I Z A Institute
of Labor Economics
Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13952

Network Sorting and Labor Market Outcomes: Evidence from the Chaotic Dispersal of the Viet Kieu

Christopher Parsons
Tyler Reysenbach
Jackline Wahba

DECEMBER 2020

I Z A Institute
of Labor Economics

Initiated by Deutsche Post Foundation

DISCUSSION PAPER SERIES

IZA DP No. 13952

Network Sorting and Labor Market Outcomes: Evidence from the Chaotic Dispersal of the Viet Kieu

Christopher Parsons
University of Western Australia and IZA

Tyler Reysenbach
Productivity Commission

Jackline Wahba
University of Southampton and IZA

DECEMBER 2020

ISSN: 2365-9793

IZA – Institute of Labor Economics

Schaumburg-Lippe-Straße 5–9
53113 Bonn, Germany

Phone: +49-228-3894-0
Email: publications@iza.org

www.iza.org

ABSTRACT

Network Sorting and Labor Market Outcomes: Evidence from the Chaotic Dispersal of the Viet Kieu*

Immigrants' social networks exert considerable influence over their labor market opportunities and yet the pre-sorting of co-nationals by ability and across space, endures as a key challenge for empiricists attempting to establish causal network effects. To surmount this issue, we leverage the chaotic dispersal of Vietnamese refugees across the U.S. in 1975, which was demonstrably exogenous in both initial network size and quality, in tandem with an absence of pre-existing networks of co-nationals, to causally identify the effects of network size and network quality on refugees': occupational outcomes, skill intensity and skill upgrading. Our administrative data provide refugee's precise initial locations and pre-placement characteristics in Vietnam, which we uniquely employ as additional controls, as well as longitudinal information about their locations and occupations six years hence. We construct instruments from the initial quasi-random refugee allocations of network size and quality and leverage refugees' geo-locations to insulate our results from the Reflection Problem. Overall, network quality is a far more important determinant of refugees' labor market outcomes when compared to network size, one interpretation of which is that the type of referrals network members receive are more important than the overall number of referrals. Blue-collar networks: increase the probability of refugees' working in blue-collar jobs, draw additional workers into more manual and less complex intensive employment and serve to up-skill individuals along the manual skill dimension. Given the protracted circumstances under which the *Viet Kieu* entered the U.S., the composition of their networks played a pivotal role in their ultimate success.

JEL Classification: F22, J61
Keywords: networks, refugees, migration, labor markets

Corresponding author:
Christopher Parsons
Economics (UWA Business School)
University of Western Australia
35 Stirling Highway
Crawley WA 6009
Australia
E-mail: christopher.parsons@uwa.edu.au

* We wish to extend our gratitude to participants of the workshop on "Forced Displacement, Asylum Seekers and Refugees" University of Queen Mary, March 2019, the Workshop on "Migration and Mobility" University of Glasgow, June 2019, the first EARN Workshop on Integration, University of Copenhagen June 2019, the Annual Conference of the Research Group on Development Economics, German Economic Association, June 2019, the 12th International Conference on Migration and Development, Madrid, June 2019, the 6th annual conference of the International Association for Applied Econometrics, Nikosia, June 2019 and seminar participants at LISER, Luxembourg, July 2019. In particular we wish to thank: Toman Barsbai, Simone Bertoli, Christian Dustmann, Francesco Fasani, Mette Foged, Dominik Hangartner, Tobias Heidland, Cağlar Özden, Oddbjørn Raaum, Steven Stillman, Alexander Teytelboym, Torben Tranæ, and Dean Yang.

1 Introduction

The success of refugees in local labor markets has occupied the minds of successive American administrations' since at least President Truman's signing of the (first) Displaced Persons Act of 1948. Given sufficient agency, refugees tend over time, to (co-)locate with their compatriots, thereby sorting into networks, enclaves and ghettos ((Bartel, 1989; Stark and Stark, 1991; Borjas, 1998). Theoretically, the influence of these agglomerations on individual's labor market outcomes is ambiguous, since they may both foster opportunities through referrals, while simultaneously discouraging individuals to acquire the local language (Edin et al., 2003). Empirically disentangling causal network effects proves challenging, not least since such agglomerations occur in terms of both size *and* quality (Grogger and Hanson (2011), Kerr et al. (2016), Kerr et al. (2017)), *both* of which are endogenous to individual's labor market outcomes.

As foreign-born individuals from similar origins co-locate, so oftentimes do they specialize in specific occupations, thereby carving out employment in particular sectors, in what is termed 'labor market niches' (see: Lieberson (1980); Waldinger (1994)). Agglomerations of (Southern) Vietnamese - the largest refugee population in the U.S. - commonly referred to as 'Little Saigons'[1] took root following the sudden arrival of 130,000 Vietnamese refugees in 1975. In the U.S, the Vietnamese overseas diaspora, the *Viet Kieu*, are associated with their cuisine that proliferated the U.S. from 1975 onward (pho and cha gio restaurants, banh mi bakeries and supermarkets), repair and maintenance companies and beauty salons among many other (typically small) businesses. In the Gulf Coast region for example, Vietnamese Americans account for 45% to 85% of the region's shrimp business, while 80% percent of California's nail technicians are Vietnamese Americans.[2]

The sorting of refugees across space and by skill poses a challenge to the empiricist, since even should individuals be randomly allocated across space, by for example a refugee dispersal policy (see: Edin et al. (2003); Damm (2009); Beaman (2011); Martén et al. (2019)), individual's labor market opportunities are necessarily going to be shaped by the economic alignment and engagement of any pre-existing and pre-sorted networks they are allocated to, since both the size and quality of networks likely matter. In other words, individual's opportunities are determined by both the *number* as well as the *type* of referrals they receive.

Causally identifying the role of network size and network quality on individual's labor market outcomes proves extremely challenging due to the endogeneity that arises from prior network formation and sorting (see Section 2). Many existing refugee dispersal policies for example, while randomly allocating individuals across space, are unlikely to result in exogenous distributions of network quality, since dispersal policies typically adhere to strict guidelines, for example placing individuals with close family

[1] Borrowing the name from the former capital of South Vietnam.

[2] https://www.latimes.com/archives/la-xpm-2008-may-05-me-nails5-story.html.

members together, else with co-nationals. The ideal thought experiment would instead involve *randomly* allocating a large number of *individuals* of varying abilities across an expansive space, one without *any* pre-existing networks, such that each individual is randomly allocated across space, while concurrently being allocated to a network of both random size *and* random quality. The empiricist would then follow such individuals over time and observe and contrast their outcomes. The natural experiment exploited in this paper almost exactly mimics these stringent conditions.

The Fall of Saigon in 1975 culminated in the largest refugee wave in American history. The unpopularity of the Vietnamese in the U.S. led Congress to mandate, for the first and only time in American history, that a refugee wave be dispersed as swiftly and as widely as possible so as to minimize impacts on host communities (Robinson, 1998; Rumbaut and Portes, 2001; Sonneborn, 2007; Thompson, 2010). Once on American soil, 130,000 Indochinese were expeditiously processed and scattered across the nation in a period of just 32 weeks, a U.S. record, (refer to Figure A.2). Voluntary Agencies (VOLAGs) were contracted to settle refugees, placing individuals as and when sponsors became available.[3]

Given the absence of pre-existing networks it was not possible to allocate the Vietnamese to pre-existing agglomerations of co-nationals, which is the rule-of-thumb followed today (Stein, 1979). The relevant authorities had little to no time to prepare the resettlement camps prior to the refugees' arrival, such that no official settlement guidelines were ever drafted. Pressure from Congress resulted in insufficient time to properly vet refugees on arrival, the process of which was complicated by poor English-language proficiency (Rumbaut and Portes, 2001).[4] Case workers were prohibited from visiting refugees, ensuring the VOLAGs were unable to shop for the best clients. Congress also mandated all refugees to be treated equally regardless of prior rank or status (Thompson, 2010). What resulted was a chaotic refugee dispersal, one particularly harrowing aspect of which was that immediate families were literally torn apart from one another, as refugees, typically from large extended families, could, at most, be settled with one immediate family member (Sonneborn, 2007). The situation was so dire, Senator Edward M. Kennedy described efforts to receive the *Viet Kieu* as a 'shambles' (Thompson, 2010). The chaotic process that ensued, catalyzed the passing of the 1980 Refugee Act, which standardized the way in which refugees were subsequently processed under the auspices of the U.S Resettlement Program, thereby ensuring such disorderly allocations of refugees were consigned to history. The chaotic dispersal of the *Viet Kieu* does however, represent the singular best opportunity for the empiricist to examine the causal impact of refugee networks on refugees' labor market outcomes in the context of the United States, since the initial allocation was the most exogenously allocated in U.S. history.

A large literature explores the role of social networks and their role in overcoming information asymmetries in the labor market, typically through providing referrals to employees and employers (Borjas,

[3]Sponsors were responsible for providing food and accommodation to refugees until they were economically self-sufficient.

[4]For many *Viet Kieu* the first word of English they learned was 'sponsor' (Thompson, 2010).

1998; Munshi, 2003; Edin et al., 2003).[5]. For immigrants, in particular, social networks shape their opportunities, choices and outcomes in the labor market, although the direction of these effects are not always positive. Munshi (2003) for example, finds that social networks yield a positive effect on the employment of Mexican immigrants, whereas Damm (2009) rather uncovers a negative effect on employment of refugees in Denmark. Network influences on occupational choices are also mixed. Immigrant and co-ethnic networks often form occupational niches locking individuals into particular occupations (Ioannides and Datcher Loury, 2004). Munshi and Rosenzweig (2006) find evidence of this effect in India, with castes locking individuals into 'traditional' occupations, while Patel and Vella (2013) find that recently arrived immigrants to the U.S. are likely to engage in the occupation of their local ethnic network. In contrast, Munshi (2003) finds that networks help Mexican immigrants improve the probability of finding work in non-agricultural occupations, suggesting that networks can improve members' occupations.

The theoretical literature provides useful insights into understanding these mixed empirical results, since the effect of social networks on the labor market outcomes is theoretically ambiguous. Although membership of a social network can be beneficial for labor market outcomes for example, as larger social networks, all else equal, are better than smaller networks at transmitting information pertaining to jobs and job referrals. Conversely, networks of lower quality might offset the advantages of larger networks as unemployed groups fail to transmit relevant job information, see e.g. Wahba and Zenou (2005). *A priori* we would therefore expect individuals to gain more from membership of a high-ability employed network than from a low-ability and unemployed network.

A few studies have examined the impact of both network size and quality on labor market outcomes. Bertrand et al. (2000) for example, use welfare dependency as a measure of network quality, although they are unable to study the direct effect of network membership, but rather the differential effect of contact. Edin et al. (2003) rather leverage ethnic income and ethnic self-employment rates as measures of quality and find that networks with less than average quality have a negative effect on refugees' earnings. We improve on previous studies that are typically unable to adequately address the endogeneity of network quality (prior network sorting), the focus of which has traditionally been on (accounting for the endogeneity of) network size. These studies therefore fail to address the endogeneity of network quality, else use the same instrument for network size and network quality, as opposed to using distinct instruments so as to control for the potential idiosyncratic endogeneity that arises separately from network size and quality. As such, prior studies are unable to distinguish empirically between any effects of network quality and their country of origin fixed effects (e.g. Edin et al. (2003) and Damm (2009). We rather circumvent this issue by focusing upon a single country of origin (Vietnam), for which we are able to exploit rich longitudinal information on both network size and quality.

One of the main challenges faced by refugee hosting countries is the limited integration of minority

[5]Please refer to Topa and Zenou (2015) and Jackson et al. (2017) for comprehensive literature reviews

groups, especially refugees. Evidence suggests that refugees might be characterized by lower labor market participation and mobility rates as when compared to immigrants and natives more broadly, see for example Brell et al. (2020). One plausible explanation for this finding, is that refugees tend to agglomerate and sort into areas with high concentrations of co-nationals, resulting in lower rates of integration and persistent inequality. Many governments have therefore adopted refugee policies that disperse individuals over various geographies with the aim of facilitating their integration although scant evidence exists on the efficacy of such policies. Brücker et al. (2019) for example, find that dispersal policies in Germany have negative effects on the labor market outcomes of refugees. Martén et al. (2019) rather conclude in the case of Switzerland that such policies only aide refugees if they are assigned to locations with dense co-national populations, where they are more likely to enter labor markets, thereby concluding that ethnic networks foster the economic integration of refugees.

Our paper contributes to this literature in several ways. First, we provide the cleanest estimates to date of network size *and* quality on individuals' labor market outcomes (measured at the commuter zone level); since we leverage the chaotic refugee allocation process in 1975, which demonstrably randomized the initial refugee allocation in terms of both network size and quality (as measured by refugees' pre-placement characteristics while in Vietnam); all in the absence of pre-existing networks (see Section 2). Policies that instead assign refugees adhere to strict guidelines, allocating for example, refugees to pre-existing, and as such, pre-sorted communities, are necessarily going to result in allocations endogenous to individuals' outcomes, see for example Edin et al. (2003) and Damm (2009).

Secondly, we examine refugees' occupational outcomes while controlling for refugees' occupations prior to arriving to the U.S, which Haines (2012) deems a 'crucial background factor' in the case of refugee resettlement. Previous related studies rather focus on earnings, while only studying positive earners. We rather examine the impact of network quality on binary occupational outcomes (blue collar and professional), skill intensities (cognitive, manual and complex) and skill-upgrading and downgrading (cognitive, manual and complex), since there is a concern that refugees tend to downgrade in the host labor market, for example Cortes (2004). We are therefore able, for the first time, to provide estimates detailing how networks of various sizes and qualities both help and hinder individual's occupational outcomes.

