations are the authors' own.

Figure 1.33: Returns to Migration: Heterogeneity



NOTES: Each plot reports coefficients from Equation (1.12) for a given type of migration and a specific time period. On each plot, the specification is run restricting the population to different subsets, and for each subset the coefficient is report with a 95% confidence interval. Panel A reports estimates for returns to out-of-state migration. Panel B reports estimates for returns to out-of-region migration. Panel C reports estimates for returns to migration to Pacific states. All standard errors are clustered at the ship-level.

References

[1] Abramitzky, R., Boustan, L., and Connor, D. S. (2024). Leaving the enclave: Historical evidence on immigrant mobility from the industrial removal office. *The Journal of Economic History*, 84(2):352–394. Publisher: Cambridge University Press.

[2] Abramitzky, R., Boustan, L., and Eriksson, K. (2020). Do immigrants assimilate more slowly today than in the past? *American Economic Review: Insights*, 2(1):125–141. Publisher: American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203.

[3] Abramitzky, R., Boustan, L., Eriksson, K., Feigenbaum, J., and Pérez, S. (2021). Automated Linking of Historical Data. *Journal of Economic Literature*, 59(3):865–918.

[4] Acemoglu, D., Autor, D., and Lyle, D. (2004). Women, War, and Wages: The Effect of Female Labor Supply on the Wage Structure at Midcentury. *Journal of Political Economy*, 112(3):497–551. Publisher: The University of Chicago Press.

[5] Aizer, A., Boone, R., Lleras-Muney, A., and Vogel, J. (2020). Discrimination and Racial Disparities in Labor Market Outcomes: Evidence from WWII.

[6] Althoff, L. and Szerman, C. (2024). The G.I. Bill and Black-White Wealth Disparities.

[7] Angrist, J. D. and Imbens, G. W. (1994). Identification and Estimation of Local Average Treatment Effect. *Econometrica*, (2):467–475.

[8] Arellano-Bover, J. (2022). Displacement, Diversity, and Mobility: Career Impacts of Japanese American Internment. *The Journal of Economic History*, 82(1):126–174.

[9] Bailey, A. K. (2011). Race, Place, and Veteran Status: Migration among Black and White Men, 1940–2000. *Population Research and Policy Review*, 30(5):701–728.

[10] Barnhardt, S., Field, E., and Pande, R. (2017). Moving to Opportunity or Isolation? Network Effects of a Randomized Housing Lottery in Urban India. *American Economic Journal: Applied Economics*, 9(1):1–32.

[11] Bedard, K. and Deschênes, O. (2006). The Long-Term Impact of Military Service on Health: Evidence from World War II and Korean War Veterans. *American Economic Review*, 96(1):176–194.

[12] Bergman, P., Chetty, R., DeLuca, S., Hendren, N., Katz, L. F., and Palmer, C. (2024). Creating Moves to Opportunity: Experimental Evidence on Barriers to Neighborhood Choice. *American Economic Review*, 114(5):1281–1337.

[13] Blumenstock, J. E., Chi, G., and Tan, X. (2023). Migration and the value of social networks. *Review of Economic Studies*, page rdad113. Publisher: Oxford University Press US.

[14] Bound, J. and Turner, S. (2002). Going to War and Going to College: Did World War II and the G.I. Bill Increase Educational Attainment for Returning Veterans? *Journal of Labor Economics*, 20(4):784–815.

[15] Boustan, L. P. (2016). *Competition in the promised land: Black migrants in northern cities and labor markets*, volume 39. Princeton University Press.

[16] Bramoullé, Y., Djebbari, H., and Fortin, B. (2020). Peer Effects in Networks: A Survey. *Annual Review of Economics*, 12(Volume 12, 2020):603–629. Publisher: Annual Reviews.

[17] Breen, C. F., Osborne, M., and Goldstein, J. R. (2023). CenSoc: Public Linked Administrative Mortality Records for Individual-level Research. *Scientific Data*, 10(1):802. Publisher: Nature Publishing Group UK London.

[18] Bureau, U. C. (2024). Resident Population in California [CAPOP]. Technical report.

[19] Büchel, K., Ehrlich, M. V., Puga, D., and Viladecans-Marsal, E. (2020). Calling from the outside: The role of networks in residential mobility. *Journal of urban economics*, 119:103277. Publisher: Elsevier.

[20] Carrell, S. E., Sacerdote, B. I., and West, J. E. (2013). From Natural Variation to Optimal Policy? The Importance of Endogenous Peer Group Formation. *Econometrica*, 81(3):855–882. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.3982/ECTA10168.

[21] Carrington, W. J., Detragiache, E., and Vishwanath, T. (1996). Migration with Endogenous Moving Costs. *The American Economic Review*, 86(4):909–930. Publisher: American Economic Association.

[22] Chetty, R., Hendren, N., and Katz, L. F. (2016). The Effects of Exposure to Better Neighborhoods on Children: New Evidence from the Moving to Opportunity Experiment. *American Economic Review*, 106(4):855–902.

[23] Chetty, R., Hendren, N., Kline, P., and Saez, E. (2014). Where is the land of opportunity? The geography of intergenerational mobility in the United States. *The quarterly journal of economics*, 129(4):1553–1623. Publisher: MIT Press.

[24] Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., and Jacob, M. (2022). Social capital I: measurement and associations with economic mobility. *Nature*, 608(7921):108–121. Publisher: Nature Publishing Group UK London.

[25] Clampet-Lundquist, S. and Massey, D. (2008). Neighborhood Effects on Economic Self-Sufficiency: A Reconsideration of the Moving to Opportunity Experiment. *American Journal of Sociology*, 114(1):107–143. Publisher: The University of Chicago Press.

[26] Collins, W. J. and Wanamaker, M. H. (2014). Selection and economic gains in the great migration of African Americans: new evidence from linked census data. *American Economic Journal: Applied Economics*, 6(1):220–252. Publisher: American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203-2425.

[27] Collins, W. J. and Zimran, A. (2024). Who Benefited from World War II Service and the GI Bill? New Evidence on Heterogeneous Effects for US Veterans.

[28] Correia, S., Guimarães, P., and Zylkin, T. (2020). Fast Poisson estimation with high-dimensional fixed effects. *The Stata Journal: Promoting communications on statistics and Stata*, 20(1):95–115.

[29] Costa, D. L. and Kahn, M. E. (2003). Cowards and heroes: Group loyalty in the American Civil War. *The Quarterly Journal of Economics*, 118(2):519–548. Publisher: MIT Press.

[30] Costa, D. L., Kahn, M. E., Roudiez, C., and Wilson, S. (2018). Persistent social networks: Civil war veterans who fought together co-locate in later life. *Regional Science and Urban Economics*, 70:289–299.

[31] Deryugina, T., Kawano, L., and Levitt, S. (2018). The economic impact of Hurricane Katrina on its victims: Evidence from individual tax returns. *American Economic Journal: Applied Economics*, 10(2):202–233. Publisher: American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203-2425.

[32] Dingel, J. I. and Tintelnot, F. (2023). Spatial Economics for Granular Settings.

[33] Dolfin, S. and Genicot, G. (2010). What Do Networks Do? The Role of Networks on Migration and “Coyote” Use. *Review of Development Economics*, 14(2):343–359. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/j.1467-9361.2010.00557.x.

[34] Einiö, E. (2019). Mixing the rich and poor: The impact of peers on education and earnings. *VATT Institute for Economic Research Working Papers*, 128.

[35] Enamorado, T. (2021). A Primer on Probabilistic Record Linkage. In *Handbook of Computational Social Science, Volume 2*. Routledge. Num Pages: 13.

[36] Fellegi, I. P. and Sunter, A. B. (1969). A Theory for Record Linkage. *Journal of the American Statistical Association*, 64(328):1183–1210.

[37] Fetter, D. K. (2013). How Do Mortgage Subsidies Affect Home Ownership? Evidence from the Mid-Century GI Bills. *American Economic Journal: Economic Policy*, 5(2):111–147.

[38] Fishback, P. and Cullen, J. (2013). Second World War spending and local economic activity in US counties, 1939–58 - Fishback - 2013 - The Economic History Review - Wiley Online Library.

[39] Fishback, P. V. and Jaworski, T. (2016). World War II and US Economic Performance. In Eloranta, J., Golson, E., Markevich, A., and Wolf, N., editors, *Economic History of Warfare and State Formation*, pages 221–241. Springer, Singapore.

[40] Furer, J. A. (1960). *Administration of the Navy Department in World War II.* US Government Printing Office.

[41] Garin, A. and Rothbaum, J. L. (2024). The Long-Run Impacts of Public Industrial Investment on Local Development and Economic Mobility: Evidence from World War II.

[42] Granovetter, M. S. (1973). The Strength of Weak Ties. *American Journal of Sociology*, 78(6):1360–1380.

[43] Guimaraes, P., Figueirdo, O., and Woodward, D. (2003). A tractable approach to the firm location decision problem. *Review of Economics and Statistics*, 85(1):201–204. Publisher: MIT Press 238 Main St., Suite 500, Cambridge, MA 02142-1046, USA journals ....

[44] Guo, Y., Jackson, M. O., and Jia, R. (2024). Comrades and Cause: Peer Influence on West Point Cadets' Civil War Allegiances.

[45] Hall, P. K. (2009). Privileged moves: migration, race and veteran status in post-World War II America.

[46] Harriman, W. A. (1948). Internal Migration in the United States: April 1940 to April 1947.

[47] Hausser, P. M. and Jaffe, A. J. (1947). The extent of the housing shortage. *Law & Contemp. Probs.*, 12:3. Publisher: HeinOnline.

[48] Hendren, N., Sprung-Keyser, B., and Porter, S. R. (2022). The Radius of Economic Opportunity: Evidence from Migration and Local Labor Markets ". *CES Working Paper, Nr. 22–27*. Publisher: CES Working Papers.

[49] Johnson, M. S. (1994). *The second gold rush: Oakland and the East Bay in World War II*. Univ of California Press.

[50] Kinnan, C., Wang, S.-Y., and Wang, Y. (2018). Access to Migration for Rural Households. *American Economic Journal: Applied Economics*, 10(4):79–119.

[51] Kling, J. R., Liebman, J. B., and Katz, L. F. (2007). Experimental Analysis of Neighborhood Effects. *Econometrica*, 75(1):83–119.

[52] Koenen, M. and Johnston, D. (2024). Social Ties and Residential Choice: Micro Evidence and Macro Implications.

[53] Ludwig, J., Duncan, G. J., Gennetian, L. A., Katz, L. F., Kessler, R. C., Kling, J. R., and Sanbonmatsu, L. (2013). Long-Term Neighborhood Effects on Low-Income Families: Evidence from Moving to Opportunity. *American Economic Review*, 103(3):226–231.

[54] Manski, C. F. (1993). Identification of Endogenous Social Effects: The Reflection Problem. *The Review of Economic Studies*, 60(3):531–542.

[55] McKenzie, D. and Rapoport, H. (2010). Self-selection patterns in Mexico-US migration: the role of migration networks. *the Review of Economics and Statistics*, 92(4):811–821. Publisher: The MIT Press.

[56] Michelman, V., Price, J., and Zimmerman, S. D. (2022). Old Boys' Clubs and Upward Mobility Among the Educational Elite*. *The Quarterly Journal of Economics*, 137(2):845–909.

[57] Molloy, R., Smith, C. L., and Wozniak, A. (2011). Internal Migration in the United States. *Journal of Economic Perspectives*, 25(3):173–196.

[58] Moretti, E., Steinwender, C., and Van Reenen, J. (2023). The Intellectual Spoils of War? Defense R&D, Productivity, and International Spillovers. *The Review of Economics and Statistics*, pages 1–46.

[59] Morison, S. E. (1963). *The Two-Ocean war: A short history of the United States navy in the second world war*. Little, Brown Boston.

[60] Munshi, K. (2003). Networks in the Modern Economy: Mexican Migrants in the U. S. Labor Market*. *The Quarterly Journal of Economics*, 118(2):549–599.

[61] Munshi, K. (2020). Social Networks and Migration. *Annual Review of Economics*, 12(1):503–524.

[62] Nakamura, E., Sigurdsson, J., and Steinsson, J. (2022). The gift of moving: Intergenerational consequences of a mobility shock. *The Review of Economic Studies*, 89(3):1557–1592. Publisher: Oxford University Press.

[63] Porcher, C. (2020). Migration with costly information. *Unpublished Manuscript*, 1(3).

[64] Porcher, C., Morales, E., and Fujiwara, T. (2024). Measuring Information Frictions in Migration Decisions: A Revealed-Preference Approach.

[65] Rosenbloom, J. L. and Sundstrom, W. A. (2004). The decline and rise of interstate migration in the United States: Evidence from the IPUMS, 1850–1990. In *Research in Economic History*, volume 22 of *Research in Economic History*, pages 289–325. Emerald Group Publishing Limited.

[66] Sacerdote, B. (2001). Peer effects with random assignment: Results for Dartmouth roommates. *The Quarterly journal of economics*, 116(2):681–704. Publisher: MIT Press.

[67] Sahai, H. and Bailey, M. (2022). Social Networks and Internal Migration: Evidence from Facebook in India. *Meta*.

[68] Sarvimäki, M., Uusitalo, R., and Jäntti, M. (2022). Habit Formation and the Misallocation of Labor: Evidence from Forced Migrations. *Journal of the European Economic Association*, 20(6):2497–2539.

[69] Shue, K. (2013). Executive Networks and Firm Policies: Evidence from the Random Assignment of MBA Peers. *The Review of Financial Studies*, 26(6):1401–1442.

[70] Sotelo, S. (2019). Practical aspects of implementing the multinomial pml estimator. *Ann Arbor: University of Michigan, mimeo*.

[71] Stuart, B. A. and Taylor, E. J. (2021). Migration networks and location decisions: Evidence from US mass migration. *American Economic Journal: Applied Economics*, 13(3):134–175. Publisher: American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203-2425.

[72] Suandi, M. (2022). *Essays in Economic History and Development Economics*. PhD thesis, UC Berkeley.

[73] Ward, Z. (2022). Internal Migration, Education, and Intergenerational Mobility: Evidence from American History. *Journal of Human Resources*, 57(6):1981–2011.

1.11 Data Appendix

1.11.1 World War II Muster Rolls

The muster rolls data, available digitally from the National Archives catalog, contains 6,459,023 scans divided between 32,101 file units.[55] This data covers the near-universe of all "activities" within the Navy between January 1, 1939 and January 1, 1949. An "activity" is any unit within the Navy including ships, training centers, stations, etc. The muster rolls report lists of enlisted personnel formally attached to each ship or activity on a quarterly basis.

I download all scans and metadata available for each of the 32,000 file units.[56] Titles of file units typically identify the name of the ship or other activity as well as dates of coverage. File units can contain anywhere from 1 scan up to 2000 scans. Each scan can have metadata which varies from containing no information, to identifying the type of scan, to occasionally including volunteer transcriptions. Information from specific ships can be located across multiple file units. I restrict attention to the approximately 14,000 file units where the "activity" is identified as being a ship. Excluded activities include construction battalions, airborne units, administrative units, hospitals, and training centers.

Muster Rolls Cleaning Procedure

I use optical character recognition (OCR) to categorize and extract data from each scan. I use Google Vision OCR software and the Layout Parser python package. Within each file unit, I am interested in scans of quarterly muster rolls that identify all enlisted personnel attached to ship at the end of the quarter and monthly reports of changes that identify any personnel changes within a month. These two types of scans typically make up 40% of a file unit. Remaining scans include title pages, blank pages, and lists of passengers Using a combination of metadata and data contained within the headers of scans, I categorize scans as belonging to one of four categories: quarterly muster roll, monthly change reports, other types of scans, and scans I cannot categorize (henceforth

[55]Record Group 24: Records of the Bureau of Navy Personnel. https://catalog.archives.gov/id/594996.

[56]Main access and download period between October 2022 and December 2022

identified as miscellaneous). I drop scans that I can confidently identify as not being a quarterly report of monthly change report. I finally extract characteristics from the tabular structure of each scan from quarterly muster rolls, monthly change reports, and miscellaneous scans. Once this process is completed, I have one csv of extracted table and header contents for each scan within a file unit.

I then clean these csvs to construct two ship-level datasets. First, for each ship, I extract the set of individuals identified by a unique seven-digit service number who were ever formally attached to that unit. For each individual, I identify name, service number, where they enlisted, when the enlisted, and when they boarded the ship. Second, I construct a quarterly panel of each person attached to ship in each quarter as well as their Navy occupational rating code at the end of the quarter.

As quality scans are highly variable and OCR encoding errors are common, I use extensive cleaning processes to identify as high quality information about each individual as possible. The high-level approach to cleaning is twofold: first I correct for common encoding errors at the observation level. Table 1.7 reports for each relevant field high-level cleaning strategies. Then I use multiple occurrences of individuals across scans to extract the highest confidence attribute for each individual. For instance, if a person with the name "Hugh T. Berry" appears on two different scans but there is a 1-digit difference in service number, I combine these two people and prioritize the version of the service number that is more common.

During this process, I also incorporate information from tags imputed by volunteers in the catalog. When available the tags while list name, date of enlistment, and place of enlistment, but they do not report service number. I use fuzzy matching on names within scans to replace encodings with tag information when possible.

In the last part of the cleaning procedure I construct a quarterly panel. For each individual, I identify the first quarter they were on a ship from a mixture of the date they boarded and the first scan I see them. I identify the last quarter on board using the last scan I see them. The date attached to each scan is recovered in one of three ways: first it can be directly reported in the

Table 1.7: Encoding Errors and Cleaning Procedures

Field	**Common Encoding Errors**	**Cleaning Procedures**
Last name	Pen along left column leads to leading character (often "X"), extra characters, duplicated characters, misspellings of less common names	Use order within scan to identify when last names are out of alphabetical order, match to last names in the 1940 census to fix common encoding errors, remove leading characters, clean up common encoding errors due to compound last names.
First name	Misspellings, extra characters	Match to first names in 1940 census to fix common encoding errors, remove any unusual character types.
Middle name	Mixed representation: middle name vs. initials vs. suffixes, extra characters	Remove suffixes, separate initials from names, ignore observations with long character lengths suggesting encoding errors.
Service number	Number encoding error, part of service number appearing in other fields, duplicated numbers	Identify 7-digit patterns using common representations, search for alternative versions of the service number in other fields, remove non-numeric characters, remove instances of duplication
Rating	Extra characters, number and letter encoding errors	Use Navy documentation on rating codes to identify and correct common transcription errors for each rating type.
Date enlistment and boarded	Misspellings of month, unclear whether a number refers to a day or month, number encoding errors	Clean up years using a known valid date range, fix common transcription errors for months, and use context within the scan to distinguish whether ambiguous numbers refer to the month or the day.
Place of enlistment	Extraneous characters, misspellings. Ambiguous city names	For the 100 most common places of enlistment, identify transcription errors and apply fuzzy string matching to associate entries with the correct location.

metadata, second it can be extracted from the OCR of the header (though often messy), or it can be extracted from the order scans existing within the file unit. To the last point, imagine that I know a ship within that file unit were active over the course of four quarters. If I know that this muster roll is placed such that it is the second quarter than I can back out the specific date belonging to that muster roll. For each person, I then interpolate between their first and last quarter aboard to construct a panel.

Once I have constructed cleaned versions of ship-level data, I proceed to bring the data together, and in particular, harmonize individual characteristics between ships. Within this process, I identify individual service numbers that appear on multiple ships and harmonize any discrepancies of characteristics about individuals that differ between ships. I also de-duplicate individuals who based off of name and service number similarity are the same but have been separated by small encoding errors. At this point, I drop any observations that are likely “false” – in particular, I drop service numbers that only appear one time in any scan.

Once completed, I have a final dataset that contains 1,450,000 people across 5,200 different ships.

1.11.2 Additional Data Sources

Place of Enlistments

Place of enlistment data was sourced from two primary methods: monthly reports of changes and service number blocks assigned to receiving stations. This dual approach allowed for comprehensive coverage and cross-validation of enlistment locations. Service number blocks were identified using several internal Navy documents:

- A Bureau of Personnel document (December 15, 1942) specifying blocks for 69 receiving stations.
- An update (August 27, 1943) providing additional ranges for stations that had exhausted their initial allocations.
- A 1990s Navy document detailing World War II service number ranges constructed from a broader set of records.

Some service number ranges (e.g., 1,000,000 to 2,000,000) could not be uniquely assigned to enlistment locations. These included numbers for World War I veterans who re-enlisted, though some were repurposed for new enlistments. The process for assigning place of enlistment was as follows:

1. Identify candidate locations using both report of changes and service number blocks.
2. When sources agreed but differed in geographic specificity, select the more granular location.
3. In cases of disagreement, prioritize the location indicated by service number blocks, as report of changes often reflected training locations rather than enlistment sites.
4. Use granular service number blocks (in 1,000-number increments) to infer additional assignments.

This methodology yielded place of enlistment data for over 95Figure 1.11 illustrates the geographic distribution of enlistment locations that processed at least 100 sailors.

Dictionary of American Naval Fighting Ships

The *Dictionary of American Naval Fighting Ships* (DANFS) was a series of reference books containing facts about the service histories of Navy ships used by the US Navy. These reference books were written between 1959 and 1991 and published in nine volumes for every Naval vessel active over the history of the Navy's service. These volumes were digitized volunteers and made available online at https://www.hazegray.org/danfs/.

The following information is usually contained within en entry for each ship: name, specifications (size, complement, ship type, etc.) , dates of activity (commissioned, laid down, launched, major battle dates), information about places the ship went (where it was laid down, launched, areas or patrol and battles), battles and commendations (number of battle stars)

I scrape the record of every Naval vessel within the muster rolls data and use textual processing to extract the following pieces of information about each ship: complement, dates launched and sunk (if available), World War II battle stars, location laid down and launched, and theater of activity during World War II.

Navy Occupation Rating Data

Within the muster rolls data occupation and occupational rank is abbreviated by alpha-numeric codes. For instance, "S1C" corresponds to Seaman 1st Class. These ratings both provide information about the nature of the job an individual was doing on a given ship as well pay grade and hierarchy of that person relative to others. For instance "1st Class" ranks are one grade above "2nd Class" ranks within the same occupation.

I use documentation from various Navy sources to construct a crosswalk of abbreviations to occupational title, paygrade, and divisional branch.[57]

[57] Sources: https://www.cem.va.gov/docs/wcag/hmm/Inscription-Abbreviations-Ranks-Navy.pdf https://www.ibiblio.org/hyperwar/USN/ref/Ranks&Rates/index.html

World War II Navy Casualty Records

I collect Navy casualty records from a few different sources. First, for each state and county, the National Archives contain scans of every individual man from that state and county who was either wounded, missing, or dead in combat.[58]

Second, I have additional information on Naval losses of particular vessels including the date and geographic location when the vessel was lost.[59]

Finally, I have records on the number of individual casualties (wounded, killed, missing) by category of ship type during the war.[60]

Full Count 1940 and 1950 Census

I use restricted versions of the 1940 Full Count Census and the public release of the 1950 Full Count Census. The 1940 Census was the first to include questions on income, supplementing existing data on occupation and employment status. It collected detailed residence information, including 1940 address and 1935 location (either the same address or a different location specified down to the town/city or county level).

The 1950 Census, while more limited in scope than its 1940 predecessor, asked 20 questions of all individuals, with additional questions in the sample-line section (20% of individuals). Reliability issues have been noted with both the veteran status question (9) and with recorded income. For this reason neither variable will be used in subsequent analysis.

Numident Social Security Death File

I use the Numident Social Security Death File provided through the CenSoc project (17)/ The records are collected from the Social Security Death Master file that has been cleaned to the Berkeley Unified Numident Mortality Database (BUNMD). There are roughly 50 million records that cover a high coverage of all deaths in the United States between 1985 and 2007.

[58] Source: https://www.archives.gov/research/military/ww2/navy-casualties/south-dakota.html
[59] https://www.navsource.org/Naval/losses.htm#ms
[60] https://apps.dtic.mil/sti/citations/ADA230803

FindAGrave

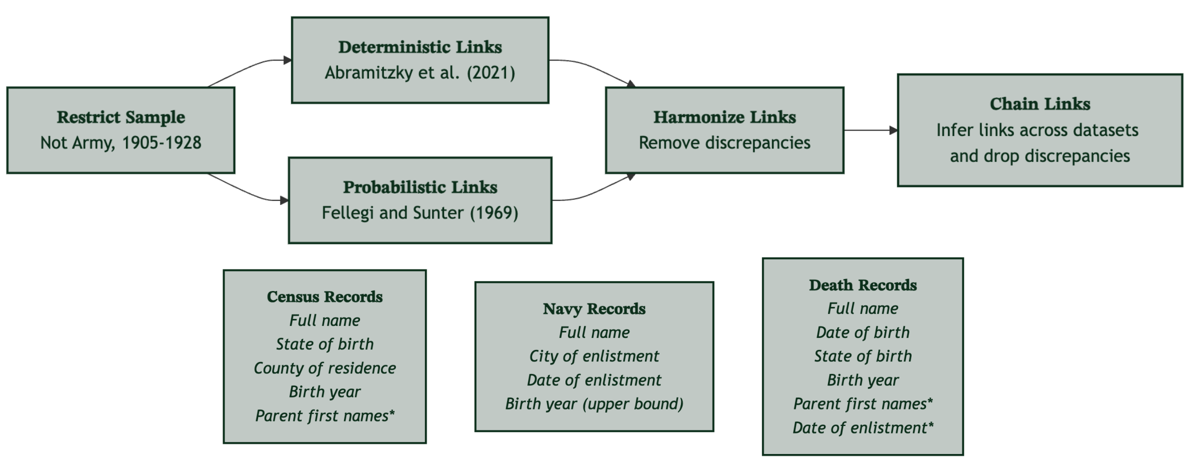
FindAGrave.com is the largest online database of gravestones. The website is primarily maintained by public contributions in which individuals will upload images of gravestones with transcriptions of information contained on those gravestones. For a portion of entries on the site, particular those belonging to veterans, information on gravestones are from large public databases maintained by military cemeteries such as Arlington or specific military agencies within the government.

I collect 2.8 million records belonging to individuals on FindAGrave.com as being born between 1900 and 1928, and additionally being tagged as a military veteran. Information almost always included on an individual include date of birth, date of death, cemetery location, and name. Very often, individual's gravestone pages will also include additional information on place of birth and place of death.

1.11.3 Additional Linking Information

This section elaborates on the four-step linking procedure discussed in Section 3.3. Figure 1.34 provides a high-level visual overview of the linking procedure and the datasets connected in this process.

Figure 1.34: Linking Procedure



Sample Restriction

To ensure accurate identification of unique matches, I first restrict the sample to individuals who could plausibly serve as rank-and-file personnel aboard a Navy ship. Across all datasets, I focus on men born between 1905 and 1928 who did not serve in the US Army. This date range encompasses the full spectrum of birth years for combat draft eligibility. According to enlistment records, over 99% of enlisted men in the Navy were born within this range. This restriction is particularly valuable for distinguishing between same-name father-son pairs.

To exclude men who served in the Army, I employ a conservative version of the linking procedure, connecting Army enlistment records to the 1940 Full Count Census, FindAGrave, and the Numident Social Security Death Records. The Army enlistment records contain additional high-quality linking fields such as state of birth and year of birth, which facilitate more accurate linking.

Following these restrictions, I proceed to construct bilateral links between each dataset.

Deterministic Linking

I implement a version of deterministic linking pioneered by (3). This approach identifies unique linked observations that match on several high-quality characteristics. Figure 1.34 illustrates the linkable fields available in each dataset. For linking between Navy records and Census records, I rely solely on unique names. I employ a fuzzy name buffering approach, allowing for unique name matches within a Jaro-Winkler string distance of 0.05. To reduce false positives, I require that matches occur within a narrower fuzziness while remaining unique in a more general fuzzy band (e.g., 0.10). When linking Navy records to death records, I incorporate date and year of enlistment as additional potential linking fields, while maintaining the same fuzzy uniqueness procedure for names. For links between death records and census records, I add standard deterministic framework linking variables: state of birth and year of birth (allowing for some fuzziness).

Probabilistic Linking

I implement the (36) method using the Python package `splink`. This framework estimates a model that predicts whether any two records belong to the same individual, given the set of characteristics they do and do not match on. This linking style is ideal when datasets have various linking variables that are either continuous or of variable quality, making the construction of a deterministic decision tree infeasible. I utilize all the linking fields from the deterministic approach, with some additions. For linking between Navy records and the 1940 Census, I include the physical distance between the county of residence in 1940 and the city of Navy enlistment. I also incorporate a probabilistic age component based on the year of enlistment. When linking Numident records to the Full Count Censuses, I add characteristics of the mother's and father's names. To address concerns about over-clustering, particularly on household and enlistment date, I remove links with large discrepancies in specific fields such as first name and last name after constructing initial candidate links. I then restrict to only unique links.

Harmonization and Chaining

After generating bilateral probabilistic and deterministic links between each dataset, I harmonize across linking procedures and perform a final round of link chaining to infer additional connections.

In the harmonization process, I remove any discrepancies where the links established between individuals disagree. Less than 0.3% of all generated links create such discrepancies. Both probabilistic and deterministic linking procedures provide informative unique links. Table 1.8 shows the share of links between each bilateral linking dataset created from either procedure or both procedures.

Finally, I chain links between datasets, using linking pairs with higher information content to infer additional links. This chaining process is particularly useful for inferring links between Navy records and the 1940 Census.

Table 1.8: Share of Link Between Datasets Generated by Each Procedure

	Probabilistic	**Deterministic**	**Both**
Navy - Numident	44%	18%	38%
Navy - 1940 Census	60%	21%	19%
Numident - 1940 Census	30%	38%	32%
FindAGrave - 1940 Census	24%	34%	42%
Navy - FindAGrave	10%	30%	59%

1.12 Additional Results

1.12.1 Network Formation: Other forces

Empirical Framework

Role of Ship Occupational Similarity:

The second specification focuses on occupational proximity:

$$\begin{aligned}
\beta^{dest}_{hd} X_{ijkd} &= \beta^{dest} Sh_{ijkd} + \beta^{dist} \log \text{dist}_{hd} Sh_{ijkd} \beta^{dest,close} Sh^{close}_{ijkd} + \beta^{dist,close} \log \text{dist}_{hd} Sh^{close}_{ijkd} \\
\beta^{home}_{d} X_{ijkd} &= \beta^{home} Sh_{ijkd} + \beta^{home,close} Sh^{close}_{ijkd}
\end{aligned} \tag{1.13}$$

In this specification, Sh^{close}_{ijkd} represents the share of shipmates from state d who operated in the same rating group. Navy ships were divided into horizontal units (e.g., deck, boiler room), and individuals were more likely to interact frequently with those in their unit. The key parameters $\beta^{dest,close}$, $\beta^{dist,close}$, and $\beta^{home,close}$ capture how these closer interactions influence the strength of network ties in predicting migration.

Role of Socioeconomic Background:

The third specification examines the role of socioeconomic background:

$$\begin{aligned}
\beta^{dest}_{hd} X_{ijkd} &= \beta^{dest} Sh_{ijkd} + \beta^{dist} \log \text{dist}_{hd} Sh_{ijkd} \beta^{dest,occ} \text{Occ}_{ijkd} + \beta^{dist,occ} \log \text{dist}_{hd} \text{Occ}_{ijkd} \\
\beta^{home}_{d} X_{ijkd} &= \beta^{home} Sh_{ijkd} + \beta^{home,occ} \text{Occ}_{ijkd}
\end{aligned} \tag{1.14}$$

Here, Occ_{ijkd} represents the average imputed occupational score of shipmates from state d. The key parameters $\beta^{dest,occ}$, $\beta^{dist,occ}$, and $\beta^{home,occ}$ capture how interactions with shipmates from higher-income households influence migration decisions. This specification is motivated by the possibility that access to individuals of higher socioeconomic status might provide better information or access to job opportunities in different locations (34).

Results

Proximity Effects

Shared shipboard roles, measured by common rating branches, only weakly influence network effects on migration. For most state pairs, the share of shipmates from a given state working in the same rating group does not meaningfully affect migration to that state. However, having more same-rating shipmates from one's home state does increase the likelihood of staying there. This effect persists long-term but is strongest in the short run.

These results suggest that sailors working closely with others from their home state may form tighter social bonds, reinforcing their ties to home. However, measuring these shared experiences presents challenges. Navy ratings change over a sailor's service, and many men spent substantial time in entry-level ratings not clearly tied to specific shipboard divisions. These measurement issues likely contribute to the weak evidence for shared role effects on migration patterns.

Socioeconomic Background Effects

The analysis of socioeconomic background, as measured by average occupational score, reveals no significant impact on migration decisions. This lack of effect is consistent across different specifications and time periods. This finding may reflect the demographic characteristics of the sample: young, non-college-educated men who were not yet established in their careers. In this context, differential exposure to individuals of slightly higher or lower socioeconomic status did not substantially influence migration choices.

Chapter 2

Demographic Preferences and Income Segregation[1]

2.1 Introduction

American life is segregated by income and race in domains ranging from residences to media diets. Given that social connections matter for economic mobility, the segregation of interactions by income raises important questions. Whether people's preferences over the demographic composition of those around them contribute to income segregation is particularly contentious. In this paper, we study how different demographic groups are exposed to high-income individuals in shared commercial spaces. We estimate individuals' preferences over the racial and income composition of co-patrons and use them to quantify sources of cross-group differences in experienced income segregation.

To measure exposure to high-income co-patrons, we use data on the movements of millions of smartphones in the United States in 2018 and 2019. Joining these movement data with building-level residential demographics, we measure the socioeconomic composition of each venue's patrons and characterize eight groups' exposure to different co-patron mixes. The eight demographic

[1]This paper is coauthored with *Victor Couture*, *Jonathan Dingel*, and *Jessie Handbury*/ Materials from this chapter were presented at NBER Summer Institute Urban, Summer 2021 and Urban Economics Association, NYC Fall 2018.

groups are four racial-ethnic categories (Asian, Black, Hispanic, and White) interacted with two income categories (split by median income).[1]

We find large differences across groups in exposure to high-income co-patrons. Unsurprisingly, within each racial group, high-income individuals have greater high-income exposure. Within income groups, Black and Hispanic individuals have lower high-income exposure than Asian and White individuals. High-income Black individuals, for example, experience nearly the same high-income exposure as low-income White individuals.

We consider three explanations for these demographic differences in experienced income segregation. First, differences in proximity to venues: low-income individuals may live far from venues with high-income patrons. Second, differences in preferences for product attributes: groups might vary in their price sensitivity or taste for particular services. Third, preferences over the demographics of co-patrons: these preferences encompass all the ways co-patron mix may affect an individual's likelihood of choosing a venue. This includes their affinities for certain groups but also, for instance, how a concentration of young professionals working in a coffee shop might create a productive ambiance.

To distinguish between these explanations, we estimate preferences using choices of venues within chain businesses. In our venue-choice model, people trade off the cost of a longer trip with the benefits of venue characteristics. Controlling for proximity, we separate preferences over co-patron mix from tastes for product attributes by contrasting venues within the same chain, which offer the same product but vary in their co-patron demographics.[2] We specify demographic preferences as a flexible function of the high-income share of co-patrons and the same-race share of co-patrons. This flexibility allows us to capture complementarities between race and income com-

[1]Following US government and social science conventions, the four racial/ethnic groups we observe are non-Hispanic Asian, non-Hispanic Black, Hispanic, and non-Hispanic White. For the sake of brevity, we often omit the adjective "non-Hispanic" and simply refer to these racial-ethnic groups as racial groups.

[2]For retail chains, brand power and economies of scale depend on a standardized offering, so product availability and service quality in a chain venue do not typically reflect the local composition of co-patrons. Even chains whose design footprint is less standardized, like Starbucks, strive to create a consistent experience and fixed menu across venues. In a robustness exercise, we restrict our estimation sample to the most standardized restaurant chains, such as Olive Garden, which is wholly owned instead of franchised and whose Google review ratings vary little across venues.

position and to distinguish preferring perfectly homogeneous co-patrons from a mere aversion to being unlike everyone else.

Our baseline estimation sample contains visits to restaurants, the business category with the largest number of chain venues. Chain restaurants are well suited as a laboratory to estimate demographic preferences, because they are frequently visited and tend to have homogeneous venues. We can therefore estimate demographic preferences, and their evolution over time, with greater precision and detail than one could in other choice settings like home purchases or school enrollments. To the extent possible, we also report demographic exposure and preference estimates within a range of other commercial and public venues.

Our estimates reveal notable regularities across demographic groups in their preferences over co-patron composition. High- and low-income individuals exhibit similar levels of racial homophily (a preference for one's own race). Black, Hispanic, and White individuals have similar levels of racial homophily (with that of Asian individuals being somewhat stronger). Members of different racial groups have broadly similar preferences for high-income exposure (with those of White individuals being somewhat weaker). Only high-income individuals, however, exhibit monotone preferences over the share of co-patrons who are high-income. Low-income individuals prefer establishments with an integrated mix of low- and high-income co-patrons.

These preferences for demographic exposure are economically large. Individuals are willing to travel two to three additional kilometers to visit a venue in the 95th percentile of either the same-race or high-income distribution rather than a venue at the 5th percentile.[3] This translates into willingness to pay of a few thousand dollars per year, close in magnitude to willingness to pay for schools with high test scores (e.g., 16; 13).

Given these regularities in demographic preferences, why does demographic exposure differ between groups? Differences in tastes for product attributes, the first of three possible explanations, are small: little income segregation arises from high-income individuals visiting different chains. In fact, within-chain income segregation resembles income segregation across all non-residential

[3]The exception to this 95th-5th comparison is low-income individuals' preference for high-income exposure, because they prefer economically integrated venues.

venues, including public and commercial venues. For example, both high- and low-income people visit McDonald's restaurants, but they tend to choose different McDonald's locations.[4] We use our estimated model to show that neighborhood of residence and demographic preferences are the factors that largely explain cross-group differences in exposure.

Racial differences in high-income exposure stem from how preferences over co-patron demographics interact with the joint distribution of income and race across venues. For instance, high-income Black and White individuals have similar willingness to travel to high-income venues. Black individuals, however, visit venues with much smaller shares of high-income co-patrons, even conditional on the distances between their residences and high-income venues. This reflects the role of racial homophily. Because majority-Black and majority-Hispanic venues generally have lower-income co-patrons, Black and Hispanic individuals face a trade-off between visiting heavily high-income venues and visiting heavily same-race venues that Asian and White individuals do not face.

The gap in high-income exposure between low- and high-income individuals within racial groups reflects differences in residential sorting and preferences. Low-income people both live in poorer neighborhoods and have weaker preferences for high-income co-patrons. High-income individuals tend to live in neighborhoods near venues with many high-income co-patrons. Conditional on where they live, their stronger income preferences lead them to choose venues with more high-income co-patrons. Overall, demographic preferences explain observed income exposure more for high- than low-income people.