Thirdly, our rich administrative data from the Office of Refugee Resettlement (ORR), identify refugees' precise initial and subsequent locations. Existing studies, for example, Edin et al. (2003) and Damm (2009) are unable to directly observe refugees, instead relying on an indirect method to identify 'likely' refugees based on their country of origin and family status. This potentially means that non-refugees, who were not subject to the dispersal policy, could be captured in their sample, and indeed that individuals in their sample could have moved before they were first observed, the time delays of which are not discussed.

Our findings suggest that network quality is a far more important determinant of refugees' labor

market outcomes when compared to network size, suggesting that the *type* of network member referrals individuals receive are more important than the overall number of referrals. Blue-collar networks increase the probability of refugees' working in blue-collar jobs, draw additional workers into more manual and less complex intensive employment and serve to upskill individuals along the manual skill dimension. In particular, lower skilled networks appear to lock lower skilled individuals into jobs lower down the occupational ladder, therein preventing them moving to more skilled or complex tasks. Another interpretation of our main results, is that the *Viet Kieu* quickly developed networks that aided individuals to find available work during a deep recession. Either way, our findings have important implications for the placement of refugees and their long term labor market performance.

The remainder of the paper proceeds as follows. Section 2 outlines the context of the Vietnamese dispersal policy and analyses whether there is any evidence of sorting. It provides stylized facts of the Vietnamese refugees post-settlement. Section 3 discusses our administrative refugee data as well as the census data used to supplement them. Section 4 discusses methodological issues, our identification strategies and the estimated models. Section 5 reports the results of the occupation and skill analyses. Section 6 summarizes and concludes.

2 Refugee settlement

2.1 Aftermath from the Fall of Saigon

Saigon fell suddenly to the Vietnamese Communist Party in April 1975. It took little time for the victors to seek retribution, leading to a mass exodus of those that had fought alongside and supported the American incursion. Millions ultimately emigrated, but of the more than 130,000 that escaped 're-education camps' and 'agricultural collectives' in 1975, 90% did so in April (Baker and North, 1984), culminating in the largest refugee wave in American history. Given the numbers involved, congress hurriedly passed the 1975 Indochina Migration and Refugee Assistance Act in order to facilitate the refugees' arrival. Whereas nowadays the President sets an annual refugee ceiling following internal deliberations. *"When Saigon fell (in 1975) there was no such consultation. The Ford administration simply decided to rescue as many of our allies as we could, and we did under Dunkirk-like conditions" (Baker and North, 1984).*

The decision to welcome the Indochinese to America's shores was not publicly well received. A 1975 Gallup poll showed that 54% of Americans didn't want to admit a single Indochinese refugee (Sonneborn, 2007), while the LA Times reported that only 36% of Americans were in favour of admitting the Vietnamese. *The Boston Globe* cited three reasons why the Vietnamese were so unpopular i) because such a seemingly large population was arriving to America's shores all at once ii) since the U.S. was experiencing a recession at a time of high unemployment (8.9 percent) and iii) since the Vietnamese were a lasting reminder of a $150 billion war, one that lasted a decade and ended in America's first military defeat (Thompson, 2010). *"A backlash of unprecedented proportions is threatening to make*

South Vietnamese the first unwanted wave of refugees to reach America's shores since the Statue of Liberty was erected in New York Harbor" it was reported (Thompson, 2010).

This unpopularity was reflected in the actions of Congress, which called for as swift and as wide a dispersal of refugees as possible. To deal with such a large influx of refugees, a special Interagency Task Force (IATF) was established on the 18th April 1975, the instructions of which called for "expeditious processing". Frank Wisner, Deputy Director of the IATF is quoted as saying *"We were under tremendous pressure to move refugees from Vietnam to Guam, from Guam to the United States, and out of the resettlement camps in the United States."* (Thompson, 2010). Congress' primary concern was to reduce the impact on any one locale, (House of Representatives, 1975; Zhou and Bankston, 1998) following the experience of the Cubans that had arrived earlier to only agglomerate in a single locale, Miami.

The refugee placement policy comprised three phases. The initial evacuation from Vietnam, which involved 140,767 refugees. Temporary care on Guam, and finally, the resettlement of refugees to the U.S. mainland (132,421), in third countries (6,632), or repatriation back to Vietnam (1,546). The refugees were resettled from four camps on the U.S. mainland: Fort Chaffee, Arkansas (50,135); Eglin Air Force Base, Florida (8,665); Fort Indiantown, Pennsylvania (21,651); and Camp Pendleton, California (48,418). Since Congress wanted to ensure that there was no *'high impact'* on any one locale from an influx of refugees, the camps were chosen to be some distance from one other (Robinson, 1998), most notably Fort Chaffee, as Arkansas represented the state with the lowest percentage of foreign-born individuals, just 0.4% compared to the United States' average of 4.7% in 1970 (Baker and North, 1984). Operating under emergency conditions all 130,000 refugees were dispersed in just a 32 week period (HEW Refugee Task Force, 1976), (refer to Figure A.2).[6]

Given the magnitude of the task, the Inter Agency Task Force (IATF) recruited nine Voluntary Agencies (VOLAGS) to help resettle the refugees. The idea was that refugees would be paired with sponsors through VOLAGs, who in turn would provide the refugees with food, shelter and clothing until they were economically self-sufficient, for which VOLAGs received $500 for every refugee paired/resettled (Thompson, 2010). To prevent the VOLAGS from shopping for the best clients among refugees their personnel were prohibited from visiting the areas where the refugees lived. *"Another decision was that all refugees were to be treated equally...there would be no more Colonels or Generals, for example, only Mr.'s"* (Thompson, 2010). In theory, the matching process between refugee and sponsor *"consisted of reviewing the refugees' occupational background against a Department of Labor's listing of labor markets needing additional workers, comparing refugees' preferences for place of resettlement against the agency's opportunities, and assigning the refugees to a sponsor in the chosen locality."* (Baker and North, 1984).

The reality on the ground was very different however. First, given the narrow window between evacuation and resettlement, there was little to no time to prepare the camps (Thompson, 2010). Sec-

[6]The last 68 refugees to be processed left in Fort Chaffee on December 20th 1975.

ondly, *"pressure came from the Congress and the Executive Branch to remove the refugees from the reception centers as quickly as possible". "Everyone worked 12-hour shifts, 7 days a week, and it was not uncommon to work 15 to 16 hours at a time"*. (Thompson, 2010) such that it wasn't possible to *"vet each and every sponsorship offer" some of which were "little more than requests for indentured servants, bed mates, or cheap labor"* (Robinson, 1998). Thirdly, the refugee arrivals had *"minimum formal education, few marketable skills, little English-language proficiency, and scant knowledge of the ways of an advanced society", such that the process of allocation was poorly understood and the process undermined* (Rumbaut and Portes, 2001). Fourthly, the ensuing chaos led to a 'mal-distribution of caseloads' among the VOLAGs, with, for example, the United States Catholic Conference resettling 48% of refugees (GAO, 1977). The VOLAGs placed refugees wherever there were willing sponsors, a translator remarked at the time *"In truth, with over 100,000 refugees to place, the placing agencies are not inclined to turn down any offers"* (Hoang, 2010). Finally, U.S. officials mandated that refugees be dispersed with at most one immediate family member (Sonneborn, 2007), inevitably meaning, at least initially, that immediate families were torn apart by the dispersal policy.

Ultimately, *"tens of thousands of Vietnamese were sent off to new lives in every state in the union and nearly every city"* (Thompson, 2010). Given the circumstances of their allocation, the refugees *"had to initially depend almost entirely upon the government and individual or institutional sponsors to determine where they would resettle"* (Rumbaut and Portes, 2001).[7]. Given the primacy of the exogeneity of the initial allocation of Vietnamese in 1975, we examine the veracity of such claims in the next section.

2.2 Evidence of Exogeneity

The central claim of the paper, is that the initial networks to which individual refugees were originally assigned were quasi-random in terms of both size and quality. Certainly the chaotic dispersal of the Vietnamese, which split immediate families apart and scattered refugees across every state in the country, across no less than 76% of counties in 1975 (van der Werf, 2019), can be considered far more likely exogenous than comparable dispersal policies (Edin et al. (2003), Damm (2009) and Beaman (2011)) that instead refugees according to pre-existing guidelines such as the availability of housing or pre-existing agglomerations of co-nationals. Whereas the historical documentation detailing the events of 1975 strongly suggests an exogenous allocation however, the policy may still have been subject to unintended sorting.

Individual characteristics may have influenced initial network size and quality. Those with a better command of English for example, could potentially have influenced their initial placement. Language ability, while an important determinant of labor market outcomes (Dustmann, 1994; Dustmann and Soest, 2001; Dustmann and Fabbri, 2003), is also positively correlated with education, meaning that if resettled

[7]The natural experiment of the chaotic dispersal was first used for causal identification in Parsons and Vézina (2018), when it was combined with an enduring trade embargo to provide a causal link of trade between U.S. states and Vietnam.

into larger networks, any results would necessarily be upwardly biased. So too could VOLAGs have been able to settle refugees (of particular abilities) into more economically favorable locations despite the chaos. Since sponsors were required to pay for food, accommodation, etc. it is not unreasonable to think that they may be allocated to more prosperous locations. Commuting zone characteristics may also have induced variation in the network size and quality in 1975, which in turn could be correlated with economic outcomes.

To investigate whether individual or commuting zone characteristics influence network size and quality, we follow the logic of the existing literature (Edin et al., 2003; Damm, 2009) and regress a rich battery of individual characteristics (education, gender, marital status, family status, age), uniquely including dummies capturing the 1-digit occupation that individuals were employed in prior to their evacuation, as well as commuting zone characteristics and state fixed effects on dummy variables that capture above average network size (mean = 4,798) and network quality (mean = 9.8%).

Reassuringly, as shown in Column 1, no individual characteristics are shown to have statistically significant correlations with network size, with the exception of being employed in a miscellaneous occupation, prior to emigrating from Vietnam, a category that we exclude in one of our robustness checks. The absence of any influence of education is particularly heartening, since that variable likely correlates best with language ability. In addition, a battery of contemporary analyses all concluded that English language proficiency had no statistically significant effect on economic outcomes conditional on both refugee's education and prior occupation (Aames et al., 1977; Dunning and Greenbaum, 1982; Caplan et al., 1985; Strand and Jones, 1985; Haines, 2012).

Perhaps unsurprisingly, a number of commuting zone characteristics do play a role. Larger and richer states hosted larger networks and so too were greater numbers of Vietnamese sent to labor markets with higher rates of unemployment. This finding tallies well with the historical record, since larger states were expected to host greater numbers of refugees (Parsons and Vézina, 2018) and despite the intentions of Congress, the dispersal policy resulted in resettlements in high unemployment areas (GAO, 1977).

Repeating the exercise with network quality in Column 2 reveals that a number of individual characteristics exhibit statistically significant correlations with networks of higher quality, as measured by the share of blue collar workers, including: being employed in clerical, processing or miscellaneous work in Vietnam, being male, being married and being dependent.

Of the three occupational categories, only processing falls into our definition of blue collar work (see Section 3). Many of the first wave of arrivals initially took up blue collar jobs in the U.S. prior experience of which would no doubt have proved useful. Interestingly however, no commuting zone characteristics are significant in determining network quality. Neither those few Vietnamese that resided in the U.S. prior to the arrival of the refugees, nor the presence of existing migrant networks of other nationalities, played any role in determining the size and quality of the refugees' networks on arrival, which lends credence to our claim that indeed there were no pre-existing Vietnamese networks in 1975.

Our central claim in this paper is therefore that the initial allocations of network size and quality were both exogenous, conditional on individual controls and commuting zone of arrival and state fixed effects. Similar assumptions are made by Edin et al. (2003) and Damm (2009) regarding the exogeneity of their network size.

2.3 Refugees' Sorting After Settlement

By December 1975, the four processing centres had resettled a total of 129,792 refugees in just eight months (HEW Refugee Task Force, 1976), see Appendix Figure A.2. Initially, the *Viet Kieu* fared poorly economically. The disorientated refugees were dispersed in the middle of an economic downturn, with few recognised skills and poor English language proficiency (ORR, 1981) into scattered communities, which often vilified them (Rumbaut and Portes, 2001). Given their almost complete absence, neither could the Vietnamese draw upon pre-existing networks since only *"A few adopted war orphans, Indochinese women married by Americans, and students at American universities comprised the Indochinese in the U.S. before April 1975"* (Stein, 1979; Rumbaut and Portes, 2001; Sonneborn, 2007; Thompson, 2010).

While the authorities went to great lengths to ensure that the initial wave of *Viet Kieu* was widely dispersed, little was done to discourage the refugees' subsequent migration. Anecdotally, the refugees migrated for a variety of reasons: some to the warmer climates of the South-West, others to reconnect with lost family and friends and others still in pursuit of better employment opportunities (Robinson, 1998).[8] Some 45% of the initial wave of Indochinese had migrated by 1980 and so *"While there was a great deal of secondary migration, we should not overlook the fact that a little more than half of the refugees stayed in the states to which they had been sent. This suggests that the government's dispersal policy, for better or worse, worked with a majority of the 1975 refugees."* (Baker and North, 1984). Whereas the 1975 refugee wave has been *"scattered widely throughout the 50 states...sprinkled in fairly small groupings within those states"* (Baker and North, 1984), by July 1980, more than half the refugees were clustered in five States: California, Texas, Washington, Pennsylvania and Illinois, while two-thirds of them resided in 10 States-those listed above plus Minnesota, Oregon, New York, Louisiana, and Virginia (Marsh, 1980). Agglomerations of *Viet Kieu* therefore took root early. Those in our sample were initially allocated to 396 different commuting zones, but following the initial secondary migrations, by 1980 these individuals lived in only 78 commuting zones.