To further examine how demographic preferences relate to neighborhood choice, we estimate how demographic preferences vary across neighborhoods with different demographic mixes. We find that people live in neighborhoods that match their preferences for demographic exposure: within demographic groups, individuals living in higher-income neighborhoods have stronger income preferences, and individuals living in more heavily own-race neighborhoods have stronger racial preferences. This alignment of individuals' demographic preferences and the dominant de-

[4]This aligns with (29), who notes the widespread popularity of large chains, with McDonald's being the most preferred restaurant across all income groups.

mographics of their residential neighborhoods suggests that demographic preferences might, in addition to determining venue choice, be a determinant of neighborhood choice.

Our model is agnostic on how individuals choose residential neighborhoods, but we can test for such sorting patterns by examining movers. Specifically, we estimate the preferences of individuals before and after they move between neighborhoods in different metropolitan areas. The estimated preferences are consistent with people sorting into neighborhoods based on their preferences over co-patron demographics. For instance, those moving to integrated neighborhoods show lower racial homophily before the move, even when controlling for their origin neighborhood's demographics. Consistent with intergroup contact theory, preferences for the local demographic mix strengthen after the move.[5] This mover analysis also validates our model specification: the estimated preferences do not shift discontinuously when an individual's choice set changes. Although this investigation of movers' preferences is limited by smaller sample sizes and a short time horizon, it demonstrates the potential for mobility data to advance our understanding of demographic preferences.

The main contribution of this paper is to a debate over the existence and importance of demographic preferences. These preferences are believed to play an important role in explaining residential segregation (e.g., Schelling 50; Card et al. 22), but a key challenge is to distinguish preferences over neighbors' income and racial demographics from tastes for other neighborhood amenities (20; 11; 35; 51; 40).[6] By studying venue choice instead of neighborhood choice, we offer the first estimate of demographic preferences in shared spaces. The business chains we study also have considerably more uniform attributes than residential neighborhoods, and we observe many more choices.[7]

[5]The intergroup contact hypothesis was originally formulated by (3). Recent work testing this hypothesis includes (41), (21), and (18).

[6]In addition to demographic preferences and differences in tastes for neighborhood amenities, persistent racial segregation of residences has also been attributed to wealth differences, prejudice, and housing-market discrimination (23; 49).

[7]In related work, (34) show that Yelp users in New York City are more likely to visit restaurants in neighborhoods with racial composition similar to their own. (8) show that anglers are less likely to fish from a site with larger Black and Hispanic populations than their neighborhood.

We also contribute to a literature documenting segregation in non-residential domains. Economists have documented the racial segregation of friendship networks (36), gender segregation of retail venues (19), and income segregation of universities (24). Closer to this paper is recent work documenting segregation in the places people visit by race (34; 6; 9), socioeconomic status (46; 59; 42; 29; 43; 47; 60), or student status (30). We document experienced segregation within commercial spaces by income and race jointly, using building-level demographic information.[8]

Finally, our paper quantifies the drivers of income segregation in shared spaces, complementing growing evidence on the economic benefit of social connections to higher-income people. Social connections help workers find jobs through referrals (15; 10). (25; 26) find that one's number of high-socioeconomic-status Facebook friends is among the strongest predictors of economic mobility. Our smartphone movement data, however, measure social exposure, not social connections. (7) use similar smartphone data to show that serendipitous encounters in Silicon Valley in the kind of venues we study produce more patent citations between the connected employers. (27) show that the introduction of Starbucks into US neighborhoods with no coffee shops increases entrepreneurship. Beyond commercial gains, (4) argues that overlapping visits to shared spaces by people of different backgrounds may be a basis for building understanding and tolerance.

2.2 Data

To measure income segregation and estimate demographic preferences, we need to know the demographic characteristics and consumption trips of a large sample of individuals. This section describes the construction of our estimation sample from smartphone movement data and building-level demographic data. Appendix 2.8 offers more details on each data source.

[8]In this vein, (57) show that residents of Black and Hispanic neighborhoods visit high-income neighborhoods less despite traveling as far as others. In the housing market, racial differences in income mean that racial minorities face a trade-off between sorting into high same-race and high-income neighborhoods (52; 12; 17; 48). We show that minorities face the same trade-off when choosing daily social spaces. Like some studies of residential decisions (14; 2), we find that racial homophily plays an important role in this setting.

2.2.1 Data sources

Our smartphone movement data are from Precisely PlaceIQ, a location data and analytics firm. Precisely PlaceIQ aggregates pings from applications that request locational services from the smartphone's operating system.[9] Pings originating from different applications on the same smartphone are linked to a unique advertising identifier, which we denote a "device." These pings are intersected with a two-dimensional map of polygons corresponding to buildings, which we denote "venues." A spatial and temporal cluster of pings by a given device in or close to a venue constitutes a "visit" to that venue. Precisely PlaceIQ uses the timing of the first and last ping in the visit ping cluster to compute a lower bound for visit duration.

The demographic characteristics of each device come from building-level data that include the income bracket and race of individuals living at an address. Precisely PlaceIQ does not disclose the third-party provider of these data, so we discuss the reliability of this demographic information later in this section. These demographic data are aggregated across all units within a building. Thus, for single-family houses we observe the demographics of the household, while for multi-unit buildings we observe building-level averages. We assign demographics to devices based on their inferred residence, which is the residential building where the device regularly spends time at night (31).

2.2.2 Estimation sample

This subsection describes the selection of devices, venues, and visits in our analysis. To estimate preferences, we create a restricted sample of devices and visits for which we confidently know demographic information and trip purpose. To measure the demographic composition of each venue, we use a sample of devices and visits that is as broad as possible. Our sample covers the 100 largest metropolitan areas from June 1, 2018 to December 31, 2019.

[9]We do not know the set of applications contributing data. Some applications collect location data only when in active use, while others collect location data while running in the background.

Device selection criteria Around 66 million devices have exactly one inferred residence over our 19-month sample period.[10] Around 46 million of these devices live in buildings for which we have demographic data. We classify building-level demographics in terms of two income groups and four racial groups: the share of a building's residents with household income above $75,000 (the bracket cutoff closest to the national median in 2019) and the shares of a building's residents who are Asian, Black, Hispanic, and White. We use visits by all devices for which we have building-level data to measure the demographic mix of co-patrons in each venue, applying device demographics probabilistically. To estimate group-specific preferences, we limit our estimation sample to the 36 million devices that reside in buildings in which at least two-thirds of the residents belong to the same income and racial group. 91% of buildings are racially homogeneous and 99% of buildings are economically homogeneous, consistent with most Americans living in single-family homes and significant sorting by residents of multi-family dwellings.

Venue selection criteria Our baseline estimation sample uses trips to restaurants, which have by far the largest number of chains, establishments, and visits. We also characterize co-patron exposure and estimate demographic preferences in banks, big box retail stores, convenience stores & gas stations, grocery stores, gyms, and pharmacies.

Table 2.8 compares the number of venues we observe in the 10 largest restaurant chains to counts from external sources. We observe on average 87% of venues within these chains, with a low of 70% for Starbucks. Since the smallest spatial unit of observation in our data is a building, we exclude venues that contain multiple establishments, such as shopping malls. To avoid measurement concerns related to venue entry and exit, we only keep venues with at least one visit prior to the beginning of our estimation window (June 1, 2018) and one visit after the end of our estimation window (December 31, 2019). This excludes around 10% of chain venues from the estimation sample.

[10]Around 10% of devices move during this period. We drop these from our baseline estimation sample. We return to studying these movers in Section 2.6.2.

Visit selection criteria Our estimation sample is restricted to trips to a venue that originate and end at home.[11] Considering only these direct trips ensures that their sole purpose is visiting a venue. This selection eliminates confounding factors due to trip chaining and allows us to identify preferences within a venue-choice model like that we introduce in Section 2.4. It means the estimation sample only includes devices that ping frequently enough to track direct trips. We also exclude visits with duration longer than three hours, as these are likely by venue employees.[12] Overall, the sample of restaurant visits we use to estimate preferences includes more than 14 million direct trips to more than 27,000 restaurant chain venues by almost 4 million devices who live in homogeneous buildings. To estimate the demographic composition of these venues, we use the sample of all 1.5 billion trips to these restaurants. While we rely on a narrower sample of visits to estimate preferences, we later show that visit patterns in our estimation sample mirror those from much broader samples.

2.2.3 Data quality and representativeness

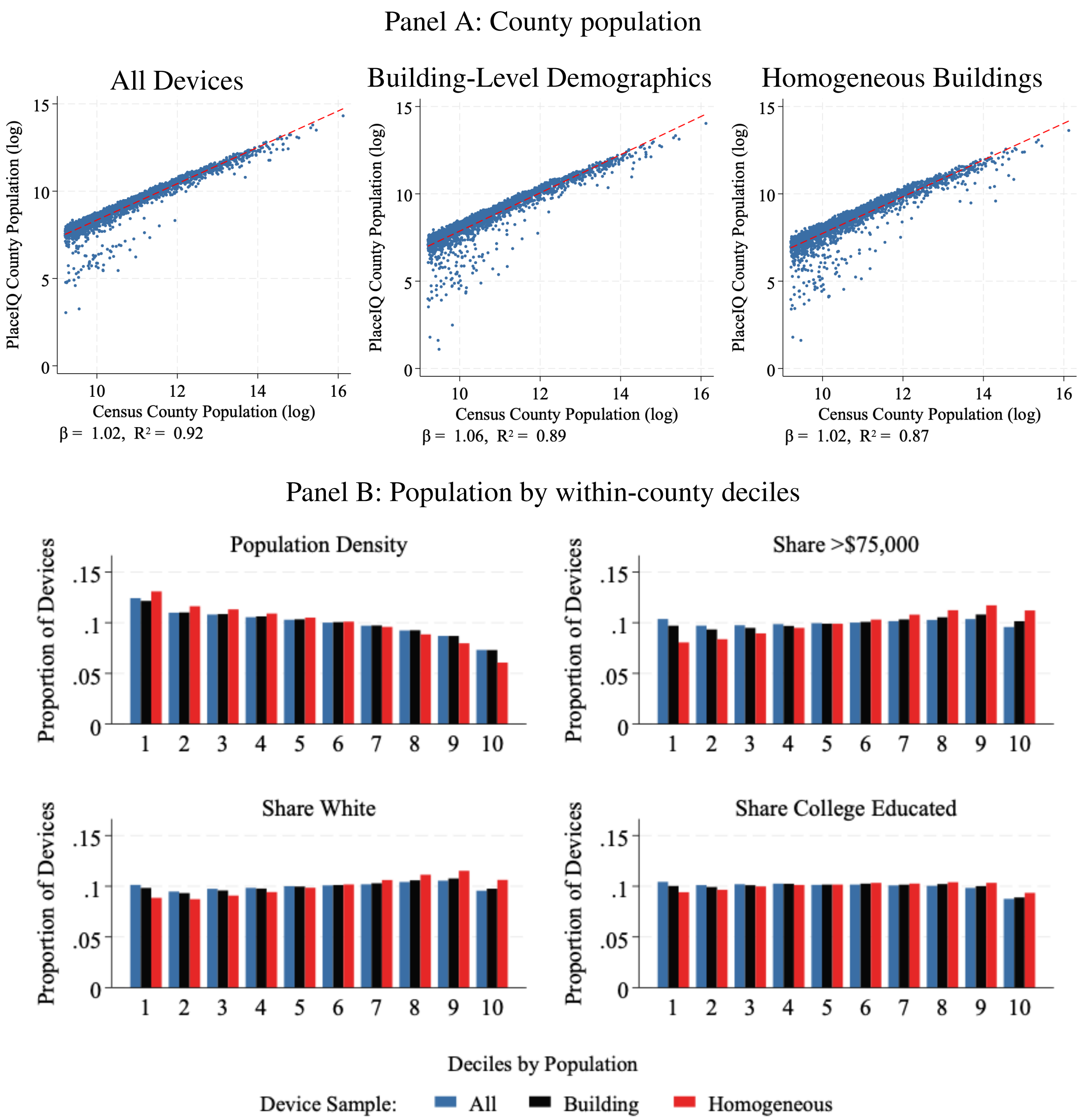
In this subsection, we first assess whether our device selection criteria bias our estimation sample. We then evaluate the reliability of the demographic information in the building-level data.

(31) show that devices active in the smartphone data are broadly spatially representative and make visits that resemble what travelers self-report in the National Household Travel Survey. Figure 2.1 shows that the additional selection criteria we impose on our estimation sample involve only limited spatial biases. Panel A plots the number of devices residing in a county for three device samples against the 2019 Census county population estimates. The “All Devices” sample includes all devices that have exactly one home assignment. The “Building” sample includes only devices with building-level demographic data. The “Homogeneous Building” sample includes only de-

[11]We define a direct trip as a visit to a venue where the preceding and succeeding visits were to a device’s home and within the same “activity day” (a 24-hour period starting at 3:00 AM). Since not every stop at a venue is observed, some trips may be mis-categorized as direct. (34) and (45) study consumption trips that can originate at workplaces.

[12]Visit duration is measured with error, but can be reasonably interpreted as a lower bound for actual duration. A visit is registered when a smartphone application collects a ping in a venue, not when the smartphone first enters the venue, so a device may spend more time at a venue than we observe.

Figure 2.1: Comparing estimation sample to full smartphone data



NOTES: Panel A compares the number of devices residing in a county (vertical axis) to the Census's estimated 2019 residential populations using three different device selection criteria: (1) all devices residing in exactly one residential building between June 1, 2018 and December 31, 2019; (2) among those devices in (1), the devices whose building-of-residence has demographic data available; (3) devices whose building-of-residence is comprised of at least 67% one income group and racial group. We exclude counties with a Census population of less than 10,000 people. Panel B depicts the share of devices living in block groups within each within-county population decile for four characteristics: population density, population share of high-income ($> \$75,000$) residents, population share of White residents, and population share of residents who have obtained a bachelor's degree. Panel B reports these decile shares for the three populations of devices shown in Panel A.

vices living in buildings in which one income-race group constitutes more than 67% of residents. Regressing the device count of a county on its Census population count yields an R^2 of at least 0.87 for each of the three device samples. Our estimation sample is nearly as representative of county population as all smartphone data.

Panel B of Figure 2.1 evaluates the spatial representativeness of our sample within counties. For instance, we show that within a county, we have about the same number of devices in block groups with the highest White share as in block groups with the lowest White share. Specifically, we compute the share of devices living in each decile of population density, high-income share, White share, and college-educated share, defined within each county using block-group data from the 2015-2019 American Community Survey. If device samples were exactly proportionate to populations in the American Community Survey within each county, each bar would be of equal height (0.10). We show these results for the three different device samples above.

The bar heights are very similar in the "All" and "Building" device samples. This alleviates concerns over spatial bias in the building-level data that we use to compute the demographic composition of venues. When we restrict attention to homogeneous buildings, as we do in the estimation sample of devices, we see a more substantial bias away from the densest block groups, with about six percent of devices living within the top density decile, and a slight bias towards more heavily White and high-income block groups. Thus, our estimation sample of devices living in homogeneous buildings is broadly spatially representative, with the exception of devices living in the top density decile (i.e., in multi-unit buildings) being somewhat underrepresented.

Figure 2.1 demonstrates that our smartphone device samples are spatially representative of residences using Census characteristics, but Appendix 2.8.4 shows that the building-level data still contain more White and high-income households. That said, the cross-county correlation between the share of a given demographic group in the Census and the share of devices in that group based on our building-level data remains above 0.8 for all demographic groups (Figure 2.7). These deviations from perfect representativeness are expected in smartphone samples, but they warrant some caution when measuring the demographic composition of co-patrons within restaurant venues. We therefore follow (30) by reporting differences in demographic exposure across groups instead of absolute levels that may overstate exposure to high-income devices.

Finally, Appendix 2.8.5 validates our building-level demographic information in three ways. In each exercise, we show that the building-level data predict individual characteristics or behavior *within* the smallest geography for which Census data are available. First, we compare the incomes in our building-level data to those of (29), which imputes income using home parcel characteristics from CoreLogic. These two sources agree whether a Census block has above-median income for 75% of blocks. The discrepancies are modest: treating blocks within $10,000 of the median as matching raises the agreement to 91%. These block-level measures are significantly correlated within block groups, suggesting that our building-level data reveal income variation below the smallest geographic level for which the Census publishes income statistics.

Second, we use voter registration data from North Carolina to show that the building-level demographic information predicts the race of voters at a given address better than one could using Census data. Our building-level data matches the race of Black voters 20% more often than one could by randomly drawing households in the same Census block group - the smallest geography for which both race and income are available - and 5% more often than one could using Census block data on race.

Third, we run an internal validation check: we show that the building-level demographic information predicts differences in the venue choices of residents of different buildings in the same Census block group. For example, residents of high-income buildings are more likely than their low-income neighbors to visit chains preferred by high-income people, such as Starbucks.[13] This final exercise delivers two conclusions that we leverage in our empirical analysis. First, behavior predicted using the building-level demographic data is consistent with behavior predicted using Census demographic data. Second, the building-level data provides more information than the

[13] To make such comparisons, we first calculate each restaurant chain's relative popularity among high-income versus low-income patrons within the same tract, using block-group-level income assignments. We then compute each chain's relative popularity at the building level, contrasting high- and low-income buildings in the same block group. These block group-level and building-level measures of a chain's relative popularity with high-income patrons have a rank correlation of 0.8 (Figure 2.8). We find similarly high correlations for racial instead of income groups and for convenience stores & gas stations (the second largest establishment category).

Census, because it allows us to predict variation in behavior within the smallest geographic unit for which both income and race are available in the Census.

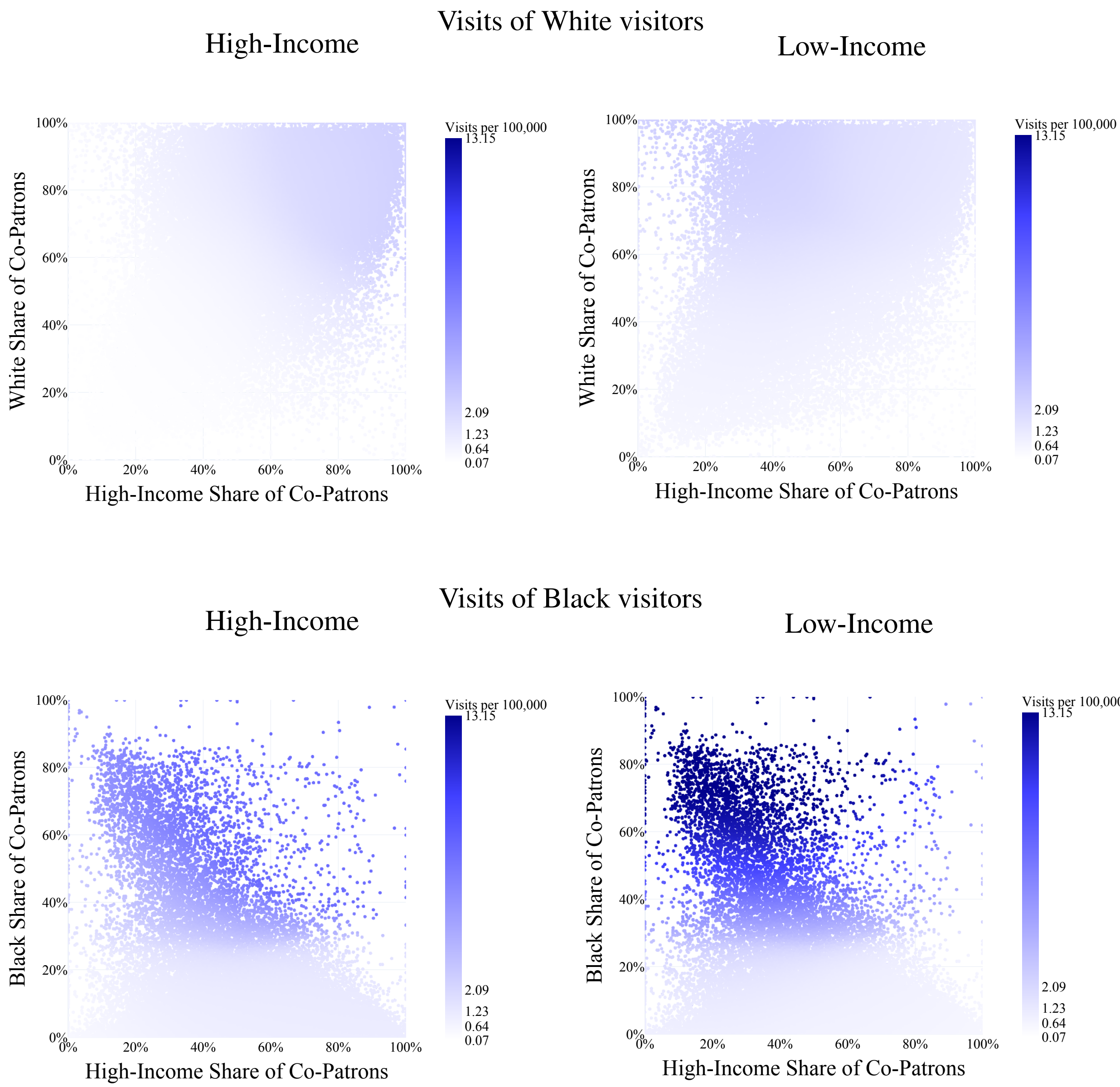
2.3 Demographic exposure in shared spaces

This section documents how exposure to different types of co-patrons varies by demographic group. We first report, for each demographic group, the full distribution of visits to chain restaurant venues by racial and income co-patron mix. We then show how these differences in visit patterns translate into disparities in income exposure across groups. Finally, we show that other categories of venues and visits have similar levels of income segregation.

Figure 2.2 shows the racial and income composition of venues and the propensity of each group in our estimation sample to visit venues by co-patron demographics. Each dot in the plot represents a chain restaurant within the 100 largest metropolitan statistical areas (MSAs). We compute visit propensity using a non-parametric kernel regression of a demographic group's share of visits to a venue on its co-patron characteristics.

These plots show notable patterns in visit propensity across demographic groups. For all groups, visit propensity is increasing in same-race share. Within each race, higher-income individuals visit venues with greater shares of high-income co-patrons than their low-income peers. However, the distribution of available co-patron demographics varies starkly across racial groups. Unsurprisingly, many more venues are predominately White than Asian, Black, or Hispanic. Heavily White venues vary significantly in their income composition, whereas heavily Black and Hispanic venues tend to be low-income. There are very few venues with a high Asian share of co-patrons, regardless of co-patron income. These visit patterns echo familiar patterns of residential segregation by race and assortative matching by income (e.g., Bayer and McMillan 14).

Figure 2.2: Exposure to high-income and same-race co-patrons



NOTES: This figure presents the results of a kernel regression of visit shares on co-patron race and income characteristics. Each panel displays the visit propensities for a specific demographic group, with each dot representing an individual restaurant. The color gradient is consistent across groups and follows a log-linear scale. This gradient is chosen to match the group with the greatest range in visit propensities between the most and least visited venues. We use an Epanechnikov kernel with a bandwidth of 0.05 and winsorize visit propensities at the 99th percentile for each group. *Continues on next page.*

Figure 2.2: Exposure to high-income and same-tace co-patrons (continued)

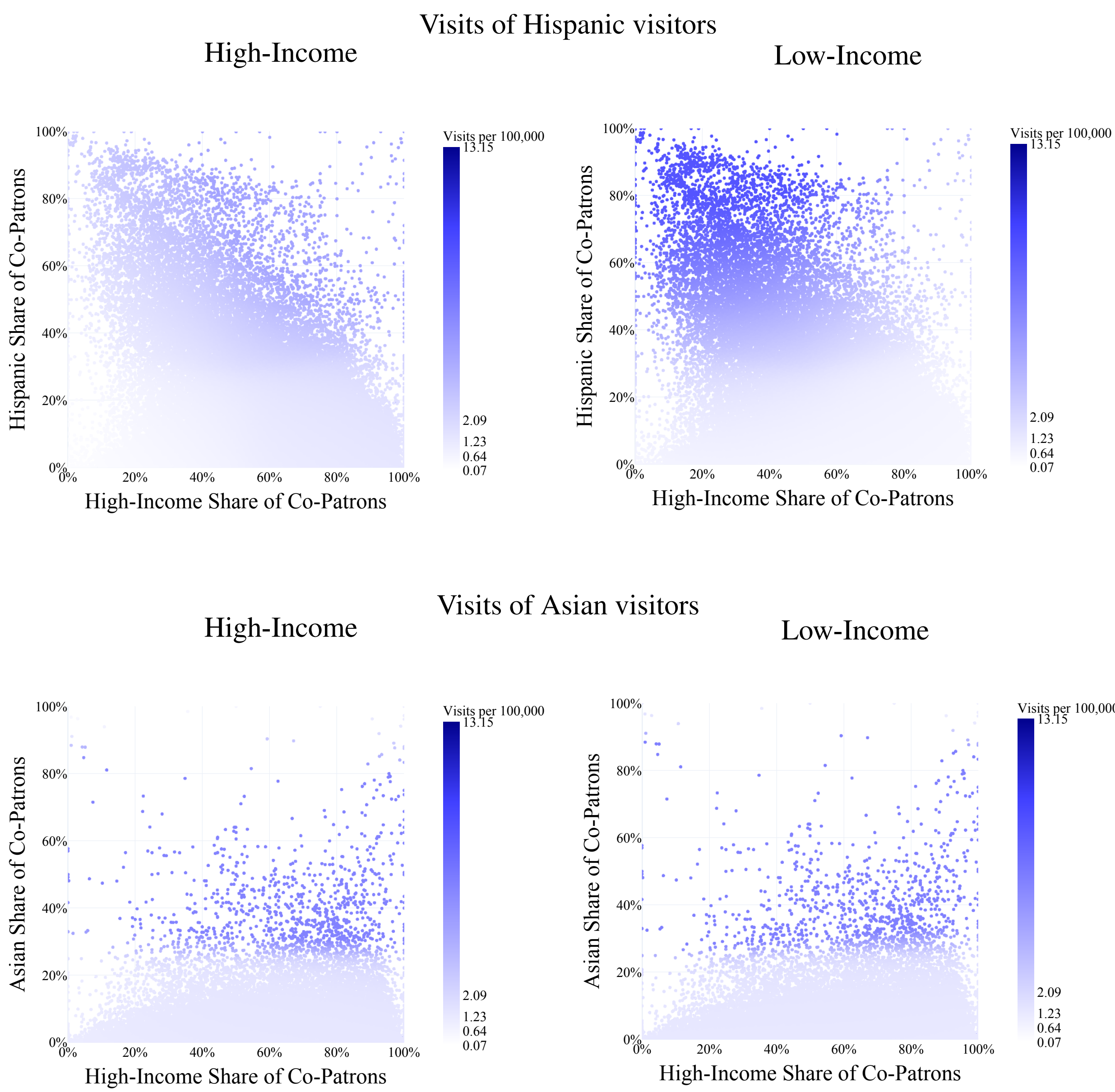


Table 2.1: Exposure to high-income co-patrons

	Low Income				High Income			
	Asian	Black	Hispanic	White	Asian	Black	Hispanic	White
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Estimation sample	-0.06	-0.17	-0.15	-0.03	0.14	-0.00	0.05	0.13
All chain-restaurant visits	-0.04	-0.13	-0.13	-0.02	0.12	-0.00	0.04	0.12
All chain-venue visits	-0.05	-0.16	-0.14	-0.03	0.12	-0.01	0.04	0.11
All non-residence venue visits	-0.03	-0.15	-0.14	-0.03	0.18	0.04	0.09	0.17
All McDonald's restaurant visits	-0.08	-0.16	-0.16	-0.03	0.09	-0.02	0.02	0.11
Census tracts	0.04	-0.12	-0.08	-0.01	0.20	0.03	0.06	0.11

NOTES: This table reports, for different visit samples, the high-income share of co-patrons that each demographic groups (eight columns) is exposed to, relative to a baseline in which all venues in that sample are visited with uniform probability. The first row shows those high-income shares for visits in the estimation sample. The second through fourth rows shows those shares for broader visit samples. The fifth row shows those shares only for visits to McDonald's restaurants. In the sixth row, those shares are computed as if each Census tract is a venue and individuals only visit the census tract that they live in.

How do these differences in visit propensity and venue availability translate into differences in income exposure? Table 2.1 reports exposure to high-income co-patrons for each demographic group across different environments. The first row reports high-income exposure computed using visits from our estimation sample. We chose this sample to suit our empirical strategy, not for its representativeness, so in subsequent rows we expand our sample to larger sets of visits and venues. The last row reports a measure of residential income exposure, computed using only Census tables, as if people's exposure equaled the income composition of their census tract of residence. This last row offers a useful benchmark to evaluate how income segregation in shared spaces compares with residential segregation.

Each column reports the income exposure of a given demographic group relative to a baseline in which these venues are visited uniformly. For example, column 8 of row 1 shows that high-income White individuals in our estimation sample choose restaurants with a share of high-income co-patron that is 13 percentage points higher than if they visited restaurants at random. We report our results in relative terms because absolute exposure levels are more sensitive to definitions of income group and sample biases.[14]

Table 2.1 yields two main results. First, there are substantial differences in exposure to high-income co-patrons across incomes and races. Within each race, high-income individuals have more high-income exposure: differences in mean exposure between high- and low-income individuals are typically 15 to 20 percentage points. Within each income group, Asian and White individuals have more high-income exposure than Black and Hispanic individuals. In fact, the average high-income exposure of a low-income White individual is only about 2 percentage points lower than that of a high-income Black individual. Second, we observe similar cross-group disparities in income exposure in a variety of settings. In particular, income segregation experienced within chain restaurants, based solely on visits in our estimation sample, resembles that experienced within all non-residential venues, both commercial and public. These broad income segregation patterns even manifest within a single chain like McDonald's. There is an approximately 25 percentage point difference between the high-income share in McDonald's locations visited by low-income Black or Hispanic individuals and those visited by high-income Asian or White individuals. These disparities in income exposure within shared spaces also mirror disparities in residential income exposure, computed using census tract demographic shares.[15]

A number of factors may explain this variation in exposure to high-income co-patrons across demographic groups. First, different groups may have different tastes for product attributes like cuisine and ambiance. Second, groups are distributed differently across cities and neighborhoods and thus might have to incur higher travel costs to patronize venues with larger high-income shares. Third, different groups may have different preferences for exposure to high-income co-patrons.

[14]The patterns in Table 2.1 are robust to different ways of weighting each device that correct for biases in the smartphone sample. There may also be differences in exposure at the extensive margin, from variation in the number of visits that each demographic group makes. Smartphone samples with a partial history of each device's movements are not well-suited to study these differences. The National Household Transportation Survey, for instance, shows that despite differences in high- and low-income individuals' propensity to visit different destination types – high-income people make more trips to restaurant – high- and low-income people make similar numbers of consumption trips and similar numbers of non-work trips.

[15]Appendix Table 2.9 shows similar regularities across environments for racial segregation.

Preferences for same-race co-patrons may generate differences in high-income exposure because venues with high-income co-patrons tend to be heavily White (Figure 2.2). In what follows, we investigate the relative importance of these explanations for demographic differences in high-income exposure.

2.4 Model

This section introduces a model of individuals' decisions to patronize venues within business chains as a function of transit costs and co-patron composition. The model delivers an estimating equation for each demographic group's preferences over co-patron demographics and travel distances. We then describe how to use the estimated model to compute counterfactual venue visit shares that quantify the contributions of various mechanisms to cross-group variation in exposure to high-income co-patrons.

2.4.1 Nested-logit preferences

We develop a nested-logit model of consumers' decisions to visit venues. We index decision makers by i, venues by j, and chains by c. A decision maker is an individual at a point in time. Denote the set of venues from which a decision maker chooses by $\mathcal{J}$. The utility that decision maker i would obtain from choosing venue j is

$$U_{ij} = V_{ij} + \epsilon_{ij},$$

where V_{ij} is a scalar that depends on preference parameters and observed covariates and ϵ_{ij} is a random component. Decision maker i chooses the venue $j \in \mathcal{J}$ that has the highest value of U_{ij}.

We assume that ϵ has an extreme-value distribution such that consumers have nested-logit preferences over business chains. We partition the set of venues into C disjoint subsets denoted by B_c (chains). Following Train (54, Ch 4.2), we denote the similarity of idiosyncratic preferences for establishments in nest B_c by $1-\lambda_c$, so that $\lambda_c = 1\ \forall c$ is the canonical multinomial logit case. The probability that decision maker i chooses venue j is

$$P_{j|i} = \frac{\exp\left(V_{ij}/\lambda_c\right)\left(\sum_{j'\in B_c}\exp\left(V_{ij'}/\lambda_c\right)\right)^{\lambda_c-1}}{\sum_{c'=1}^{C}\left(\sum_{j'\in B_{c'}}\exp\left(V_{ij'}/\lambda_{c'}\right)\right)^{\lambda_{c'}}}. \tag{2.1}$$

2.4.2 Within-chain choice probabilities

If the utility shifter V_{ij} depends on preference parameters Γ, the log likelihood function associated with the choice probability (2.1) is

$$LL(\Gamma) = \sum_i \sum_j I_{ij} \ln P_{j|i},$$

where $I_{ij} = 1$ if i chooses j.

Following Train (54, p.82), the choice probability (2.1) can be rewritten as the product of within-chain and between-chain components: $P_{j|i} = P_{j|ic} \times P_{c|i}$. Thus, we can rewrite the log likelihood function as

$$LL(\Gamma) = \sum_i \sum_j I_{ij}\left(\ln P_{j|ic} + \ln P_{c|i}\right) = \sum_i \sum_j I_{ij} \ln P_{j|ic} + \sum_i \sum_j I_{ij} \ln P_{c|i}.$$

While a model of $P_{c|i}$ must incorporate the parameters appearing in $P_{j|ic}$ via an "inclusive value" term, we can maximize the first likelihood component, $\sum_i \sum_j I_{ij} \ln P_{j|ic}$, alone. We estimate preference parameters of interest using only within-chain variation, leveraging the conditional choice probability

$$P_{j|ic} = \frac{\exp\left(Y_{ij}/\lambda_c\right)}{\sum_{j'\in B_c}\exp\left(Y_{ij'}/\lambda_c\right)},$$

where Y_{ij} denotes the component of V_{ij} that varies across venues within chain c.[16] We allow Y_{ij} to depend on observable attributes with coefficients that are common across chains. In order to identify parameters common across chains, we assume a common within-chain correlation of idiosyncratic preference shocks such that $\lambda_c = \lambda \; \forall c$.

2.4.3 Mean utility specification

In the baseline specification, we assume that the mean utility of patronizing a venue within a chain depends on the distance to the venue and its co-patron composition. These preferences may vary across demographic groups, indexed by g. Preferences over distance and co-patron composition are additively separable. In particular, the component of utility that varies across venues within a chain, Y_{ij}, depends on the distance from the consumer's home to the venue (distance_{ij}), the high-income share of co-patrons (s_j^{highinc}), and the same-race share of co-patrons (s_j^{samerace}):

$$Y_{ij} = f_1(\ln \text{distance}_{ij}; \delta^g) + f_2(s_j^{\text{samerace}}, s_j^{\text{highinc}}; \beta^g),$$

where $f_1(\ln \text{distance}_{ij}; \delta^g)$ is a polynomial of log distance, $f_2(s_j^{\text{samerace}}, s_j^{\text{highinc}}; \beta^g)$ is a polynomial of the two co-patron shares, and δ^g and β^g are group-specific coefficient vectors on travel distance and co-patron composition, respectively.

Choosing the degrees of the polynomials $f_1()$ and $f_2()$ involves a trade-off between parametric flexibility and statistical power. Our baseline specification uses second-degree polynomials:

$$\begin{aligned} Y_{ij} = {} & \delta_1^g \ln \text{distance}_{ij} + \delta_2^g \left(\ln \text{distance}_{ij}\right)^2 \\ & + \beta_1^g s_j^{\text{highinc}} + \beta_2^g \left(s_j^{\text{highinc}}\right)^2 + \beta_3^g s_j^{\text{samerace}} + \beta_4^g \left(s_j^{\text{samerace}}\right)^2 + \beta_5^g s_j^{\text{samerace}} \times s_j^{\text{highinc}}. \end{aligned} \tag{2.2}$$

[16]Chain-level attributes Y_{ic} that are common across establishments within a chain, such as menu items and prices, do not affect $P_{j|ic}$:

$$P_{j|ic} = \frac{\exp\left(\left[Y_{ij} + Y_{ic}\right]/\lambda_c\right)}{\sum_{j' \in B_c} \exp\left(\left[Y_{ij'} + Y_{ic}\right]/\lambda_c\right)} = \frac{\exp\left(Y_{ic}/\lambda_c\right) \exp\left(\left[Y_{ij}/\lambda_c\right)\right.}{\exp\left(Y_{ic}/\lambda_c\right) \sum_{j' \in B_c} \exp\left(Y_{ij'}/\lambda_c\right)} = \frac{\exp\left(Y_{ij}/\lambda_c\right)}{\sum_{j' \in B_c} \exp\left(Y_{ij'}/\lambda_c\right)}.$$

Second-degree polynomials are flexible enough to fit observed choice patterns well and parsimonious enough to be precisely estimated for the demographic groups with modest numbers of observations.

We can express preferences over co-patron composition in terms of willingness to travel. We define group g's willingness to travel for the co-patron composition $(s^{\text{samerace}}, s^{\text{highinc}})$ as the incremental distance Δ^g that equates the mean utility of a venue at the average distance with co-patron composition $(s^{\text{samerace}}, s^{\text{highinc}})$ and a venue at the average distance plus the incrementΔ^g with the average co-patron composition:

$$\begin{aligned} & f_1\left(\ln \overline{\text{distance}}; \delta^g\right) + f_2\left(s^{\text{samerace}}, s^{\text{highinc}}; \beta^g\right) \\ & \quad = f_1\left(\ln\left(\overline{\text{distance}} + \Delta^g(s^{\text{samerace}}, s^{\text{highinc}})\right); \delta^g\right) + f_2\left(\overline{s}^{\text{samerace}}, \overline{s}^{\text{highinc}}; \beta^g\right), \end{aligned} \tag{2.3}$$

where $(\overline{s}^{\text{samerace}}, \overline{s}^{\text{highinc}})$ and $\overline{\text{distance}}$ denote the characteristics of the average venue. The function $\Delta^g(s^{\text{samerace}}, s^{\text{highinc}})$ is group g's willingness to travel for co-patron composition $(s^{\text{samerace}}, s^{\text{highinc}})$.[17]

2.4.4 Maximum likelihood estimation

We estimate the preference coefficients by maximizing the likelihood component $\sum_i \sum_j I_{ij} \ln P_{j|ic}$. The optimization problem is

$$\max_{\delta^g, \beta^g} \sum_i \sum_j I_{ij} \ln\left(\frac{\exp\left(\left[f_1(\ln \text{distance}_{ij}; \delta^g) + f_2(s_j^{\text{samerace}}, s_j^{\text{highinc}}; \beta^g)\right]/\lambda\right)}{\sum_{j' \in B_c} \exp\left(\left[f_1(\ln \text{distance}_{ij'}; \delta^g) + f_2(s_{j'}^{\text{samerace}}, s_{j'}^{\text{highinc}}; \beta^g)\right]/\lambda\right)}\right). \tag{2.4}$$

Since each parameter is g-specific, the model can be estimated separately by demographic group.