In the absence of pre-existing networks, *Viet Kieu* networks first formed from 1975 onward, being ever strengthened by subsequent waves of Vietnamese refugees (Parsons and Vézina, 2018). On arrival,

[8] With a temperate climate and generous welfare regime, many chose to move to California. California received 30,495 refugees in 1975, with the commuting zone containing Los Angeles receiving 10,421 of those, which we omit as a robustness check.

Table 1: Exogeneity of Initial Refugee Network Size and Quality

	(1) Average Network Size b/se	(2) Average Network Quality b/se
	Pre-placement Controls	
Vietnamese Occupations		
Professional	-0.085	0.117
	(0.123)	(0.204)
Clerical	-0.025	0.545**
	(0.158)	(0.249)
Service	-0.126	0.030
	(0.110)	(0.214)
Agriculture	-0.280	0.190
	(0.555)	(0.727)
Processing	-0.123	0.680**
	(0.182)	(0.329)
Machine Trade	0.032	0.490
	(0.229)	(0.344)
Benchwork	0.027	0.042
	(0.144)	(0.218)
Construction	-0.019	-0.004
	(0.130)	(0.136)
Miscellaneous	0.789*	-0.999*
	(0.434)	(0.590)
	Individual Controls	
Education		
Primary Education	0.035	-0.092
	(0.042)	(0.082)
Secondary Education	0.040	-0.109
	(0.045)	(0.079)
Tertiary Education	0.039	-0.116
	(0.042)	(0.083)
Gender		
Male	-0.010	0.044***
	(0.011)	(0.014)
Marital Status		
Divorced	-0.021	-0.060
	(0.056)	(0.130)
Married	0.001	-0.035**
	(0.011)	(0.017)
Widowed	0.035	0.170
	(0.047)	(0.114)
Family Status		
Dependent	-0.011	0.033***
	(0.007)	(0.011)
Unattached Individual	0.007	0.015
	(0.013)	(0.016)
Age		
Age	-0.001	-0.003
	(0.003)	(0.005)
Age × Age	0.000	0.000
	(0.000)	(0.000)
	Commuting Zone of Arrival Controls	
Employment Rate in 1970	-11.071***	0.793
	(3.075)	(3.004)
Average income in 1970	1.111***	-1.478
	(0.307)	(0.935)
% of Migrants in 1970	-1.460	-4.450
	(1.889)	(2.670)
% of Vietnamese in 1970	-112.550	-225.297
	(269.556)	(612.793)
Population in 1970	0.123***	0.244
	(0.033)	(0.147)
Constant	-9.607***	8.682
	(2.099)	(5.539)
State FE	Y	Y
Observations	4,522	4,522

Notes: Mean network size in 1975 is 4,524. Standard errors clustered by State and shown in parentheses.

many of the first wave found employment in blue collar jobs.[9] The *Viet Kieu* found gainful employment in electronic engineering[10] and machine assembly, restaurants, supermarkets and nail and hair salons (Galens et al., 1995); concentrations of which are today referred to as 'Little Saigons'.

The *Viet Kieu* networks therefore sorted across space both by size and ability. Whole communities formed around single occupations. In Santa Clara County, California, for example, many Vietnamese became electronic technicians with enrollments in technician classes comprising 40-90% Vietnamese, with some training schools even offering courses in Vietnamese (Finnan, 1981). Similarly, in the Houston-Brazoria metropolitan area, 0.68% of hairdressers were Vietnamese in 1980, a number that grew 6.21% by 1990 (Patel and Vella, 2013). In their analysis of the 2014 ACS, (Eckstein and Peri, 2018) highlight the role of the Vietnamese in significantly expanding and subsequently dominating the manicure sector through developing their formal and informal networks.[11] Somewhat less known are the Indochinese fishing communities that established themselves in the 1980s from the Mississippi Delta to the Florida Panhandle, where the muggy weather is similar to the climatic conditions in Vietnam. Today in the small Alabaman town of Bayou La Batre, one-third of the population is Vietnamese, after early Vietnamese pioneer migrants rescued the then ailing local crabbing industry[12]. Thus agglomerations of *Viet Kieu* therefore formed by both size and quality.

Since we ultimately control for commuting zone of arrival fixed effects in our estimation, the identifying variation exploited in our analysis pertains to those that subsequently moved commuting zone. It proves instructive therefore to compare the movers and the stayers to substantiate if any selection occurred in relation to those secondary migrations. Table 3 therefore compares the characteristics of stayers (column 2) and movers (column 3) for those in our sample. Overall, the characteristics of the two groups are strikingly similar. Those that moved proved slightly less educated,[13] fractionally more likely male, somewhat less likely to be married or a dependent and approximately two years younger. In terms of *commuting zone of arrival* characteristics, those that moved were initially allocated to slightly poorer and more populous commuting zones with marginally less unemployment and fewer migrants. Interestingly, from the perspective of our analysis, movers were first allocated to slightly smaller networks that were employed marginally more in blue collar work.

The historical accounts agree that the first wave of *Viet Kieu* initially experienced significant skill downgrading, although the estimates vary. In part, this was because many refugees did not have their foreign skills recognized. Appendix Figure A.1, drawn based on contemporary data on the entire incom-

[9] https://www.latimes.com/local/california/la-me-ff-0501-saigon-fall-reaction-20150501-story.html.

[10] Electronics work in particular was perceived by the *Viet Kieu* as professional as opposed to blue collar work (Finnan, 1981), which in turn fostered a more general orientation toward more technical occupations (Zhou and Bankston III, 1994).

[11] In the year 2000, some 13% of all Vietnamese in the U.S. were employed in the 'hairdresser and other grooming services' sector, which represented 34% of the total foreign-born in the sector.

[12] See for example: https://www.nytimes.com/1991/02/24/us/asians-spread-across-a-land-and-help-change-it.html

[13] This finding dovetails with a contemporary study by McInnis (1981) that found that education was uncorrelated with secondary migration.

ing population of *Viet Kieu* in 1975, shows that 46% of those comprising the initial refugee wave had white-collar work in Vietnam, whereas only 21% had managed to find white-collar work seven months after the Fall of Saigon. By 1977, 75% of household heads reporting occupational downgrading (GAO, 1977). Marsh (1980) later writes that *"In November-December 1978, more than 3 years after the original group of Indochinese refugees were resettled, about 56 percent of the former white-collar household heads were craftsmen, transportation workers, farm managers, laborers, or engaged in miscellaneous blue collar pursuits."* This is not uncommon as most migrants and refugees find themselves undergoing some level of occupational downgrading upon arrival (Cortes, 2004).

Table 2 is rather constructed for our sample from our administrative data and shows the distribution of occupations in Vietnam for our sample in 1975, prior to emigrating and those in the U.S. in 1981, six years hence.[14] While 43 percent of the sample experienced some occupational downgrading, 33% percent rather upskilled, while the remainder stayed in a comparable occupation in the U.S. This tallies with anecdotal evidence. (GAO, 1977) for example states that *"We also found some refugees with jobs comparable to those they held in Indochina. One man, who speaks English fluently, was the chief of a beverage distributor center in Vietnam. Currently, he has been doing similar work as a sales manager in a manufacturing company"* (GAO, 1977). But English language was not the sole determinant, since the same report to Congress found that *"Another refugee, who did not speak English, was able to continue his job as a camera repair technician in the United States"* (GAO, 1977).

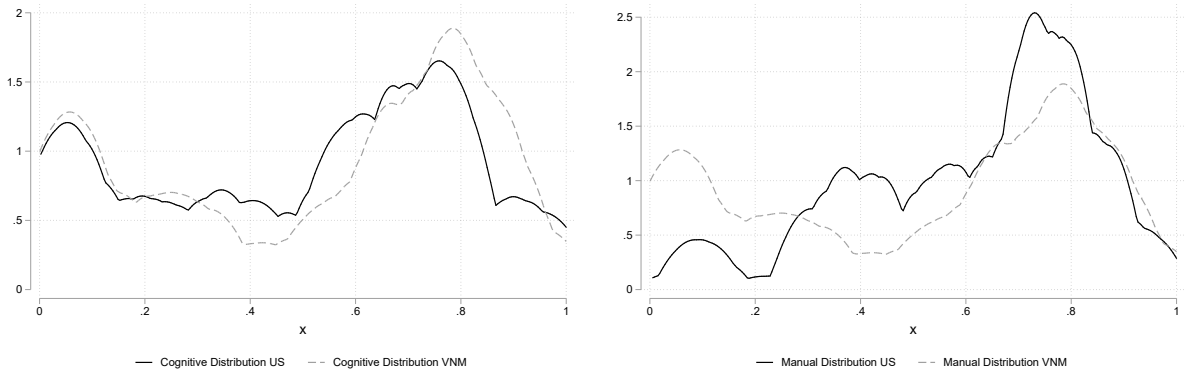
Table 2: Distribution of Occupations in Vietnam before Arrival in 1975 and in the U.S. in 1981

	Occupation in Vietnam %	Occupation in the U.S.%
Miscellaneous	13%	3%
Structural and Construction	6%	21%
Benchwork	14%	14%
Machine Trades	9%	16%
Processing	1%	3%
Agricultural and Fishing	3%	1%
Service	19%	15%
Clerical and Sales	17%	15%
Professional and Managerial	20%	13%
Total	100%	100%

Alternatively, we can highlight the differences in skills employed between Vietnam and the U.S., by examining which skills are required to undertake specific occupations, as detailed in O*NET, the details of which are provided in the following section. To that end, Figures 1a and 1b show the changes in intensities along the continuum either cognitive or manual skills, respectively. Figure 1a shows that those in the middle of the distribution of cognitive skills in Vietnam upgraded along that dimension on

[14] Tables A3 and A4 in the Appendix are the transition matrices for occupations and task complexity respectively.

Figure 1: Distributions of Skill Intensities: 1975 and 1980



(a) Panel A: Changes in Cognitive Intensities

(b) Panel B: Changes in Manual Intensities

arrival in the U.S., while those toward the upper end of the distribution of cognitive skills in Vietnam downgraded having migrated to the U.S. In contrast, those in the middle and at the higher end of the distribution of manual skills in Vietnam upgraded along the manual dimension, while those at the lower end significantly downgraded along the manual dimension.

3 Data

The Office of Refugee Resettlement (ORR) is an office within the Department of Health and Human Services in the U.S. government. It records personal information on every refugee that arrives to the U.S. (day of arrival, age, gender, marital status, family status, education and occupation prior to migration) and the city, county and state a refugee is settled in. Our ORR archive data, deriving from digitized files obtained from the National Archives, are also matched to the Immigration and Naturalization Service's (INS) Alien Registration Program data, which comprise four follow-up samples spanning 1978 to 1981. As detailed in an ORR report to Congress, data were collected in: 1976, 1978, 1979, 1980 and 1981 (ORR, 1981), while the data in our sample pertain to follow-up surveys from 1978 to 1981, which was verified by an examination of our year of arrival variable. The INS, under the auspices of the Alien Registration program, mandated refugees to provide employment and location details every January, with a question pertaining to occupation fielded in 1981.[15] January 1981 was the final year in which these data were collated, since the requirement for annual reporting was abolished with the enactment of the Immigration and Nationality Act of 1981, the so-called INS Efficiency Act. A geo-referenced version of our ORR data set comprising the micro-data of the universe of refugees that arrived into the U.S. between 1975 and 2008, are available from the authors on request.

[15] If an occupation was listed the refugee was regarded as working (Baker and North, 1984).

Table 3: Comparison of Movers' and Stayers' Characteristics

	(1) Stayers	(2) Movers
	Vietnamese Occupations	
Blue Collar Job	0.29	0.30
	(0.45)	(0.46)
Professional Job	0.20	0.19
	(0.40)	(0.40)
	Individual Characteristics	
Secondary Education	0.55	0.52
	(0.50)	(0.50)
Male	0.58	0.62
	(0.49)	(0.49)
Married	0.61	0.52
	(0.49)	(0.50)
Dependent	0.50	0.48
	(0.50)	(0.50)
Age	37.35	35.85
	(9.72)	(9.21)
	Commuting Zone of Arrival Characteristics	
Unemployment Rate in 1970	0.22	0.021
	(0.007)	(0.006)
Average income in 1970	3,456	3,417
	(409.89)	(392.31)
% of Migrants in 1970	0.72	0.68
	(0.37)	(0.37)
% of Vietnamese in 1970	0.00005	0.00004
	(0.37)	(0.037)
Population in 1970	4,502,871	4,199,861
	(4,084,677)	(3,914,427)
Ethnic Share	0.001	0.001
	(0.000)	(0.000)
	Network Measures	
1980 Network Size	20,438	20,122
	(19,916)	(18,428)
1975 Network Size	4,798	4,336
	(4,219)	(4,135)
1980 Blue Share	0.177	0.186
	(0.070)	(0.059)
1975 Blue Share	0.098	0.099
	(0.023)	(0.019)
State FE	Y	Y
Observations	1,837	2,685

Notes: Standard deviations in parentheses.

Commuting zones as opposed to counties are employed as our geographical unit of analysis as they better capture local labor market dynamics, in that residents both live and work within the same commuting zones. County borders are often based on the political or geographical aspects of an area, leading to labor markets spanning county borders. Commuting zones, instead, use commute-to-work data to capture both individuals place of residence and work. Designed by Tolbert et al. (1987) in 1987, this paper uses the 741 commuting zones in the 1996 update (Tolbert and Sizer, 1996). As a result, individuals must be mapped to a commuting zone. Counties are not always entirely contained within a commuting zone and so we rely on the crosswalk file provided by Autor and Dorn (2013). We restrict our sample to those of working age i.e. 18 to 64.

Throughout this paper, we refer to two distinct commuting zones. *Commuting zones of arrival* refer to those commuting zones to which refugees were quasi-randomly assigned to on arrival to the U.S. in 1975. *Commuting zones of destination* rather refer to those commuting zones in which individuals reside in 1981, which may or may not be the same as their commuting zone of arrival in 1975, depending upon whether secondary migration took place.

Whereas Edin et al. (2003) define an 'ethnic enclave' as *"a municipality where ethnic concentration (the size of the ethnic group relative to the population in each municipality) was at least twice as large as the share of the ethnic group in the entire population"*, this definition is inapplicable in our case due to the absence of pre-existing Vietnamese communities. In other words, our study focuses on the initial role of migrant networks in fostering individuals' entry into local labor markets, as opposed to the impact of pre-existing (and therefore necessarily pre-sorted) ethnic enclaves. We therefore adopt a definition of a 'migrant network' as a *commuting zone* in which the concentration of Vietnamese is at least as large as the share of Vietnamese in the entire population, although this choice is subject to a battery of robustness checks.

Our measure of network quality is based upon blue collar work given the historical role blue collar work played for the *Viet Kieu*. Our occupation data recorded by the ORR importantly record both refugees' occupations in Vietnam and the U.S., both according to U.S. occupational nomenclatures, which pertain to the Dictionary of Occupational Titles (DOT) occupation classification system. DOT comprises 83 occupations grouped into nine categories. The broad occupations classified as blue-collar occupations following ORR (1981); Helwig (2001) include DOT codes: *Processing Occupations (5)*, *Machine Trade Occupations (6)*, *Benchwork Occupations (7)* and *Structural Work Occupations (8)*. We construct a dummy variable equal to 1 if an individual is engaged in blue-collar work and 0 if they are otherwise employed. Since our 'Miscellaneous' occupational category is ambiguous, we subject those that fall within this category to one of our robustness checks.

To measure skills yet further, we use the productive task components of each occupation following Ottaviano et al. (2013) and Foged and Peri (2016). We use two primary skill dimensions: manual (defined as the intensity of eye-hand-foot use) and cognitive (mathematical-analytical tasks), as well as an

occupational complexity score following Ottaviano et al. (2013), which we define as cognitive intensity divided by manual intensity. The skill intensity measure is a weighted average of O*NET ability data. O*NET superseded DOT in 1991, as a database on occupational tasks, skills and requirements, using a more precise collection methodology, one based on surveys, as opposed to relying on manual collection by job analysts (Handel, 2016). The abilities data record the importance of each ability on a scale of 1-5 for 52 different abilities for over 12,000 occupations. We construct our *manual* variable from 19 different physical skill measures such as strength, dexterity and reaction time. Our *cognitive* variable rather comprises 16 skills including: reasoning, perception and numeracy. For a full list of categorizations please refer to the Appendix. For each DOT code, a weighted average is constructed due to O*NET containing more occupations than DOT. An equal weighting to every O*NET classification regardless of the frequency they may occur within the DOT code may create measurement error. We use the 1980 census sample to determine the occupational frequency.[16] The weighted average of each score is reported in the appendix. Each score is standardized to a range of 0-1, as the scale of O*NET is arbitrary.

We use the 1970 Integrated Public Use Microdata Series (IPUMS) 5% census samples from Ruggles et al. (2017) to construct commuting zone characteristics in 1970 before the arrival of the refugees and the 1980 census sample to construct our 1980 network variables. For our 1980 measures of network size and quality variables, we identify Vietnamese individuals from their birthplace variable, while reweighting the census person weight for commuting zones. Network size is measured as the log of the total number of Vietnamese in a *commuting zone of destination* in 1980. Our measure of network quality is simply the share of blue collar work in a particular *commuting zone of destination* in 1980, although we also interact this with our measure of network size for the sake of robustness. We crosswalk the census occupation classification to DOT, to arrive at the number of blue-collar workers in each commuting zone. Our comparable instruments for network size and quality are rather constructed for *commuting zones of destination* using our 1975 ORR data. These are defined as the size of the initial network allocated to a specific *commuting zone of destination*, else the initial share of blue collar workers in the *commuting zone of destination*, as measured by the *Viet Kieu's* pre-placement occupations in Vietnam. Summary statistics are provided in Table 4.

4 Empirical Strategy

4.1 Identification Issues

In order to estimate the causal effects of social networks' size and quality on occupational choice, skill intensity and skill upgrading in 1981, we need to address the sorting of our 1980 refugee network measures both across space i.e. network size, as well as by ability i.e. network composition. To that end,

[16] This is an imperfect measure since the census also comprises fewer occupations than O*NET. It does, however, create a more precise measurement of the skill intensity of each DOT occupation.

Table 4: Summary Statistics

Variable	Observations	Mean	Std. Dev.	Min	Max
Outcome Measures					
Blue Occupation (U.S.)	4,522	0.537	0.499	0	1
Professional Occupation (U.S.)	4,522	0.127	0.333	0	1
Blue Occupation (VNM)	4,522	0.297	0.457	0	1
Professional Occupation (VNM)	4,522	0.196	0.397	0	1
Cognitive Intensity (U.S.)	4,522	0.505	0.307	0	1
Manual Intensity (U.S.)	4,522	0.615	0.227	0.005	1
Complexity Intensity (U.S.)	4,522	0.437	0.242	0.001	1
Cognitive Intensity (VNM)	4,522	0.524	0.320	0	1
Manual Intensity (VNM)	4,522	0.565	0.263	0.005	1
Complexity Intensity (VNM)	4,522	0.477	0.270	0.001	1
Cognitive Upgrade	4,522	0.406	0.491	0	1
Cognitive Downgrade	4,522	0.489	0.500	0	1
Manual Upgrade	4,522	0.490	0.500	0	1
Manual Downgrade	4,522	0.405	0.491	0	1
Complex Upgrade	4,522	0.400	0.490	0	1
Complex Downgrade	4,522	0.494	0.500	0	1
Network Measures					
Network Size 1980	4,522	20,250	19,045	5	47,118
Network Size 1975	4,522	4,524	4,175	15	10,421
Blue Share 1980	4,522	0.182	0.064	0	0.6
Blue Share 1975	4,522	0.099	0.021	0	0.201
CZone Arrival Characteristics					
Employment Rate 1970	4,522	0.050	0.15	0.022	0.010
Average Income 1970	4,522	3,433	400	1,881	4,077
Migrant % 1970	4,522	0.069	0.037	0.007	0.119
Vietnamese % 1970	4,522	0.000	0.000	0	0.001
Individual Characteristics					
Male	4,522	0.603	0.489	0	1
Age	4,522	36.455	9.445	19	64
Secondary Education	4,522	0.532	0.499	0	1
Married	4,522	0.558	0.497	0	1
Dependent	4,522	0.490	0.500	0	1

we construct two instruments, which capture the initial size and quality of the quasi-randomly allocated initial 1975 *Viet Kieu* networks, crucially in those commuting zones in which the individual's in our sample were located in 1981 i.e. the *commuting zones of destination*. Specifically, our instrument for network size is the log of the total 1975 network size in the *commuting zones of destination*; while our instrument for network quality is measured as the share of those initially placed in the *commuting zone of destination* employed in blue collar work *in Vietnam*, i.e. before the refugee dispersal. The richness of our data therefore facilitates the creation of exogenous measures of network quality, as opposed to having to interact (a potentially mis-measured) network size variable with an endogenous measure of network quality as per Edin et al. (2003) and Damm (2009). Unlike previous studies that rely on other examples of dispersal policies to estimate the impact of networks, our context is unique as we do not have a pre-existing network before the arrival of *Viet Kieu* in 1975. The dispersal policy that we leverage is likely far more exogenous therefore when compared to comparable policies. So too are we are able to employ rich longitudinal data to estimate the causal effects of network quality.

The specification of our first stages is:

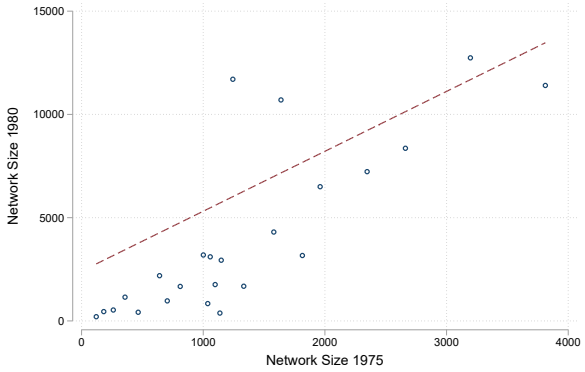
$$network_{1980d} = \beta_0 + \beta_1 network_{1975d} + \boldsymbol{\beta} X_i + \delta_a + \varepsilon_{ida} \tag{1}$$

where d is the commuting zone where an individual i resides in 1981 i.e. *commuting zone of destination* and a is an individual's commuting zone in 1975 i.e. *commuting zone of arrival*. X_i comprises a full battery of individual controls (dummies capturing one-digit occupations in Vietnam prior to emigration, highest educational attainment in Vietnam, gender, marital status, family status, age and age squared). *Commuting zone of arrival* fixed effects are included to control for the endogeneity that arises from individual's labor market opportunities on arrival *a la* Edin et al. (2003). We avoid the collinearity of our instrumental variable and our *commuting zone of arrival* fixed effect, since our instrumental variable is the quasi-randomly allocated network of an individual's *commuting zone of destination* i.e. the network allocated to the commuting zone where the *Viet Kieu* resided in 1981. Given that just under half of the *Viet Kieu* emigrated, it is this secondary migration that occurred between 1975 and 1981 that creates our identifying variation, even in the presence of our *commuting zone of arrival* fixed effects.

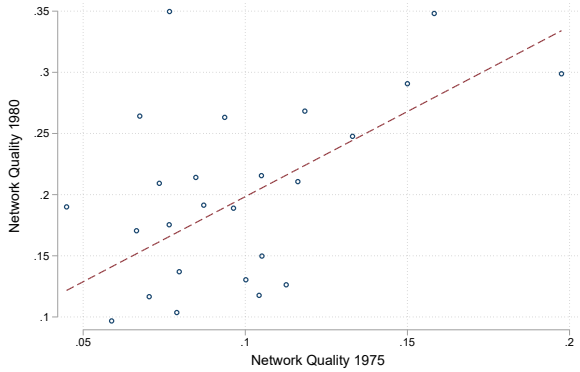
Consider the following example pertaining to two refugees, both of whom arrive in the *commuting zone of arrival* that includes Miami-Dade county. One subsequently migrates to Orange County, while the other remains in Miami-Dade. Crucially, both refugees in our example are associated with different instrumental variables, since our instruments capture those refugees quasi-randomly allocated to Miami-Dade's and Orange's commuting zones in 1975 and it is this identifying variation that allows us to concurrently implement *commuting zone of arrival* fixed effects.

For our results to have a causal interpretation, the instrument must be conditionally random, satisfy the exclusion restriction and be sufficiently predictive of the endogenous variable (Angrist and Imbens, 1995). The first condition was discussed in Section 2. We also assume that assigned commuting zones do

Figure 2: Binned Scatter Plots of IV and Endogenous Variables



(a) Panel A: Correlation network size



(b) Panel B: Correlation network quality

not have permanent effects on outcomes i.e. commuting zone effects are local to that time and place in order to satisfy the exclusion restriction. This is reasonable as commuting zones are designed to capture local labor market dynamics and therefore the effects on outcomes are specific to that commuting zone and time period. The correlations between our endogenous variables and instruments are shown in Figures 2a and 2b. Network size and quality in 1975 must also sufficiently predict network variables in 1980. As 41% of refugees stayed in their commuting zone of assignment, the network size in 1975 should have at least some power to predict network size five years later. As the refugees experienced significant occupational downgrading (Stein, 1979), our network quality measures may not be sufficiently strong however. In the 1981 sample though, 23% of refugees that arrived in 1975 are recorded as having the same occupation as they did in Vietnam.

Our results can be interpreted as the average treatment effect, if network effects are homogeneous across the population (Angrist et al., 1996). This is unlikely to be the case as the returns to networks vary with the skills of members (Borjas, 1998; Edin et al., 2003). Instead, a useful heuristic would be to interpret the results as a weighted average of individual returns among those who were influenced by the 'treatment'.

Estimating the causal impacts of networks raises several further empirical challenges. As first highlighted by Manski (1993), there are three potential threats to the identification of what undergirds the relationship between the outcomes of individuals' that interact with one another. First, the behaviour of an individual varies with the behaviour of their network, although that behaviour is endogenous. Secondly, the behaviour of an individual varies exogenously with the behaviour of the network, the so-called contextual effect. Thirdly, individuals in the same network tend to behave similarly since face similar environments or common shock. Disentangling the contextual and endogenous effects in particular, what is otherwise known as the Reflection Problem constitutes a major empirical challenge. Does net-

work behavior affect individual behavior, else is network behaviour simply the aggregation or reflection of individual behaviours? Several excellent surveys on the topic, also proffer alternative identification strategies see for example: Blume et al. (2011) and Jackson et al. (2017).