[17] The equation implicitly defines this function. Given the functional form used in equation (2.2), there is a closed-form expression for $\Delta^g(s^{\text{samerace}}, s^{\text{highinc}})$.

2.4.5 Empirical implementation

We estimate consumer preferences using 19 months of data on devices in the 100 most populous US metropolitan areas, as described in Section 2.2.2. We estimate the model separately by demographic group and business category, so our baseline estimates of δ^g and β^g are specific to both demographic group g and the restaurant category. We assume that consumers consider all venues within their metropolitan area, so the nest B_c is the set of venues that belong to both the same business chain and metropolitan area. We exclude metro-chain pairs with very few observations and randomly sample a subset of observations from those with very many.[18]

We estimate consumer preferences using within-chain comparisons in order to distinguish consumer preferences over co-patron composition from other venue characteristics. Co-patron composition may correlate with other traits in the very broad set of (potentially unobserved) characteristics entering V_{ij}, such as service quality, comfort, or product offering. The set of characteristics contributing to Y_{ij}, which vary across venues within a chain business and metropolitan area, is substantially smaller. Venues in the same chain are similar because brand power and economies of scale depend on a standardized offering. In particular, the food products and quality of service typically are not a reflection of the local composition of co-patrons within each venue. In robustness checks, we focus on a subset of the most standardized chains based on franchising terms and dispersion in establishment-level characteristics.

One may worry that co-patron composition might predict venue choice if patrons of nearby venues are demographically similar because of residential segregation. Our specification addresses this concern, because we control for bilateral distance as an individual-by-venue-specific cost shifter.[19] Moreover, our estimation sample of direct trips from home is a small share of the to-

[18]In particular, we keep group-MSA-chain triplets that have at least 25 direct visits. When a group-MSA-chain triplet has more than 75 venues, we randomly sample only 75 venues. We randomly sample 20,000 devices from group-MSA pairs that have more than 20,000 devices and only keep group-MSA pairs with at least 200 visits. From this set, we randomly sample 750,000 visits per demographic group to reduce computational burden.

[19]We could extend our venue choice model to account for residential choice as in (34) Appendix C.6. This shows that residential sorting does not bias estimates of preferences for venues as long as individuals choose a residence based on the expected utility of their venue choice set (and not on their idiosyncratic preference for a specific venue).

tal visits to venues, so the observed co-patron composition is not driven by the choice shares in our estimation sample.

For brevity, we refer to all mechanisms that cause co-patron composition to predict consumer decisions as preferences over co-patron composition. Of course, co-patron composition may predict decisions not because consumers have preferences over co-patron demographics but because co-patron demographics predict other elements of the decision. For example, a consumer who is indifferent to strangers' demographics may choose a venue in order to meet up with their demographically similar friends.[20] This behavior could generate the same observed outcomes as a consumer who has homophilic preferences over anonymous co-patrons. We need not separate homophily among strangers and homophily in friendship networks to quantify the importance of homophily in explaining cross-group differences in experienced income exposure. Similarly, if consumers are not aware of all the venues in their choice set, co-patron demographics could predict consideration sets. Our estimation approach would infer homophilic preferences if demographically similar venues are more likely to be considered. While this distinction would be important when considering some counterfactual scenarios, our decomposition of observed exposure to high-income co-patrons will not distinguish preferences over co-patron demographics from consideration sets that vary with demographics.

2.4.6 Decomposition of exposure to high-income co-patrons

We decompose exposure to high-income co-patrons by contrasting the distribution of visits across venues in our fitted model with the distributions resulting from various counterfactual market shares. We summarize exposure to high-income co-patrons for members of group g by fitting a density $f^g(\cdot)$ using kernel $K(\cdot)$ and bandwidth h to the high-income share in each venue:

$$\hat{f}^g\left(s^{\text{highinc}}\right) = \frac{1}{h}\sum_{j\in\mathcal{J}} K\left(\frac{s^{\text{highinc}} - s_j^{\text{highinc}}}{h}\right) P_{j|g}. \tag{2.5}$$

[20] We also replicate our preference estimation in categories other than restaurants — like banks and big box stores — where meeting friends is less likely.

To compute our benchmark 'model-predicted' distribution of exposure to high-income co-patrons, we use the model-predicted share of visits to each venue j by group g,

$$P_{j|g} = P_{j|ic} \times P_{ic|g}, \tag{2.6}$$

where $P_{j|ic}$ comes from our estimated model of within-chain venue choice and $P_{ic|g}$ comes from the observed distribution of visits to each chain by members of a demographic group.

To quantify the contributions of various mechanisms to income exposure, we report the distributions resulting from various counterfactual market shares $P_{j|g}$.[21] A simple starting point is the observed distribution of high-income co-patrons across all venues. This is the density that results from evaluating equation (2.5) using a uniform probability of visiting venues, $P_{j|g} = \frac{1}{|\mathcal{J}|}$. To quantify the contribution of between-MSA variation in demographics to exposure to high-income co-patrons, we use a uniform probability conditional on metropolitan area m, $P_{j|g} = \frac{1}{|\mathcal{J}_m|} P_{m|g}$. The difference between the nationwide uniform probability $\frac{1}{|\mathcal{J}|}$ and this measure captures the contribution of demographic differences between MSAs to income exposure. To quantify the contribution of between-chain variation in demographics, we use a uniform probability conditional on metropolitan area and business chain, $P_{j|g} = \frac{1}{|\mathcal{J}_{mc}|} P_{mc|g}$. To quantify the contribution of residential proximity, we compute market shares with counterfactual probabilities $P_{j|ic}$ using the estimated distance parameters $\hat{\delta}^g$ absent any co-patron preferences ($\beta^g = \mathbf{0}$). To quantify the contribution of preferences over co-patron composition, we compute market shares with counterfactual probabilities $P_{j|ic}$ using the estimated co-patron preference parameters $\hat{\beta}^g$ absent any disutility of travel distance ($\delta^g = \mathbf{0}$).[22]

[21] This is similar to the approach (40) uses to quantify the role of preferences versus constraints in generating residential racial segregation in the U.S. and (44) use to quantify the role of the exogenous school environment versus preferences in the formation of social links.

[22] The last two counterfactual scenarios in which we set $\beta^g = \mathbf{0}$ or $\delta^g = \mathbf{0}$ are non-nested scenarios. Nested scenarios, such as the stratified uniform-probability scenarios (e.g., within-MSA to within chain-MSA), lend themselves to simple comparisons because they differ in only one respect. Non-nested scenarios, such as alternatively setting preference parameters for distance or for co-patron composition to zero, must be interpreted carefully. These alternative scenarios do not provide an additive decomposition of the observed outcomes, because marginal effects are non-linear functions of the coefficients and covariates. We address this issue further in our discussion of these results in Section 2.6.

2.5 Estimation results

In this section, we estimate preferences over travel distance and co-patron composition. We document notable regularities in demographic preferences: different income and racial groups display similar levels of racial homophily. Preferences for high-income co-patrons are also similar across racial groups, but lower-income individuals have less pronounced tastes for co-patron income. These preference patterns are consistent across a number of robustness checks, including restricting our estimation sample to the most standardized restaurant chains.

2.5.1 Estimated preference parameters

Table 2.2 reports estimates of the preference parameters in equation (2.2) for each of the eight demographic groups. Panel A reports estimates of the distance coefficients δ^g, and Panel B reports estimates of the co-patron composition coefficients β^g.

Our estimates of distance coefficients δ^g imply distance elasticities around -2.2. These distance elasticities capture both the cost of longer travel distances and the substitutability of venues within a restaurant chain. Higher-income individuals have larger distance elasticities, consistent with empirical evidence that the value of time spent traveling rises with income (53). Our distance elasticities are larger than previous estimates from studies that consider venue choice among all restaurants, consistent with venues within the same chain being closer substitutes.[23]

Table 2.2 Panel B reports, for each of the eight demographic groups, estimates of all five coefficients in β^g, which together govern preferences for the high-income share of co-patrons and the same-race share of co-patrons. Figures 2.3 and 2.4 present two different representations of these preferences. Figure 2.3 depicts preferences over both the income and race of co-patrons, with the preferences of each demographic group in a separate heatmap. Each point represents a chain restaurant venue in the 100 largest metropolitan areas. The color of the venue captures willingness to travel to that venue in co-patron composition space, $\Delta^g(s^{\text{samerace}}, s^{\text{highinc}})$ from equation (2.3). To

[23](5), (34), and (33) estimate elasticities between -1.0 and -1.5 when considering substitution between all restaurants or non-tradable services in a given city.

Table 2.2: Preference estimates

Panel A. Estimates of distance coefficients δ^g

		Estimates		Distance Elasticity at		
		Linear δ_1^g	Quadratic δ_2^g	Mean	25p	75p
Low-Income	Asian	-1.29	-0.23	-2.15	-1.59	-2.24
Low-Income	Black	-1.25	-0.23	-2.11	-1.54	-2.20
Low-Income	Hispanic	-1.26	-0.27	-2.24	-1.59	-2.35
Low-Income	White	-1.11	-0.32	-2.29	-1.51	-2.42
High-Income	Asian	-1.25	-0.26	-2.21	-1.58	-2.32
High-Income	Black	-1.11	-0.30	-2.24	-1.50	-2.36
High-Income	Hispanic	-1.22	-0.30	-2.34	-1.60	-2.46
High-Income	White	-1.00	-0.36	-2.33	-1.46	-2.48

Panel B. Estimates of co-patron composition coefficients β^g

		β_1^g Linear High-Income	β_2^g Quadratic High-Income	β_3^g Linear Same-Race	β_4^g Quadratic Same-Race	β_5^g Interaction Term
Low-Income	Asian	4.13	-3.20	12.19	-9.93	-4.02
		(0.08)	(0.07)	(0.12)	(0.16)	(0.14)
Low-Income	Black	5.06	-4.20	4.94	-3.86	-2.18
		(0.07)	(0.05)	(0.05)	(0.04)	(0.06)
Low-Income	Hispanic	6.08	-4.21	7.08	-4.99	-4.23
		(0.07)	(0.05)	(0.08)	(0.06)	(0.07)
Low-Income	White	2.94	-4.13	2.17	-1.70	2.35
		(0.07)	(0.06)	(0.09)	(0.08)	(0.09)
High-Income	Asian	5.55	-3.21	12.78	-12.31	-2.21
		(0.07)	(0.05)	(0.08)	(0.10)	(0.09)
High-Income	Black	5.35	-3.18	4.15	-3.61	-0.29
		(0.07)	(0.05)	(0.05)	(0.04)	(0.05)
High-Income	Hispanic	7.72	-4.18	7.14	-5.10	-4.49
		(0.09)	(0.06)	(0.09)	(0.08)	(0.08)
High-Income	White	4.21	-3.65	4.47	-3.70	2.59
		(0.09)	(0.07)	(0.11)	(0.09)	(0.10)

NOTES: This table reports estimates of the preference parameters in equation (2.2). Panel A reports estimates of the distance coefficients δ^g, and Panel B reports estimates of the co-patron composition coefficients β^g. Panel A reports the distance elasticity at the mean and 25$^{\text{th}}$ and 75$^{\text{th}}$ percentiles of trip distance, which are 6.4, 1.9, and 7.8 kilometers, respectively.

facilitate comparisons across groups, Figure 2.4 depicts the preferences of all eight demographic groups over one dimension of co-patron composition on the same plot, fixing the other dimension at its median value for each group. This is akin to looking at variation along one horizontal or vertical slice of Figure 2.3.

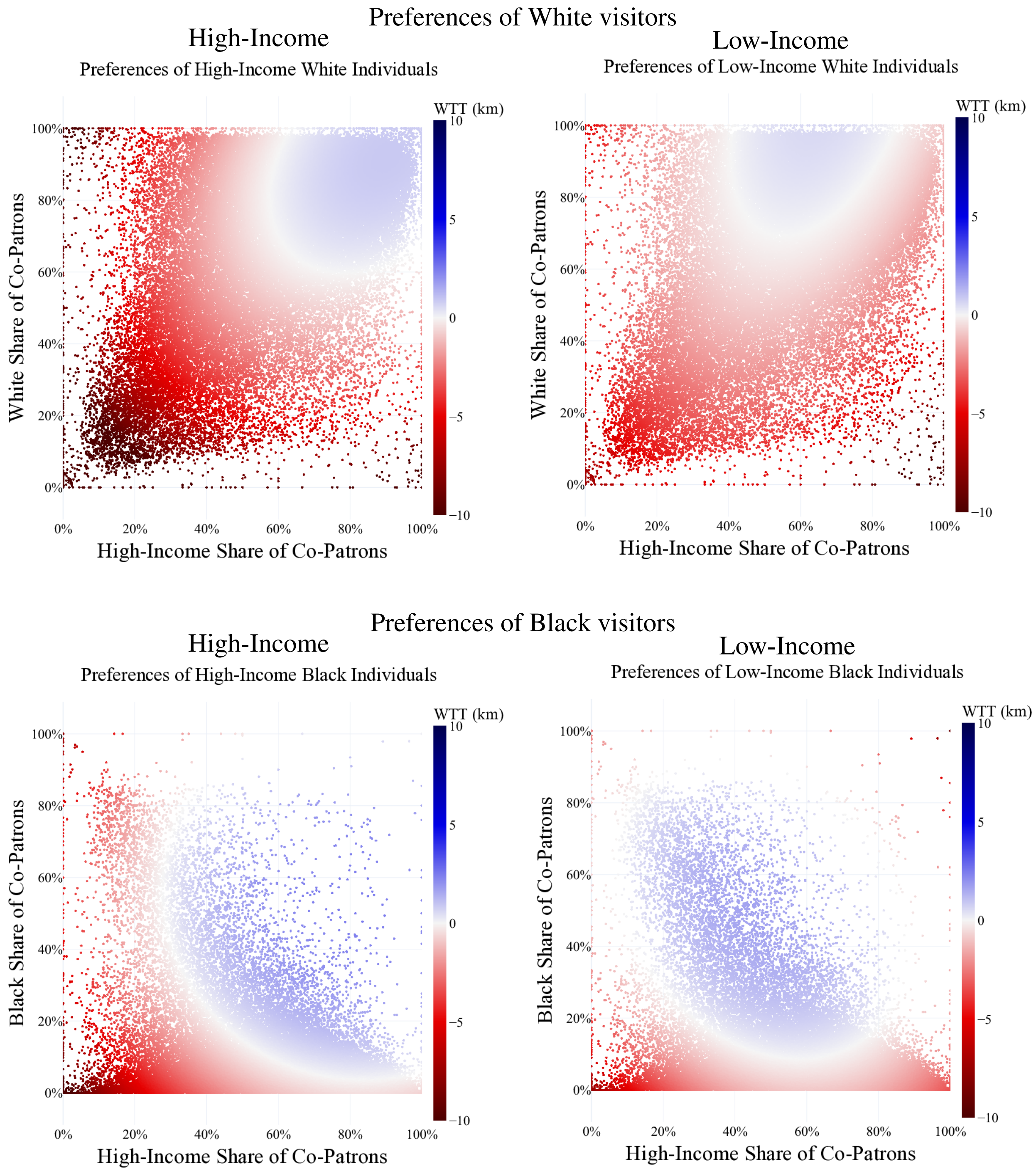
We find that preferences for co-patron income are remarkably similar across racial groups, but high-income individuals have a stronger taste for high-income co-patrons. Figure 2.3 shows that high-income individuals have monotone preferences over high-income co-patron share, increasing from left to right in each plot, whereas their low-income counterparts have non-monotone preferences for co-patron income. Figure 2.4 Panel A quantifies the strength of income preferences for each group by depicting variation in willingness to travel for establishments with the median same-race share. Across all racial groups, high-income individuals are willing to travel 2.3 to 3.1 additional kilometers to visit a venue at the 95th percentile of high-income co-patron share, relative to a venue at the 5th percentile. Low-income individuals have less pronounced income preferences, with the most preferred share of high-income co-patron between 50 to 60 percent for all racial groups. Low-income Asian, Black and Hispanic individuals are willing to travel around 1.4 additional kilometers to visit a venue with their most preferred income mix, while low-income White individuals have an even lower willingness to travel of around 0.7 kilometers.

Turning to preferences for same-race co-patrons, we find that all demographic groups exhibit substantial racial homophily. In all eight panels of Figure 2.3, the most preferred venues have high same-race shares. Figure 2.4 Panel B shows that the strength of this racial homophily does not vary by income. Comparing levels of racial homophily across racial groups is difficult because the observed same-race shares differ greatly. White individuals are the racial majority in most venues, while non-White individuals have few venues in their choice sets with large same-race shares. That said, Black, Hispanic, and White individuals have similar same-race preferences, in the sense that all are willing to travel about 2.1 km farther to visit a venue in the 95th percentile of same-race share rather than one in the 5th percentile. The racial homophily of Asian individuals appears to be stronger, albeit on very limited support. Finally, we note that racial and income preferences are roughly similar in magnitude. This similarity is important for how people trade-off income exposure for racial exposure, as we show in Section 2.6.

Our paper emphasizes broad patterns of demographic preferences, but our estimates offer some finer insights into what individuals prefer. Here, we briefly discuss two additional features of these preferences. First, Figure 2.4 shows that preferences are concave in co-patron shares. People are most willing to travel to avoid being an overwhelming racial minority or in a venue heavily patronized by low-income people. Second, we estimate significant interactions between racial and economic composition. Table 2.2 reports positive interaction terms for White individuals, indicating that they value high-income co-patrons more when surrounded by White co-patrons. For other racial groups, the reverse holds: the high-income share matters less as same-race share increases. This suggests that all four racial groups have sharper income preferences when co-patrons are White.[24]

[24]Seven out of eight interaction terms are between 2.2 and 4.5 in absolute value. For example, for high-income Hispanic individuals when visiting an establishment in the top quartile of Hispanic co-patron share, they are willing to travel an extra 1.73 kilometers to visit a higher-income venue (comparing venues at the 95th vs 5th percentile of high-income share). This effect is even stronger at establishments in the bottom quartile of Hispanic co-patron share, where high-income Hispanic individuals are willing to travel an additional 3.85 kilometers.

Figure 2.3: Preferences over co-patron demographics



NOTES: Each plot visualizes the preference estimates over co-patron composition reported by equation (2.2) for each demographic group. Preferences are expressed in willingness to travel in kilometers relative to the average venue, $\Delta^g(s^{\text{samerace}}, s^{\text{highinc}})$, as defined in equation (2.3). Each dot corresponds to an individual venue. *Continues on next page*.

Figure 2.3: Preferences over co-patron demographics (continued)

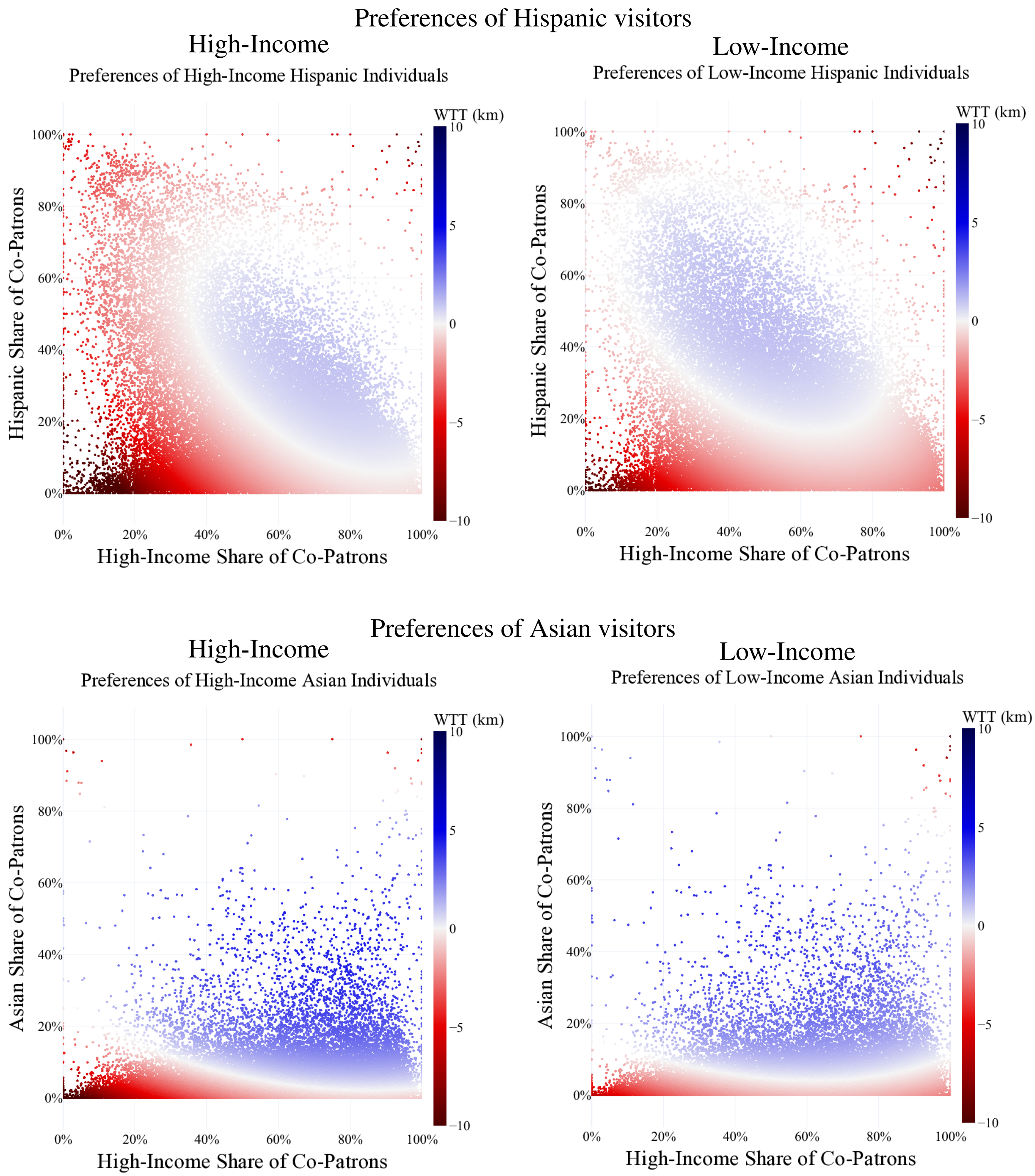
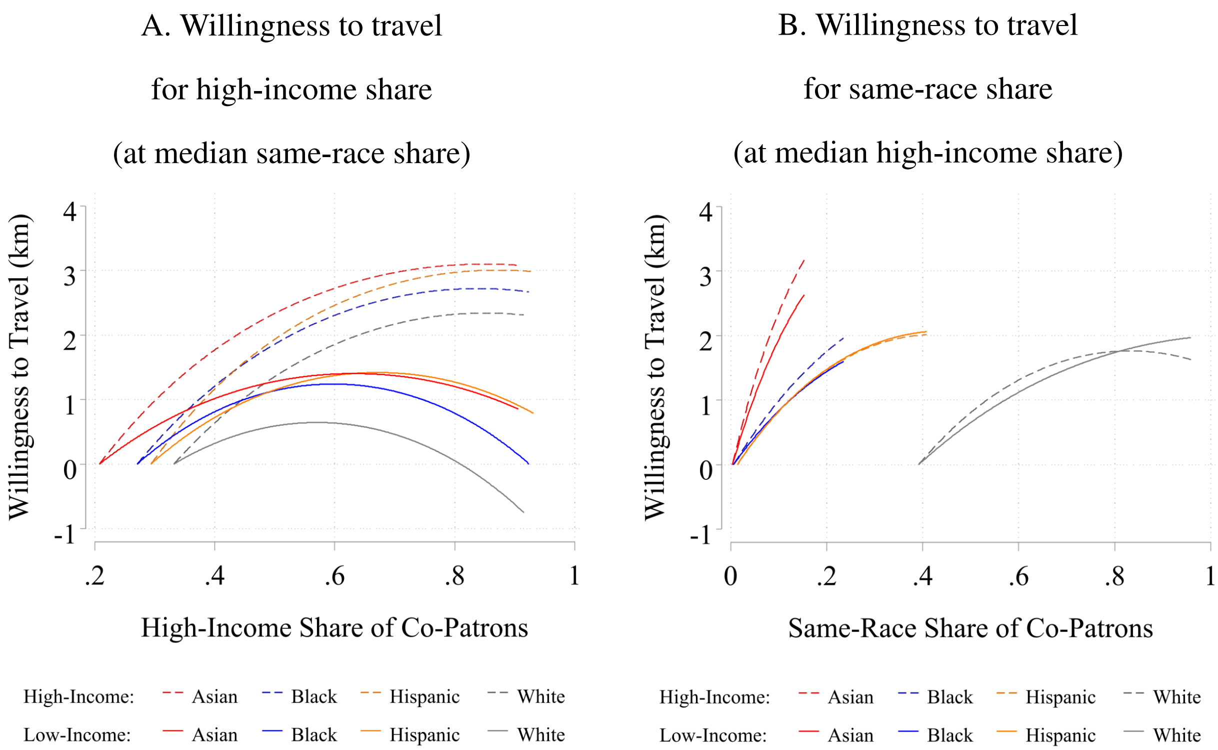


Figure 2.4: Preferences for high-income and same-race co-patrons



NOTES: This figure depicts the preference estimates reported in Table 2.2 Panel B along each dimension of co-patron composition while fixing the other. Panel A displays preferences over the high-income share of co-patrons. It shows willingness to travel relative to a venue at the 5th percentile of the high-income share distribution across all venues, holding same-race share fixed at its median value. Panel B displays preferences over the same-race share. It shows willingness to travel relative to a venue at the 5th percentile of the same-race share distribution across all venues, holding high-income share fixed at its median value.

Translated to dollars, these parameter estimates imply substantial willingness to pay for preferred demographic exposure. At typical travel speeds and values of time, an additional kilometer translates to about one dollar, so high-income individuals traveling 2.5–3.0 km farther to visit a venue in the 95th percentile of high-income co-patron share rather than one in the 5th percentile implies a \$2.5–\$3 difference in willingness to pay per trip.[25] The 2km difference in willingness to travel between the 5th and 95th percentile of same-race share for Black, Hispanic, and White individuals translates to a \$2 difference per trip. The average US driver makes more than 500 con-

[25] Assuming households make roundtrips from home at an average speed of about 40km/hour (32) and that they value time at \$19 per hour (37), willingness to travel an additional kilometer is equivalent to about one dollar.

sumption trips per year (32), so these estimates imply a \$1,000–1,500 annual willingness to pay to span the range of available demographic exposure in either income or race. For context, (16) estimates that the marginal resident is willing to pay approximately 2.1 percent of the mean house price to access schools with one standard deviation higher test scores, which amounts to \$3948, amortized over the years during which one lives in that house. Like school quality, demographic preferences may therefore be an important determinant of neighborhood choice.

2.5.2 Robustness

This section addresses two potential sources of estimation bias: within-chain heterogeneity in venue characteristics and mis-specification of the travel-cost function $f_1(\ln \text{distance}_{ij}; \delta^g)$. Restaurant chains offer standardized settings and products, but there still may be within-chain variation in venue characteristics that correlate with co-patron composition. To address this, we estimate specifications in which we (i) restrict the estimation sample to the chains with the most standardized venues and (ii) control for more venue and neighborhood characteristics. Travel costs might also differ from a quadratic polynomial in log distance in a way that is correlated with co-patron composition. To address this, we estimate specifications in which we (i) restrict the estimation sample to cities in which trips are overwhelmingly made by car and (ii) use more flexible functions of distance.

To facilitate comparisons across samples and specifications, we use a parsimonious specification in which $f_2(s_j^{\text{samerace}}, s_j^{\text{highinc}}; \beta^g)$ is a first-degree polynomial so there is one coefficient β_y^g for high-income share and one coefficient β_r^g for same-race share:

$$Y_{ij} = \delta_1^g \ln \text{distance}_{ij} + \delta_2^g (\ln \text{distance}_{ij})^2 + \beta_y^g s_j^{\text{highinc}} + \beta_r^g s_j^{\text{samerace}}. \tag{2.7}$$

Broadly, these robustness checks deliver preference estimates that are quite similar across the various samples and specifications.

Within-chain heterogeneity

We select the most standardized chains based on two characteristics. The first is the coefficient of variation in the Google Places star rating across venues within the chain.[26] Less variation in reviewer ratings across venues suggests a more standardized service. The second characteristic is ownership structure. Following (58)'s argument that franchising facilitates local adaptation, we expect chains with franchisees to be less standardized than owner-operated chains.[27] Out of 76 restaurant chains in our sample, we classify the 10 chains with fewer than 5 percent franchised venues as "entirely wholly owned" chains, the 7 chains with between 5 and 20 percent of franchised venues as "almost wholly owned", and the remaining chains as "franchised."[28]

These two measures of chain standardization are consistent with one another. The Google Places star rating variation of a franchised chain is on average twice as large as that of a wholly-owned chain. Four of the five chains with the least variation in star ratings are wholly-owned, and all ten entirely wholly-owned chains are among the third of chains with the least variation in star ratings. None of the largest and perhaps most familiar chains – like McDonalds, Subway, Starbucks, and Burger King – are among the most standardized using these metrics.[29]

Figure 2.13 reports estimation results for all eight demographic groups within four different samples of restaurant chains: the baseline sample with all chains, only entirely wholly-owned chains, only almost wholly-owned chains, and the bottom quartiles of chains (weighted by number of venues) with the lowest coefficient of variation in star rating. The preference estimates are qualitatively similar across all these chain samples, albeit noisy for some groups due to small

[26]The Google Places data on restaurant venue location and characteristics comes from Akbar et al. (1). We were able to match 41 percent of Precisely PlaceIQ venues in our estimation sample to a venue in that Google Places data. Appendix 2.8.6 provides more details on the Google Places data and variable construction.

[27](39) makes the related argument that franchisees may shirk on quality by free-riding on brand reputation. In a meta-analysis of 44 studies on franchising, (28) find support for the hypothesis that agency theory explains franchising. For instance, more geographically dispersed chains have higher franchising rates.

[28]We collect the franchise data from multiple sources: annual reports to investors for the 2020 fiscal year, company websites, franchise disclosure documents, and a franchise database compiled by Entrepreneur magazine.

[29]McDonalds, Subway, Starbucks, and Burger King are all in the top third of chains for star rating variation, and all are franchised (except Starbucks, which is hybrid with about 45 percent of franchised venues as of September 2019 based on the company's 10-K filing). The five chains with the lowest coefficients of variation are Culvers, Longhorn Steakhouse, Olive Garden, MOD Pizza, and Cracker Barrel. Of these, only Culvers is franchised.

visit counts. Overall, the preference patterns highlighted in Section 2.5.1 hold within the most standardized restaurant chains, which are less subject to concerns about variation across venues in product attributes and service quality.

Our second set of robustness checks addresses within-chain heterogeneity by controlling for more venue and neighborhood characteristics. Figure 2.14 depicts the results of adding three venue characteristics: the Google Places star rating, the Google Places number of reviews, and the venue square footage from Precisely PlaceIQ.[30] Adding these covariates has little impact on estimated preferences for same-race co-patrons, but it raises our estimates of preference for high-income co-patrons for all groups. Finally, we control for the demographic composition of the residents of or visitors to the neighborhood in which the venue is located.[31] The coefficients on the co-patron composition of the venue itself are similar in sign and magnitude when we add controls for the shares of same-race and high-income residents in the venue's census tract or the shares of same-race and high-income co-patrons within all other commercial venues located in a venue's census tract (Figure 2.15).[32]

Travel-cost specification

The estimates reported in Section 2.5.1 could be biased by some component of travel costs that is not captured by a quadratic polynomial in log distance and correlated with co-patron composition. For example, venues with more high-income co-patrons may be in locations better accessed by walking than driving. Or individuals may have a particular taste for very short trips, which would tend to be to demographically similar venues, given residential sorting by income and race. We

[30] We include variation in square footage, but these data are often mis-measured (it sometimes includes parking lots for instance) and we are not confident that it captures true variation across chains.

[31] The correlations between a venue's co-patron shares and the demographic shares of the block group in which it is located are quite high: 0.58 (high income), 0.63 (Asian), 0.75 (Black), 0.81 (Hispanic), and 0.77 (White). If we were to use only neighborhood-level residential demographics, as (34) do, we would estimate similar but smaller coefficients than those on the venue-level co-patron demographics. When we include both demographic covariates, the coefficients on the venue-level shares are substantially larger than those on the block-group-level shares, suggesting that individuals care primarily about venue composition, not neighborhood composition.

[32] One exception is the income preferences of low-income individuals, which weaken after adding controls for visitors to other venues in the same Census tract. The income preferences of low-income individuals are quadratic and hardest to capture with a single coefficient, so these coefficient estimates are less stable.

address these concerns by restricting attention to car-dominated cities and using more flexible functions of distance.

To address varying transport-mode choices, Figure 2.16 reports preference estimates for the eight demographic groups for subsets of the 100 largest MSAs based on car usage. The preference estimates for all 100 MSAs are similar to those obtained when restricting the estimation sample to MSAs in which at least 90% or 95% of trips to commercial venues are by car.[33]

Figure 2.17 reports preference estimates using three alternative functions of distance. The first uses a linear function of log distance, and the second uses a cubic polynomial of log distance, which is more flexible than our baseline quadratic polynomial. The third introduces a dummy variable indicating the venue closest to the individual's residence, which would capture a preference for very short trips or a particular salience of the nearest establishment. These specifications both yield coefficients on co-patron shares very similar to our baseline specification.

A final piece of evidence suggesting that the travel-cost function is well-specified comes from event studies of moves between demographically-distinct neighborhoods reported in Section 2.6 below. If preference estimates were biased by neighborhood demographics co-varying with distances to venues, estimated preferences would shift when individuals move between neighborhoods with different demographics. Figure 2.6 does not show any discontinuous shift in preferences around such moves.

Other categories of commercial venues

Our baseline analysis reported results for the largest venue category, restaurants. Restaurants and coffee shops have been singled out as a plausible setting for demographic exposure by other studies (6; 7; 43), and restaurant chains generally strive to provide a consistent experience across venues within the same city. It is possible, however, to estimate demographic preferences within other kinds of commercial venues that have chains. A priori, it is unclear whether to expect weaker or

[33] These MSA-level statistics come from the 2017 National Household Travel Survey (NHTS). Almost 90% of trips to commercial venues in the United States are by car. Appendix 2.8.7 provides more detail on the NHTS data and variable construction.

stronger preference estimates in other settings. For instance, preferences for co-patrons within retail chains may be weaker if demographic exposure is less salient in that environment. Conversely, preference estimates may be biased upward if stores tailor their product offering to the characteristics of their clientele, for instance by offering higher quality products in richer neighborhoods or more shelf space for Asian food in a predominantly Asian neighborhood. Figure 2.18 reports estimated preferences for each demographic group for eight distinct venue categories: banks, big-box stores, convenience stores & gas stations, grocery stores, gyms, pharmacies, restaurants, and all business categories pooled together.[34] We find some variation in the magnitude of preferences across categories, but the results confirm that the preference patterns documented in Section 2.5.1 are not unique to restaurants. For all business categories, individual exhibit racial homophily and prefer venues with more high-income co-patrons, with this inclination being again more muted for low-income people. A notable exception is banks, where low-income individuals avoid branches with high-income co-patrons. This exception is unsurprising: given the nature of banking services, different branches of the same bank must tailor their services and advisory expertise to the income of their clientele.

2.6 Determinants of demographic exposure

This section examines how preferences over co-patron demographics contribute to realized exposure to high-income individuals. The preferences estimated in Section 2.5 show that, conditional on where they live, people are willing to travel significant distances for different co-patrons demographics. Section 2.6.1 shows that these choices impact income segregation directly. For example, high-income Asian, Hispanic, and White individuals' preferences for high-income co-patrons are sufficient to single-handedly explain most of their exposure to high-income co-patrons. Beyond their direct effects, preferences over co-patron demographics are also informative about residential

[34]Table 2.8 shows the five largest chains in each category.

sorting. Section 2.6.2 shows that, within demographic groups, individuals reside in and move to neighborhoods with demographics that align with their preferences over co-patron demographics.

2.6.1 Decomposition of exposure to high-income co-patrons

Table 2.1 documented large differences across demographic groups in their exposure to high-income co-patrons. To quantify the importance of various factors in explaining these disparities, we compute model-predicted visits under the different counterfactual scenarios described in Section 2.4.6.

Table 2.3: Mean exposure to high-income co-patrons

		Low Income				High Income			
		Asian (1)	Black (2)	Hispanic (3)	White (4)	Asian (5)	Black (6)	Hispanic (7)	White (8)
(1)	Uniform within MSA	-0.02	-0.01	-0.05	-0.01	0.01	0.01	-0.02	0.01
(2)	Uniform within MSA-chain pair	-0.02	-0.02	-0.06	-0.00	0.04	0.02	-0.01	0.04
WITHIN MSA-CHAIN NESTS									
(3)	Residential proximity	-0.06	-0.16	-0.15	-0.03	0.11	-0.01	0.04	0.11
(4)	Demographic preferences only	-0.00	-0.05	-0.07	0.01	0.09	0.04	0.03	0.10
(5)	Income preferences only	-0.00	-0.01	-0.02	-0.00	0.09	0.06	0.06	0.09
(6)	Race preferences only	-0.01	-0.04	-0.06	0.01	0.05	-0.01	-0.01	0.06
(7)	Model-predicted visits	-0.06	-0.16	-0.15	-0.03	0.14	-0.00	0.05	0.13
(8)	Estimation-sample visits	-0.06	-0.16	-0.15	-0.02	0.14	-0.00	0.05	0.14

NOTES: This table shows the average high-income share of co-patrons an individual of each group would be exposed to under different counterfactual visit scenarios. All rows are differenced from a "uniform" counterfactual scenario in which devices visit venues with uniform probability nationwide. The first row reports the counterfactual scenario in which devices visit venues uniformly within their MSA of residence. while the second row reports the scenario with uniform visits within their MSA of residence and choice of chain. The third row reports the counterfactual scenario in which devices consider their distance dis-utility while ignoring preferences for co-patron characteristics. The fourth row reports the counterfactual scenario in which devices consider their preferences for co-patron characteristics while ignoring their distance dis-utility. The fifth row reports the counterfactual scenario in which devices only consider their preferences for high-income co-patrons. The sixth row reports the counterfactual scenario in which devices only consider their preferences for same-race co-patrons. The seventh row reports the counterfactual scenario in which devices consider both their preferences for co-patron compositions and distance dis-utility. Finally, the eighth row shows the actual exposure based on the home-venue-home visits in our estimation sample.