In our case, we deal with the correlated effects due to common local shocks or local omitted variables through the implementation of our *commuting zone of arrival* (in 1975) fixed effects. Individual's occupations in Vietnam, as captured by binary variables at the one-digit occupation level are included as regressors in order to account for any omitted personal characteristics, something that previous studies have been unable to do.

To address the potential Reflection Problem, we rely on the dynamic structure of the network, meaning that we employ lagged network measures (see for example Brock and Durlauf (2001)) in tandem with our measures of network size and quality in 1980, while our outcome variable rather refers to 1981. Our network measures derive from the 1980 census, while our individual level occupation variables derive from our ORR data. Since our pre-dispersal measures of network quality are necessarily exogenous to our 1981 labor market outcomes, our two-sample-two-stage-least squares approach isolates the exogenous variation in our network measures so as to examine its effect on the individual occupations in 1981.

As a further robustness check however, and so as to categorically insulate our results from any remaining endogeneity concerns pertaining to the Reflection Problem, we undertake two additional exercises. First, we omit the individual from our instruments, although we are unable to exclude the individual from our network measure since we rely on two alternative data sources. As such, we instead rely on the precise geo-locations of our refugees in 1981 from our INS data, which in turn allow us to exclude differing proportions of the *commuting zone of destination* network based on buffers of particular radii, 10km, 25km etc. The 3,218 individuals in our sample for which our administrative data provide sufficient detail to precisely geolocate them are captured as residing in some 429 U.S. urban agglomerations. In effect this exercise is a more extreme version of simply excluding the individual for our network measure, see for example Topa and Zenou (2015).

4.2 Occupational Outcomes

To identify the effects of networks on refugees' occupational choices we estimate several linear models using 2SLS as follows:

$$occupation_{1981idj} = \beta_0 + \beta_1 \widehat{networksize}_d + \beta_2 \widehat{networkoccupation}_d + \boldsymbol{\beta} X_i + \delta_a + \varepsilon_{id} \quad (2)$$

where occupation is the occupation of individual i in 1981. j=1 for a blue-collar job, while j=2 for a professional job. *d* refers to the *commuting zone of destination* i.e. the commuting zone where individual *i* resides in 1981, as opposed to the *commuting zone of arrival*, the commuting zone individual *i* was

assigned to in 1975, which is instead denoted by the subscript *a*. *Commuting zone of arrival* fixed effects are used to control for labor market opportunities on arrival.

Network size in the *commuting zone of destination d* in 1980 is instrumented using the number of assigned Vietnamese to the *commuting zone of destination in 1975*, when it was a *commuting zone of arrival*. Network quality in the *commuting zone of destination d* in 1980 is similarly instrumented using the shares of blue-collar else professional networks in the commuting zone of destination *in 1975*. X is a set of controls for observed individual characteristics including: age, age squared, education, gender, martial status and uniquely pre-assignment occupations in Vietnam at the one-digit level. All errors are clustered at the *commuting zones of arrival* and then the *commuting zones of destination* for robustness.

4.3 Skill Intensity Outcomes

Having examined the role of network size and quality on binary occupation outcomes, we continue by asking if network size and quality deferentially affect the position of individual's occupations within the distributions of our manual, cognitive and complex skills, for which we run the following specification:

$$skill_{US1981idk} = \beta_0 + \beta_1 \widehat{networksize}_d + \beta_2 \widehat{networkoccupation}_d + \boldsymbol{\beta} X_i + \delta_a + \varepsilon_{id} \qquad (3)$$

where the dependent variable refers to the skill intensity (which can loosely be interpreted as a percentile, noting above that we normalise the O*NET data to be on a scale of 0-1) of the refugee's occupation in the U.S. in 1981. k=1 for cognitive skills, k=2 for manual skills and k=3 for complex skills. We control for individual characteristics as before and for individual skill intensity of occupation in Vietnam. We also instrument for network size and network quality as described above.

4.4 Changes in Skill Intensities

We continue further by examining the difference in skill intensities between those occupations performed in Vietnam and the U.S., which proves possible since refugees' occupations in Vietnam were recorded on arrival according to comparable American occupational nomenclatures. To this end, we estimate the following equation using 2SLS:

$$skill_{US1981ik} - skill_{Vietik} = \beta_0 + \beta_1 \widehat{networksize}_d + \beta_2 \widehat{networkoccupation}_d + \boldsymbol{\beta} X_i + \delta_a + \varepsilon_{id} \quad (4)$$

where the dependent variable is the difference in the skill intensity of the refugee's occupation in Vietnam and the skill intensity of occupation in the U.S. in 1981. We examine upgrading and downgrading along the three (k) occupational dimensions of skill namely: manual, cognitive and complex tasks.

In addition, as before, we control for individual characteristics and for individual initial skill intensity of occupation in Vietnam. Network size and network quality are instrumented as before.

5 Results

Tables 5 to 9 present our estimates of the effect of network size and quality on binary occupational outcomes (blue and professional) and our skill intensity (cognitive, manual and complex) measures that are continuous over the [0,1] interval. In each table, columns (1) and (4) include our measure of network size alone, while columns (2) and (5) include our network quality measure, namely the share of blue collar workers in a *commuting zone of destination*, while columns (3) and (6) rather include the interaction of our network size and quality measures. Given the preponderance of *Viet Kieu* that arrived into the LA commuting zone in 1975, we omit all those initially placed in LA from our regressions in columns (4)-(6), for the sake of robustness which reduces our sample size from 4,522 to 3,777. Similarly we cluster our standard errors to ensure that our results are robust to clustering at both the commuting zones of arrival and destination, as presented in square and curly brackets respectively. In columns (2), (3), (5) and (6) for which we include two instruments, the rule-of-thumb mandating that our first-stage F-statistics be above 10 is no longer valid, since instead one needs refer to the critical values provided by (Stock and Yogo, 2002). Our instruments prove sufficiently strong across all specifications used in this paper.

5.1 Network Size and Quality and Binary Occupational Outcomes

As shown in Table 5, the effect of network size on blue collar work alone is close to zero. Once we account for network quality as measured by the share of blue collar workers however, the estimates show that both larger networks as well as higher shares of blue collar workers have a positive effect of the probability of obtaining blue collar work. The estimated effects of network quality however, are quantitatively far larger, with our sample regression coefficients suggesting that a one unit increase in the share of blue collar workers leads to a 1.8% increase in the probability of being engaged in a blue-collar job. This is equivalent to a 3 percentage point increase in the mean (54%) of being engaged in blue-collar work. Using our alternative quality measure, a 10% increase in the blue collar network, equivalent to approximately 400 persons, is associated with a 2.4% increase in the probability of being employed in a blue-collar job. Measuring network quality as the interaction of the share of blue collar workers and network size, our estimates suggest a positive role for network quality and conversely a negative role for network size on blue collar work, though the overall impact is still positive.

In terms of professional occupational outcomes, Table 6, shows that for network quality we find that the converse is true, namely that larger blue-collar networks are associated with somewhat lower probabilities of obtaining professional employment, although the magnitudes of this disincentive effect are notably smaller than the estimated positive effects on blue-collar occupational outcomes. We also

find that network size has a positive impact on professional work, though this effect is much smaller than the negative impact of network quality.

Tables A19 to A23 in the Appendix, provide the corresponding OLS results to those presented in this section, a comparison of which is informative with regards the overall presence of endogeneity. Focusing on our measure of network quality, since in the current paper, that is largely where the action takes place, our OLS estimates are consistently downward biased. In the case of our blue-collar outcome measure, this bias is equivalent to 1.806-0.944=0.862 or over 90% of the estimated biased coefficient. It is important to note the degree to which our results change when we are able to meaningfully control for the *additional* endogeneity arising from the sorting of networks i.e. network composition or quality, which our results serve to highlight can otherwise lead to falsely assigning signs, significance and magnitudes to any estimated coefficient. Similar patterns are observed for example by Edin et al. (2003) who find that there is a downward bias in estimates that do not account for network sorting and that the main effect of network size is much smaller once network quality is controlled for. Previous studies have also shown that immigrants may actually lose from being members of low quality networks, not withholding the limitations of the network quality measures and the unique circumstances of the chaotic dispersal of the *Viet Kieu* that we focus upon.

5.2 Network Size and Quality and Skill Intensity Outcomes

Turning to our continuous skill intensity outcome measures, we find small but non-negligible effects of network size on cognitive, manual and complex occupational outcomes: positive for cognitive work, negative for manual work and so overall positive for complex work. Given the relative magnitudes of our estimated coefficients however, network quality again proves pivotal, since although our measure plays no significant role in determining individual's cognitive outcomes, they foster individuals being able to find more manual and less complex intensive occupations. Our estimates show that a one unit increase in the share of blue-collar network increases the probability of manual work by 0.8 which is equivalent to 1.3 percentage point increase in the mean probability of being engaged in manual work (=61.5%). As was the case previously, our estimated 2SLS coefficients are approximately twice as large as the comparable OLS coefficients, which again highlights the role of endogeneity arising from network quality, see Tables A21-A23 in the Appendix. These results underscore the role played by network quality in shaping occupational outcomes (blue versus professional) and skills.

5.3 Network Size and Quality and Skill Up/Downgrading

Another important outcome that could be shaped by networks is occupational mobility. We study the impact of network size and quality on several occupational mobility using the three skill intensity measures above but distinguish between upgrading and downgrading for each of those skill measures by

Table 5: Blue Collar Work and Networks

	1	2	3	4	5	6
VARIABLES	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
log network	-0.003	0.023	-0.016	-0.002	0.027	-0.013
	[0.007]	[0.009]**	[0.008]**	[0.008]	[0.010]***	[0.008]
	(0.009)	(0.009)**	(0.008)*	(0.010)	(0.009)***	(0.008)
blue share		1.806			1.832	
		[0.390]***			[0.406]***	
		(0.431)***			(0.444)***	
network*share			0.241			0.244
			[0.040]***			[0.042]***
			(0.043)***			(0.044)***
Vietnam Controls	Y	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y	Y
Observations	4,522	4,522	4,522	3,777	3,777	3,777
R-squared	[0.071]	[0.073]	[0.078]	[0.07]	[0.072]	[0.078]
	(0.07)	(0.072)	(0.078)	(0.07)	(0.072))	(0.078)
KP F-stat	[1707.6]	[18.3]	[32.8]	[256.5]	[7.8]	[10.03]
	(2080.3)	(16.5)	(29.3)	(290.2)	(7.5)	(9.7)

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 6: Professional Work and Networks

	1	2	3	4	5	6
VARIABLES	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
log network	0.007	0.002	0.009	0.006	0.00	0.008
	[0.004]*	[0.005]	[0.004]**	[0.004]	[0.006]	[0.004]*
	(0.004)*	(0.006)	(0.003)***	(0.003)*	(0.005)	(0.003)***
blue share		-0.339			-0.349	
		[0.220]			[0.232]	
		(0.190)*			(0.187)*	
network*share			-0.045			-0.046
			[0.026]*			[0.028]*
			(0.023)**			(0.022)**
Vietnam Controls	Y	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y	Y
Observations	4,522	4,522	4,522	3,777	3,777	3,777
R-squared	[0.037]	[0.038]	[0.0738]	[0.037]	[0.038]	[0.038]
	(0.035)	(0.037)	(0.037)	(0.035)	(0.037)	(0.037)
KP F-stat	[1707.6]	[18.3]	[32.8]	[256.5]	[7.8]	[10.03]
	(2080.3)	(16.5)	(29.3)	(290.2)	(7.5)	(9.7)

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 7: Cognitive Work and Networks

	1	2	3	4	5	6
VARIABLES	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
log network	0.017	0.002	0.017	0.019	0.018	0.018
	[0.003]***	[0.005]***	[0.004]***	[0.004]***	[0.005]***	[0.005]***
	(0.006)***	(0.008)**	(0.006)***	(0.005)***	(0.007)***	(0.005)***
blue share		-0.029			-0.014	
		[0.287]			[0.300]	
		(0.283)			(0.281)	
network*share			-0.006			0.010
			[0.036]			[0.037]
			(0.034)			(0.033)
Vietnam Controls	Y	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y	Y
Observations	4,522	4,522	4,522	3,777	3,777	3,777
R-squared	[0.154]	[0.154]	[0.154]	[0.154]	[0.154]	[0.154]
	(0.159)	(0.159)	(0.159)	(0.159)	(0.159)	(0.159)
KP F-stat	[1788]	[18.1]	[36]	[277.5]	[7.8]	[10.5]
	(2361.5)	(16.4)	(32.4)	(302.5)	(7.3)	(9.9)

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 8: Manual Work and Networks

	1	2	3	4	5	6
VARIABLES	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
log network	-0.009	-0.002	-0.017	-0.008	-0.000	-0.016
	[0.003]**	[0.004]	[0.005]***	[0.004]**	[0.004]	[0.004]***
	(0.003)***	(0.004)	(0.004)***	(0.003)**	(0.004)	(0.004)***
blue share		0.800			0.752	
		[0.228]***			[0.222]***	
		(0.279)***			(0.274)***	
network*share			0.103			0.097
			[0.024]***			[0.023]***
			(0.026)***			(0.026)***
Vietnam Controls	Y	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y	Y
Observations	4,522	4,522	4,522	3,777	3,777	3,777
R-squared	[0.115]	[0.109]	[0.114]	[0.115]	[0.109]	[0.114]
	(0.113)	(0.108)	(0.113)	(0.113)	(0.108)	(0.113)
KP F-stat	[992.5]	[17]	[33.1]	[134.3]	[7.6]	[10.7]
	(1118.4)	(15.8)	(30.3)	(144.8)	(7.4)	(10.5)