Table 2.3 reports how the mean exposure to high-income co-patrons varies across these counterfactuals. Each cell describes the visit-weighted average high-income share of co-patrons that a given demographic group experiences in a given counterfactual scenario. All rows show changes in mean exposure relative to a "uniform benchmark": the exposure each group would receive if all venues were visited with uniform probability. Our venue-choice model fits the data well, as the model-predicted visits generate exposures very close to those in the estimation sample and all observed visits to chain restaurants.

The other rows of Table 2.3 introduce different counterfactual scenarios. The "uniform within MSA" and "uniform within MSA-chain pair" rows reveal the importance of (unmodeled) variation across MSA-chain nests for explaining differences in exposure. Then, we compare how the two within-nest factors included in our model of venue choice—residential proximity and demographic preferences—determine model-predicted visits.

The "uniform within MSA" row depicts the change in mean exposure to high-income co-patrons if individuals visited venues uniformly within their MSA of residence, relative to a national uniform exposure benchmark. Accounting for metropolitan differences barely shifts exposure relative to the national average benchmark. The only substantial shift is for low-income Hispanic individuals, who tend to reside in poorer cities. For them, differences across MSAs explain one-third of their experienced income exposure relative to the benchmark (-0.05 of -0.15).

The "uniform within MSA-chain pair" row reports the change in mean exposure if individuals visited venues uniformly within the chains that they visit in the MSA where they reside, again relative to the benchmark of national uniform high-income exposure. Accounting for chains has little effect: most exposure differences arise within MSA-chain pairs. The largest shifts in exposure are among high-income Asian and White individuals, whose choice of chain shifts their high-income exposure up by 3 percentage points, less than one-quarter of the overall difference between their experienced exposure and the national average. While there are statistically significant differences in the types of chains visited by different groups—for instance, high-income individuals are more likely to visit Starbucks than Dunkin' Donuts—these differences are not large enough to meaningfully impact experienced income segregation.[35] In summary, differences in the distribution of

[35] Chain choice depends on tastes for chain product offerings, preferences for average co-patron composition within each chain, and proximity to chain venues. Given the small importance of chain choice, however, we conclude that most of demographic exposure is determined by venue choice within, not across, chains. (47) suggest that the availability of venues catering to richer people is the main reason why larger cities feature more experienced segregation.

groups across cities and chains fail to explain most of the differences in income exposure across groups.

Thus, most of these differences in income exposure must be explained by proximity to venues or demographic preferences. The "residential proximity" and "demographic preferences only" rows of Table 2.3 summarize the contribution of each factor to predicted income exposure. Specifically, the "residential proximity" row reports the income exposure outcomes that would result if individuals chose which of a chain's venue to visit based only on distance parameters δ^g absent any co-patron preferences ($\beta^g = \mathbf{0}$). Conversely, the "demographic preferences only" row reports the income exposure outcomes that would result if individuals chose which of a chain's venue to visit based only on preferences for co-patron composition absent any disutility of distance ($\delta^g = \mathbf{0}$).

Residential distance dramatically reduces the high-income exposure of low-income individuals from all racial groups. The neighborhoods in which low-income individuals reside are the most important factor in explaining why they are less exposed to high-income co-patrons. In fact, the "residential proximity" and the "model-predicted visits" rows of columns 1 through 4 in Table 2.3 are nearly identical. In this sense, the income segregation of consumption venues reflects the income segregation of residences. For high-income individuals, however, residential proximity increases high-income exposure, because they live in neighborhoods near high-income venues. The exception is high-income Black individuals, who live in much poorer neighborhoods than their high-income counterparts in other racial groups. For them, residential proximity diminishes income exposure relative to a uniform visit benchmark.

Demographic preferences also play an important role in explaining racial differences in income exposure. In particular, racial homophily impacts income exposure differently across racial groups. As an example, Black and White individuals share similar preferences for high-income co-patrons and similar levels of racial homophily, but these same preferences drive Black individuals to visit venues with much smaller shares of high-income co-patrons. This is because Black and Hispanic individuals, who have lower income on average, must choose between visiting high-income venues

We find that this hypothesis does not explain differences in income exposure across demographic groups, because these demographic differences persist *within* business chains.

or heavily same-race venues.[36] Asian and White individuals do not face this trade-off, and eliminating same-race preferences, in row 5 of Table 2.3, considerably narrows the income exposure gap between racial groups.

Overall, residential proximity dominates demographic preferences for low-income individuals, whose experienced income exposure is lower than what their demographic preferences alone would suggest. For high-income Asian, Hispanic and White individuals, however, either residential proximity or demographic preferences suffice to account for much of their observed high-income exposure. These groups live in neighborhoods where venues that suit their demographic preferences—those with large shares of high-income and same-race co-patrons—are located nearby. These results are consistent with either individuals choosing neighborhoods based on their demographic preferences or with residential experiences in those neighborhoods affecting preferences. The next section investigates this further.

2.6.2 Residential sorting on demographic preferences

We now use our model of venue choice to study the link between demographic preferences and residential sorting. We perform two exercises. First, we estimate the model for individuals in the same demographic group who reside in neighborhoods with different demographic mixes. We find that individuals are sorted across neighborhoods in a way that is correlated with their demographic preferences. Second, we study the evolution of demographic preferences around residential moves. When individuals move, they move to neighborhoods that suit their pre-move demographic preferences, and, in turn, their demographic preferences partially converge to those of their new neighborhood over the next five months.

[36](14) suggest that the scarcity of high-income majority Black neighborhoods may explain why Black households live in poorer neighborhoods than White households with similar incomes.

The extent of sorting on demographic preferences

Do people living in neighborhoods with more high-income or same-race residents have stronger preferences for high-income or same-race co-patrons? We divide Census tracts into income terciles and same-race terciles based on the composition of their residents. Leveraging our building-level demographic data, we estimate the model of equation (2.2) separately for devices residing in each income tercile and race tercile.[37] The results reveal preference heterogeneity within demographic groups. Figure 2.5 depicts these preference estimates for high- and low-income White individuals (dashed and solid lines, respectively).[38]

Figure 2.5 shows that White individuals' preferences over co-patron demographics are aligned with their neighborhoods' demographics. The upper-left panel shows that White residents of higher-income neighborhoods have stronger preferences for high-income co-patrons. This is true for both high- and low-income individuals. These differences in preferences across terciles of neighborhood income are large: a low-income resident of a top tercile neighborhood (solid red) has preferences for co-patron income similar to a high-income resident of a middle tercile neighborhood (dashed gray).

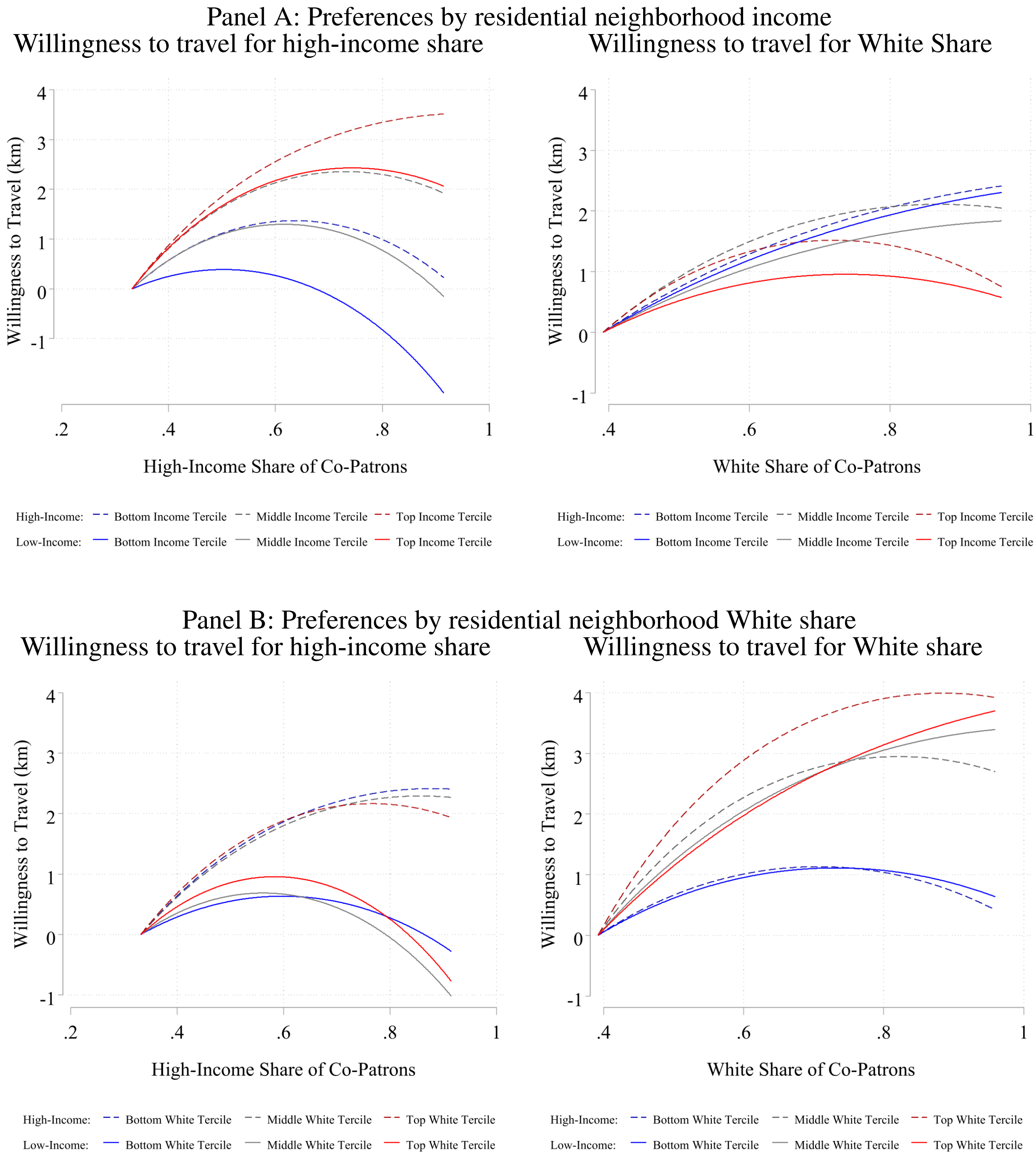
The lower-right panel of Figure 2.5 shows a similar alignment of same-race preferences with neighborhood demographics: White individuals who reside in heavily White neighborhoods exhibit stronger preferences for same-race co-patrons. Those residing in the top tercile of neighborhoods with the highest White share would travel five times farther than those in the bottom tercile to visit a venue in the 95th percentile of White co-patron share rather than a venue at the 5th percentile.

Overall, we find that within race-income groups, residential demographics and preferences over co-patron demographics are meaningfully correlated. These patterns could arise because individuals select neighborhoods with residents matching their preferred co-patron composition or

[37]This design requires building-level demographic data. We could not observe variation in preferences across neighborhoods within demographic groups if these groups were defined using neighborhood characteristics.

[38]Analogous figures for other racial groups are in Appendix 2.9. They show generally similar patterns, but are noisier due to smaller samples. For example, few low-income Black individuals live in the upper tercile of the tract income distribution.

Figure 2.5: Preference heterogeneity across White visitors by residential neighborhoods



NOTES: These plots are analogous to those in Figure 2.4 except that each line depicts preferences for high- or low-income White visitors who reside in neighborhoods (Census tracts) with different demographics. Residential tract high-income terciles are defined using tract-level high-income share weighted by high-income tract population. Similarly for same-race terciles. We estimate preferences for a randomly selected 100,000 visits by devices in a given residential tercile. To compare across neighborhood terciles, we evaluate willingness to travel relative to the same fixed demographic composition in all terciles. On the left, we show willingness to travel relative to a venue at the 5th percentile of the high-income share distribution across all venues, holding same-race share fixed at its median value. On the right, we show willingness to travel relative to a venue at the 5th percentile of the same-race share distribution across all venues, holding high-income share fixed at its median value.

because residential experiences affect those preferences (e.g., the intergroup contact hypothesis). The movers analysis in the next subsection speaks to these links.

Origins of sorting on demographic preferences

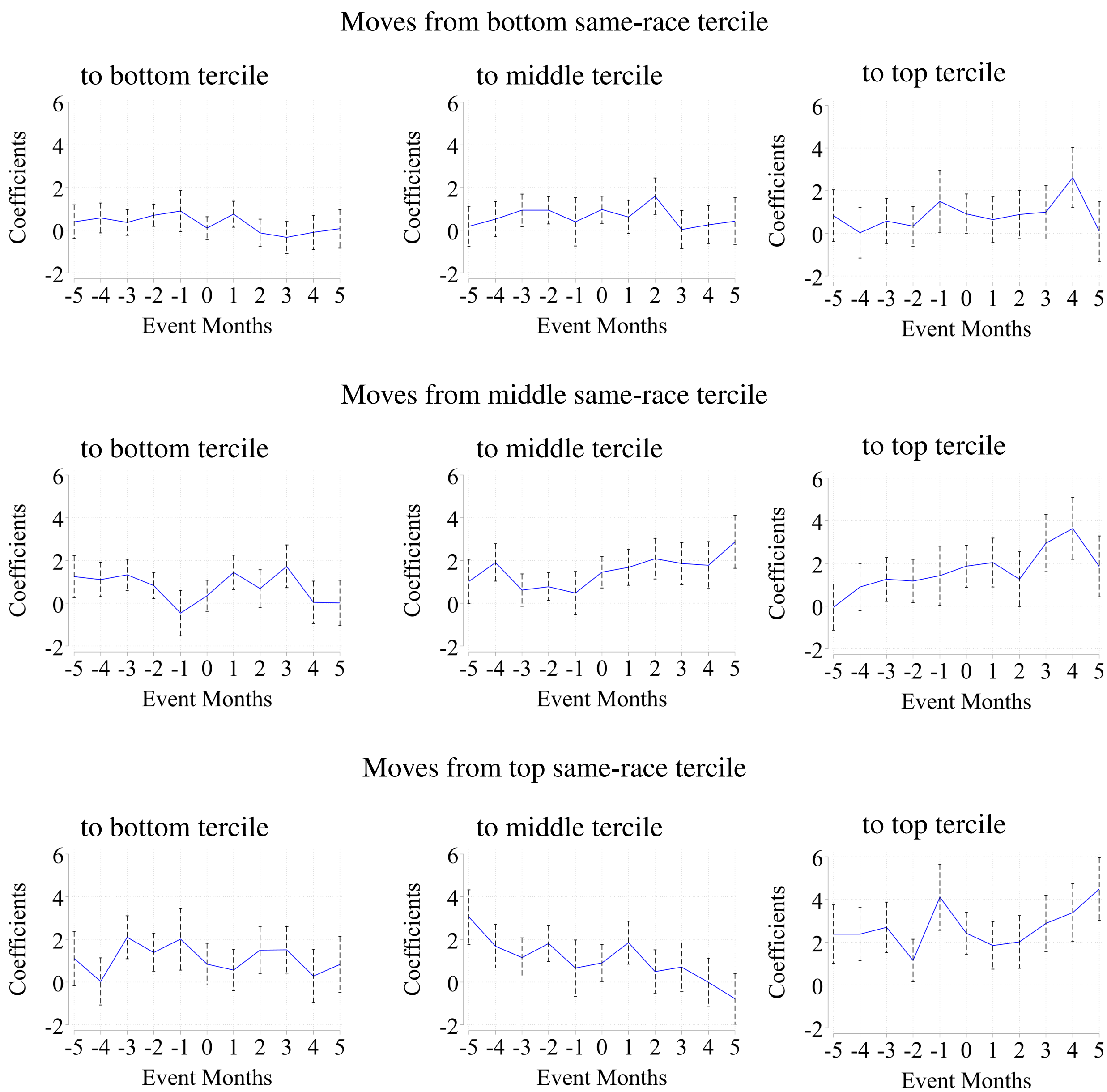
To study the relationship between residential demographics and preferences for venue co-patrons, we estimate preference parameters at a monthly frequency before and after moves. To achieve sufficient sample sizes, we study high-income White individuals who move between MSAs for five months before and five months after their change in residence.[39] We examine how preference parameters evolve when an individual moves from an origin neighborhood in same-race tercile o to a destination neighborhood in same-race tercile d. We focus on moves across same-race terciles because moves across income terciles, shown in Appendix 2.10, may arise from income shocks that could affect preferences. We estimate the parsimonious linear specification of our model—used for all robustness exercises in Section 2.5.2—in which preferences over co-patron composition depend on only the same-race share of co-patrons and the high-income share of co-patrons, but allow preference coefficients over both travel costs ($\delta_k^{g,od}$) and co-patron composition ($\beta_k^{g,od}$) to vary depending on the origin and destination same-race terciles (o and d) and the month relative to the move (k). Specifically, we re-estimate the model outlined in Section 2.4.2, replacing the utility component that varies across venues within chain (Y_{ij}) with a time-varying counterpart:

$$Y_{ijt} = \sum_{k=-5}^{5} \mathbf{1}\{t = k\} \left(\delta_{1k}^{g,od} \ln \text{distance}_{ij} + \delta_{2k}^{g,od} \ln \text{distance}_{ij}^2 + \beta_{rk}^{g,od} s_j^{\text{samerace}} + \beta_{yk}^{g,od} s_j^{\text{highinc}} \right), \tag{2.8}$$

where $\beta_{rk}^{g,od}$ and $\beta_{yk}^{g,od}$ are the preference coefficients on same-race share and high-income share, respectively, for group g in month k when moving from tercile o to tercile d.

[39] We focus on between-MSA moves so that there is a stark change in the choice set and no scope for venue-specific habits to drive behavior. We follow devices for only ten months because most devices appear in the data for less than a year. The estimation sample is an unbalanced panel: not every mover is in the sample for all ten months and not every mover makes a home-venue-home restaurant visit in every month. High-income White individuals are by far the largest sample of cross-MSA movers.

Figure 2.6: Preferences over same-race share by month by tercile-to-tercile move



NOTES: Each plot in this figure depicts event-month-specific, origin-race-tercile-destination-race-tercile-specific estimates of preference for same-race co-patrons $\beta_{rk}^{g,od}$ in equation (2.8). Each dot depicts a point estimate and the bands depict 95% confidence intervals. The estimation samples contain home-venue-home visits to restaurants by high-income White individuals who move between MSAs, split by the same-race tercile of the origin residence and the same-race tercile of the destination residence.

Table 2.4: Average same-race preference of high-income White devices before & after moves

	O1			O2			O3		
	D1	D2	D3	D1	D2	D3	D1	D2	D3
Pre-Move	0.59	0.60	0.65	0.81	0.96	0.94	1.33	1.67	2.54
	(0.17)	(0.20)	(0.27)	(0.19)	(0.20)	(0.26)	(0.27)	(0.25)	(0.29)
Post-Move	0.06	0.65	1.03	0.71	1.96	2.27	0.92	0.53	2.84
	(0.15)	(0.18)	(0.25)	(0.19)	(0.21)	(0.27)	(0.23)	(0.22)	(0.26)

NOTES: This table reports pooled estimates of $\beta_{rk}^{g,od}$ from equation (2.8). For each origin-destination pair, coefficients are pooled for event-months prior to move ($t = -5$ to $t = -1$), and post-move ($t = 0$ to $t = 5$). The estimation samples contain home-venue-home visits to restaurants by high-income White individuals who move between MSAs, split by the same-race tercile of the origin residence (with O1 being the origin tercile with the lowest share White) and the same-race tercile of the destination residence (with D1 being the destination tercile with the lowest share White).

Figure 2.6 shows estimated preferences for same-race exposure before and after moves across terciles of the neighborhood same-race distribution.[40] Given the small sample of movers, the month-specific parameter estimates are noisy. Table 2.4 reports the pre-move and post-move five-month averages of these coefficients for each event study.

Figure 2.6 and Table 2.4 reveal three patterns of interest. First, we find some evidence that individuals move to neighborhoods with demographics that suit their pre-move demographic preferences. Conditional on the demographic tercile of the origin, individuals with stronger pre-move racial preferences tend to move to destination neighborhoods with higher same-race shares. These differences are not always statistically significant, but they are especially large – a near doubling in the strength of racial homophily – when contrasting individuals moving from a third-tercile origin to a third-tercile destination (O3 to D3) with individuals moving from a third-tercile origin to a

[40]Table 2.10 replicates Table 2.4 for income preferences following moves across neighborhood income terciles. It reports substantial increases in income preferences following moves to higher income terciles, providing additional evidence that preferences change in line with the demographic of one's new neighborhood. As expected, we do not find that individuals sort across income terciles due to their income preferences. Such moves may be explained by income shocks. Appendix 2.9.2 reports the full set of event studies for income exposure around moves to different same-race terciles, and for same-race and income exposure around moves to different income terciles. Out of these 36 event studies, we note one unexplained jump in preferences following moves from the highest tercile to the bottom tercile of the income distribution. These movers to substantially poorer neighborhoods appear to experience an immediate drop in their preferences for high-income exposure upon moving. This result is difficult to interpret because such movers are rare and may have experienced a negative income shock.

first-tercile destination (O3 to D1). In other words, high-income White individuals moving from a high- to a low-share White neighborhood have substantially weaker same-race preferences before the move.

Second, estimated preference coefficients do not jump discontinuously when individuals move. The nine panels of Figure 2.6 depict event studies for moves between the nine pairs of terciles of the same-race neighborhood distribution. The top-right (O1 to D3) and bottom-left (O3 to D1) plots show the greatest contrasts between origin and destination terciles. In all cases, there are neither pre-trends in preferences prior to a move nor noticeable jumps in preferences right after moving.

Third, demographic preferences partially converge to the local demographic mix after a move. The changes in estimated preference parameters, while not always statistically significant, are consistent with neighborhoods shaping preferences. Individuals who move to neighborhoods with a higher same-race share tend to exhibit stronger racial homophily after moving. For instance, the same-race preferences of movers from O1 to D3 rise by almost 60% post-move, but remain weaker than the preferences of those already living in O3. Similarly, the preferences of individuals who move to a neighborhood with a lower same-race share evolve to exhibit weaker racial homophily. Given the limited time window, we cannot assess whether preferences fully converge, but the point estimates constitute evidence for neighborhoods shaping preferences.

Overall, our examination of movers suggests that residential sorting explains some of the documented spatial heterogeneity in demographic preferences and that exposure to new neighbors may in turn shape these preferences. Sorting appears relevant because preferences before a move predict the dominant demographic of the destination neighborhood. Residential experiences appear relevant because after a move preferences for the dominant demographic of the destination neighborhood strengthen. These results, while limited by sample sizes and time spans, would be consistent with extended contact with one's neighbors early in life influencing preferences for demographic exposure, which then evolve slowly with new experiences.

2.7 Conclusion

Our study offers new insights into the demographic fragmentation of American life. We measure exposure to high-income co-patrons in shared commercial spaces by combining data from millions of smartphones with building-level demographic characteristics. We estimate preferences for the racial and income composition of these spaces by studying visits to large chains in which variation in venue attributes is limited.

We find large disparities across groups in their exposure to high-income co-patrons. Black, Hispanic, and low-income individuals experience lower high-income exposure. Demographic preferences, however, are broadly shared across groups and economically large. For instance, racial homophily translates into a willingness to travel two additional kilometers to visit a venue at the 95th percentile of the same-race distribution rather than the 5th percentile, and these preferences do not attenuate at higher incomes.

Our findings have important implications for the future of American segregation. Since demographic preferences alone can explain a substantial portion of income segregation in shared spaces, eliminating differences in product tastes and residential proximity may not suffice to close the gaps in exposure to high-income individuals across demographic groups. Moreover, if these demographic preferences influence behavior in other settings, such as neighborhoods and schools, removing structural barriers alone is unlikely to fully integrate American society.

Our analysis, however, reveals settings in which demographic preferences are less pronounced. In particular, preferences are weaker in more integrated neighborhoods and weaken following a move to a more integrated area. Consistent with the intergroup contact hypothesis, integrated spaces may therefore promote further integration over time by attenuating demographic preferences.

References

[1] Akbar, P. A., Couture, V., Duranton, G., and Storeygard, A. (2023). The fast, the slow, and the congested: Urban transportation in rich and poor countries. Processed, University of British Columbia.

[2] Aliprantis, D., Carroll, D. R., and Young, E. R. (2022). What explains neighborhood sorting by income and race? *Journal of Urban Economics*, page 103508.

[3] Allport, G. (1954). *The Nature of Prejudice*. Number v. 10 in The Nature of Prejudice. Addison-Wesley Publishing Company.

[4] Anderson, E. (2011). *The Cosmopolitan Canopy: Race and Civility in Everyday Life*. W.W. Norton & Company.

[5] Athey, S., Blei, D., Donnelly, R., Ruiz, F., and Schmidt, T. (2018). Estimating heterogeneous consumer preferences for restaurants and travel time using mobile location data. *AEA Papers and Proceedings*, 108:64–67.

[6] Athey, S., Ferguson, B., Gentzkow, M., and Schmidt, T. (2021). Estimating experienced racial segregation in us cities using large-scale gps data. *Proceedings of the National Academy of Sciences*, 118(46):e2026160118.

[7] Atkin, D., Chen, K., and Popov, A. (2022). The returns to face-to-face interactions: Knowledge spillovers in silicon valley.

[8] Backstrom, J. and Woodward, R. T. (2021). Racial distance matters too: Racial and ethnic reference dependence in recreation decision making. *Available at SSRN 3818265*.

[9] Baldenius, T., Koch, N., Klauber, H., and Klein, N. (2023). Heat increases experienced racial segregation in the United States. Papers 2306.13772, arXiv.org.

[10] Barwick, P. J., Liu, Y., Patacchini, E., and Wu, Q. (2023). Information, mobile communication, and referral effects. *American Economic Review*, 113(5):1170–1207.

[11] Bayer, P., Casey, M. D., McCartney, W. B., Orellana-Li, J., and Zhang, C. S. (2022). Distinguishing causes of neighborhood racial change: A nearest neighbor design. Working Paper 30487, National Bureau of Economic Research.

[12] Bayer, P., Fang, H., and McMillan, R. (2014). Separate when equal? racial inequality and residential segregation. *Journal of Urban Economics*, 82:32–48.

[13] Bayer, P., Ferreira, F., and McMillan, R. (2007). A unified framework for measuring preferences for schools and neighborhoods. *Journal of Political Economy*, 115(4):588–638.

[14] Bayer, P. and McMillan, R. (2005). Racial sorting and neighborhood quality. NBER working paper.

[15] Bayer, P., Ross, S. L., and Topa, G. (2008). Place of work and place of residence: Informal hiring networks and labor market outcomes. *Journal of Political Economy*, 116(6):1150–1196.

[16] Black, S. (1999). Do better schools matter? parental valuation of elementary education. *The Quarterly Journal of Economics*, 114(2):577–599.

[17] Bruch, E. E. (2014). How population structure shapes neighborhood segregation. *American Journal of Sociology*, 119(5):1221–1278.

[18] Bursztyn, L., Chaney, T., Hassan, T. A., and Rao, A. (2024). The immigrant next door. *American Economic Review*, 114(2):348–384.

[19] Caetano, G. and Maheshri, V. (2019). Gender segregation within neighborhoods. *Regional Science and Urban Economics*, 77:253–263.

[20] Caetano, G. and Maheshri, V. (2021). A unified empirical framework to study segregation.

[21] Cantoni, E. and Pons, V. (2022). Does context outweigh individual characteristics in driving voting behavior? evidence from relocations within the united states. *American Economic Review*, 112(4):1226–1272.

[22] Card, D., Mas, A., and Rothstein, J. (2008). Tipping and the dynamics of segregation. *The Quarterly Journal of Economics*, 123(1):177–218.

[23] Charles, C. Z. (2003). The dynamics of racial residential segregation. *Annual Review of Sociology*, 29(1):167–207.

[24] Chetty, R., Friedman, J. N., Saez, E., Turner, N., and Yagan, D. (2020). Income segregation and intergenerational mobility across colleges in the united states. *The Quarterly Journal of Economics*, 135(3):1567–1633.

[25] Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., et al. (2022a). Social capital i: measurement and associations with economic mobility. *Nature*, 608(7921):108–121.

[26] Chetty, R., Jackson, M. O., Kuchler, T., Stroebel, J., Hendren, N., Fluegge, R. B., Gong, S., Gonzalez, F., Grondin, A., Jacob, M., et al. (2022b). Social capital ii: determinants of economic connectedness. *Nature*, 608(7921):122–134.

[27] Choi, J., Guzman, J., and Small, M. L. (2024). Third places and neighborhood entrepreneurship: Evidence from starbucks cafés. Technical report, National Bureau of Economic Research.

[28] Combs, J. G. and Ketchen Jr, D. J. (2003). Why do firms use franchising as an entrepreneurial strategy?: A meta-analysis. *Journal of management*, 29(3):443–465.

[29] Cook, C. (2023). Heterogeneous preferences for neighborhood amenities: Evidence from gps data. *Available at SSRN 4212524*.

[30] Cook, C., Currier, L., and Glaeser, E. L. (2022). Urban mobility and the experienced isolation of students. Working Paper 29645, National Bureau of Economic Research.

[31] Couture, V., Dingel, J. I., Green, A., Handbury, J., and Williams, K. R. (2021). Jue insight: Measuring movement and social contact with smartphone data: a real-time application to covid-19. *Journal of Urban Economics*, page 103328.

[32] Couture, V., Duranton, G., and Turner, M. A. (2018). Speed. *Review of Economics and Statistics*, 100(4):725–739.

[33] Couture, V., Gaubert, C., Handbury, J., and Hurst, E. (2023). Income Growth and the Distributional Effects of Urban Spatial Sorting. *The Review of Economic Studies*, page rdad048.

[34] Davis, D. R., Dingel, J. I., Monras, J., and Morales, E. (2019). How segregated is urban consumption? *Journal of Political Economy*, 127(4):1684–1738.

[35] Davis, M. A., Gregory, J., and Hartley, D. A. (2023). Preferences over the racial composition of neighborhoods: Estimates and implications.

[36] Echenique, F. and Fryer, R. G. J. (2007). A measure of segregation based on social interactions. *The Quarterly Journal of Economics*, 122(2):441–485.

[37] Goldszmidt, A., List, J. A., Metcalfe, R. D., Muir, I., Smith, V. K., and Wang, J. (2020). The value of time in the united states: Estimates from nationwide natural field experiments. Working Paper 28208, National Bureau of Economic Research.

[38] Klopack, B. (2020). One size fits all? the value of standardized retail chains. Texas A&M University working paper.

[39] Krueger, A. B. (1991). Ownership, Agency, and Wages: An Examination of Franchising in the Fast Food Industry. *The Quarterly Journal of Economics*, 106(1):75–101.

[40] Li, N. Y. (2023). Racial sorting, restricted choices, and the origins of residential segregation in u.s. cities". mimeo.

[41] Lowe, M. (2021). Types of contact: A field experiment on collaborative and adversarial caste integration. *American Economic Review*, 111(6):1807–44.

[42] Magontier, P., Schläpfer, M., and v. Ehrlich, M. (2022). The fragility of urban social networks-mobility as a city glue.

[43] Massenkoff, M. and Wilmers, N. (2023). Rubbing shoulders: Class segregation in daily activities. https://ssrn.com/abstract=4516850.

[44] Mayer, A. and Puller, S. L. (2008). The old boy (and girl) network: Social network formation on university campuses. *Journal of public economics*, 92(1-2):329–347.

[45] Miyauchi, Y., Nakajima, K., and Redding, S. J. (2021). Consumption access and agglomeration: Evidence from smartphone data. Discussion Paper 6741-1613580728, CEPR.

[46] Moro, E., Calacci, D., Dong, X., and Pentland, A. (2021). Mobility patterns are associated with experienced income segregation in large us cities. *Nature communications*, 12(1):4633.

[47] Nilforoshan, H., Looi, W., Pierson, E., Villanueva, B., Fishman, N., Chen, Y., Sholar, J., Redbird, B., Grusky, D., and Leskovec, J. (2023). Human mobility networks reveal increased segregation in large cities. *Nature*.

[48] Reardon, S. F., Fox, L., and Townsend, J. (2015). Neighborhood income composition by household race and income, 1990–2009. *The ANNALS of the American Academy of Political and Social Science*, 660(1):78–97.

[49] Rothstein, R. (2017). *The Color of Law: A Forgotten History of How Our Government Segregated America*. Liveright.

[50] Schelling, T. C. (1971). Dynamic models of segregation. *The Journal of Mathematical Sociology*, 1(2):143–186.

[51] Schönholzer, D. (2023). Measuring preferences for local governments.

[52] Sethi, R. and Somanathan, R. (2004). Inequality and segregation. *Journal of Political Economy*, 112(6):1296–1321.

[53] Small, K. A. and Verhoef, E. T. (2007). *The economics of urban transportation*. Routledge.

[54] Train, K. (2009). *Discrete Choice Methods with Simulation*. Cambridge University Press.

[55] USDOT, F. H. A. (2017). 2017 national household travel survey.

[56] Waldfogel, J. (2008). The median voter and the median consumer: Local private goods and population composition. *Journal of Urban Economics*, 63(2):567–582.

[57] Wang, Q., Phillips, N. E., Small, M. L., and Sampson, R. J. (2018). Urban mobility and neighborhood isolation in america's 50 largest cities. *Proceedings of the National Academy of Sciences*, 115(30):7735–7740.

[58] Williamson, O. E. (1991). Comparative economic organization: The analysis of discrete structural alternatives. *Administrative Science Quarterly*, 36(2):269–296.

[59] Xu, Y., Belyi, A., Santi, P., and Ratti, C. (2019). Quantifying segregation in an integrated urban physical-social space. *Journal of The Royal Society Interface*, 16(160):20190536.

[60] Yabe, T., Bueno, B. G. B., Dong, X., Pentland, A., and Moro, E. (2023). Behavioral changes during the covid-19 pandemic decreased income diversity of urban encounters. *Nature Communications*, 14(2310).

2.8 Data Appendix

2.8.1 Smartphone visits data

As described in (31), each observed visit consists of a device, a venue, a timestamp, and an attribution score. Precisely PlaceIQ's attribution scores are larger when a device is more likely to have been within a venue, based on the number and density of pings, data source of pings, and proximity of the pings to the polygon defining the venue. We retain all visits with an attribution score greater than a threshold value recommended by Precisely PlaceIQ based on their experience correlating their data to a diverse array of truth sets, including consumer spending data and foot-traffic counts. Precisely PlaceIQ also reports a lower bound for the visit's duration based on the time between consecutive pings at the same venue.

When two venues are in close geographic proximity, a single visit may have an attribution score exceeding Precisely PlaceIQ's recommended threshold value for multiple venues. Following the methodology outlined in (31), in these cases, we retain only the visit to the venue with the highest attribution score. In other cases, the polygons of two different venues overlap.[41] When two polygons overlap, we retain polygons with an identified business category over those lacking a category. If both polygons have identified business categories or neither have identified business categories we drop those visits.

We include all visits between June 1, 2018 and December 31, 2019. On the average day, there were 167 million visits produced by 38 million devices visiting 40 million residential and non-residential venues. The average device appears in the data for 159 days over the 19-month window, but a notable number appear on only one day. After we restrict attention to devices in our estimation sample (one permanent home assignment over the 19-month window) there are 104 million visits from 18 million devices visiting 30 million venues on an average day.

2.8.2 Home assignments

We construct home assignments using a procedure introduced in (31), which we repeat here for convenience. Residential venues are a distinct category in the Precisely PlaceIQ data. This allows us to construct a weekly panel of home locations for a subset of devices using the following assignment methodology:

[41]The most common case of such an overlap happens when the basemap contains one polygon representing a business establishment and a second polygon representing both that building and the accompanying parking lot.

1. For each week, we assign a device to the residential venue where its total weekly visit duration at night (between 5pm and 9am) is longest, conditional on that device making at least three nighttime visits to that venue within the week.[42] If a device does not visit any residential location on at least three nights, then on initial assignment that device-week pair has a missing residential location.

2. After this preliminary assignment, we fill in missing weeks and adjust for noisiness in the initial panel using the following interpolation rules:

Rule 1: *Change "X · X" to "X X X"*: If the residential assignment for a week is missing and the non-missing residential assignment in the weeks before and after is the same, we replace the missing value with that residential assignment.

Rule 2: *"a X Y X b" to "a X X X b" where $a \neq Y$ and $b \neq Y$*: If a device has a residential assignment Y that does not match the assignment X in the week before or after, we replace Y with X as long as Y was not the residential assignment two weeks before or two weeks after.[43]

3. After step 2's interpolation, for any spells of at least four consecutive weeks where a device is assigned the same residential venue, we assign that venue as a device's "home" for those weeks. Spells of less than four weeks are set to missing.

4. If a device has more than one home assignment and the pairwise distance between them is less than 0.1 kilometers, then we keep the home that appears for the most weeks.

5. If a device has the same home assignment in two non-consecutive periods and no other home assignments in between, then we assign all weeks in between to that home assignment.

[42]Since we only observe minimum duration, there are instances where total duration is 0 across all residential locations. In these cases, we assign the residential venue as the venue where a device makes the most nighttime visits.

[43]For cases where a device's residential location is bouncing between two places ("Y X Y X X") we are not able to ascertain whether Y or X is more likely to be a device's residence in a given week.

Table 2.5: Homogeneous Buildings by Race and Income

Category	Group	Buildings	Percent
Race/Ethnicity	All	34,547,538	100
	Asian	1,013,998	3
	Black	2,198,715	6
	Hispanic	3,936,421	11
	White	24,203,483	70
Income	All	34,547,538	100
	Low Income	14,369,436	42
	High Income	19,839,966	57

NOTES: This table shows the number of buildings for we which we have information on race/ethnicity and income. The "All" rows show that number for all buildings, and other rows show the number of buildings that are "nearly homogeneous" ($> 67\%$) for the four racial groups and two income groups.

2.8.3 Building-level demographics

Precisely PlaceIQ provides us with demographic data at the building level for around 36 million residential buildings. This includes information on standard demographic information such as education, income, race, gender, and age. Each category is reported in discretized buckets, and a building is assigned weights across buckets reflecting the share of people who live in the building who fall into each bucket. For income, we aggregate the provided bins to low-income ($< \$75,000$, the bracket cutoff closest to the national median in 2019), and high-income ($> \$75,000$). For racial/ethnic categories, we aggregate the provided bins to non-Hispanic Asian, non-Hispanic Black, non-Hispanic White, Hispanic, and all other racial/ethnic groups not covered by our study.[44]

Table 2.5 shows the number of buildings that contain information on race/ethnicity and income. The table also shows the number of buildings that are "nearly homogeneous" ($> 67\%$) for the four race/ethnicity groups and two income groups. 99% of buildings are at least 67% low- or high-income. 91% of buildings are at least 67% Asian, Black, Hispanic, or White.

[44]Specifically, Asian includes "Central and Southwest Asian", "Far Eastern", "Southeast Asian", Black includes "African American", White includes "Eastern European", "Jewish", "Mediterranean", "Scandinavian", "Western European", Hispanic includes "Hispanic", and the remaining groups not covered by our study are "Middle Eastern", "Native American", "Other" and "Polynesian".

2.8.4 Building-level data representativeness

In the main text, we show that the smartphone sample for which we have building-level data is spatially representative. For instance, within a county, we have about the same number of devices in block groups with a high White share as in block groups with a low White share. In this appendix, we show that the building-level demographic data are highly correlated with publicly available Census demographic data when aggregated to larger spatial units. This exercise shows, for example, that we have more high-income devices in the building-level data relative to the census. Figure 2.7 compares four county-level demographic shares in the building-level data to those in the 2015-2019 American Community Survey (ACS): share of non-Hispanic Black residents, share of non-Hispanic White residents, share of Hispanic residents, and share of residents whose household income is less than $75,000. The two county-level measures are highly correlated: the R^2 exceeds 0.81 for all four demographic shares. Given Panel B of Figure 2.1 showing that the smartphone sample with building demographics is broadly spatially representative across block groups within counties, the gaps between observations and the 45-degree line in Figure 2.7 largely reflect differences within, rather than across, block groups.

Overall, we find that aggregating the building-level demographic information to counties yields more White and high-income households than found in the Census data. Figure 2.7 shows that regressing the high-income household share in the Census on the same share in the Precisely PlaceIQ building-level data yields a coefficient of 1.36 and an R^2 of 0.84. Similar regressions using share of Black, Hispanic, and White households yield coefficients of 0.84, 0.87, and 0.89, and R^2s of 0.81, 0.97, and 0.92.

These differences also vary in intensity across counties. The top-left plot of Figure 2.7 shows that the share of low-income devices in the building-level data is smaller than the share of low-income residents in the Census, except in counties with the largest low-income shares. This means that low-income households are under-represented in the building-level data, but less so in counties with more low-income households. Finally, the three other plots of Figure 2.7 show that, compared to the Census data, Hispanic households are proportionally represented while White households are over-represented and Black households are under-represented in the Precisely PlaceIQ data. We note that the disparities between demographics in our smartphone sample and Census data are more pronounced for income than for race. Given these differences and the arbitrary nature of a dichotomous definition of "high-income", we avoid reporting income exposure shares in absolute terms. This decision is based on the notion that, although we can identify venues with higher shares of high-income individuals, we may systematically overestimate the income of individuals in these venues.

2.8.5 Building-level data reliability

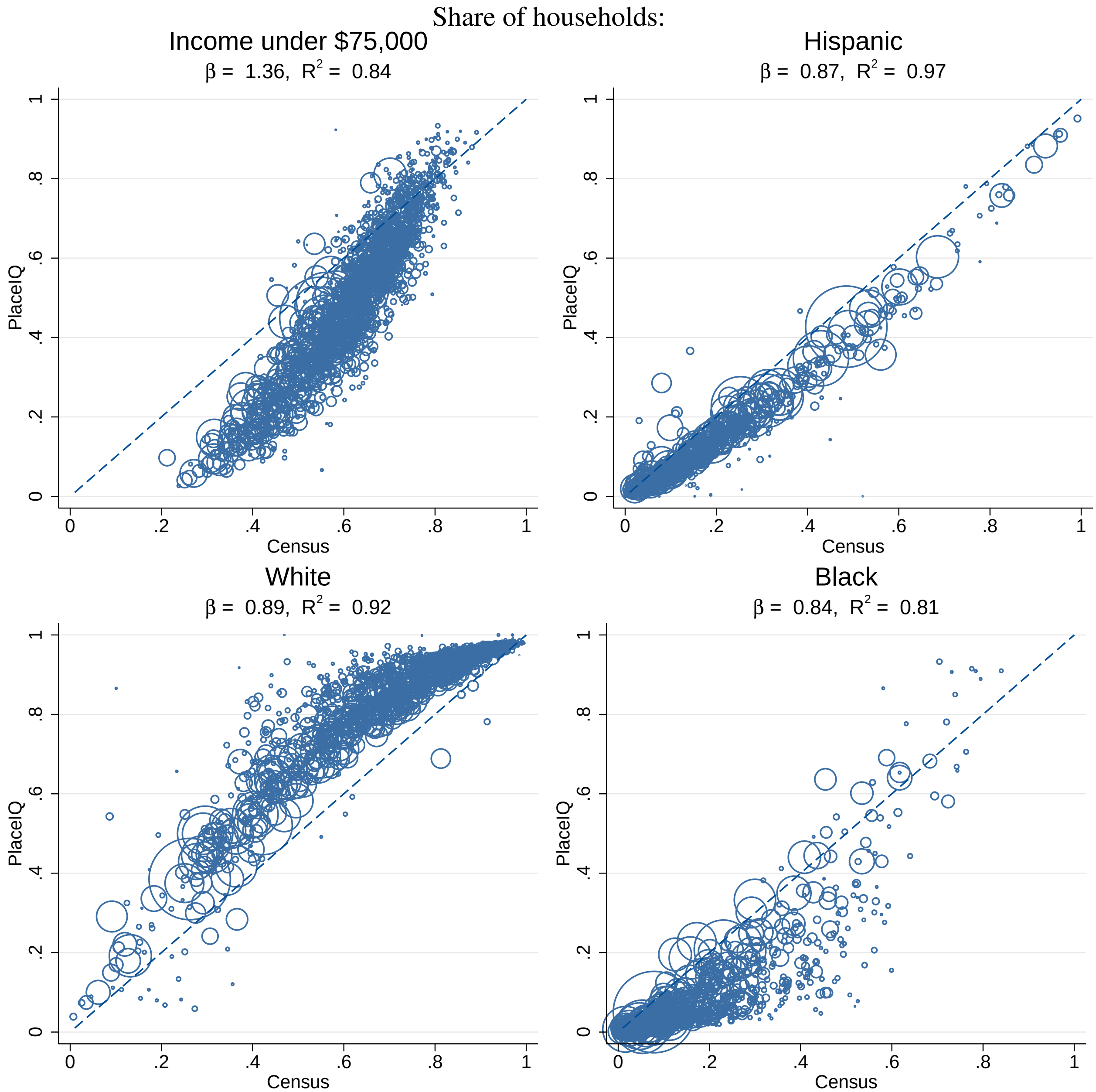
We now seek to validate the accuracy of the demographic information in the building-level data. First, we compare the Precisely PlaceIQ building-level data to building-level incomes inferred from house prices in (29) and address-level race and ethnicity data from the North Carolina State Board of Elections (NCSBE) voter registration data. Then, we demonstrate that the Precisely PlaceIQ data reliably predicts differences in chain visit propensity between residents of neighboring houses (i.e., two households living in the same block group).

Comparison with imputed building-level income from (29)

We compare our building-level data to address-level income data from (29), who kindly shared block-level averages with us. The (29) income data is imputed from home parcel characteristics from CoreLogic – including market value, size, and location – and Census tables on block group income. To perform this comparison, we create a corresponding Census block-level dataset by averaging our building-level data over all buildings in a Census block and assuming a uniform distribution of income within each income bracket. The income data in (29) correlates with our own with a slope of 0.92 and an R-squared fit of 0.43, for Census blocks below the Precisely PlaceIQ topcode of $150,000.

Table 2.6 focuses only on whether both datasets agree on above-median income classification. We regress a dummy for above-median income in our PlaceIQ Precisely data to a similar dummy

Figure 2.7: Comparison of county-level demographics in Census and building-level data



NOTES: These plots compare county-level demographic composition in the 2015-2019 American Community Survey and that of devices for which we have building-level demographic data. The diameter of each marker is proportionate to the county's population in the ACS. The regression coefficient and R^2 reports the result of regressing Precisely PlaceIQ county shares on ACS shares, weighted by the ACS population.

Table 2.6: Agreement on block-level incomes in Precisely PlaceIQ and (29) data

	High-Income in Precisely PlaceIQ	
	(1)	(2)
High-Income in Cook (2023)	0.470	0.153
	(0.001)	(0.001)
Observations	2248835	2245469
Block Group FE	No	Yes
R-squared	0.212	0.510
Within R-squared	0.212	0.011

NOTES: This table reports the results of regressing a dummy for above-median income from our PlaceIQ Precisely data on a similar dummy from (29). Each observation is a Census block.

in Cook's data. The R^2 of that regression, in column 1, is 0.21. The unmatched blocks rarely reflect wide discrepancies. Our above median-income classification agrees with that in (29) for 75% of census blocks, but allowing for a \$10,000 buffer around median income raises the match rate to 91%. So even at the narrowest Census geography, our income data rarely disagrees with Cook's data by more than \$10,000.

The second column of Table 2.6 adds a block group fixed effect. The regression coefficient remains highly significant (t-statistic of 154), but the R-squared is much smaller at 1.1%. So while we are highly confident that the two datasets are related below the block-group level, the smallest geographic unit for which Census income data are available, we cannot tell from this comparison how noisy our building-level dataset is at small geographical scales. The data in (29) comes from home parcel data that is imperfectly correlated with income. Our building-level data only contains a fraction of the buildings in any given census block, so it is also measured with error.

Comparison with address-level race in North Carolina voter registration data

We now compare the Precisely PlaceIQ building-level race data to the address-level race data from the North Carolina State Board of Elections (NCSBE) voter registration data. In particular, we ask whether the Precisely PlaceIQ data can match racial composition at the address-level in the NCSBE data better than one could using the most detailed information available in the Census.

We first compute the share of buildings in which the race of residents reported in Precisely PlaceIQ and NCSBE are matching. To simplify this comparison, we compute this match rate using only monoracial buildings. So we identify buildings in the Precisely PlaceIQ data with all residents either White or Black, and we match them to addresses in the NCSBE data whose Census coordinates are within one meter of the Precisely PlaceIQ building. We limit attention to single-family homes in the NCSBE data and buildings with 10 or fewer devices in the Precisely PlaceIQ data. We then calculate the share of monoracial Black buildings in the Precisely PlaceIQ data in which only Black voters report residing in the NCSBE data (and, similarly, the share of monoracial White buildings in which only White voters report residing). This is a demanding exercise because we only record a match if all residents of a building match in both datasets, and the voter registration data is prone to errors.

The first column of Table 2.7, 'Building-level' reports the share of monoracial buildings in the Precisely PlaceIQ data in which all NCSBE voters are of that same race. This match rate is 85% for White buildings and 61% for Black buildings. This implies, for instance, that in 39 percent of buildings, there is a least one non-Black registered voter in a building that Precisely-PlaceIQ reports as monoracial Black. These discrepancies could be due to errors in either the Precisely-PlaceIQ data or the voter registration data, errors in geocoding, variation in reporting of Hispanic status (we report Black as non-Hispanic Black) or of multiracial individuals, or the fact that both datasets were collected at the same time.

In other columns, we show that our ability to match NCSBE voters race would be worse if we used Census data instead of our building level data. For instance, if we simply allocated race to buildings in proportion to racial shares in the entire state of North Carolina, the share of Black building in which we correctly match voters racial composition would be only 0.18. If we used instead Census data at the block group level, the smallest Census geography at which the interaction of race and income we use in the paper is available, the match rate rises to 0.49, still lower than 0.61 achieved with our building level data. Unlike income, race is available at the Census block level,

Table 2.7: Match rate in racial composition of Precisely PlaceIQ and NCSBE Data

	Match rate when using data at:			
	Building level	Block level	Block Group level	State level
White	0.85	0.84	0.82	0.77
		(0.0003)	(0.0004)	(0.0009)
Black	0.61	0.58	0.49	0.18
		(0.0015)	(0.0026)	(0.0048)

NOTES: The table shows the match rate between race of voters in the NCSBE data and our building-level data, and compares that match rate with predictions from Census tables at various geographies. The first row reports this match rate for monoracial White buildings and the second row for monoracial Black buildings. The first column shows the share of monoracial buildings in the Precisely PlaceIQ data in which a all voters in the NCSBE data report being in that same race. The second, third, and fourth columns report the mean and standard deviation (in parentheses) of these shares calculated for 50 simulated versions of the Precisely PlaceIQ data, in which building demographics are randomly re-assigned within Census block, Census block group and state, respectively. White and Black are defined as non-Hispanic White and non-Hispanic Black, respectively.

but even if we allocate race to buildings using this most geographically detailed Census data—as in smartphone papers like (6)—the match rate for Black buildings only rises to 0.58.

Overall, we match address-level racial composition in the North Carolina voters data better with our building level data than one could using Census data. For monoracial Black buildings, the match rate is (0.61-0.49)/0.61 = 20% better than what can be achieved using the Census block group data that we would otherwise have to use, and 5% better than using Census block data, the narrowest Census geography at which race (but not income) is available.

Internal validation from chain visit propensity

The Precisely PlaceIQ data reliably predicts differences in chain visit propensity between residents of neighboring houses (i.e., two households living in the same block group). For example, we find that residents of high-income buildings are more likely than their neighbors in low-income buildings to visit chains preferred by high-income individuals, such as Starbucks.

To show this, we compare visit patterns between buildings within the same block group. This exercise establishes that the building-level data is informative at a finer level than the most granular demographic data publicly available from the Census Bureau, which is at the block group level.

Our methodology is based on the idea that building-level demographic differences should generate observable differences in visit patterns between people of different demographic groups living in the same block group (and therefore facing the same choice set). Consistent with this idea, research by (56) and (38) shows that race and income correlate with heterogeneity in preferences for different types of venues and chains.

These observable differences in behavior predicted using only building-level demographic data should, in turn, correspond to similar observable differences in behavior predicted using only demographic information from the Census. Therefore, we compare how across-building variation in demographics within a block group predicts chain popularity to how across-block-group variation in demographics within a tract predicts chain popularity. If demographics predict chain patronage, we should find similar ranking of chains using these two different data sources.

We proceed with this comparison as follows. First, we compute the within-block group chain popularity ranking using only building-level demographic information. For each chain in each block group, we compute the ratio of the average number of visits by devices living in high-income buildings (at least 67% of residents earning more than $\$75,000$) to the average number of visits by devices living in all other buildings. We then take a weighted average of this ratio across block groups to obtain a ranking of the chains by popularity with high-income relative to non-high-income individuals.[45]

Second, we compute the within-tract chain popularity ranking using only Census block group demographic information. To do this, we compute, for each chain in each census tract, the average ratio of visits for devices living in block groups that are at least 67% high-income to visits from

[45] Let g_i be an indicator if device i belongs to demographic group g, and $\bar{g}_b$ be the share of devices in block group b belonging to demographic group g. We weight using the number of devices living in a block group (N_b) and a variance weight ($\sum_{i \in b} (g_i - \bar{g}_b)^2$):

$$w_b = N_b \sum_{i \in b} (g_i - \bar{g}_b)^2.$$

Adding the second variance term produces a statistic that exactly matches the ranking of chains produced by the OLS estimate of γ_{cg} for each chain:

$$\log y_{ic} = \gamma_{cg} z_c g_i + \delta_{cb} z_c d_{b(i)} + \epsilon_{ibc},$$

where y_{ic} is the number of visits from device i to chain c, g_i is again an indicator if device i belongs to demographic group g, z_c is an indicator for chain c, and $d_{b(i)}$ is an indicator if a device i lives in block group b.

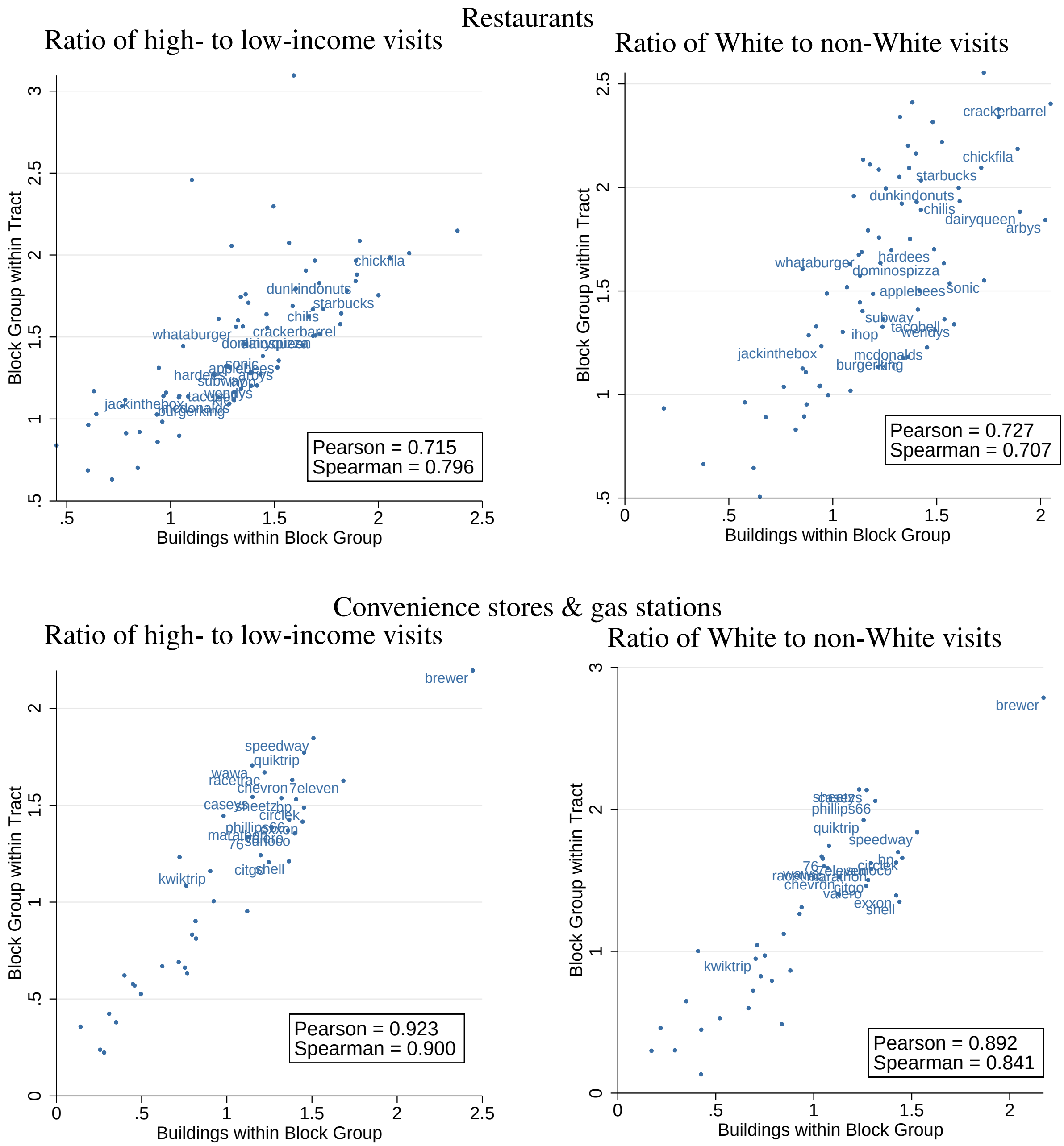
devices living in all other block groups. Finally, we compute an analogous set of ratios for visits by White versus non-White devices..

Figure 2.8 depicts the results of these comparisons for restaurants and convenience stores & gas stations, the two business categories with the greatest number of chains. Each observation in the plot is a chain in those categories. Plots on the left-hand side show the relative propensity of high-income devices to visit a chain, and plots on the right-hand side shows the relative propensity of White devices to visit a chain. The ranking obtained using only building-level demographic variation within a block group is very similar to the ranking obtained using only Census demographic information. We find Spearman rank correlations between 0.7 and 0.9 for restaurant and convenience stores, for both income and race. For example, both building-level and Census-block-group-level demographic information suggest that high-income individuals make relatively more visits to Starbucks than Dunkin' Donuts.

Table 2.8 summarizes information on the size of our venue sample for the five largest chains in each category of establishment (ten largest for restaurants). The table compares the actual number of establishments in each chain (gathered from various sources including company websites and investor reports) with the number of establishments in the Precisely PlaceIQ data. It also reports the total number of visits to each chain. The Precisely PlaceIQ basemap of venues is close to comprehensive, and contains upward 80 percent of all venues for most chains, except for gyms where the basemap is less comprehensive.[46] Bank and Gym chains receive fewer visitors than other categories, so preference estimates are noisier for these categories.

[46]Three chains have more venues in the basemap than were open circa 2019 according to company records. For Rite Aid and Walmart, this reflects store closures rather than wrongly identified locations. We are less confident about why Safeway has more venues in our basemap, but it could reflect rebranding.

Figure 2.8: Using building-level data to predict visit propensity within block groups



NOTES: The top-left panel depicts visits by high-income devices relative to visits by all other devices for restaurant chains. The vertical axis is this ratio for devices living in high-income relative to all other block groups within the same Census tract. The horizontal axis is this ratio for devices living in high-income relative to all other buildings within the same Census block group. The top-right panel depicts analogous ratios for White visits relative to non-White visits. The bottom two panels depict these comparisons using the "convenience store & gas stations" category rather than restaurants. The largest 20 chains by number of visits within each category are labeled on each plot. Data on block group income and race is from the 2015-2019 American Community Survey. High-income block groups are defined as having at least 67% of residents earning more than $75,000. White block groups are at least 67% White. Data on building-level income and race is described in this appendix. High-income buildings are defined as having at least 67% of residents earning more than $75,000. White buildings are at least 67% White.

Table 2.8: Coverage of chain venues in Precisely PlaceIQ basemap

Category	Chain	Actual (#)	PIQ (#)	(%)	PIQ Visits (M)	Source
	Bank Of America	4300	3124	73	0.68	
	Wells Fargo	5352	3035	57	0.66	
Bank	Chase	4976	3225	65	0.57	
	PNC	2296	1687	73	0.51	
	Citizens Bank	1105	823	74	0.34	
	Walmart	3947	4429	112	393.25	
	Target	1868	1360	73	66.89	
Big Box	Costco	542	505	93	61.00	
	Sams Club	599	555	93	47.54	
	Tractor Supply Co	1844	1709	93	18.60	
	Shell	12846	12635	98	280.94	
	7-Eleven	9267	7712	83	204.16	
Convenience	Circle K	6339	6061	96	163.07	
	Exxon	10830	7633	70	153.78	
	Chevron	7892	7742	98	147.02	
	Kroger	2757	2247	82	103.73	
	Safeway	895	1315	147	35.56	
Grocery	Ahold Delhaize	1973	1524	77	31.50	
	Walmart Market	809	692	86	30.67	
	Publix	1239	855	69	28.20	
	Planet Fitness	1934	809	42	0.88	
	LA Fitness	623	472	76	0.69	
Gym	Orange Theory Fit.	1150	373	32	0.54	
	Anytime Fitness	2469	837	34	0.53	
	24 Hour Fitness	445	380	85	0.39	
	CVS Pharmacy	9895	8656	87	187.11	
Pharmacies	Walgreens	9168	8380	91	135.51	
	Rite Aid	2461	2649	108	23.40	
	Subway	24154	21693	90	992.49	
	Starbucks	15041	10598	70	618.57	
	McDonalds	13846	13050	94	580.77	
	ChickFilA	2493	2030	81	224.12	
Restaurant	Dunkin' Donuts	9630	7719	80	218.47	
	Burger King	7346	6789	92	153.43	
	Wendys	5852	5475	94	139.08	
	Taco Bell	6766	6743	100	138.03	
	Pizza Hut	7280	6085	84	129.12	
	Panda Express	2161	1630	75	101.41	

NOTES: This table reports the number of venues within the five largest chains in all commercial venue categories: "Bank", "Big Box", "Convenience Stores & Gas Stations", "Grocery", "Gym", "Pharmacies", and ten largest chains for "Restaurant". The 'Actual (#)' column reports the number of U.S. chain locations as reported in various sources, which are hyperlinked in the last 'Source' column. The 'PIQ (#)' column reports the total number of venues in the Precisely PlaceIQ basemap, including those excluded from the estimation sample. The 'PIQ Visits (M)' column reports all visits to the chain between June 1, 2018 through December 31, 2019, by all devices in Precisely PlaceIQ, in millions.

2.8.6 Google Places data

We use data from Google Places to assign characteristics to restaurant chain venues for the robustness exercise in Section 2.5. The Google Places venue data was collected by (1) in 2019.[47] That Google data is available in 98 of the 100 largest MSAs. The Google data identifies the name of the venue (e.g., McDonald's), as well as its exact geo-location, number of reviews, star rating, and price in four categories ($, $$, $$$, $$$$). We match a Google venue to a Precisely PlaceIQ restaurant chain venue if its location falls within the Precisely PlaceIQ polygon for that venue (enlarged by a factor 1.3 to account for uncertainty in the exact locations of establishments) and if the name of the chain matches. 53 percent of restaurant chain venues in Precisely PlaceIQ correspond to a Google establishment. In 86 percent of these cases Precisely PlaceIQ and Google establishment agree that the same chain is at the location. This match rate reflects in part the fact that the city boundaries in (1) are smaller than those of MSAs, but also that the Google Places venue sample in (1) does not capture all venues available on Google Places.

2.8.7 National Household Travel Survey data

We use the 2017 NHTS (55) to identify MSAs with large shares of consumption trips by car for the robustness exercises in Section 2.5. We match each NHTS household to our MSA boundary using their county of residence.[48] We only keep individuals aged 18 and over in the sample. We define consumption trips as including all trips with NHTS trip purpose code 11 (Buy goods: groceries, clothes, appliances, gas), 12 (Buy services: dry cleaners, banking, service a car, pet care), and 13 (Buy meals: go out for a meal, snack, carry-out). We only keep the 92 MSAs (out of the 100 largest MSAs) with at least 100 consumption trips in the NHTS. We define 'car' trips as all trips with NHTS transport mode code 3 (Car), 4 (SUV), 5 (Van) and 6 (Pick Up Trucks). We then use NHTS trip-level weight to compute the share of consumption trips by car. Among the MSAs with

[47]That data comes from querying one of 109 keywords (like "restaurants") on Google Maps, at a large number of random locations within city boundaries defined by (1). The 109 keywords were designed to capture the universe of possible trip purpose.

[48]This geographic matching requires the confidential geo-coded version of the NHTS. We thank Gilles Duranton, who has access to it, for producing this list of MSAs for us.

more than 100 consumption trips, there are 56 MSAs in which more than 90% of consumption trips are by car and 14 MSAs in which more than 95% of consumption trips are by car.

2.9 Appendix tables and figures

2.9.1 Patterns of demographic exposure

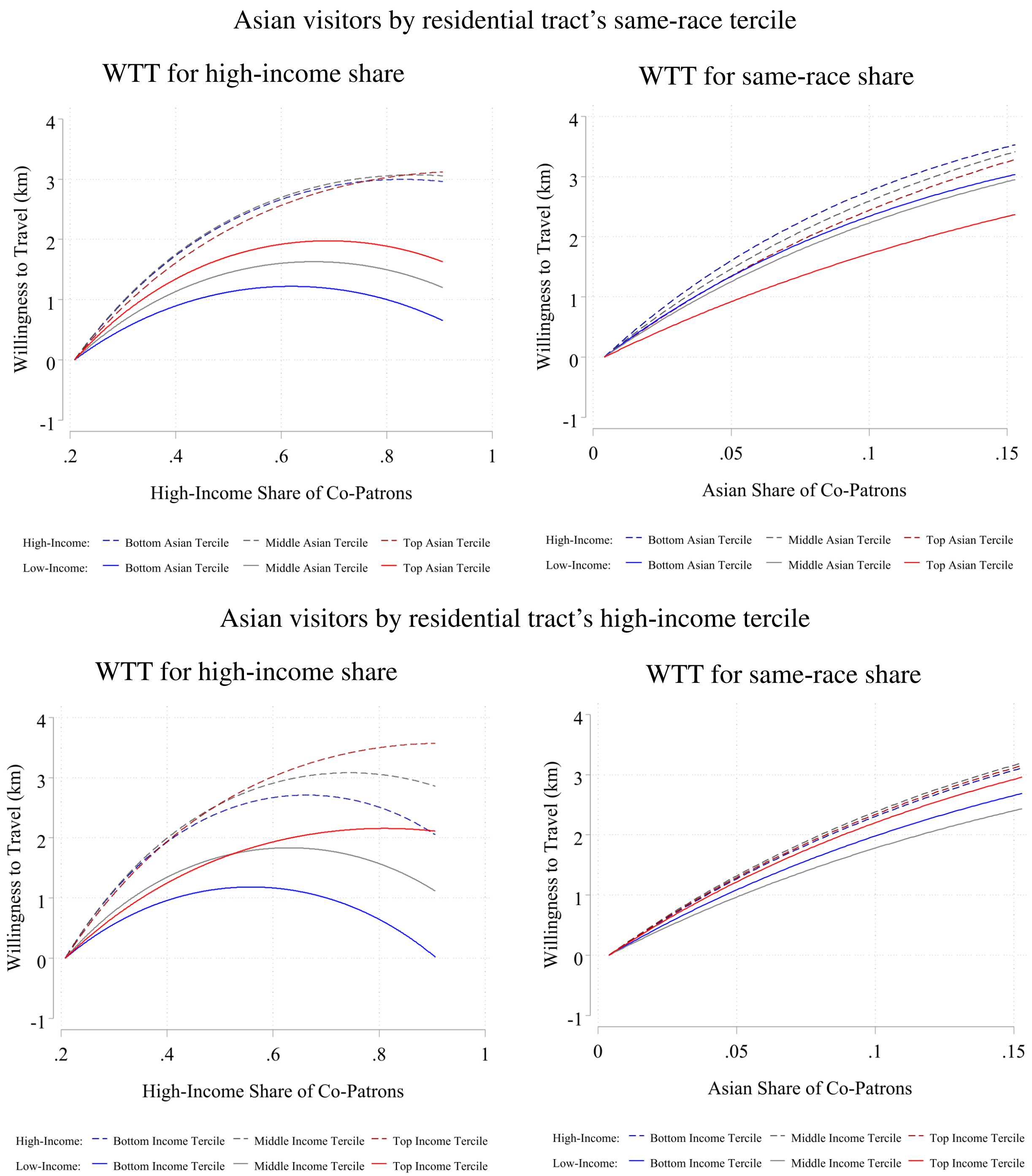
Table 2.9: Mean exposure to same-race co-patrons

	Low Income				High Income			
	Asian (1)	Black (2)	Hispanic (3)	White (4)	Asian (5)	Black (6)	Hispanic (7)	White (8)
Estimation sample	0.08	0.27	0.25	0.08	0.08	0.19	0.13	0.11
All chain-restaurant visits	0.08	0.23	0.24	0.07	0.08	0.17	0.14	0.10
All chain-venue visits	0.09	0.25	0.24	0.07	0.09	0.18	0.14	0.09
All non-residence venue visits	0.15	0.30	0.32	0.02	0.15	0.21	0.19	0.05
All McDonald's restaurant visits	0.08	0.26	0.26	0.08	0.08	0.20	0.16	0.11
Census tracts	0.16	0.32	0.30	0.13	0.18	0.24	0.23	0.13

NOTES: This table is analogous to Table 2.1 but reports same-race exposure instead of high-income exposure. The table reports, for different visit samples, the same-race share of co-patrons that each demographic groups (eight columns) is exposed to, relative to a baseline in which all venues in that sample are visited with uniform probability. The first row shows those same-race shares for visits in the estimation sample. The second through fourth rows shows those shares for broader visit samples. The fifth row shows those shares only for visits to McDonald's restaurants. In the sixth row, those shares are computed as if each Census tract is a venue and individuals only visit the census tract that they live in.

2.9.2 Determinants of demographic exposure

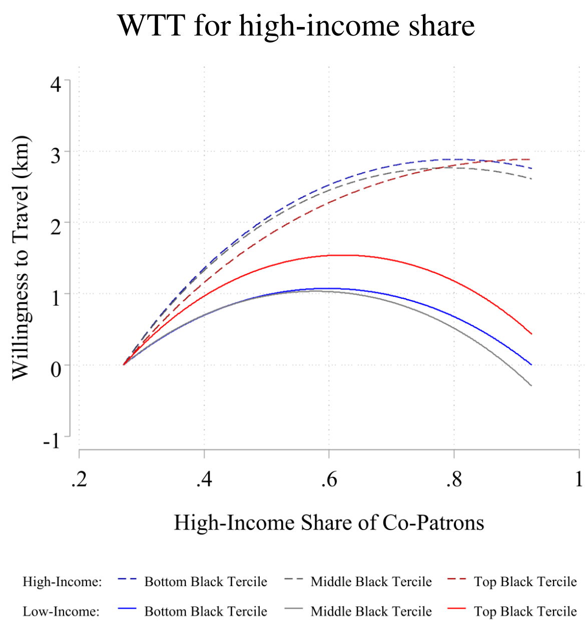
Figure 2.9: Preference heterogeneity by residential neighborhoods

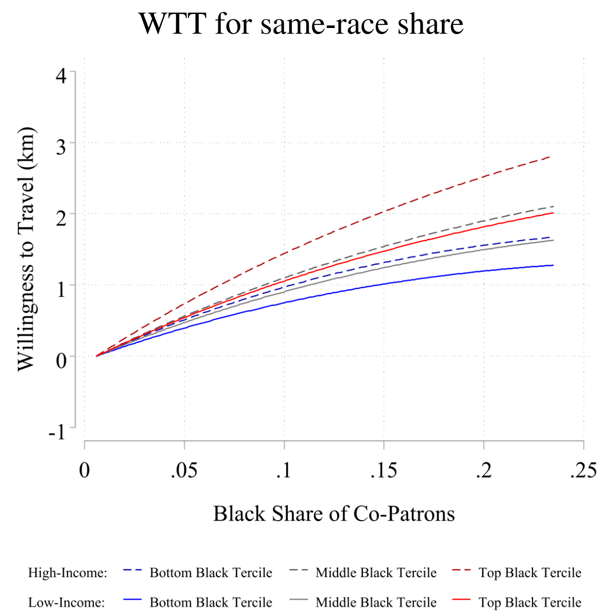


NOTES: This figure is analogous to Figure 2.5. except that it shows preference estimates for Asian, Black, and Hispanic visitors. Residential tract high-income terciles are defined using tract-level high-income share weighted by high-income tract population. Same for same-race terciles. The terciles for high-income residents are consistent across 8 demographic groups. *Continues onto next page.*

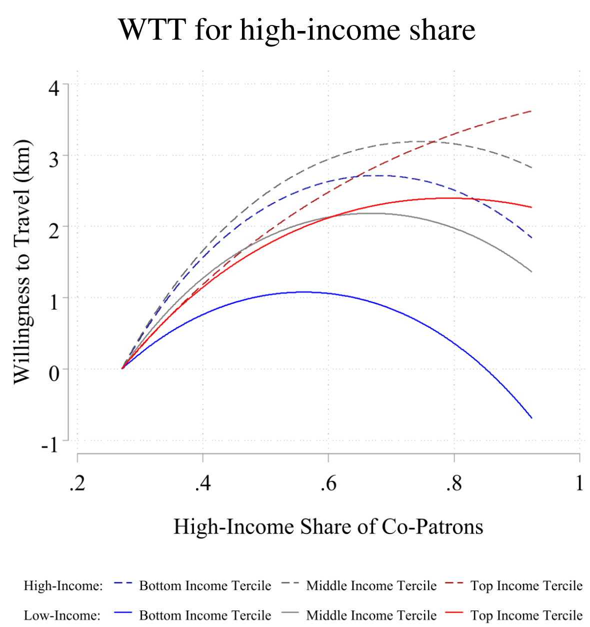
Preference heterogeneity by residential neighborhoods (continued)

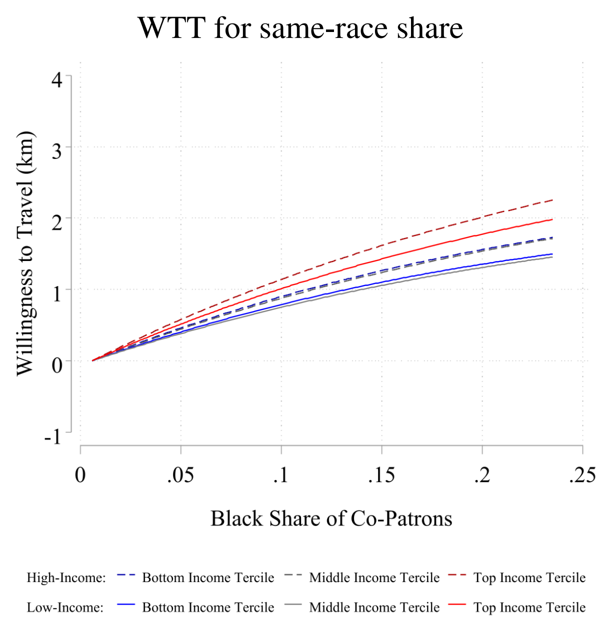
Black visitors by residential tract's same-race tercile