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table 9: Complex Work and Networks

	1	2	3	4	5	6
VARIABLES	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
log network	0.012	0.007	0.017	0.001	0.007	0.016
	[0.003]***	[0.004]**	[0.004***	[0.004]***	[0.004]*	[0.004]***
	(0.003)***	(0.003)**	(0.004)***	(0.002)**	(0.003)**	(0.003)***
blue share		-0.444			-0.417	
		[0.186]**			[0.189]**	
		(0.187)**			(0.184)**	
network*share			-0.061			-0.056
			[0.021]***			[0.022]***
			(0.019)***			(0.018)***
Vietnam Controls	Y	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y	Y
Observations	4,522	4,522	4,522	3,777	3,777	3,777
R-squared	[0.101]	[0.103]	[0.104]	[0.101]	[0.103]	[0.104]
	(0.098)	(0.102)	(0.102)	(0.098)	(0.102)	(0.102)
KP F-stat	[1167.3]	[19.7]	[39.1]	[159.5]	[9]	[12.1]
	(1401.4)	(18.2)	(35.4)	(178.5)	(8.5)	(11.7)

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

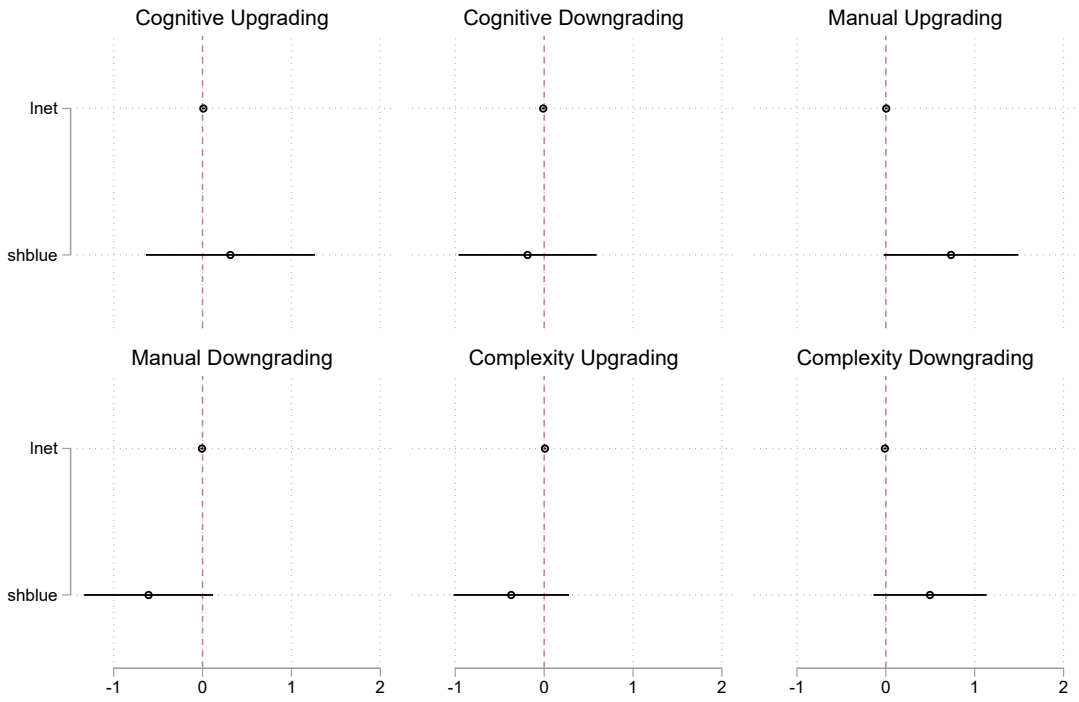


Figure 3

comparing skills in Vietnam relative to skill in the U.S. Figure 3 and Tables A13-A18 in the Appendix show that network quality has positive significant effects on manual work upgrading as well as negative impact on manual downgrading. On the other hand, blue collar networks do not have any significant impact on cognitive and complex upgrading or downgrading. Examining the effects of network quality on manual work upgrading, the estimates show that a one unit increase in the share of blue-collar network increases the probability of manual work upgrading by 0.7, which is equivalent to 1.4 percentage point increase in the mean probability of upgrading of manual work. On the other hand, a one unit increase in the share of blue-collar network increases the probability of downgrading from manual work by 0.06, which is equivalent to 0.15 percentage point increase in the mean. Overall, the results suggest that a larger shares or greater numbers of blue collar workers serve to facilitate occupational mobility but only within manual jobs, thereby potentially locking individuals into manual work.

5.4 Additional Robustness

To further examine the validity of our results, we conduct a raft of robustness checks, specifically in relation to further insulating our results from concerns of the Reflection Problem, assessing the validity of our definition of networks and finally by excluding those employed in miscellaneous occupations.

Disentangling the contextual and endogenous effects or what is typically referred to as the Reflection Problem represents a significant challenge to empirical analyses of networks. Establishing that individuals' network behavior affects individual behavior, as opposed to network behavior simply reflecting the aggregation or reflection of individual behaviors is at the heart of the issue. We control for the possibility of correlated effects due to common local shocks else local omitted variables related to networks, through the implementation of our *commuting zone of arrival* (in 1975) fixed effects. We also control for occupations in Vietnam to deal with any omitted personal characteristics. To address the potential Reflection Problem however, we rely on the dynamic structure of the network i.e. lagged network measures; measuring network size and quality in 1980 and outcomes in 1981. We also rely on two-sample-two-stage-least squares as our network measures are based on the 1980 census, while our individual level occupation variables derive instead from the ORR. As a further robustness check therefore, we exclude differing proportions of the *commuting zone of destination* network based on buffers of particular radii of 10km, 25km and 50km, which are placed over the precise geo-locations of our refugees in 1981 as detailed in our INS data and which on average remove 1.5%, 8.5% else 26.5% of our 1980 network measures. The 3,218 individuals in our administrative data for which this proves possible, are captured as residing in some 429 distinct locations. This exercise represents a more extreme version of simply excluding the individual from our 1980 network measures, in which we instead exclude all people within those radii from our sample to reduce any potential reflection concerns. As seen in Tables A10- A12, all our previous results are robust to these exclusions. The quality of networks (blue collar share) has a positive significant effect on blue collar work and manual work even after reducing network coverage by

subtracting buffers of various radii.

Aware that we somewhat arbitrarily defined a 'network' as a Vietnamese concentration equal to the average across the U.S. population as a whole, we further subject this choice to a raft of robustness checks. To this end, we implement various thresholds ranging from 50% to 150%. Tables A5- A8 show that all our results are robust and indeed qualitatively similar to our previous estimates. Finally, we exclude those employed in miscellaneous occupations from our regressions, since it is unclear which individuals comprise this group. Table A9 shows that all our previous results hold following their omission.

6 Conclusion

In this paper, we seek to provide the cleanest estimates of network size and quality on refugees' occupational choices and skill upgrading, in recognition that the endogenous location decision of migrants in general, and refugees in particular, is determined by the economic orientation of any existing network, as well as the size of that network. To this end, we exploit the natural experiment of the dispersal of the first wave of overseas Vietnamese the *Viet Kieu* in 1975, which not only constitutes the largest refugee wave in U.S. history but also the most exogenously allocated. Unlike comparable dispersal policies that have since proliferated across the globe, which typically allocate refugee adhering to strict guidelines, the *Viet Kieu* were chaotically allocated, with individuals being placed with at most one close family member, meaning that initially *Viet Kieu* families were torn apart. The end result was that the *Viet Kieu* were initially placed with networks of random size and quality.

Importantly, these events occurred in the absence of pre-existing and therefore pre-sorted networks. We construct instruments from refugees' pre-placement characteristics from our rich administrative data, while also uniquely controlling for refugees' pre-placement ability. Network quality is found to constitute a far more important determinant of refugees' labor market outcomes when compared to network size, which might be suggestive that the *type* of network member referrals individuals receive are more important than the overall number of referrals. Blue-collar networks increase the probability of refugees' working in blue-collar jobs, draw additional workers into more manual and less complex intensive employment and serve to upskill individuals along the manual skill dimension. While blue collar networks are found to marginally increase the probability of finding a complex occupation, since blue collar workers upskill refugees proportionally more along the manual dimension, over time those in our sample are found to be employed in less complex occupations. One interpretation of this finding is that lower skilled networks lock individuals into jobs lower down the occupational ladder, therein preventing them moving to more skilled or complex occupations. An alternative perspective would be that the *Viet Kieu* arrived during a deep recession at a time of significant animosity toward them and in the absence of a counterfactual that the composition of *Viet Kieu* networks proved pivotal in their ultimate success.

As global geopolitical events renew debate in settlement policy, policymakers are faced with decisions as to whether to place refugees into networks or not. Our findings highlight the pivotal role of network quality in particular. In turn, this suggests that policy interventions designed to improve network quality can yield positive spillover effects.

References

Aames, J. S., Aames, R. L., Jung, J., and Karabenick, E. (1977). *Indochinese Refugee Self-sufficiency in California: A Survey and Analysis of the Vietnamese, Cambodians and Lao and the Agencies that Serve Them*.

Angrist, J. and Imbens, G. (1995). Identification and estimation of local average treatment effects.

Angrist, J. D., Imbens, G. W., and Rubin, D. B. (1996). Identification of causal effects using instrumental variables. *Journal of the American statistical Association*, 91(434):444–455.

Autor, D. H. and Dorn, D. (2013). The growth of low-skill service jobs and the polarization of the US labor market. *American Economic Review*, 103(5):1553–97.

Baker, R. P. and North, D. S. (1984). *The 1975 refugees: Their first five years in America*. New TransCentury Foundation.

Bartel, A. P. (1989). Where do the new us immigrants live? *Journal of Labor Economics*, 7(4):371–391.

Beaman, L. A. (2011). Social networks and the dynamics of labour market outcomes: Evidence from refugees resettled in the US. *The Review of Economic Studies*, 79(1):128–161.

Bertrand, M., Luttmer, E. F., and Mullainathan, S. (2000). Network effects and welfare cultures. *The Quarterly Journal of Economics*, 115(3):1019–1055.

Blume, L. E., Brock, W. A., Durlauf, S. N., and Ioannides, Y. M. (2011). Identification of social interactions. In *Handbook of social economics*, volume 1, pages 853–964. Elsevier.

Borjas, G. J. (1998). To ghetto or not to ghetto: Ethnicity and residential segregation. *Journal of Urban Economics*, 44(2):228–253.

Brell, C., Dustmann, C., and Preston, I. (2020). The labor market integration of refugee migrants in high-income countries. *Journal of Economic Perspectives*, 34(1):94–121.

Brock, W. A. and Durlauf, S. N. (2001). Discrete choice with social interactions. *The Review of Economic Studies*, 68(2):235–260.

Brücker, H., Hainmueller, J., Hangartner, D., Jaschke, P., and Kosyakova, Y. (2019). Refugee migration to germany revisited: Some lessons on the integration of asylum seekers. In *XIth European Conference of the fRDB on âHow to Manage the Refugee Crisisâ, Reggio Calabria*.

Caplan, N., Whitmore, J., and Bui, Q. (1985). Southeast asian refugee self-sufficiency study (report prepared by the institute for social research, university of michigan). *Washington, DC: US Department of Health and Human Services, Office of Refugee Resettlement*.

Cortes, K. E. (2004). Are refugees different from economic immigrants? some empirical evidence on the heterogeneity of immigrant groups in the united states. *Review of Economics and Statistics*, 86(2):465–480.

Damm, A. P. (2009). Ethnic enclaves and immigrant labor market outcomes: Quasi-experimental evidence. *Journal of Labor Economics*, 27(2):281–314.

Dunning, B. B. and Greenbaum, J. (1982). *A systematic survey of the social, psychological and economic adaptation of Vietnamese refugees representing five entry cohorts, 1975-1979*. Bureau of Social Science Research.

Dustmann, C. (1994). Speaking fluency, writing fluency and earnings of migrants. *Journal of Population economics*, 7(2):133–156.

Dustmann, C. and Fabbri, F. (2003). Language proficiency and labour market performance of immigrants in the uk. *The Economic Journal*, 113(489):695–717.

Dustmann, C. and Soest, A. v. (2001). Language fluency and earnings: Estimation with misclassified language indicators. *Review of Economics and Statistics*, 83(4):663–674.

Eckstein, S. and Peri, G. (2018). Immigrant niches and immigrant networks in the us labor market. *RSF: The Russell Sage Foundation Journal of the Social Sciences*, 4(1):1–17.

Edin, P.-A., Fredriksson, P., and Åslund, O. (2003). Ethnic enclaves and the economic success of immigrants-Evidence from a natural experiment. *The Quarterly Journal of Economics*, 118(1):329–357.

Finnan, C. R. (1981). Occupational assimilation of refugees. *International Migration Review*, 15(1-2):292–309.

Foged, M. and Peri, G. (2016). Immigrants' effect on native workers: New analysis on longitudinal data. *American Economic Journal: Applied Economics*, 8(2):1–34.

Galens, J., Sheets, A., and Young, R. (1995). Gale encyclopedia of multicultural america (vol. 2). *Detroit, MI: Gale Research Publications*.