Black visitors by residential tract's high-income tercile

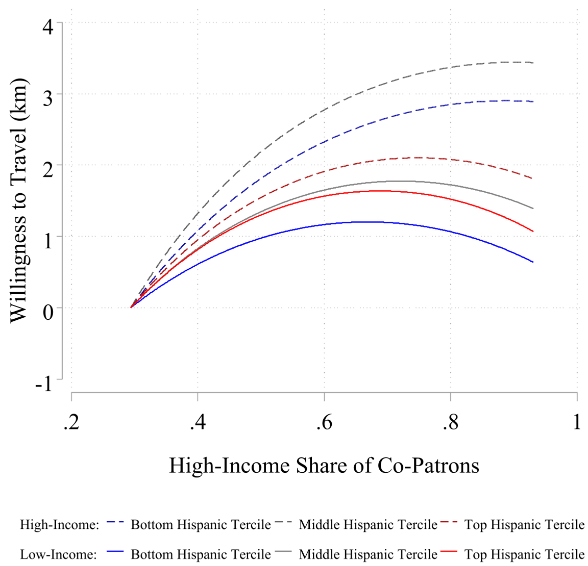




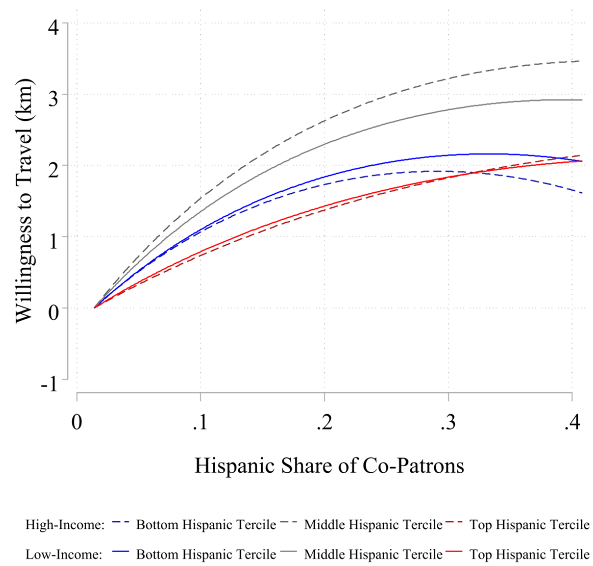
Preference heterogeneity by residential neighborhoods (continued)

Hispanic visitors by residential tract's same-race tercile

WTT for high-income share

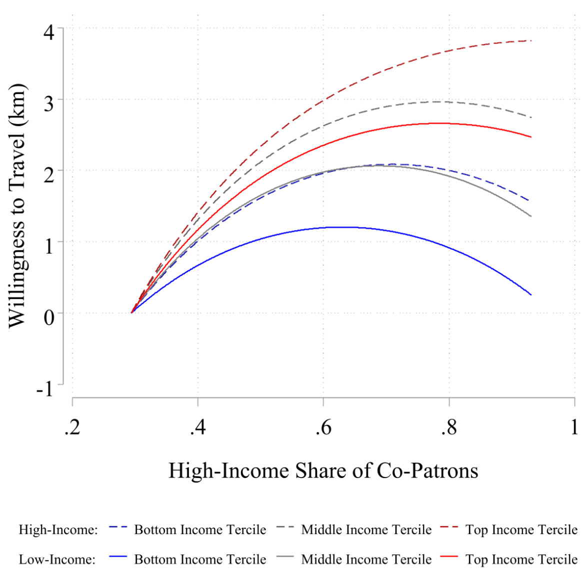


WTT for same-race share



Hispanic visitors by residential tract's high-income tercile

WTT for high-income share



WTT for same-race share

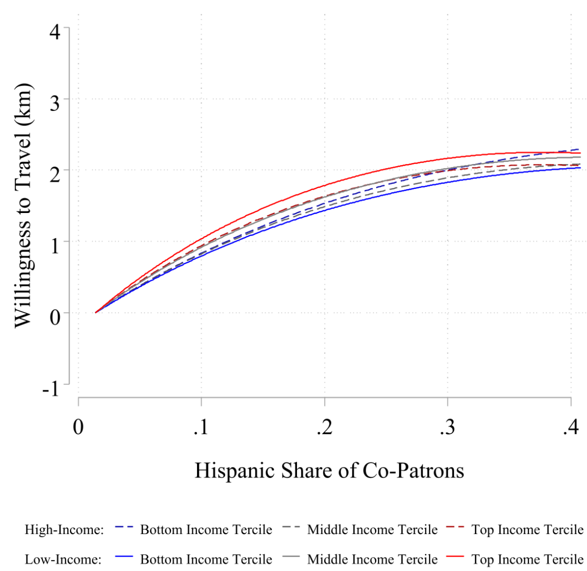
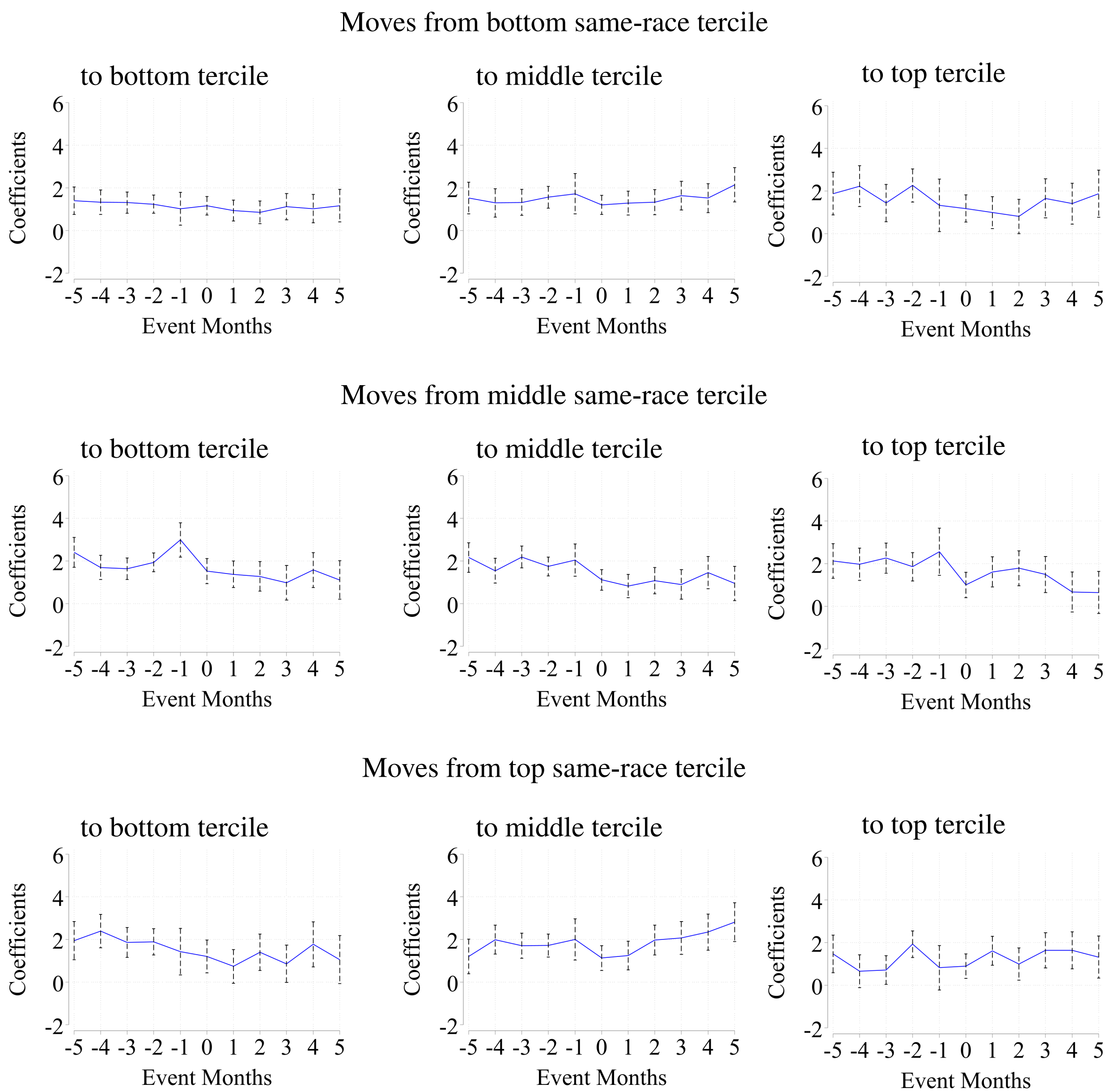
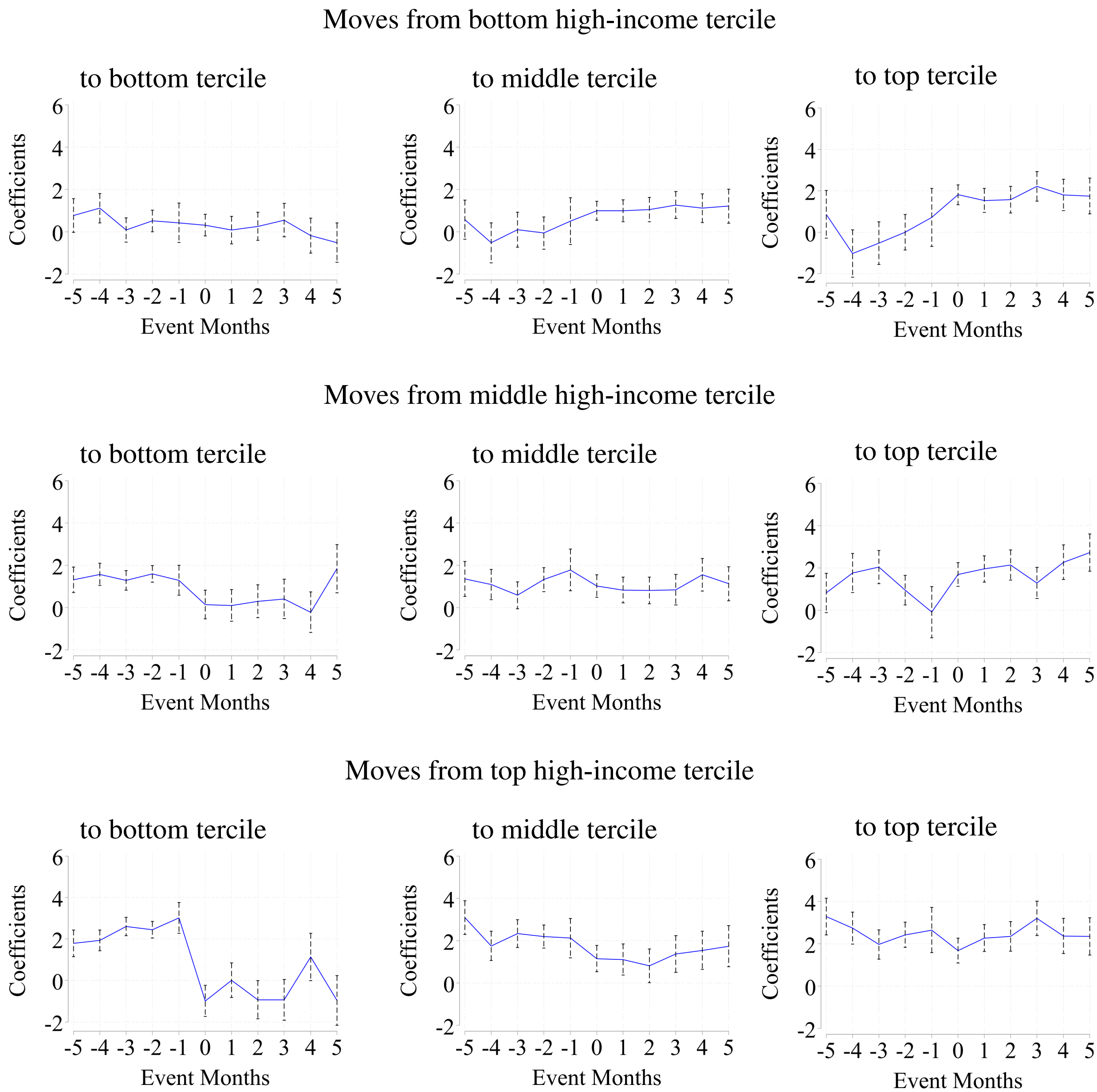


Figure 2.10: Movers result across same-race terciles: High income preference



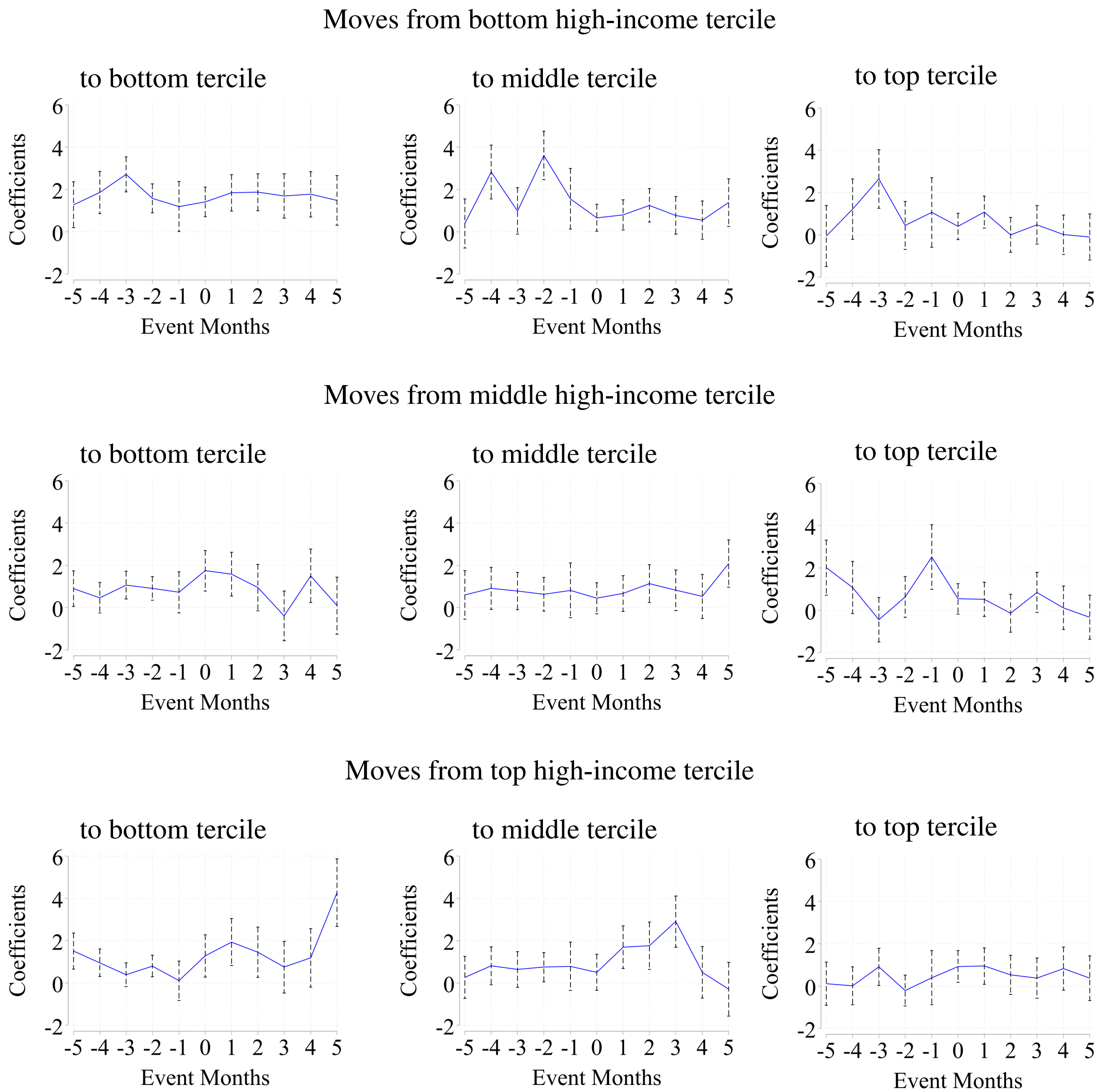
NOTES: This figure is analogous to Figure 2.6 except it shows estimates for high-income co-patrons $\beta_{yk}^{g,od}$ in equation (2.8).

Figure 2.11: Movers result across high-income terciles: High income preference



NOTES: This figure replicates the exercise in Section 2.6.2 for moves between income terciles. It is analogous to Figure 2.6 except it showing estimates over high-income co-patrons $\beta_{yk}^{g,od}$ in equation (2.8) for cross-MSA moves between income terciles. Residential tract high-income terciles are defined using tract-level high-income share weighted by high-income tract population.

Figure 2.12: Movers result across high-income terciles: Same-race preference



NOTES: This figure replicates the exercise in Section 2.6.2 for moves between income terciles. It is analogous to Figure 2.6 except it shows estimates over same-race co-patrons $\beta_{rk}^{g,od}$ in equation (2.8) for cross-MSA moves between income terciles. Residential tract high-income terciles are defined using tract-level high-income share weighted by high-income tract population.

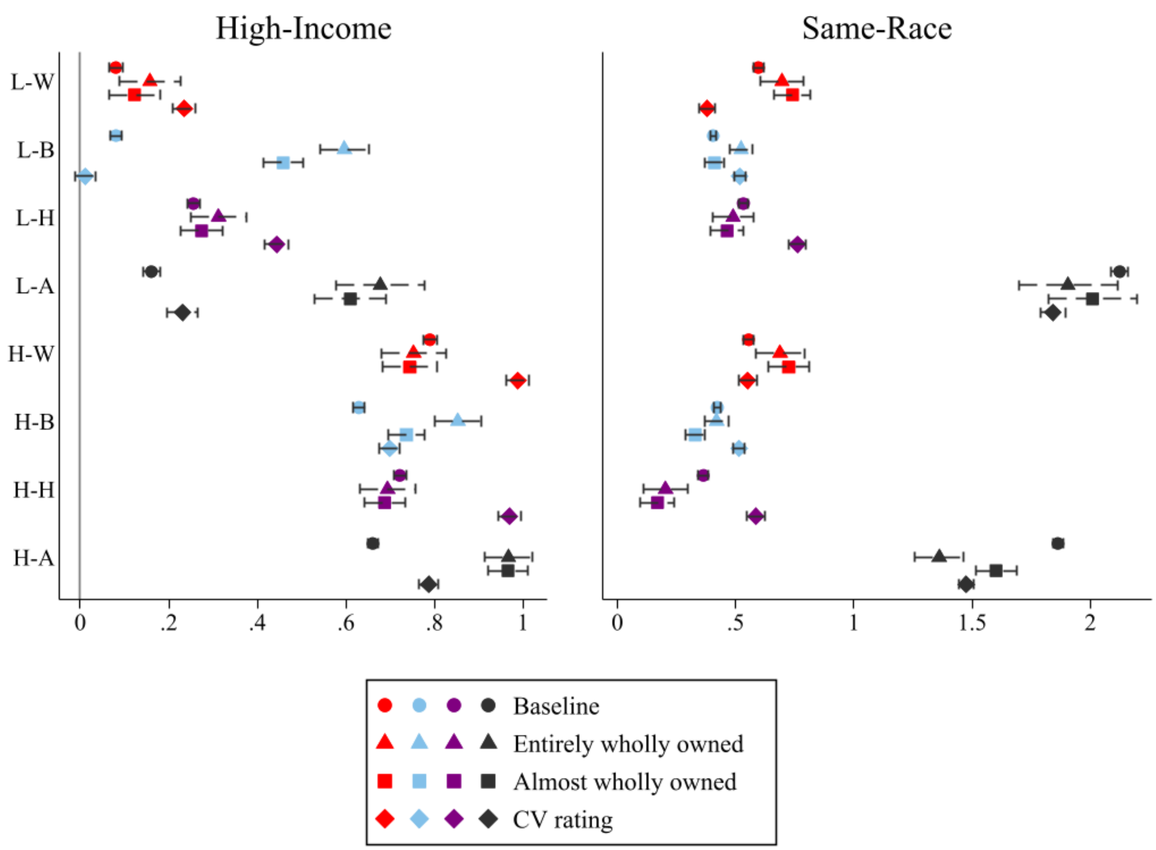
Table 2.10: Average preferences of high-income White individuals before & after move across income terciles

	O1			O2			O3		
	D1	D2	D3	D1	D2	D3	D1	D2	D3
Pre-Move	0.58	0.12	0.01	1.42	1.23	1.10	2.36	2.30	2.61
	(0.16)	(0.21)	(0.26)	(0.12)	(0.17)	(0.21)	(0.13)	(0.17)	(0.19)
Post-Move	0.09	1.11	1.78	0.43	1.03	2.01	-0.45	1.29	2.37
	(0.15)	(0.13)	(0.14)	(0.18)	(0.14)	(0.15)	(0.20)	(0.17)	(0.16)

NOTES: This table is analogous to Table 2.4 except it shows estimates for moves between income terciles. The table reports pooled estimates of $\beta_{yk}^{g,od}$ from equation (2.8). For each origin-destination pair, coefficients are pooled for event-months prior to move ($t = -5$ to $t = -1$), and post-move ($t = 0$ to $t = 5$). The estimation samples contain home-venue-home visits to restaurants by high-income White individuals who move between MSAs, split by the high-income tercile of the origin residence (with O1 being the origin tercile with the lowest high-income share) and the high-income tercile of the destination residence (with D1 being the destination tercile with the lowest high-income share).

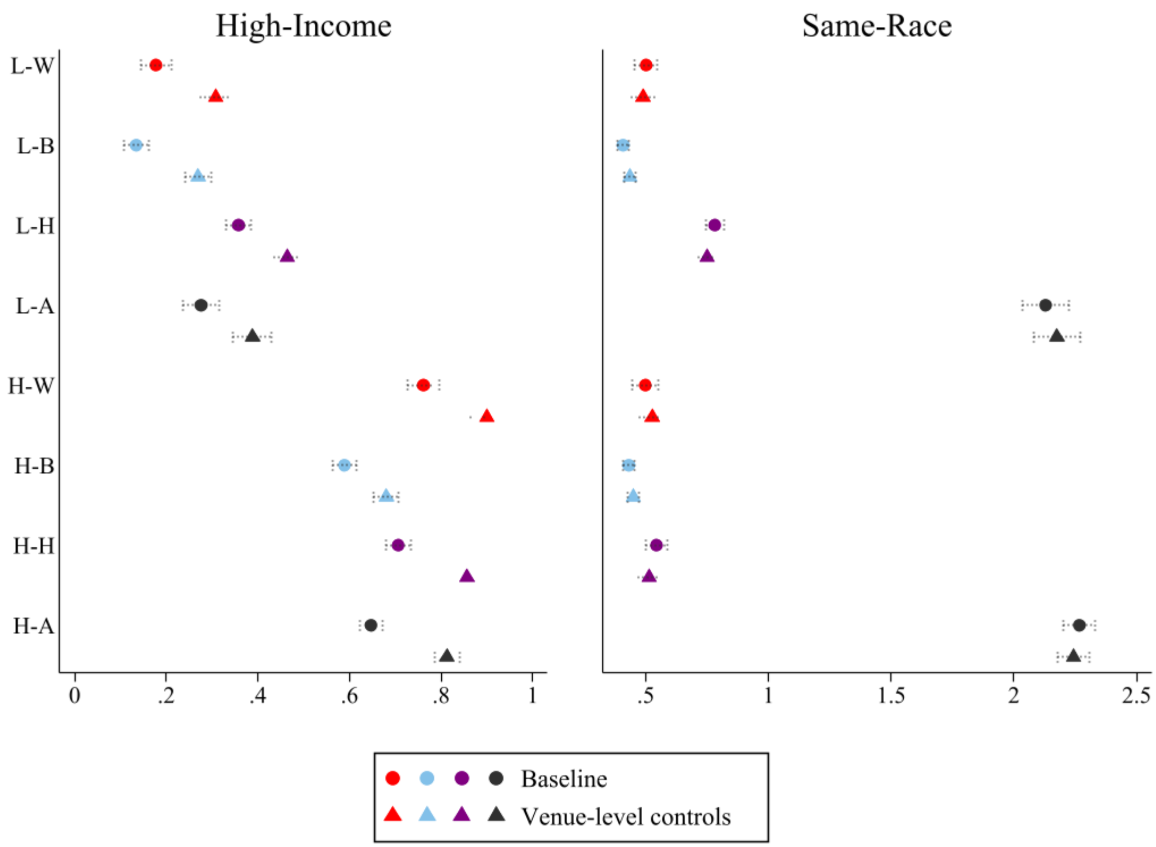
2.9.3 Robustness checks

Figure 2.13: Robustness to standardized chains categories



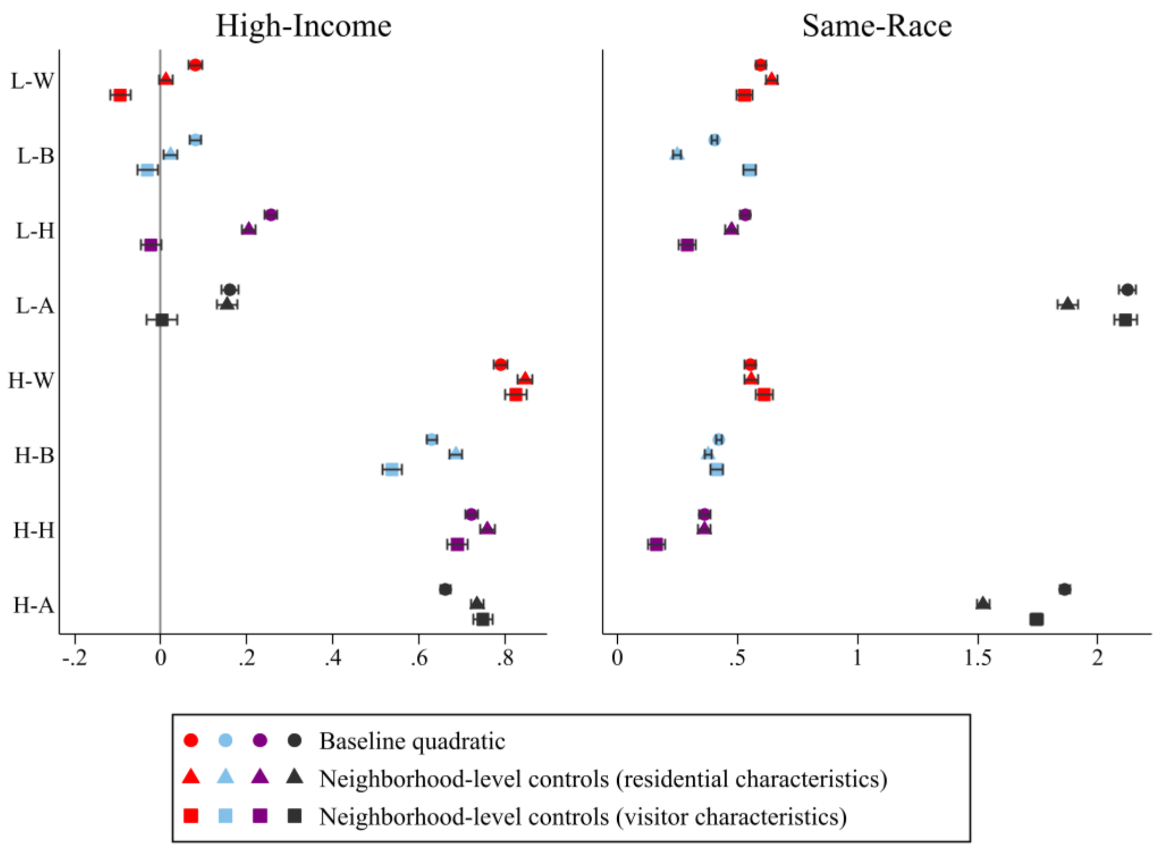
NOTES: These figures show results for restaurant chains restricted by chain standardization metrics. The preference estimates over co-patron composition use the linear shares specification from Equation (2.7) for each demographic group. Preferences are expressed in willingness to travel in kilometers relative to the average venue, $\Delta^g(s^{\text{samerace}}, s^{\text{highinc}})$, as defined in equation (2.3). The baseline sample includes preference estimates over all restaurant chains. "Entirely wholly owned" restricts to chains with 5% or fewer franchised venues; "Almost wholly owned" restricts to chains with 20% or fewer. "CV Rating" limits to the 25% of chains with the lowest variation in Google Places star rating.

Figure 2.14: Robustness to adding venue-level controls



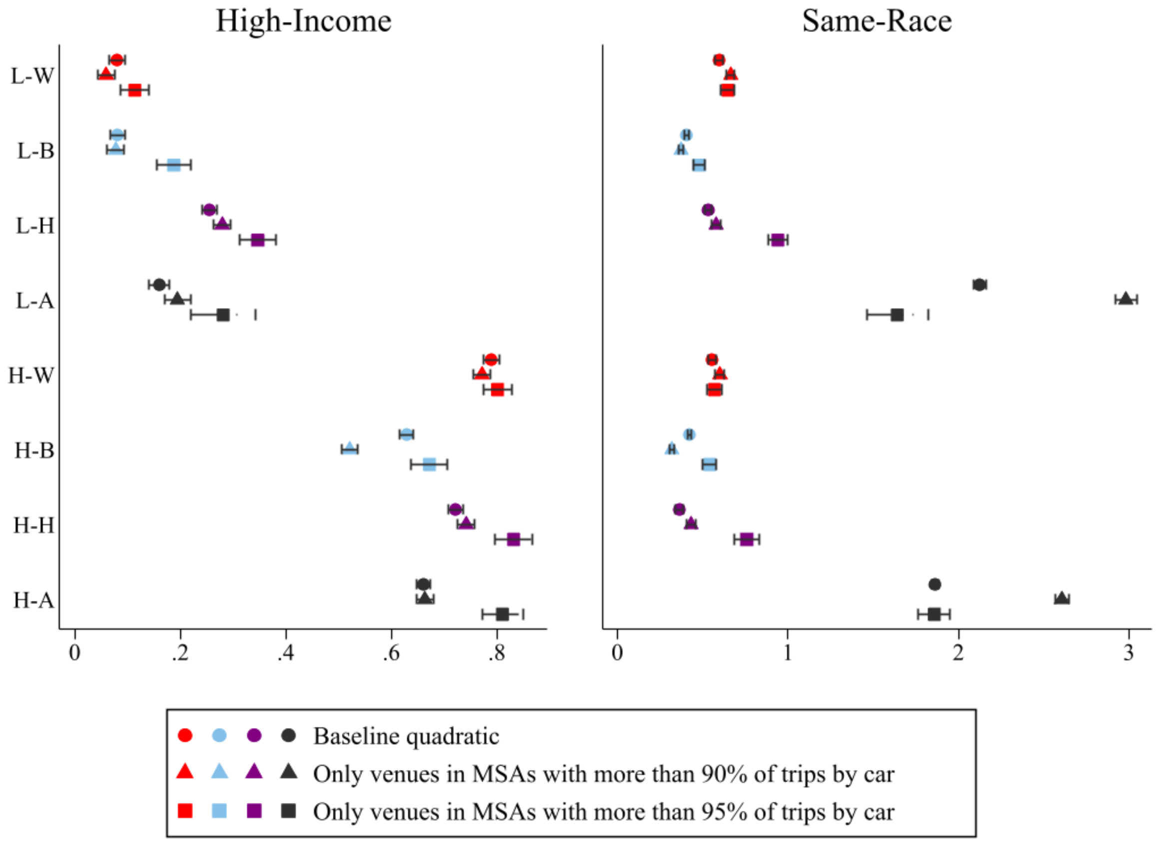
NOTES: These figures are analogous to Figure 2.13 except they add venue-level controls to the baseline estimation. Preference estimates over co-patron composition are reported from the linear shares specification from Equation (2.7). The venue-level controls are Google Places star rating, Google Places number of reviews, and venue square footage as reported by PlaceIQ Precisely.

Figure 2.15: Robustness to adding neighborhood-level controls



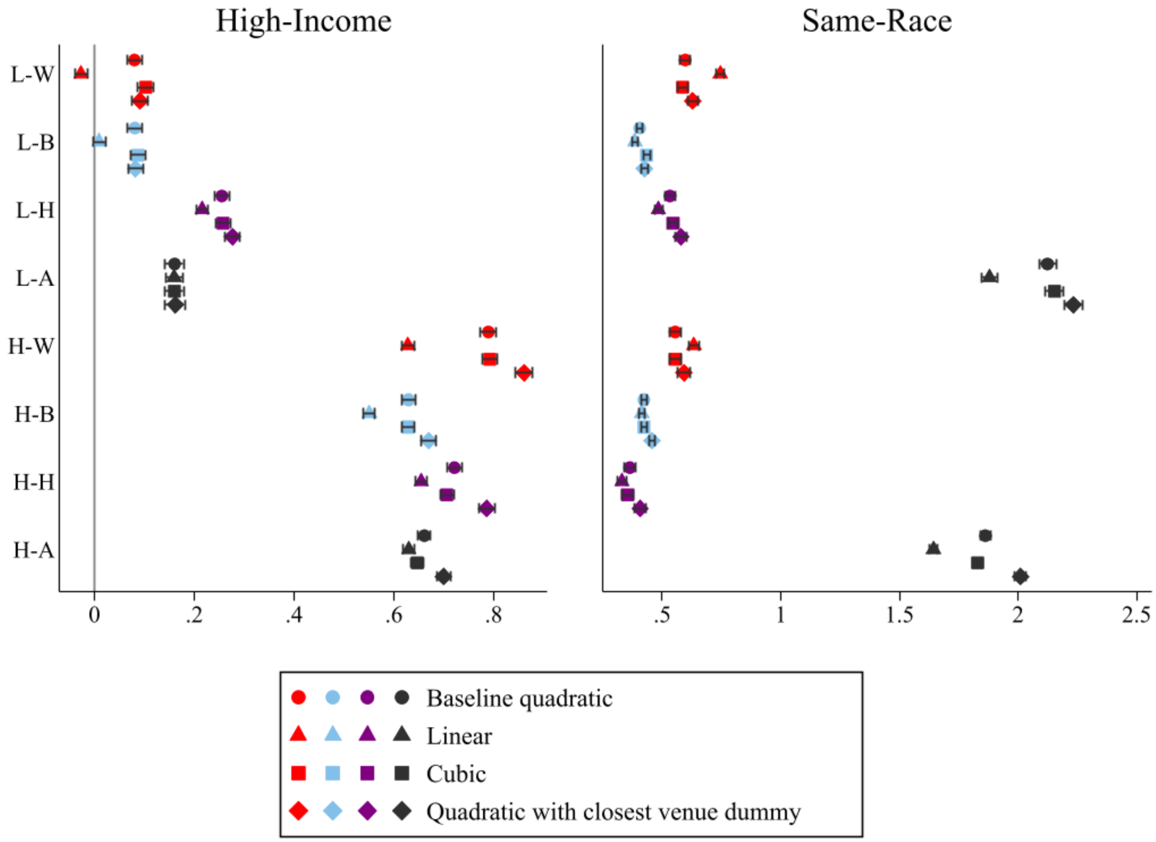
NOTES: These figures are analogous to Figure 2.13 except they add neighborhood-level controls to the baseline estimation. The preference estimates over co-patron composition use the linear shares specification from Equation (2.7) for each demographic group. Residential characteristics include the shares of same-race residents and high-income residents in the census tract where each venue is located. Visitor characteristics include the same-race share of co-patrons and high-income share of co-patrons to all other commercial venues within the census tract where each venue is located. Commercial venues are identified by PlaceIQ Precisely.

Figure 2.16: Robustness to transportation mode



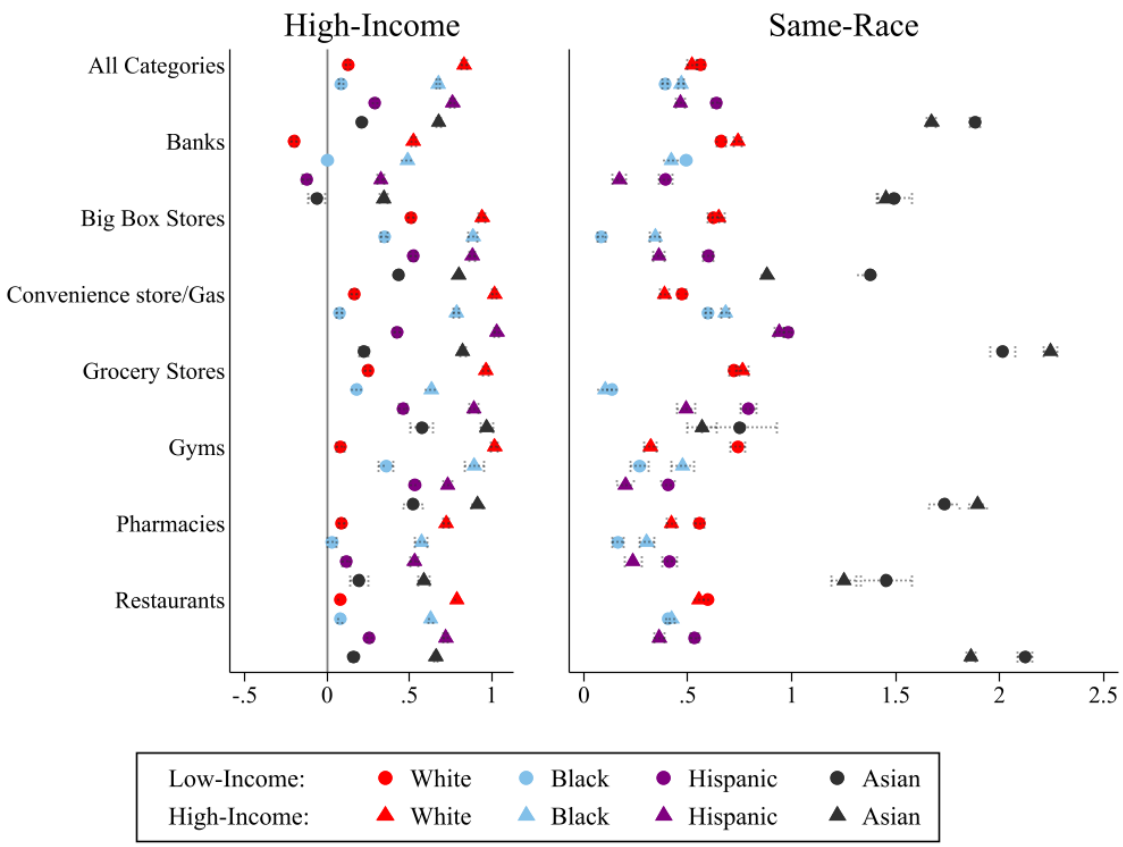
NOTES: These figures are analogous to Figure 2.13 but restrict the set of MSAs in the baseline restaurant sample by car usage. Preference estimates over co-patron composition are reported from the linear shares specification from Equation (2.7). Using data from the 2017 National Household Travel Survey (NHTS), the sample of venues is limited to MSAs where at least 90% or 95% of trips to commercial venues are by car. For the 90% threshold, 79 MSAs remain, and for the 95% threshold, 23 MSAs remain.

Figure 2.17: Robustness to distance specifications



NOTES: These figures are analogous to Figure 2.13 but report preference coefficients when varying the specification on distance in the baseline restaurant sample. Preference estimates over co-patron composition are reported from the linear shares specification from Equation (2.7). The baseline quadratic specification is the same as reported in equation (2.2). The cubic specification adds a cubic log-distance term. The closest venue specification adds a dummy for the venue within each chain that is closest to a visitor's residence.

Figure 2.18: Robustness to different categories



NOTES: These figures are analogous to Figure 2.13 but report preference coefficients for various business categories. The preference estimates over co-patron composition use the linear shares specification from Equation (2.7). Each point represents the coefficient on the high-income share of co-patrons (left) or same-race share of co-patrons (right) for a specific chain category.