GAO (1976). Evacuation and temporary care afforded indochinese refugees–operation new life. Report to the congress, US General Accounting Office.

GAO (1977). Domestic Resettlement Of Indochinese Refugees — Struggle For Self-Reliance. *Report to the congress, US General Accounting Office*.

Grogger, J. and Hanson, G. H. (2011). Income maximization and the selection and sorting of international migrants. *Journal of Development Economics*, 95(1):42–57.

Haines, D. W. (2012). *Safe haven?: A history of refugees in America*. Kumarian Press.

Handel, M. J. (2016). The O*NET content model: Strengths and limitations. *Journal for Labour Market Research*, 49(2):157–176.

Helwig, R. T. (2001). Worker displacement in a strong labor market.

HEW Refugee Task Force (1976). HEW Task Force for Indochina Refugees Report to the Congress June 15, 1976.

Hoang, C. (2010). *Boat People: Personal Stories from the Vietnamese Exodus 1975-1996*. Carina Hoang Communications.

House of Representatives (1975). Hearings before the Subcommittee on Immigration, Citizenship, and International Low of the Committee on the Judiciary: First Session on Indochina Refugees. May 5 and 7, 1975.

Ioannides, Y. M. and Datcher Loury, L. (2004). Job information networks, neighborhood effects, and inequality. *Journal of economic literature*, 42(4):1056–1093.

Jackson, M. O., Rogers, B. W., and Zenou, Y. (2017). The economic consequences of social-network structure. *Journal of Economic Literature*, 55(1):49–95.

Kerr, S. P., Kerr, W., Özden, Ç., and Parsons, C. (2016). Global talent flows. *Journal of Economic Perspectives*, 30(4):83–106.

Kerr, S. P., Kerr, W., Özden, Ç., and Parsons, C. (2017). High-skilled migration and agglomeration. *Annual Review of Economics*, 9:201–234.

Lieberson, S. (1980). *A piece of the pie: Blacks and white immigrants since 1880*. Univ of California Press.

Manski, C. F. (1993). Identification of endogenous social effects: The reflection problem. *The review of economic studies*, 60(3):531–542.

Marsh, R. E. (1980). Socioeconomic status of indochinese refugees in the united states: Progress and problems. *Soc. Sec. Bull.*, 43:11.

Martén, L., Hainmueller, J., and Hangartner, D. (2019). Ethnic networks can foster the economic integration of refugees. *Proceedings of the National Academy of Sciences*, 116(33):16280–16285.

McInnis, K. (1981). Secondary migration among the indochinese. *Journal of Refugee Resettlement*, 1(3):36–42.

Munshi, K. (2003). Networks in the modern economy: Mexican migrants in the US labor market. *The Quarterly Journal of Economics*, 118(2):549–599.

Munshi, K. and Rosenzweig, M. (2006). Traditional institutions meet the modern world: Caste, gender, and schooling choice in a globalizing economy. *American Economic Review*, 96(4):1225–1252.

ORR (1981). Refugee Resettlement Program, Report to Congress 1981.

Ottaviano, G. I., Peri, G., and Wright, G. C. (2013). Immigration, offshoring, and American jobs. *American Economic Review*, 103(5):1925–59.

Parsons, C. and Vézina, P.-L. (2018). Migrant networks and trade: The Vietnamese boat people as a natural experiment. *The Economic Journal*, 128(612):F210–F234.

Patel, K. and Vella, F. (2013). Immigrant networks and their implications for occupational choice and wages. *Review of Economics and Statistics*, 95(4):1249–1277.

Robinson, W. C. (1998). *Terms of refuge: The Indochinese exodus & the international response*. Zed Books.

Ruggles, S., Genadek, K., Goeken, R., Grover, J., and Sobek, M. (2017). Integrated public use microdata series: Version 7.0 [dataset]. Minneapolis: University of Minnesota.

Rumbaut, R. G. and Portes, A. (2001). Ethnogenesis: Coming of age in immigrant america. *Ethnicities: Children of immigrants in America*, pages 1–20.

Sonneborn, L. (2007). *Vietnamese Americans*. Infobase Publishing.

Stark, O. and Stark, O. (1991). The migration of labor.

Stein, B. N. (1979). Occupational adjustment of refugees: The Vietnamese in the United States. *International Migration Review*, 13(1):25–45.

Stock, J. H. and Yogo, M. (2002). Testing for weak instruments in linear iv regression. Technical report, National Bureau of Economic Research.

Strand, P. J. and Jones, W. (1985). *Indochinese refugees in America: Problems of adaptation and assimilation*, volume 1. Duke Univ Pr.

Thompson, L. C. (2010). *Refugee workers in the Indochina exodus, 1975-1982*. McFarland.

Tolbert, C. M. et al. (1987). Labor Market Areas for the United States.

Tolbert, C. M. and Sizer, M. (1996). US commuting zones and labor market areas.

Topa, G. and Zenou, Y. (2015). Neighborhood and network effects. In *Handbook of regional and urban economics*, volume 5, pages 561–624. Elsevier.

van der Werf, C. (2019). The impact of refugees on native studentsâ academic achievement.

Wahba, J. and Zenou, Y. (2005). Density, social networks and job search methods: Theory and application to egypt. *Journal of Development Economics*, 78(2):443–473.

Waldinger, R. (1994). The making of an immigrant niche. *International migration review*, 28(1):3–30.

Zhou, M. and Bankston, C. (1998). *Growing up American: How Vietnamese children adapt to life in the United States*. Russell Sage Foundation.

Zhou, M. and Bankston III, C. L. (1994). Social capital and the adaptation of the second generation: The case of vietnamese youth in new orleans. *International migration review*, 28(4):821–845.

Appendix

Supplementary Contemporaneous Evidence

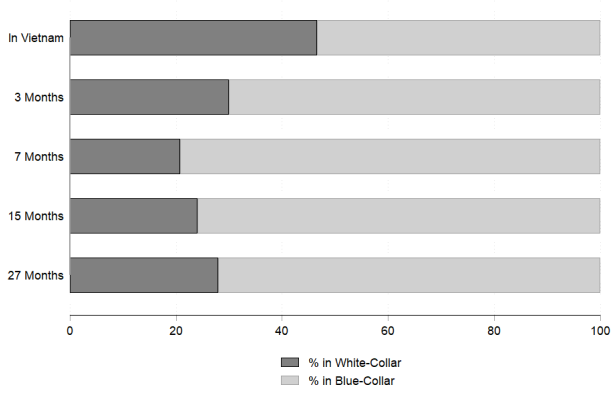


Figure A.1: Vietnamese Refugees in White-Collar and Blue-Collar Work. Source: Stein (1979). Note: Classifications differ from those used in the occupation analysis.

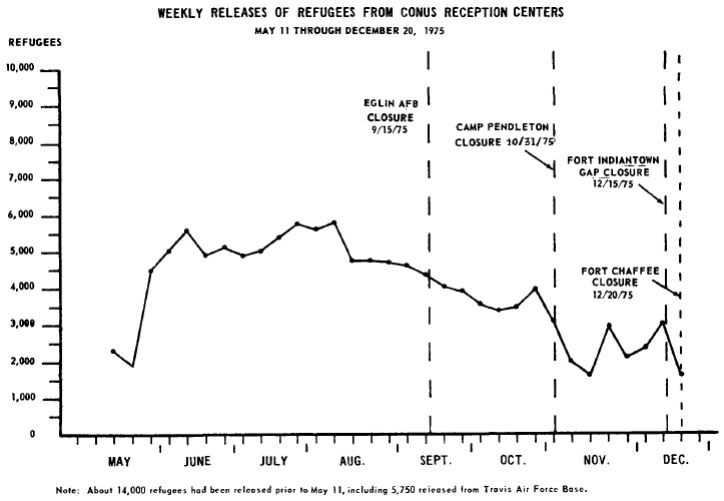


Figure A.2: Weekly Indochinese departures and Camp Closure Dates: April-December 1975. Source: GAO (1976)

Skills

Table A1: Classification of 52 O∗NET Skills.

Skill Name	Classification	Skill Name	Classification
Arm-Hand Steadiness	Manual	Response Orientation	Manual
Auditory Attention		Selective Attention	Cognitive
Category Flexibility	Cognitive	Sound Localization	
Control Precision	Manual	Spatial Orientation	Cognitive
Deductive Reasoning	Cognitive	Speech Clarity	
Depth Perception		Speech Recognition	
Dynamic Flexibility	Manual	Speed of Closure	Cognitive
Dynamic Strength	Manual	Speed of Limb Movement	Manual
Explosive Strength	Manual	Stamina	Manual
Extent Flexibility	Manual	Time Sharing	Cognitive
Far Vision		Trunk Strength	Manual
Finger Dexterity	Manual	Visual Color Discrimination	
Flexibility of Closure	Cognitive	Visualization	Cognitive
Fluency of Ideas	Cognitive	Wrist-Finger Speed	Manual
Glare Sensitivity		Written Comprehension	
Gross Body Coordination	Manual	Written Expression	
Gross Body Equilibrium	Manual		
Hearing Sensitivity			
Inductive Reasoning	Cognitive		
Information Ordering	Cognitive		
Manual Dexterity	Manual		
Mathematical Reasoning	Cognitive		
Memorization	Cognitive		
Multi-limb Coordination	Manual		
Near Vision			
Night Vision			
Number Facility	Cognitive		
Oral Comprehension			
Oral Expression			
Originality	Cognitive		
Perceptual Speed	Cognitive		
Peripheral Vision			
Problem Sensitivity			
Rate Control	Manual		
Reaction Time	Manual		

Table A2: Weighted Average of Occupation Skills

DOT	Manual	Cognitive	C/M
Professional, Technical and Managerial Occupations in:			
Architecture, Engineering, And Surveying	1.68	3.05	1.81
Mathematics And Physical Sciences	1.93	3.11	1.61
Computer-Related Occupations	1.29	2.91	2.25
Life Sciences	1.46	2.94	2.02
Social Sciences	1.07	2.97	2.78
Medicine And Health	1.97	2.92	1.49
Education	1.35	2.87	2.13
Museum, Library, And Archival Sciences	1.58	2.87	1.82
Law And Jurisprudence	1.20	3.00	2.50
Religion And Theology	1.08	2.59	2.39
Writing	1.29	2.74	2.13
Art	1.79	2.93	1.64
Entertainment And Recreation	1.86	2.53	1.36
Administrative Specializations	1.26	2.92	2.32
Managers And Officials, Not Elsewhere Classified	1.64	2.93	1.78
Miscellaneous Professional, Technical, And Managerial Occupations	1.70	2.92	1.72
Clerical, Sales and Related Occupations in:			
Stenography, Typing, Filing, And Related	1.35	2.52	1.87
Computing And Account-Recording	1.68	2.51	1.49
Production And Stock Clerks And Related	2.09	2.54	1.21
Information And Message Distribution	1.56	2.32	1.49
Miscellaneous Clerical	1.39	2.40	1.72
Sales of Services	1.28	2.71	2.12
Sales of Commodities Not Elsewhere Classified	1.59	2.62	1.65
Miscellaneous Sales	0.32	0.54	1.70
Service and Related Occupations in:			
Domestic Service	2.24	2.16	0.96
Food And Beverage Preparation And Service	2.28	2.54	1.11
Lodging And Related Services	1.98	2.55	1.29
Barbering, Cosmetology, And Related Services	2.14	2.66	1.24
Amusement And Recreation Services	1.96	2.17	1.11
Miscellaneous Personal Service	1.77	2.35	1.33
Apparel And Furnishings Services	2.30	2.21	0.96
Protective Services	2.11	2.72	1.29
Building And Related Services	2.33	2.13	0.91
Agricultural, Fishery, Forestry and Related Occupations in:			
Plant Farming	2.71	2.51	0.93
Animal Farming	2.24	2.67	1.19
Fishery And Related	3.03	2.61	0.86
Forestry	2.76	2.72	0.98
Processing Occupations in:			
Metal Products	2.63	2.59	0.99
Ore Refining And Foundry	2.82	2.52	0.89
Food, Tobacco, And Related Products	2.33	2.67	1.15
Chemicals, Plastics, Synthetics, Rubber, Paint, And Related Products	2.69	2.64	0.98
Leather, Textiles, And Related Products	2.64	2.01	0.76
Machine Trade Occupations in:			
Metal Machining	2.67	2.71	1.02
Metalworking Occupations, Not Elsewhere Classified	1.58	1.52	0.96
Mechanics	2.88	2.88	1.00
Machinery Repairers	2.83	2.79	0.98
Paperworking	2.70	2.60	0.96
Printing	2.53	2.66	1.05
Wood Machining	1.85	1.68	0.91
Textile	2.80	2.48	0.89
Machine Trades Occupations, Not Elsewhere Classified	2.46	2.34	0.95
Benchwork Occupations in Fabrication, Assembly and Repair of:			
Fabrication, Assembly, And Repair Of Metal Products, Not Elsewhere Classified Related Products	2.23	2.67	1.20
Fabrication And Repair Of Scientific, Medical, Photographic, Optical, Horological, And Related Products	2.19	2.74	1.25
Assembly And Repair Of Electrical Equipment	2.42	2.58	1.07
Fabrication And Repair Of Products Made From Assorted Materials	2.56	2.56	1.00
Fabrication And Repair Of Wood Products	2.93	2.77	0.95
Fabrication And Repair Of Sand, Stone, Clay, And Glass Products	2.74	2.65	0.97
Fabrication And Repair Of Textile, Leather, And Related Products	2.27	2.25	0.99
Structural Work and Related Occupations in:			
Metal Fabricating, Not Elsewhere Classified	2.45	2.59	1.06
Welders, Cutters, And Related Occupations	2.56	2.47	0.97
Electrical Assembling, Installing, And Repairing Occupations	2.60	2.98	1.14
Painting, Plastering, Waterproofing, Cementing, And Related Occupations	2.62	2.39	0.91
Excavating, Grading, Paving, And Related Occupations	2.73	2.66	0.98
Construction Occupations, Not Elsewhere Classified	2.73	2.67	0.98
Structural Work Occupations, Not Elsewhere Classified	2.70	2.71	1.00
Miscellaneous Occupations in:			
Motor Freight	2.65	2.65	1.00
Transportation, Not Elsewhere Classified	2.50	2.75	1.10
Packaging And Materials Handling	2.52	2.22	0.88
Extraction Of Minerals	2.94	2.83	0.96
Production And Distribution Of Utilities	2.46	2.79	1.14
Graphic Art Work	2.03	2.69	1.32