Chapter 3

Annexation and the Making of the Sun Belt: Municipal Boundary Expansions and Local Public Finance[1]

3.1 Introduction

Between 1945 and 2000, the average Sun Belt city grew in land area by nearly a factor of six, while the average city elsewhere grew by a factor of two. This growth was not merely relative catch-up—the average Sun Belt city surpassed the average non-Sun Belt city in area by the 1950s, and by 2000, all ten of America's largest cities by land area were in the Sun Belt. This divergence stemmed from state annexation laws that allowed cities in Sun Belt states to expand into unincorporated areas, while municipalities in Northern states had few legal options to grow. As a result, some urban areas developed large central city-governments, while others maintained relatively small central cities fenced in by a large number of surrounding municipalities. City planners believed that the spatial growth of a central city's boundaries was consequential for local governance: proponents of annexations argued they would lower costs, increase revenue, and prevent

[1] Materials from this chapter were presented at the...

suburbanites from free-riding on central city public goods ((11)). As post-war suburbanization shocks threatened the financial stability of many cities, the ability of Sun Belt cities to expand and retain their tax bases may have allowed them to remain attractive environments in terms of municipal services and taxes ((28)).

In this paper, we examine large annexation shocks to investigate how changes in municipality size affected local public finance in the United States and whether the difference in city growth contributed to the rise of the Sun Belt. To estimate the effects of boundary expansions, we digitize annual annexation data published in the Municipal Yearbooks between 1948 and 1968 and employ a staggered difference-in-difference design around large annexation events. The data come from annual surveys on the square mileage and population of annexed areas provided directly by municipal governments to the International City Planner's Association. We supplement the annexation data with data from the Annual Survey of Governments and construct an annual panel of city boundary changes and municipal revenues, expenditures, and employment. Our empirical strategy allows us to estimate the short- to medium-run effects of these boundary expansions.

We find that the average large annexation in our sample increases a city's population by 45,000 people—a 24% increase. Our findings suggest that these large increases in population allow costs to be dispersed across the larger population: our estimates suggest a 20%–25% decrease in per capita current expenditures, a 12% decrease in total employment per capita, and a 18% decrease in revenues per capita, all of which persist nine years out from the annexation event. These reductions in per capita expenditures may reflect economies of scale, where similar or better levels of public goods are provided more efficiently, or alternatively, they could indicate reduced public service quality, particularly if newly incorporated areas receive lower levels of public good provision after annexation.

The persistent decrease in aggregate current expenditure contrasts with the patterns for policing and fire protection, where spending in annexing cities declines in the short run but returns to the level of similar non-annexing cities nine years out. These results suggest that after an annexation, cities were unable to immediately scale operations for two categories of municipal services, but within the medium-term had adjusted to providing services to newly connected areas. Since spending gaps close for public goods that are likely to be labor intensive and have closer to constant returns to scale, the overall reduction we see in spending might be generated by scale effects rather than a reduction in provision.

Finally, results for total debt issued suggest that city planners may have used these expanded tax bases to increase access to credit, although the results are noisily estimated. We plan to augment our analysis of municipal credit by digitizing data from Moody's Municipal Manuals in future work. This data includes information on all bonds issued by municipalities, their credit ratings, and local property tax rates.

The size of municipal boundaries relative to urban areas may affect government efficiency through several channels: namely, scale effects, local public good spillovers, and heterogeneity in constituents[1] Empirical estimates of these forces are rare, in the case of public good spillovers and heterogeneity, or inconclusive, in the case of scale effects. As a result, our estimates have bearings for the literature that finds mixed returns to scale in the similar but distinct reform of inter-municipal co-operation ((32), (26), (7) (9), (13) , (8), (3), (19)). Inter-municipal co-operation refers to the agglomeration of municipal services across several municipalities, such as the formation of a joint police department: while co-operation is not exactly the same as the enlargement of single municipality's boundaries, the effects of these policies should jointly depend on the scale effects present in local public good provision.

Our estimates differ from the existing estimates in several respects: first, the estimates in the existing literature come from settings outside the United States where local public finance rules may differ significantly; second, they reflect estimates for municipalities of significantly larger populations. For example, in (32), which considers intermunicipal co-operation in France, the average municipality size is 1,800. The discontinuity around which the effect of municipality size is estimated in (Narasimhan and Weaver) is 1,000 people. In our sample, the average municipality is the central city of a metropolitan area and has a population of 215,000. And, in contrast to the

[1]For examples, see the models of (31), (2), and (5).

aforementioned papers, we find evidence that increased size allowed for cities to realize economies of scale.

This paper also contributes to the literature on the historical development of cities and municipal services in the United States, particularly in the post-war period. Prior work has documented the productivity advantages of cities driven by agglomeration forces ((6), (17), (27)), and the municipal services necessary to resolve the urban externalities that arise from this agglomeration ((15), (14), (12), (10) (4)). This paper extends this work by examining a regional divergence in municipal scope that took place after city governments had already introduced municipal police, fire protection, water, sewage, and waste management, and asks how this divergence in municipal scale affected the provision of these services.

This paper contributes to the ongoing debate over the causes of the rise of Sun Belt cities in post-war America. Previous research has attributed the Sun Belt's robust economic performance and population growth in the latter half of the 20th century to various factors: a relatively business friendly environment relative to a high-wage, high-union, and high-tax North ((25)); the flow of federal investment in the form of defense contracting ((20)); amenity improvements through technological change (i.e. air conditioning); and housing supply ((16)). We add to this literature by exploring the role of municipal finance and public good provision in the development of the Sun Belt.

This paper proceeds as follows: in Section 3.2, we describe in further detail the institutions and rules that governed annexation in this period and present some stylized descriptives; Section 3.3 explains our data sources; Section 3.5 describes the empirical design; Section 3.6 provides our main estimates; and Section 3.7 concludes and poses future avenues to expand the work in this paper.

3.2 Background

> *"... [Oklahoma City] had to expand or be hemmed in by small satellite communities... With annexation, the city tax base would continue to expand, and the city could control wealthy suburbs and outlying industries"*[2]

A central challenge of post-war municipal finance was "the Metropolitan problem": the question of how to structure local governance, taxation, and public service provision in increasingly sprawling and potentially fragmented metropolitan areas ((23)). The fear of a metropolitan area composed of many small municipalities was one of duplicated costs, limited returns to scale, coordination issues and inequality induced by strategic sorting across jurisdictions. One solution to such a problem is to liberally allow for the expansion of municipal boundaries through annexation or consolidation. Annexation had been a popular policy in the late 19th and early 20th centuries: Philadelphia consolidated to its county borders in 1854, Manhattan and the Bronx joined with the three remaining boroughs of Brooklyn, Queens, and Staten Island in 1898, Los Angeles became the largest city in the United States in land area through annexation in the 1920s. However, cities, at various points, encountered resistance to annexations, halting their expansions outward: Boston was rebuffed by Brookline in 1874, Chicago by Evanston in 1894, and Los Angeles by Beverly Hills in 1923.

As the Supreme Court would affirm in its decision in *Hunter v. Pittsburgh (1907)*, a city's ability to annex places without their approval was entirely a decision of the state legislature ((29)). These state-level rules were rules regarding two factors: (1) the legal ability of existing municipalities to expand their boundaries into surrounding territory; and (2) the ease by which surrounding areas could incorporate into municipalities of their own and thereby protect themselves from annexation. While states had initially been willing to allow unilateral annexation (i.e. annexation without the approval of a majority in the *annexed* area), the string of wealthy suburbs resisting annexation by central cities that began with Brookline in 1874 slowly shifted states, particularly in the Northeast and Midwest, towards severely restricting the annexation powers of cities ((29)).

[2] (30)

In their review of state annexation laws, (11) survey the variety of annexation rules that had been set in place by the mid-20th century: in Texas, municipalities had the right to *unilaterally* annex territory into the city; in six states, cities had no legal annexation method; in most other states, municipalities were only allowed to expand into territory with approval from a majority of residents in the *annexed* territory. The necessity for approval in the annexed territory in particular drew the ire of urban and municipal planners who viewed the policy as privileging the welfare of a select population over potential gains for the entire population ((28)). As an example of these skewed incentives, we can look at Boston's failed 1874 annexation of Brookline: 6,291 of 7,775 voters in Boston approved Brookline's annexation, only 292 of 1,006 voters in Brookline approved the annexation ((1)). Annexation's proponents argued that Brookline's vote reflected the preferences of a much smaller population than the city of Boston, whose population would have stood to gain from the annexation.

When annexation resurged as a policy tool, after a lull during the Great Depression and Second World War, it emerged into the policy context that had been shaped by annexation debates earlier in the 20th century. Therefore, only city leadership in certain states found themselves capable of annexing. Nevertheless, there was a resurgence in the popularity of annexation. In 1948, 145 square miles were annexed into American cities; by 1960, roughly 1,000 square miles were being annexed annually ((24)). This "boom" was however, not evenly distributed across the country, although there were several possible pathways to expansion: Texan cities, with their legal right to unilateral annexation, expanded swiftly and liberally; cities elsewhere triumphed in the face of more restrictive laws, as was the case with Columbus, whose pro-annexation mayor leveraged the politics of access to the water to convince surrounding territory to fold into the city; and finally certain cities benefited from legal intervention from the state legislatures—as was the case with Atlanta's 1952 annexation.

This paper focuses on the specific post-war period in the history of annexation in the United States for two reasons. The first is practical: data on municipal finance and boundary changes become readily available only in the post-War period. The second is that, as argued above, this period reflects an important period for annexation as a policy in the United States: as most cities faced large suburbanization shocks, only certain cities were able to expand outward to continue to incorporate the suburban tax base. In fact, while annexation laws remained liberal in certain states until the 2010s, annexations never achieved a similar popularity: in terms of magnitude, the annexation shocks of the immediate post-war period, with cities adding dozens and sometimes hundreds of square miles in the span of a couple years, would remain significantly larger than most later annexations.

3.3 Data

3.3.1 Data Sources

Annexations We digitize data on annual boundary and population changes from the Municipal Yearbooks, which were published by the International City Planners Association annually between 1934 and 2014. Between 1948 and 1968, given the resurgence of annexations, the Municipal Yearbooks collected annual surveys of cities and their annexation activity, including questions about the square mileage and population of annexed areas. In certain years, they also asked about services needed in the annexed areas, and the city planners' expectations of short or long-run returns to the annexation. In most years, almost all cities responded to the surveys: in the cases in which a survey was not responded to, it is noted in the municipal yearbooks.

Annual Population Because we lack annual data on population, we impute annual population series using the population data included in the Municipal Yearbooks, as well as the decadal Census data. The underlying assumption to this process is that annexation induced large jumps in population in a given year and that all other population growth would have occurred gradually over time. To do this, we assign non-annexation population growth to a linear annual trend, and the observed annexed population to the annexation event year. To illustrate this process, take the hypothetical example of City X: lets say we observe City X with a population of 100,000 in 1950 and 150,000 in 1960, and an annexation of 30,000 people in 1955. The annual series will then include linear growth of 2,000 people a year, with a jump of 30,000 people in 1955.

City Government Finances We use annual data on municipal spending and revenue from the Annual Survey of Governments, which begins collecting data in 1952. These data include information on spending and revenue by category and, because our sample of annexing cities includes only relatively large cities, covers the entire universe of cities in our sample.

City Government Employment We use annual data on municipal employment from the Census Bureau that is compiled for cities between 1952 and 2012. Although these data cover different cities over time, they are always available for cities in our sample, because they are sufficiently large to consistently be in the sample.

3.3.2 Sample Construction

We restrict our sample to all cities with a population above 75,000 in 1990, comprising 296 cities. We identify treatment cities by several criteria: first, we restrict to events that would have a sufficient pre- and post-treatment window between 1952 and 1968, where we can impute annual population series; second, we exclude cities that have multiple treatment events in the window, to avoid muddling the estimated treatment events. We do permit short clusters of years where several large annexations take place, and in these cases, we take the first year as the treatment year. We only consider large annexations of more than 20 square miles.

These criteria leave twenty-one treatments in our sample, concentrated mostly in the South, West, and Mid-West. For each of these treatment events, we identify cities that do not report annexation events within a 15-year window and include them as candidate donors to the synthetic control.

3.4 Descriptive Facts

This section provides descriptive evidence for two key aspects of post-war municipal expansions: the annexation behavior of cities and trends in municipal revenues and spending during the annexation boom.

Figure 3.1 compares annexation patterns between Sun Belt and non-Sun Belt cities.[3]

Panel A displays the average land area of cities over time, while Panel B shows the percentage of a metropolitan area's population living within its largest municipality. Before World War II, cities across regions shared similar characteristics: they averaged around 30 square miles in area, and approximately 40%–45% of metropolitan populations resided in the largest municipality. Between 1944 and 1990, however, cities underwent significant territorial expansion, following a period of relatively stable boundaries from 1934-1944.

This expansion was highly uneven across regions: Sun Belt cities expanded dramatically, increasing their land area sixfold, while cities outside the Sun Belt experienced minimal growth. In the Sun Belt, the average city grew to 188 square miles by 2000, making the average Sun Belt city in 2000 larger than nearly all cities in land area in 1944.[4] Similarly, although the share of population living in the largest municipality declined in both regions, the decrease was substantially greater outside the Sun Belt, dropping from 46 percent in 1950 to 31 percent in 2000, while the Sun Belt share only declined from 42 to 39 percent, and in fact slightly increased during the peak of annexations in 1960. These patterns demonstrate that annexation was concentrated in Sun Belt cities, was large in scale, and enabled these cities to maintain a greater proportion of their metropolitan population during suburbanization.

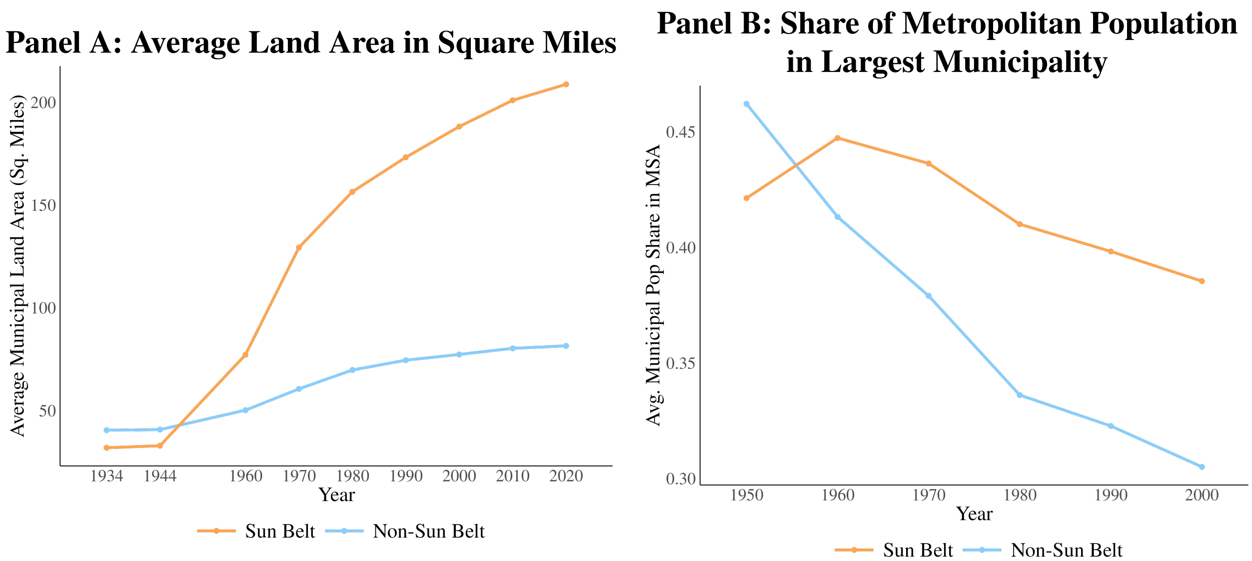
In Table 3.1, we decompose 1940 to 1990 population growth for select cities into growth within the 1940 city boundaries and within territory annexed between 1940 to 1990.[5] We document that

[3] We define a Sun Belt city as any city in the following states: Alabama, Arizona, Arkansas, California, Florida, Georgia, Louisiana, Mississippi, Nevada, New Mexico, North Carolina, Oklahoma, South Carolina, Tennessee, and Texas. These correspond to the states mostly south of the Parallel 3630 north.

[4] Only four cities: Los Angeles, New York, Chicago and New Orleans had a 1944 land area exceeding 188 square miles.

[5] We identify 1940 city boundaries using data from the Urban Transition Historical GIS project ((18))

Figure 3.1: Municipal Annexation by Sun Belt Region



NOTES: This figure shows differences in annexation for cities in the Sun Belt and outside of the Sun Belt. Sun Belt city is any city in the following states: Alabama, Arizona, Arkansas, California, Florida, Georgia, Louisiana, Mississippi, Nevada, New Mexico, North Carolina, Oklahoma, South Carolina, Tennessee, and Texas. Panel A shows the average land area in square miles of cities over time, while Panel B shows the share of a metropolitan area's population contained within its largest municipality. The sample includes the a balanced panel of the most populous cities within each MSA in 1990, where the city has a 1990 population of at least 75,000 and an MSA population of at least 200,000. This corresponds to 64 Sun Belt cities/MSAs and 66 non-Sun Belt cities/MSAs. Sources: Municipal Year Book (1934-1960), Boundary and Annexation Survey (1970-2020), and Decennial Censuses.

Table 3.1: Population Changes in 1940 City Boundaries versus Annexed Areas, 1940-1990

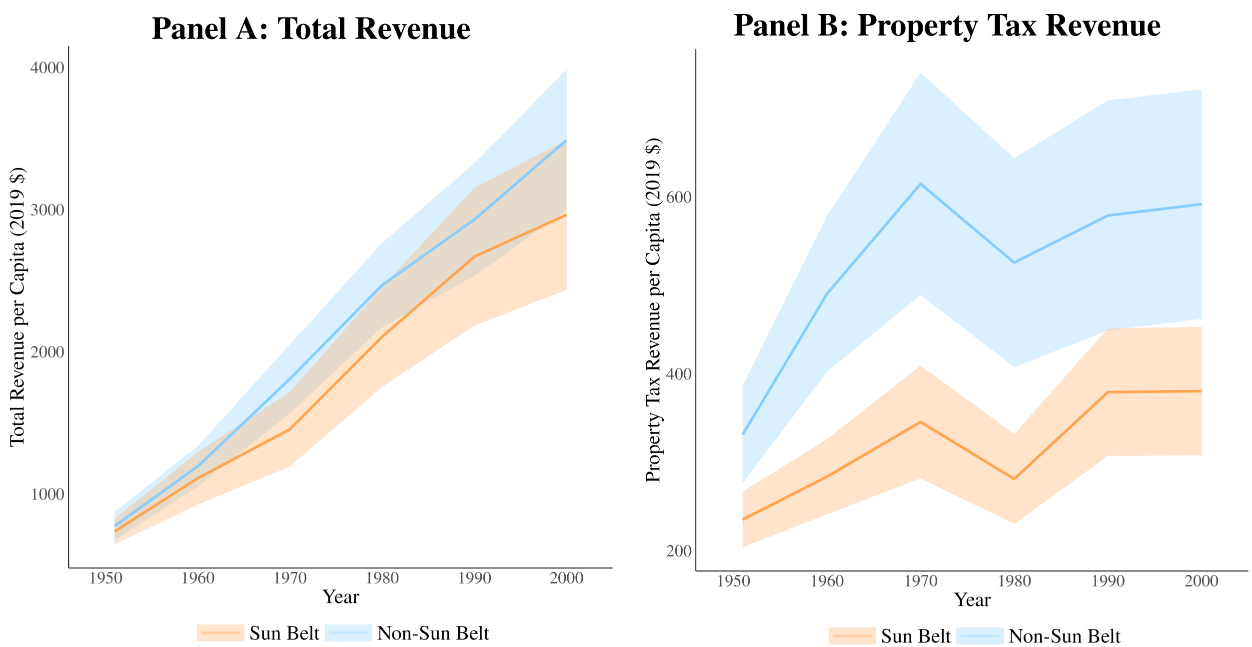
	1940	1990			
	1940 City	1940 City	Pct. Δ City	Annex	Total
Houston	384,514	331,704	-13.8%	1,303,777	1,631,766
Dallas	294,734	219,063	-25.5%	799,906	1,006,877
Oklahoma City	204,424	142,817	-30.4 %	318,453	444,719
Detroit	1,623,452	1,027,974	-36.7 %		1,027, 974
Boston	770, 816	574,283	-25%		574, 283

Houston, Dallas, and Oklahoma City all lost population within their 1940 city boundaries, and that population growth occurred mostly through annexed territories. These patterns are striking considering that a naive comparison of 1940 to 1990 population growth would suggest tremendous growth for cities like Houston and Dallas—growing from 380,000 to 1.6 million and 220,000 to 1 million, respectively. However, our analysis suggests that annexation accounted for more than the entirety of this growth. By 1990, 1.3 million of Houston's 1.6 million residents resided in territory that had been annexed between 1940 and 1990. Notably, the population associated with annexed areas in 1990 is more than even the total of Houston's population growth between 1940 and 1990—a product of the fact that Houston, Dallas, and Oklahoma City all lost population within their 1940 city boundaries. To give a sense of scale to these patterns, the change in central 1940 city population for Dallas and Oklahoma City, at 25 and 30% respectively, are not remarkably different from the 25% and 36% declines that took place in cities like Boston and Detroit, places that did not have similar opportunities to expand outwards.

While, of course, the patterns of growth and decline across American cities in this period was determined by the interaction of varied and complex forces—among them white flight, de-industrialization, and technological change such as air conditioning—these results suggest that naive comparisons of city populations mask true underlying trends. While we cannot observe what the central city populations would have been had Dallas, Houston, or Oklahoma City not been able to annex, the descriptive patterns documented here do suggest that accounting for municipal boundary changes reveals a pattern of central city flight that was more prevalent than sheer populations would suggest. Only some cities, however, had a policy lever—annexation—that could be employed to combat the declines in municipal revenue that would come from central city flight—which is the divergence we examine in this paper. However, we can only conduct this exercise for cities for which we have digitized 1940 city maps, courtesy of Urban Transition Project ((18)), and therefore cannot document similar trends for the entirety of our sample.

Figure 3.2 illustrates municipal revenue trends for Sun Belt and non-Sun Belt cities during and after the annexation boom of the 50s and 60s. Panel A demonstrates per capita revenue growth over time, revealing that average city revenue more than tripled in real terms by 2000, with similar growth patterns both in and outside the Sun Belt. Panel B shows that non-Sun Belt cities began with substantially higher per capita property tax revenue in 1950, which peaked around 1970 before leveling off. This figure highlights three key insights: first, this period witnessed dramatic expansion in municipal government size; second, despite the annexation boom in the Sun Belt, revenue growth remained comparable to the rest of the country; and third, Sun Belt cities relied on different revenue streams than their non-Sun Belt counterparts.

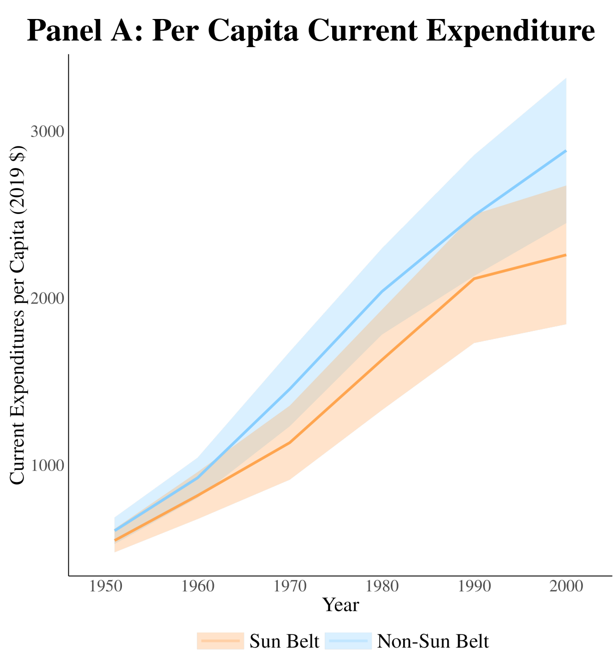
Figure 3.2: Trends in Municipal Per Capita Revenue by Sun Belt Region

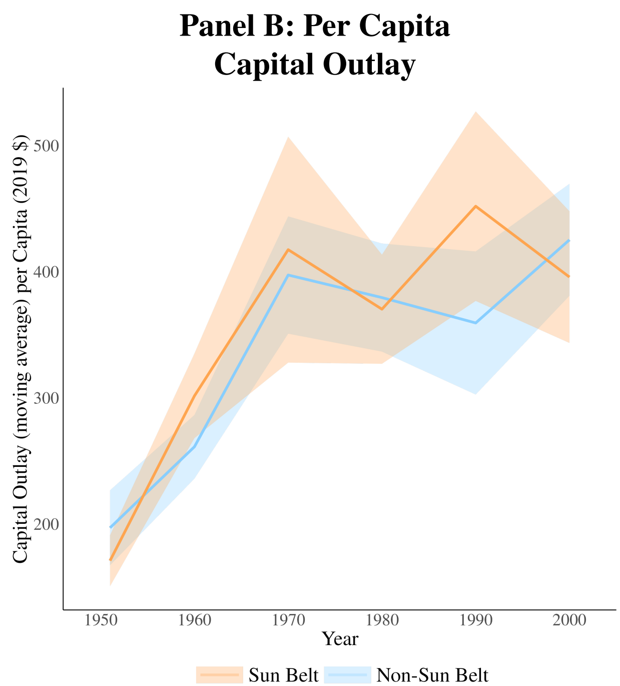


NOTES: This plot shows trends in average municipal revenue in 2019 dollars. Panel A shows per capita total revenue (dollars) from 1950-2000. Panel B shows per capita revenue from property taxes. For both plots, average trends are broken out by Sun Belt and non-Sun Belt cities. A Sun Belt city is any city in the following states: Alabama, Arizona, Arkansas, California, Florida, Georgia, Louisiana, Mississippi, Nevada, New Mexico, North Carolina, Oklahoma, South Carolina, Tennessee, or Texas. Each plot shows the mean and 95% confidence interval of the mean across cities in that year. The sample includes the most populous cities within each MSA in 1990, where the city has a 1990 population of at least 75,000 and an MSA population of at least 200,000. This corresponds to 64 Sun Belt cities/MSAs and 66 non-Sun Belt cities/MSAs. Source: Census of Governments.

Figure 3.3 documents trends in municipal expenditure patterns across Sun Belt and non-Sun Belt states. Panel A shows the growth in per capita current expenditures (variable costs); while both regions had similar expenditures in 1950, non-Sun Belt cities experienced somewhat higher growth over time. In contrast, Panel B reveals that per capita capital outlay (fixed costs) followed remarkably similar trends across both regions, with a notable spike between 1950 and 1970 (during the height of the annexation boom) before leveling off. Figure 3.4 then looks expenditures across four major subcategories of spending: police, fire, utilities, and highways. In 1950 the average city had similar patterns of spending in both levels and spending across categories. However, this figure highlights distinct regional differences in spending growth patterns. While police and fire protection expenditures closely track each other both in and outside the Sun Belt, though growing somewhat faster outside the Sun Belt, spending on utilities spiked in Sun Belt states between 1970 and 1990, whereas highway expenditure growth was consistently growing faster in non-Sun Belt cities.

Figure 3.3: Trends in Municipal Expenditure by Sun Belt





NOTES: This plot shows trends in average municipal expenditure in 2019 dollars. Panel A shows per capita current expenditure (dollars) from 1950-2000. Panel B shows per capita capital outlay (3-year moving average). For both plots, average trends are broken out by Sun Belt and non-Sun Belt cities. A Sun Belt city is any city in the following states: Alabama, Arizona, Arkansas, California, Florida, Georgia, Louisiana, Mississippi, Nevada, New Mexico, North Carolina, Oklahoma, South Carolina, Tennessee, or Texas. Each plot shows the mean and 95% confidence interval of the mean across cities in that year. The sample includes the most populous cities within each MSA in 1990, where the city has a 1990 population of at least 75,000 and an MSA population of at least 200,000. This corresponds to 64 Sun Belt cities/MSAs and 66 non-Sun Belt cities/MSAs. Source: Census of Governments. Source: Census of Governments.

This section highlights the dramatic growth of American cities in the postwar period, both in territory and municipal budgets. Sun Belt cities expanded their borders substantially, enabling them to maintain a larger share of metropolitan populations during a period of mass suburbanization. However, the implications for municipal finance are unclear. While cities were growing everywhere, per capita growth rates in revenue and expenditure were somewhat greater outside the Sun Belt. This pattern supports multiple interpretations: if local public goods exhibit increasing returns to scale across certain categories, then Sun Belt cities—where a greater share of metropolitan populations resided within central city boundaries—would face lower per capita costs in providing public amenities compared to regions where central cities contained smaller portions of the metropolitan population. Alternatively, residents in the Sun Belt might simply have different preferences regarding public amenities. To distinguish between these interpretations, data on the actual level of public good provision is necessary.

3.5 Empirical Strategy

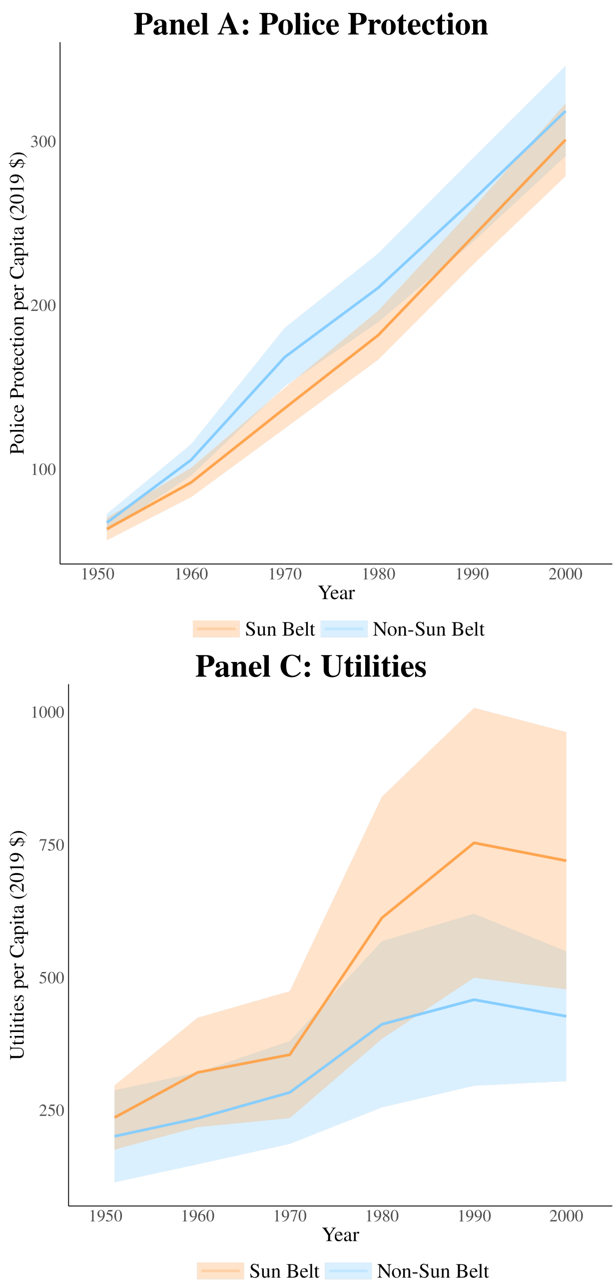
We now turn to estimating the impact of large annexations on municipal finance outcomes within individual cities.

Our main estimating equation is a stacked synthetic control difference-in-differences around large annexation events. The estimating equation is as below.

$$Y_{it} = \sum_{n=-T}^{T} \beta_n \cdot EventTime_n \cdot Annex_i + \alpha_i + \mu_t + \epsilon_{it} \tag{3.1}$$

Here, Y_{it} reflects the outcome of interest, for example, municipal tax revenue per capita. $Annex_i$ is a dummy that identifies treated cities. $EventTime_n$ are standard event-time indicators. To control for secular trends in municipal spending, we include year fixed-effects μ_t. Finally, α_i identifies group fixed-effects where the group is identified as the treatment-synthetic control pairing.

Figure 3.4: Trends in Municipal Expenditure by Sun Belt Region and Spending Type



NOTES: This plot shows trends in average municipal expenditure in 2019 dollars by spending sub-category. Average trends are broken out by Sun Belt and non-Sun Belt cities. Sun Belt city is any city in the following states: Alabama, Arizona, Arkansas, California, Florida, Georgia, Louisiana, Mississippi, Nevada, New Mexico, North Carolina, Oklahoma, South Carolina, Tennessee, and Texas. Each plot shows the mean and 95% confidence interval of the mean across cities in that year. The sample includes the most populous cities within each MSA in 1990, where the city has a 1990 population of at least 75,000 and an MSA population of at least 200,000. This corresponds to 64 Sun Belt cities/MSAs and 66 non-Sun Belt cities/MSAs. Source: Census of Governments. Source: Census of Governments.

Synthetic Control Construction

For each treatment city-year in our sample, we construct a synthetic control city by matching, within region, on pre-period population three and four years prior to the event, as well as baseline 1950 demographics—share Black, share native, and share native-born. We then use the weighting scheme generated by the synthetic control to construct all other outcomes for the synthetic control city.

It should be noted that this differs from the standard approach: usually the synthetic control is constructed on the outcome of interest, often matching on that outcome on the pre-period. Population is not our outcome interest, but rather functions as a "first-stage" of the treatment. As a result, while our method functionally imposes parallel trends in the "first-stage", there is no guarantee that the weighting scheme will also generate parallel trends for other outcomes.

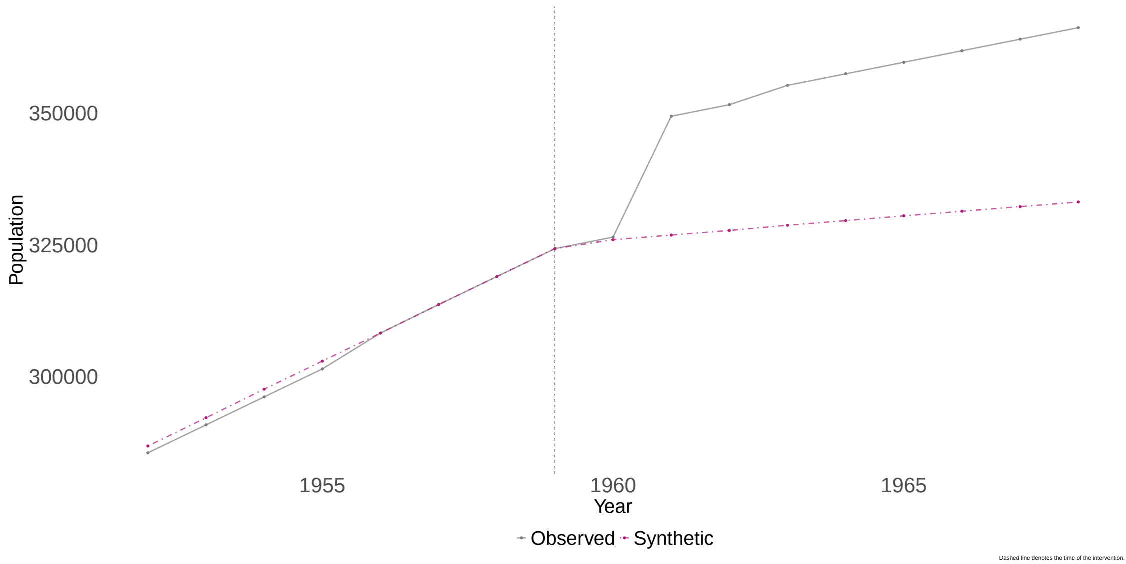
To give an example of a weighting scheme generated by this process: synthetic Oklahoma City is a composite of outcomes from Atlanta, Louisville, Tuscaloosa, Birmingham, and New Orleans, as well as other cities in much smaller shares.

Figure 3.5 plots the generated annual population series for synthetic and observed Oklahoma City: as mentioned above, because pre-period populations are include in the synthetic control construction, the tightness of parallel trends in years -4 and -1 before the event are a direct result of the synthetic control construction. However, in this case, the fact that the series approximates the observed Oklahoma City population in the 0 year before the annexation population jump is validation that the synthetic control is constructing a reasonable comparison city. The uptick in population post-event documents the divergence in Oklahoma City's population as a result of the annexation.

3.6 Results

This section reports our primary estimates stacking the twenty-one events in our samples on the impact of major annexation events on population and municipal finance and operation. As the vast

Figure 3.5: Observed vs. Synthetic Oklahoma City Population Series



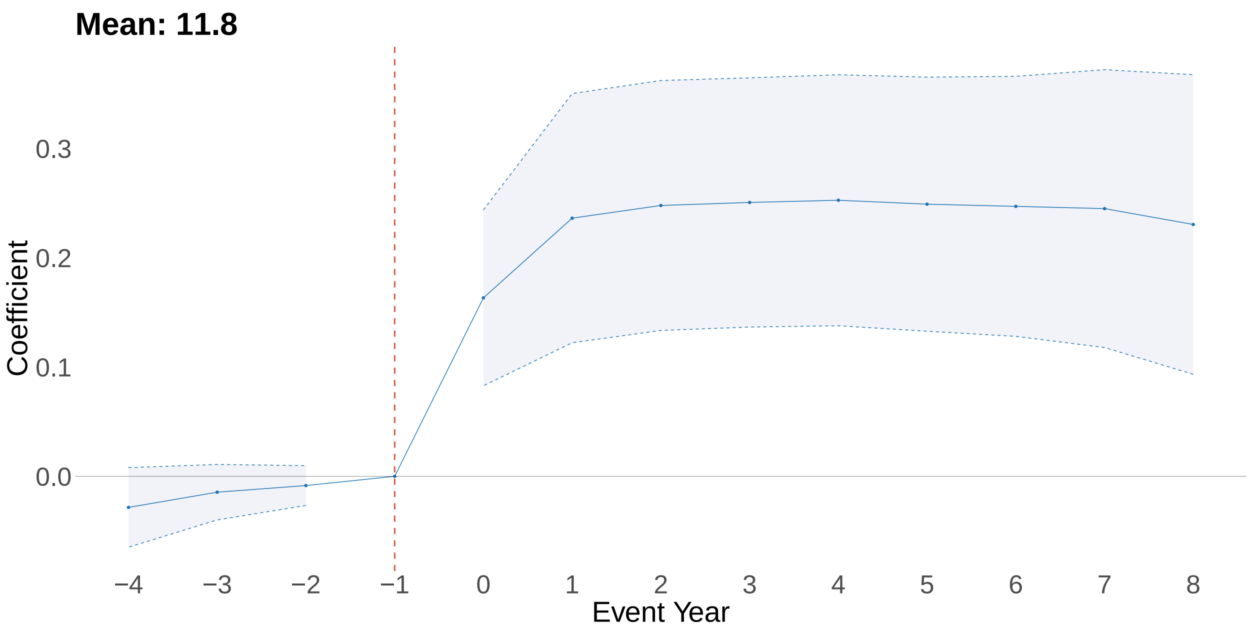
NOTES: This figure shows the population series for actual and synthetic Oklahoma city around its major annexation event in 1959. The dashed red line reflects the constructed synthetic control for Oklahoma City, while the smooth black line represents the actual population series. The vertical dashed line reflects the annexation event. Annual population changes are imputed using estimated population changes from annexation events Municipal Year Books and interpolating between population counts in decadal Census data.

majority of these treatment cities fall within the Sun Belt, the analysis speaks to the ways in which the annexation boom contributing to divergence in municipal outcomes between the Sun Belt and the rest of the country.

3.6.1 Population

After stacking our twenty-one treatment events, the estimates reflect an increase of roughly 45,000 people, which corresponds to a 24% increase in population from the baseline year. These results are reported in Figure 3.6. As a result of allowing for multiple annexations to occur in concentrated spurts, the effect almost entirely occurs in the first two years, and then stabilizes around the 24% increase in population. As described in the description of the synthetic control, the close match in the pre-period is not surprising and is a by-product of how the synthetic controls were constructed.

Figure 3.6: Log Population Event Study Estimates



NOTES: This plot shows the estimated event study coefficients for log population, pooling event years into two year bins. The shaded gray area and dashed blue lines reflect the 95% confidence intervals for the estimated coefficients. The baseline mean is reported in the top right of the figure. The vertical redline reflects event year -1.

3.6.2 Revenues

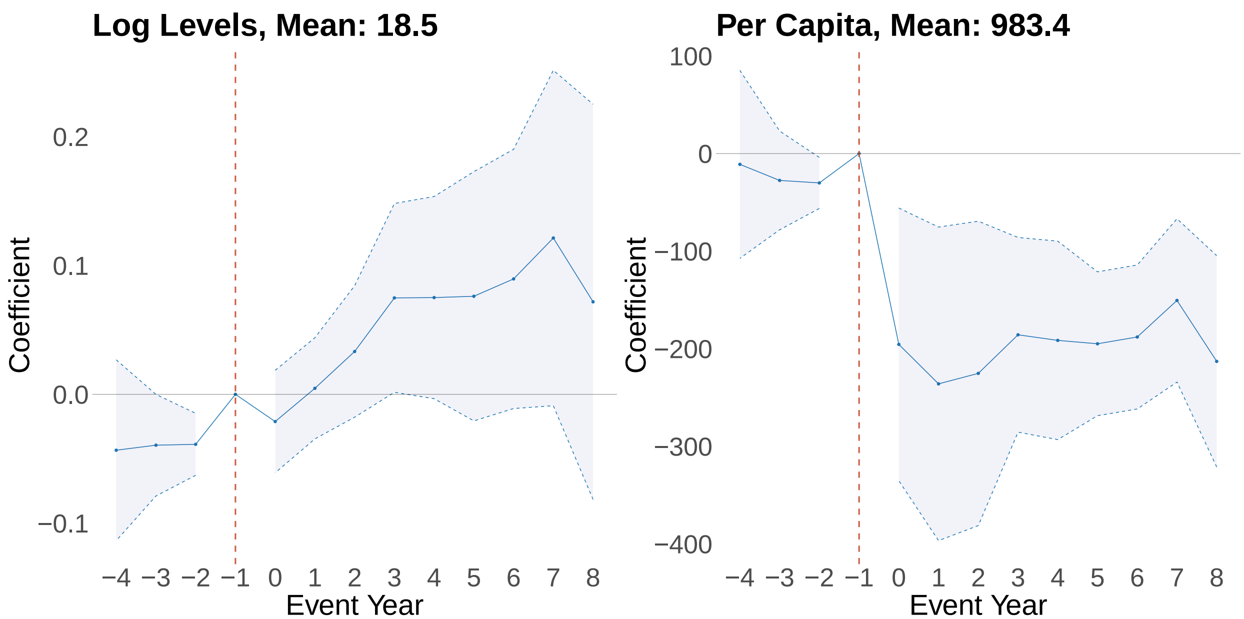
In this section, we examine how large annexation events impact total municipal revenues and property tax revenues. For all plots subsequently presented, the left hand panels reflect estimates in log levels and right hand panel reflects the estimates in per capita terms.

Figure 3.7 shows noisily estimated point estimates suggesting that total revenue increases by roughly 10% following an annexation. We observe a statistically significant reduction in revenues per capita of approximately 20%. Total revenues per capita reflect the average sum paid by residents of the city to the city government. In this sense, this reduction in per capita revenue is particularly notable given that the annexed areas were typically wealthier than the pre-existing municipal core, suggesting that either cities reduced tax rates to pass on efficiency gains to residents or that newly annexed populations were not immediately fully incorporated into the tax base ((21)).[6]

[6]Ongoing work incorporating property tax rates and assessments from Moody's will help distinguish between these two competing explanations.

Property tax revenue, a key municipal funding source likely to be impacted directly by annexation, shows similar but noisier patterns to overall municipal revenue, as shown in Figure 3.8. Property taxes slowly increase over three to four years after an annexation, stabilizing at a 20% increase. The slow increase in property tax revenue is plausibly explained by the delay caused by cities' needs to assess new areas. However, several years out from this event, growth in revenue from property taxes in annexing cities stabilize. These results suggest that while there seems to be a slight delay in taxation of newly annexed parts of the city, this alone does not explain the gap in per capita property tax revenue between annexing cities and similar non-annexing cities.

Figure 3.7: Total Revenue Event Study Estimates

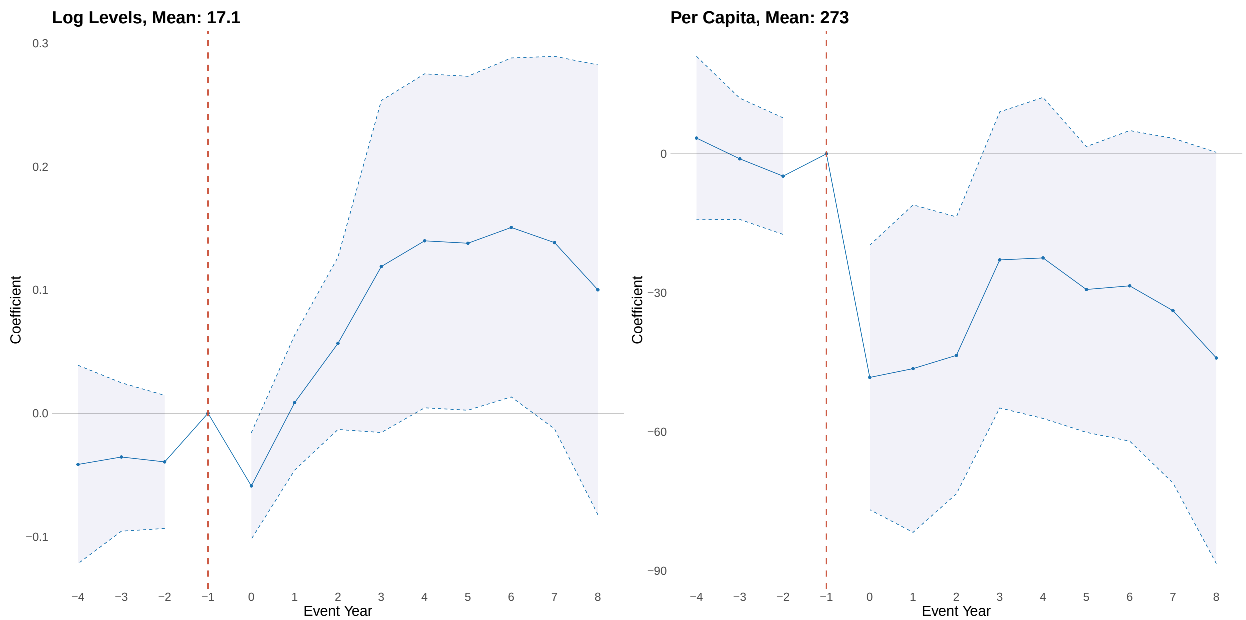


NOTES: This figure shows the impact of annexation on city-level total revenue. The left-hand panel reflects estimated event-study coefficients for log total revenue. The right-hand panel reflects the estimated coefficients for total revenue per capita. The shaded gray area and dashed blue lines reflect the 95% confidence intervals for the estimated coefficients. The baseline mean is reported in the top right of the plots. The vertical red lines reflect event years -1.

3.6.3 Expenditures

In this section, we study how these large annexations impact total expenditures and spending on current expenditures (variable costs) and capital outlay (fixed costs). We find little evidence of reductions in total per capita expenditures following annexation, but observe persistent decreases

Figure 3.8: Property Tax Revenue Event Study Estimates



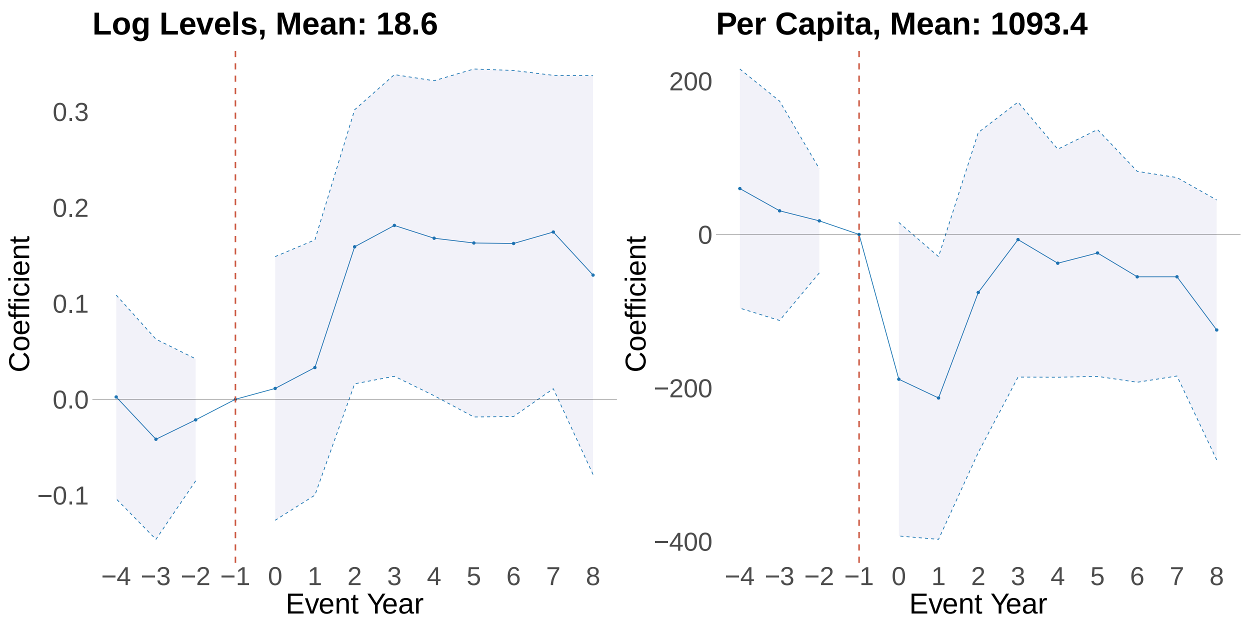
NOTES: This figure shows the impact of annexation on city-level revenue from taxes. The left-hand panel reflects estimated event-study coefficients for log property tax revenue. The right-hand panel reflects the estimated coefficients for property tax revenue per capita. The shaded gray area and dashed blue lines reflect the 95% confidence intervals for the estimated coefficients. The baseline mean is reported in the top right of the plots. The vertical red lines reflect event years -1.

in per capita current expenditures. This difference is explained by significant increases in capital outlays, which offset the reductions in current expenditures. The estimates for capital outlay, while somewhat noisy, suggest a substantial 40% increase per capita three years after annexation. These results indicate that while overall spending per capita remains stable, this stability masks two countervailing effects: variable municipal costs decrease as they spread over a larger population, while fixed costs increase through major new capital investments required to serve newly annexed territories.

Figure 3.9 presents the estimates for total expenditures. In log levels, we find evidence that total expenditures increased in levels for the treated cities: this result is in line with cities increasing spending to accommodate newly annexed territories. Point estimates are positive starting two years out from annexation, and are statistically significant in certain years. The point estimates suggest a 20% increase in total expenditures, which is roughly proportional to the increase in population.

In per capita terms, when the increase in expenditure arises after the first two years, there do not appear to be noticeable decreases in total expenditures.

Figure 3.9: Total Expenditure Event Study Estimates



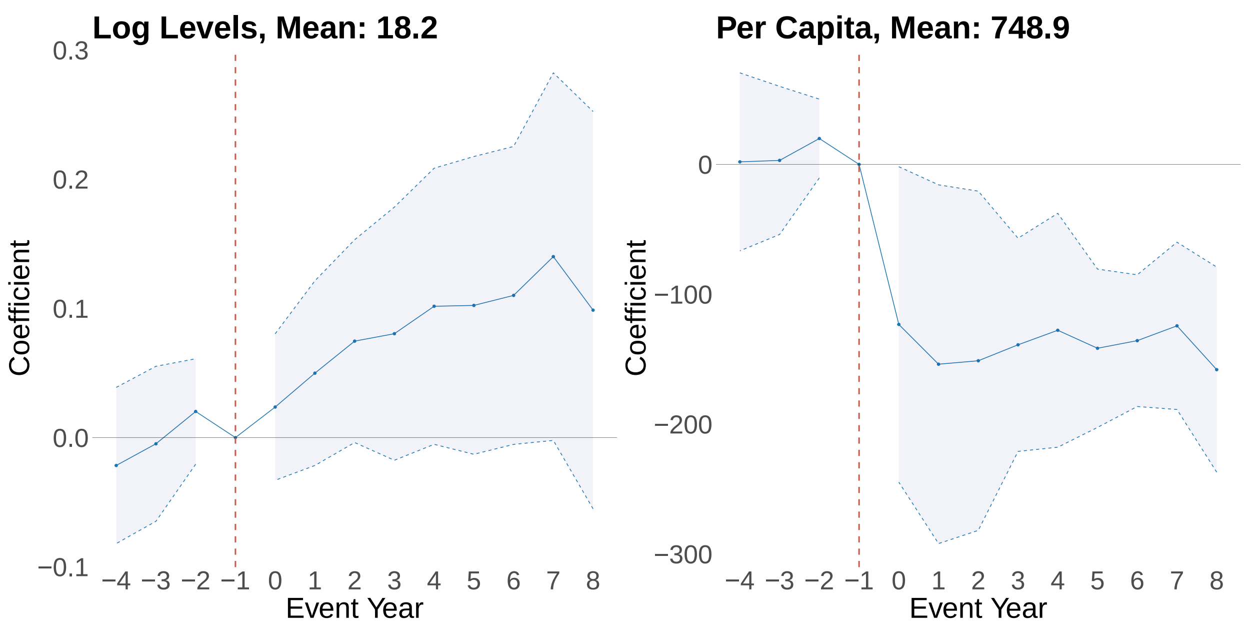
NOTES: This figure shows the impact of annexation on city-level total expenditure. The left-hand panel reflect estimated event-study coefficients for log total expenditures, pooling event years into two year bins. The right-hand panel reflects the estimated coefficients for total expenditure per capita. The shaded gray area and dashed blue lines reflect the 95% confidence intervals for the estimated coefficients. The baseline mean is reported in the top right of the plots. The vertical red lines reflects event years -1.

However, in Figure 3.10, when examining current expenditures, we observe a persistent decrease in per-capita costs. The mean current expenditure per capita the year before the annexation event is $748, and the point estimates suggest that after annexation, costs persistently decrease by 20%. Because these costs reflect annual spending, we interpret these results as suggestive of these cities being able to spread costs over a large population. A competing explanation might be that city governments simply failed to provide services to newly annexed areas, and this would also appear as decreased per capita expenditures. We lean against this interpretation for several reasons: first, the point estimates do suggest an increase in total and current expenditures; second, the persistent and stable decline in current expenditures nine years out from annexation, suggests that the decreases are not being driven by failures on the municipal governments to accommodate the new population.

The reduction in current expenditures is offset by large increases in capital outlay, as presented in Figure 3.11. The estimates, though noisy, suggest a 50% increase in annual capital outlays relative to similar non-annexing cities. These increases in levels translate to per capita increases of around 30%. These increases translate to per capita increases of approximately 30%. This pattern indicates that some, but not all, cities make capital investments following annexations that are proportionally larger than their population increases.

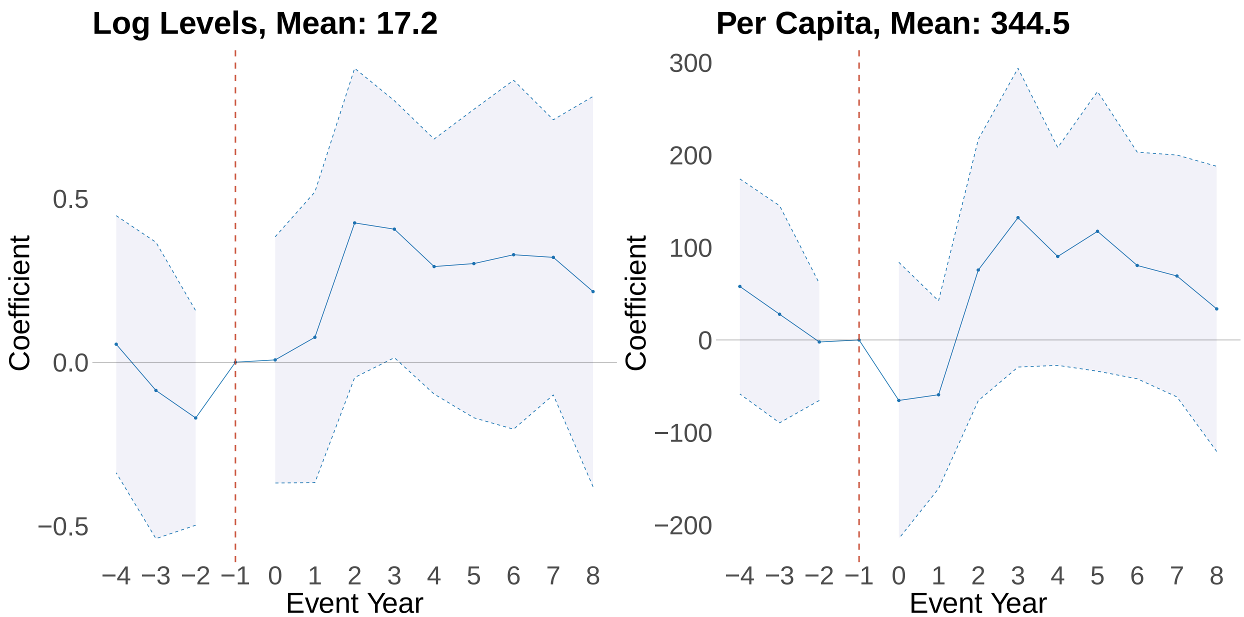
Together, these results suggest that cities leverage economies of scale to provide municipal services at lower per capita costs after annexation, while simultaneously making significant capital investments to support the expanded territory. This combination of reduced operational costs and increased capital investment maintains overall expenditure levels while potentially improving service delivery efficiency.

Figure 3.10: Current Expenditures Event Study Estimates



NOTES: This figure shows the impact of annexation on city-level current expenditure. The left-hand panel reflect estimated event-study coefficients for log current expenditures. The right-hand panel reflects the estimated coefficients for current expenditure per capita. The shaded gray area and dashed blue lines reflect the 95% confidence intervals for the estimated coefficients. The baseline mean is reported in the top right of the plots. The vertical red lines reflects event years -1.

Figure 3.11: Capital Outlay Event Study Estimates



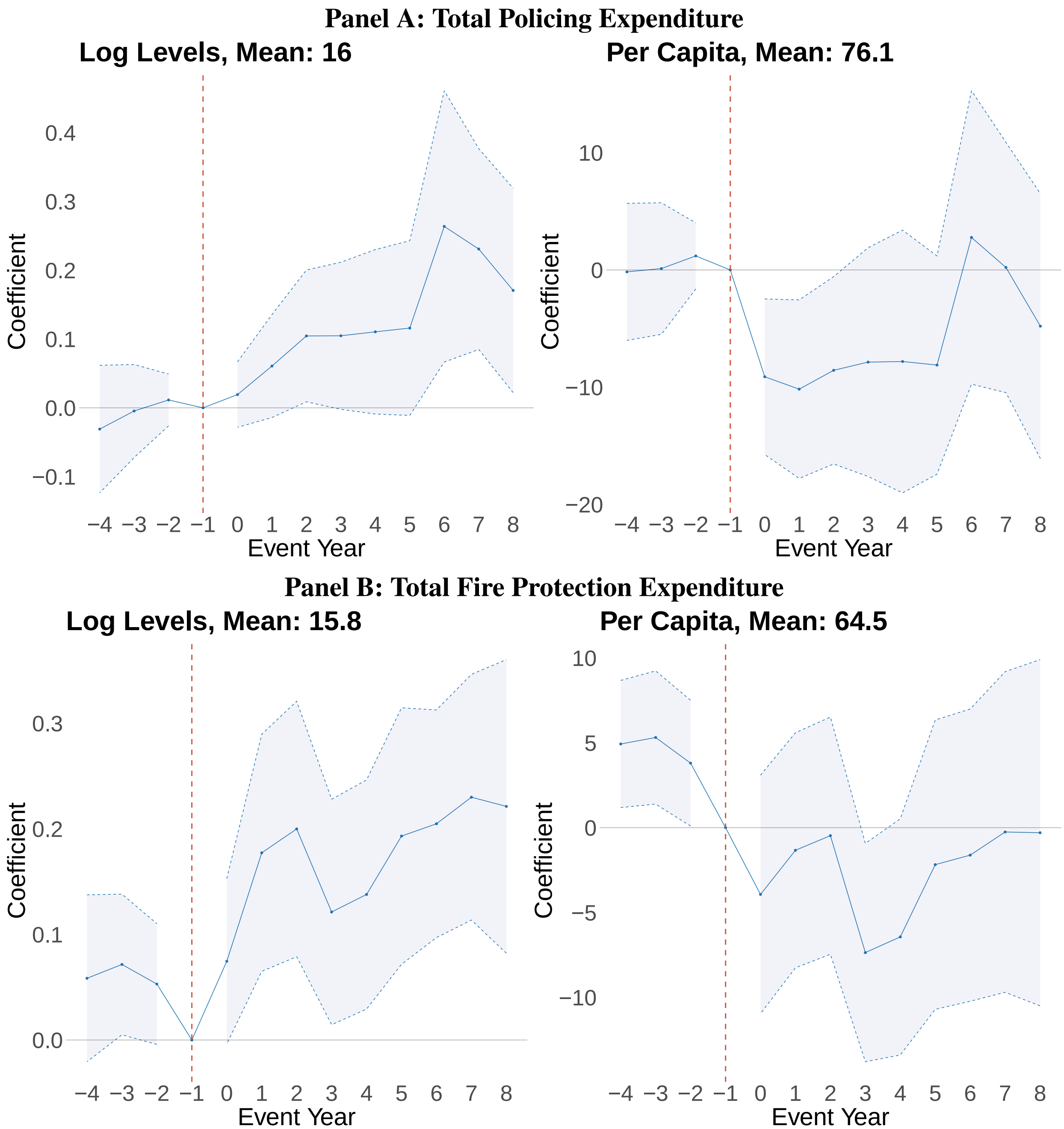
NOTES: This figure shows the impact of annexation on city-level capital outlay. The left-hand panel reflect estimated event-study coefficients for logtotal capital outlay. The right-hand panel reflects the estimated coefficients for total capital outlay per capita. The shaded gray area and dashed blue lines reflect the 95% confidence intervals for the estimated coefficients. The baseline mean is reported in the top right of the plots. The vertical red lines reflects event years -1.

Police and Fire Protection

We identify distinct patterns for high-labor cost public goods: policing and fire protection (Figure 3.12). These municipal services may require investments that scale proportionally to larger populations and areas, and therefore if municipal services are being provided to newly connected areas we would expect expenditures to scale with population after an annexation. Our estimates confirm these priors: while per capita expenditures on policing and fire protection appear to decrease immediately after an annexation, within nine years of an annexation, spending on both municipal services is similar to non-annexing comparison cities.

These results on expenditures are consistent with cities expanding services to accommodate new populations, and would suggest that the reduction in current expenditures that we observe does not arise from high labor-cost municipal services, but rather from other categories of municipal spending that depend on large fixed cost investments, such as water or sewage treatment.

Figure 3.12: Police and Fire Expenditure Event Study Estimates

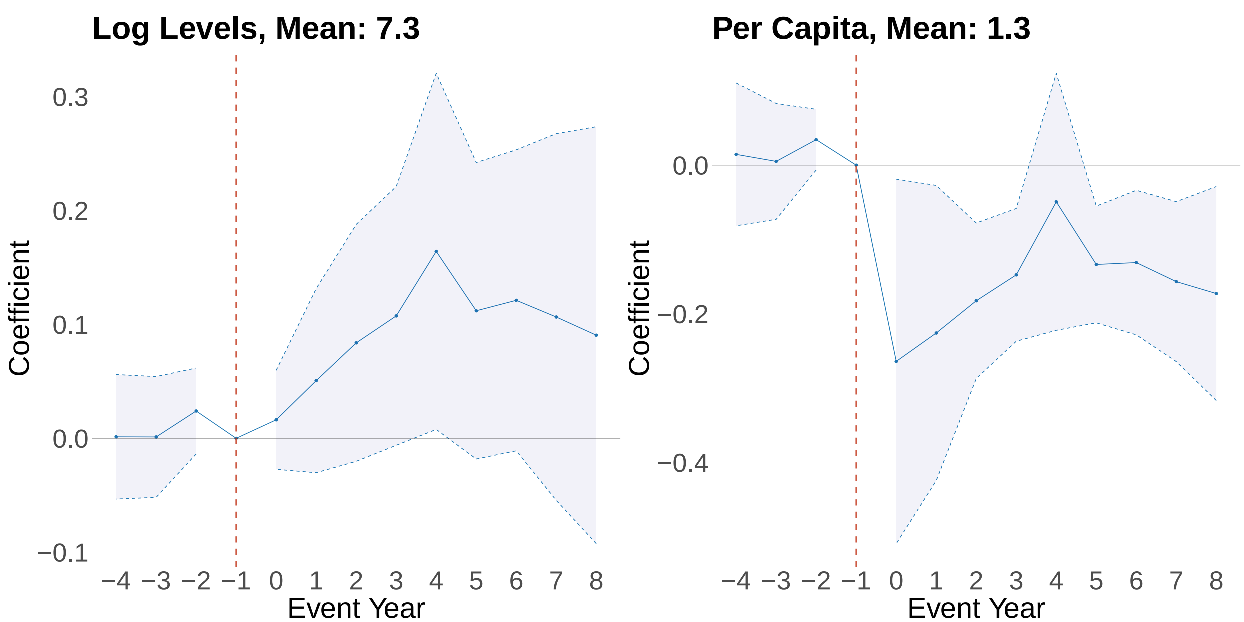


NOTES: This figure shows the impact of annexation on city-level expenditures for policing and fire protection municipal services. Panel A displays policing expenditures, while Panel B shows fire protection expenditures. For each panel, the left-hand side reflects estimated event-study coefficients for log total expenditure. The right-hand side reflects the estimated coefficients for expenditure per capita. The shaded gray area and dashed blue lines reflect the 95% confidence intervals for the estimated coefficients. The baseline mean is reported in the top right of each plot. The vertical red lines reflect event years -1.

3.6.4 Aggregate Municipal Employment

Furthering evidence on returns to scale, we also find persistent reductions in aggregate municipal employment per capita. These results are reported in Figure 3.13. Similar to some other outcomes, total municipal employment increases in levels in the point estimates, suggesting that cities expand employment as they expand services. Total employment per 100 citizens, however, seems to follow the same pattern as current expenditures and revenues per capita, exhibiting a decrease that persists up to nine years out from the annexation event. The magnitude of the point estimates for total employment per capita are slightly lower than those for current expenditures: they reflect a roughly 15% decrease in total employment per capita. For ease of interpretation, these estimates have been re-scaled to reflect employment per 100 residents, as opposed to employment per capita.

Figure 3.13: Total Municipal Employment Event Study Estimates

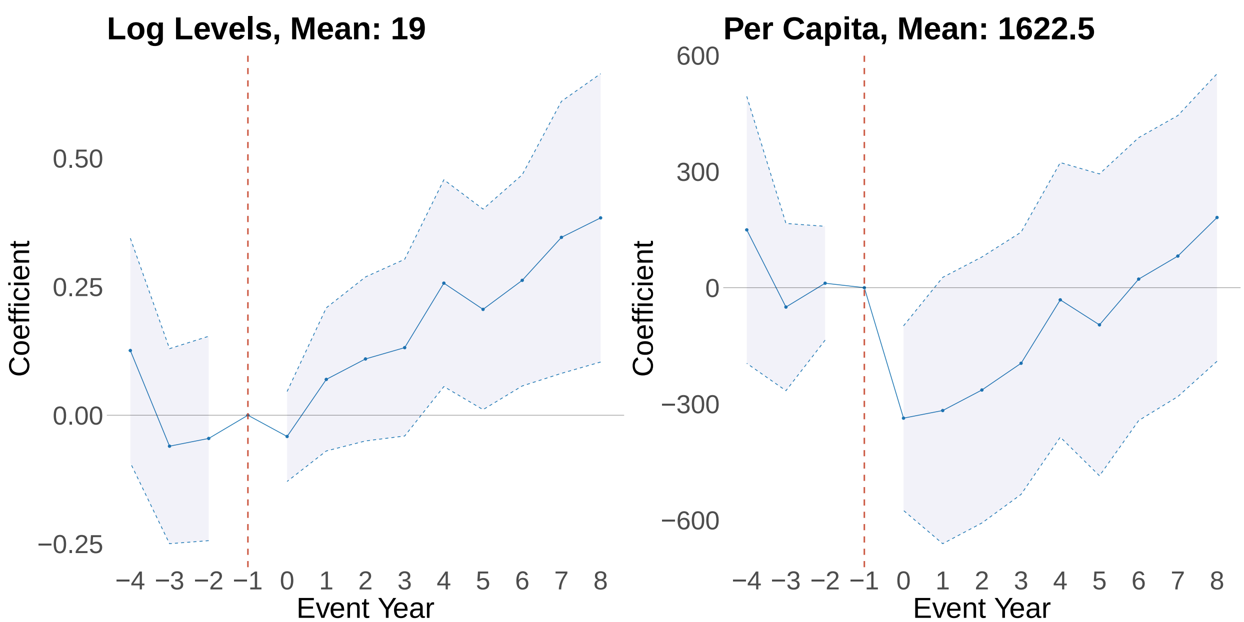


NOTES: This figure shows the impact of annexation on city-level municipal employment. The left-hand panel reflects estimated event-study coefficients for log total municipal employment. The right-hand panel reflects the estimated coefficients for total municipal employment per capita. The shaded gray area and dashed blue lines reflect the 95% confidence intervals for the estimated coefficients. The baseline mean is reported in the top right of the plots. The vertical red lines reflect event years -1.

vi. Total Debt

Finally, we turn to the total debt issued by cities. The left-hand panel shows that the debt levels increase gradually and consistently following the annexation. Here, the point estimates suggest that the increase in debt is larger than the increase in population and in total expenditures. In fact, in per capita terms, the point estimates suggest an increase in per-capita debt. These results are consistent with increases in municipal size increasing the ability of cities to secure bonded debt to finance larger investments.

Figure 3.14: Total Debt Event Study Estimates



3.7 Conclusion

Our results suggest that Sun Belt cities made use of liberal annexation laws to significantly expand their boundaries and populations. In doing so, Sun Belt cities were able to realize modest gains from scale, dispersing current expenditures and municipal employment over larger populations. We also find evidence that lower costs per capita may have translated to lower average bills paid into the municipality. At the same time, we find evidence that these cities made large capital investments following annexations, and some suggestive evidence that these investments were financed by greater access to bonded debt.

Future work will extend the analysis in this paper by digitizing data on local public good levels (i.e. stocks of police and fire equipment) from the Municipal Yearbooks to examine levels of public goods, as well as data on credit ratings and local property tax rates as outcomes of interest from Moody's Municipal Manuals.

Nevertheless, our current estimates provide some insights into urban development of Sun Belt cities. We find that Sun Belt states, by permitting their municipalities to expand their boundaries outward, also enabled them to spread some costs over a larger population and thereby reduce the total tax bill of residents into the city.

References

[1] (1873). Annexation: Result of Yesterday's Voting. *Boston Globe*.

[2] Alesina, A. and Spolaore, E. (1997). On the Number and Size of Nations. *The Quarterly Journal of Economics*, 112(4):1027–1056. Publisher: Oxford University Press.

[3] Allers, M. A. and de Greef, J. (2018). Intermunicipal cooperation, public spending and service levels. *Local Government Studies*, 44(1):127–150. Publisher: Routledge _eprint: https://doi.org/10.1080/03003930.2017.1380630.

[4] Anderson, D. M., Charles, K. K., and Rees, D. I. (2022). Reexamining the Contribution of Public Health Efforts to the Decline in Urban Mortality. *American Economic Journal: Applied Economics*, 14(2):126–157.

[5] Besley, T. and Coate, S. (2003). Centralized versus decentralized provision of local public goods: a political economy approach. *Journal of Public Economics*, 87(12):2611–2637.

[6] Bleakley, H. and Lin, J. (2012). Portage and Path Dependence *. *The Quarterly Journal of Economics*, 127(2):587–644.

[7] Blesse, S. and Baskaran, T. (2016). Do municipal mergers reduce costs? Evidence from a German federal state. *Regional Science and Urban Economics*, 59:54–74.

[8] Blom-Hansen, J., Houlberg, K., Serritzlew, S., and Treisman, D. (2016). Jurisdiction Size and Local Government Policy Expenditure: Assessing the Effect of Municipal Amalgamation. *American Political Science Review*, 110(4):812–831.

[9] Cobban, T. W. (2019). Bigger Is Better: Reducing the Cost of Local Administration by Increasing Jurisdiction Size in Ontario, Canada, 1995–2010. *Urban Affairs Review*, 55(2):462–500. Publisher: SAGE Publications Inc.

[10] Cutler, D. and Miller, G. (2005). The role of public health improvements in health advances: the twentieth-century United States. *Demography*, 42(1):1–22.

[11] Dixon, R. G. and Kerstetter, J. R. (1959). *Adjusting municipal boundaries: the law and practice in 48 states*. American Municipal Association, Washington, D.C, tentative, preliminary, abbreviated edition edition.

[12] Ferrie, J. P. and Troesken, W. (2008). Water and Chicago's mortality transition, 1850–1925. *Explorations in Economic History*, 45(1):1–16.

[13] Frere, Q., Leprince, M., and Paty, S. (2014). The Impact of Intermunicipal Cooperation on Local Public Spending. *Urban Studies*, 51(8):1741–1760. Publisher: Urban Studies Journal Limited.

[14] Geddes, R. and Troesken, W. (2003). Municipalizing American Waterworks, 1897-1915.

[15] Glaeser, E. L. (2013). Chapter 4 - Urban Public Finance. In Auerbach, A. J., Chetty, R., Feldstein, M., and Saez, E., editors, *Handbook of Public Economics*, volume 5 of *handbook of public economics, vol. 5*, pages 195–256. Elsevier.

[16] Glaeser, E. L. and Tobio, K. (2008). The Rise of the Sunbelt. *Southern Economic Journal*, 74(3):609–643.

[17] Hall, J. C., Matti, J., and Zhou, Y. (2020). The economic impact of city–county consolidations: a synthetic control approach. *Public Choice*, 184(1):43–77.

[18] Logan, J. R., Minca, E., Bellman, B., Kisch, A., and Carlson, H. J. (2024). From Side Street to Ghetto: Understanding the Rising Levels and Changing Spatial Pattern of Segregation, 1900–1940. *City & Community*, 23(2):155–179.

[19] Luca, D. and Modrego, F. (2021). Stronger together? Assessing the causal effect of inter-municipal cooperation on the efficiency of small Italian municipalities. *Journal of Regional Science*, 61(1):261–293. _eprint: https://onlinelibrary.wiley.com/doi/pdf/10.1111/jors.12509.

[20] Markusen, A. R. (1991). *The Rise of the Gunbelt: The Military Remapping of Industrial America*. Oxford University Press. Google-Books-ID: dM_qhG6rUacC.

[21] Miller, A. R. and Varon, B. (1962). Adjustment For Area Comparability Of Statistics On Family Income From The 1960 And 1950 Censuses Of Population: Major Cities And Their Standard Metropolitan Statistical Areas. (7).

[Narasimhan and Weaver] Narasimhan, V. and Weaver, J. Polity Size and Local Government Performance: Evidence from India. *American Economic Review*.

[23] Nolting, O. F. and Arnold, D. S. (1952). *The Municipal Yearbook 1952*. The International City Managers Association, Chicago.

[24] Nolting, O. F. and Arnold, D. S. (1965). *The Municipal Yearbook 1965*. The International City Managers Association, Chicago.

[25] Perry, D. C. and Watkins, A. J. (1977). *The Rise of the Sunbelt Cities*. SAGE Publications. Google-Books-ID: tVJPAAAAMAAJ.

[26] Reingewertz, Y. (2012). Do municipal amalgamations work? Evidence from municipalities in Israel. *Journal of Urban Economics*, 72(2):240–251.

[27] Rosenthal, S. S. and Strange, W. C. (2008). The attenuation of human capital spillovers. *Journal of Urban Economics*, 64(2):373–389.

[28] Rusk, D. (2013). *Cities without Suburbs*. Johns Hopkins University Press.

[29] Shoked, N. (2023). ANNEXATION'S LONG GOOD-BYE.

[30] Smallwood, J. (1977). *Urban builder : life and times of Stanley Draper*. Norman: Published for the Oklahoma Heritage Association by the University of Oklahoma Press, 1st ed. edition.

[31] Tiebout, C. M. (1956). A Pure Theory of Local Expenditures. *Journal of Political Economy*, 64(5):416–424. Publisher: University of Chicago Press.

[32] Tricaud, C. (2021). Better Alone? Evidence on the Costs of Intermunicipal Cooperation. *SSRN Electronic Journal*.