Transition Matrices

Table A3: Transition Matrix of Occupations in Vietnam in 1975 and those in the U.S. in 1981

U.S. Occupation Occupation in Viet	Misc.	Const	Benchwork	Trades	Processing	Agric	Services	Clerical& Sales	Prof & Manag	Total
Misc	**4%**	21%	14%	17%	2%	1%	13%	14%	14%	566
Construction	4%	**31%**	8%	26%	5%	0%	11%	5%	10%	280
Benchwork	2%	16%	**34%**	7%	4%	1%	21%	9%	7%	617
Trades	2%	27%	6%	**38%**	2%	0%	10%	5%	9%	403
Processing	2%	27%	20%	20%	**0%**	0%	24%	0%	7%	41
Agricultural	3%	18%	15%	14%	6%	**3%**	28%	4%	8%	120
Service	2%	21%	15%	9%	3%	0%	**24%**	15%	11%	841
Clerical & Sales	2%	16%	14%	11%	3%	1%	14%	**25%**	13%	767
Prof and Manag	3%	21%	7%	19%	2%	0%	8%	19%	**21%**	887

Notes: Robust standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A4: Transition Matrix of Complex Skills in Vietnam in 1975 and in the U.S. in 1981

U.S.	1	2	3	4	5	6	7	8	9	10	Total
Vietnam											
1	13%	5%	36%	9%	18%	3%	4%	3%	5%	4%	462
2	12%	4%	32%	9%	27%	2%	3%	4%	6%	2%	112
3	7%	2%	45%	10%	20%	2%	3%	3%	4%	2%	1083
4	10%	2%	35%	6%	26%	3%	2%	6%	6%	3%	156
5	6%	3%	23%	8%	26%	6%	8%	8%	7%	5%	829
6	6%	2%	25%	5%	21%	19%	4%	9%	6%	4%	485
7	5%	3%	26%	6%	21%	6%	10%	10%	5%	7%	500
8	4%	1%	21%	7%	17%	4%	13%	13%	7%	12%	212
9	4%	2%	20%	5%	24%	6%	5%	8%	17%	9%	212
10	5%	2%	20%	7%	17%	4%	9%	16%	8%	13%	471

Robustness

Table A5: Alternative Ethnic Concentration Threshold - 50%

	1	2	3	4	5
VARIABLES	Blue Collar	Professional	Cognitive	Manual	Complex
log network	0.032	0.006	0.017	0.000	0.005
	[0.010]***	[0.006]	[0.006]***	[0.004]	[0.004]
blue share	1.999	-0.022	0.008	0.610	-0.442*
	[0.503]***	[0.319]	[0.378]	[0.251]**	[0.245]
Vietnam Controls	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y
Observations	5,205	5,205	5,205	5,205	5,205
R-squared	[0.057]	[0.04]	[0.162	[0.109]	[0.102]
KP F-stat	[7.4]	[7.4]	[7.7]	[9.4]	[10.5]

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A6: Alternative Ethnic Concentration Threshold - 75%

	1	2	3	4	5
VARIABLES	Blue Collar	Professional	Cognitive	Manual	Complex
log network	0.028	0.001	0.015	-0.000	0.005
	[0.008]***	[0.005]	[0.005]***	[0.003]	[0.003]
blue share	1.844	-0.157	0.074	0.658	-0.340
	[0.397]***	[0.249]	[0.300]	[0.239]***	[0.212]
Vietnam Controls	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y
Observations	4,755	4,755	4,755	4,755	4,755
R-squared	[0.07]	[0.038]	[0.154]	[0.11]	[0.101]
KP F-stat	[13.7]	[13.7]	[12.9]	[11.9]	[13.6]

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A7: Alternative Ethnic Concentration Threshold - 125%

	1	2	3	4	5
VARIABLES	Blue Collar	Professional	Cognitive	Manual	Complex
log network	0.019	-0.002	0.011	-0.002	0.004
	[0.008]**	[0.005]	[0.004]**	[0.003]	[0.004]
blue share	1.821	-0.428	-0.219	0.917	-0.700
	[0.365]***	[0.191]**	[0.253]	[0.222]***	[0.194]***
Vietnam Controls	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y
Observations	3,795	3,795	3,795	3,795	3,795
R-squared	[0.08]	[0.042]	[0.152]	[0.11]	[0.102]
KP F-stat	[29.5]	[29.5]	[31.04]	[29.8]	[29.9]

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A8: Alternative Ethnic Concentration Threshold - 150%

	1	2	3	4	5
VARIABLES	Blue Collar	Professional	Cognitive	Manual	Complex
log network	0.030	0.001	0.025	0.002	0.006
	[0.008]***	[0.006]	[0.005]***	[0.004]	[0.005]
blue share	1.623	-0.239	-0.156	0.703	-0.686
	[0.365]***	[0.222]	[0.226]	[0.203]***	[0.199]***
Vietnam Controls	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y
Observations	2,860	2,860	2,860	2,860	2,860
R-squared	[0.091]	[0.042]	[0.152]	[0.13]	[0.104]
KP F-stat	[28.2]	[28.2]	[33.4]	[25]	[29.5]

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A9: Dropping Miscellaneous Occupations

	1	2	3	4	5
VARIABLES	Blue Collar	Professional	Cognitive	Manual	Complex
log network	0.022	0.001	0.018	-0.003	0.009
	[0.010]**	[0.006]	[0.005]***	[0.004]	[0.004]**
blue share	1.935	-0.338	0.071	0.700	-0.326
	[0.379]***	[0.228]	[0.294]	[0.196]***	[0.167]*
Vietnam Controls	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y
Observations	4,394	4,394	4,394	4,394	4,394
R-squared	[0.078]	[0.040]	[0.165]	[0.113]	[0.107]
KP F-stat	[17.6]	[17.6]	[17]	[16.9]	[19.1]

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A10: Removing Individuals from Network Measures - 10km buffer

	1	2	3	4	5
VARIABLES	Blue Collar	Professional	Cognitive	Manual	Complex
log network	0.091	-0.006	0.045	-0.003	0.015
	[0.038]**	[0.020]	[0.023]*	[0.004]	[0.016]
blue share	2.209	-0.435	0.054	0.596	-0.26
	[0.717]***	[0.359]	[0.516]	[0.306]*	[0.333]
Vietnam Controls	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y
Observations	3,218	3,218	3,218	3,218	3,218
R-squared	[0.06]	[0.044]	[0.151]	[0.121]	[0.106]
KP F-stat	[9.6]	[9.1]	[8.5]	[9.5]	[9.6]

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A11: Removing Individuals from Network Measures - 25km buffer

	1	2	3	4	5
VARIABLES	Blue Collar	Professional	Cognitive	Manual	Complex
log network	0.098	-0.007	0.048	-0.003	0.016
	[0.042]**	[0.022]	[0.025]*	[0.016]	[0.017]
blue share	2.262	-0.438	0.074	0.600	-0.248
	[0.746]***	[0.369]	[0.530]	[0.318]*	[0.344]
Vietnam Controls	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y
Observations	3,218	3,218	3,218	3,218	3,218
R-squared	[0.057]	[0.044]	[0.149]	[0.121]	[0.106]
KP F-stat	[8.7]	[8.7]	[8.3]	[8.7]	[9]

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A12: Removing Individuals from Network Measures - 50km buffer

	1	2	3	4	5
VARIABLES	Blue Collar	Professional	Cognitive	Manual	Complex
log network	0.089	-0.006	0.043	0.003	0.015
	[0.040]**	[0.020]	[0.023]*	[0.016]	[0.016]
blue share	2.200	-0.434	0.039	0.600	-0.253
	[0.706]***	[0.358]	[0.505]	[0.318]*	[0.340]
Vietnam Controls	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y
Observations	3,218	3,218	3,218	3,218	3,218
R-squared	[0.059]	[0.044]	[0.149]	[0.121]	[0.107]
KP F-stat	[9.9]	[9.9]	[9.3]	[9]	[9.5]

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Supplementary Regression Tables

Table A13: Cognitive Upgrading and Networks

	1	2	3	4	5	6
VARIABLES	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
log network	0.005	0.009	0.003	0.003	0.007	0.001
	[0.007]	[0.009]	[0.008]	[0.007]	[0.010]	[0.008]
	(0.009)	(0.012)	(0.010)	(0.008)	(0.011)	(0.009)
blue share		0.314			0.281	
		[0.485]			[0.510]	
		(0.546)			(0.566)	
network*share			0.039			0.036
			[0.055]			[0.058]
			(0.063)			(0.066)
Vietnam Controls	Y	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y	Y
Observations	4,522	4,522	4,522	3,777	3,777	3,777
R-squared	[0.019]	[0.017]	[0.017]	[0.019]	[0.017]	[0.017]
	(0.022)	(0.019)	(0.02)	(0.022)	(0.019)	(0.02)
KP F-stat	[1173.8]	[18.4]	[32.8]	[258.3]	[7.8]	[10]
	(2094.7)	(16.6)	(29.3)	(294.4)	(7.4)	(9.6)

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A14: Cognitive Downgrading and Networks

	1	2	3	4	5	6
VARIABLES	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
log network	-0.008	-0.011	-0.007	-0.006	-0.007	-0.005
	[0.007]	[0.008]	[0.007]	[0.007]	[0.009]	[0.007]
	(0.008)	(0.010)	(0.008)	(0.007)	(0.010)	(0.007)
blue share		-0.187			-0.069	
		[0.396]			[0.395]	
		(0.472)			(0.487)	
network*share			-0.023			-0.01
			[0.044]			[0.045]
			(0.055)			(0.058)
Vietnam Controls	Y	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y	Y
Observations	4,522	4,522	4,522	3,777	3,777	3,777
R-squared	[0.021]	[0.02]	[0.02]	[0.021]	[0.02]	[0.02]
	(0.023)	(0.022)	(0.022)	(0.023)	(0.022)	(0.022)
KP F-stat	[1731.8]	[18.4]	[32.8]	[258.3]	[7.8]	[10]
	(2094.7)	(16.6)	(29.3)	(294.4)	(7.4)	(9.6)

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A15: Manual Upgrading and Networks

	1	2	3	4	5	6
VARIABLES	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
log network	-0.006	0.005	-0.011	-0.007	0.003	-0.011
	[0.007]	[0.008]	[0.008]	[0.008]	[0.009]	[0.008]
	(0.008)	(0.009)	(0.009)	(0.008)	(0.008)	(0.008)
blue share		0.734			0.653	
		[0.386]*			[0.384]*	
		(0.412)*			(0.399)	
network*share			0.101			0.090
			[0.046]**			[0.045]**
			(0.047)**			(0.045)**
Vietnam Controls	Y	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y	Y
Observations	4,522	4,522	4,522	3,777	3,777	3,777
R-squared	[0.034]	[0.03]	[0.031]	[0.034]	[0.03]	[0.031]
	(0.037)	(0.035)	(0.35)	(0.037)	(0.035)	(0.035)
KP F-stat	[1173.8]	[18.4]	[32.8]	[258.3]	[7.8]	[10]
	(2094.7)	(16.6)	(29.3)	(294.4)	(7.4)	(9.6)

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A16: Manual Downgrading and Networks

	1	2	3	4	5	6
VARIABLES	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
log network	0.003	-0.006	0.007	0.005	-0.002	0.007
	[0.007]	[0.009]	[0.007]	[0.007]	[0.009]	[0.007]
	(0.008)	(0.008)	(0.009)	(0.007)	(0.008)	(0.008)
blue share		-0.061			-0.044	
		[0.370]			[0.337]	
		(0.356)*			(0.336)	
network*share			-0.086			-0.064
			[0.045]*			[0.041]
			(0.041)**			(0.039)*
Vietnam Controls	Y	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y	Y
Observations	4,522	4,522	4,522	3,777	3,777	3,777
R-squared	[0.044]	[0.041]	[0.041]	[0.044]	[0.041]	[0.041]
	(0.042)	(0.041)	(0.041)	(0.042)	(0.041)	(0.041)
KP F-stat	[1173.8]	[18.4]	[32.8]	[258.3]	[7.8]	[10]
	(2094.7)	(16.6)	(29.3)	(294.4)	(7.4)	(9.6)

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A17: Complexity Upgrading and Networks

	1	2	3	4	5	6
VARIABLES	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
log network	0.014	0.009	0.017	0.014	0.010	0.016
	[0.006]**	[0.008]	[0.007]***	[0.007]**	[0.009]	[0.007]**
	(0.008)**	(0.007)	(0.007)***	(0.006)**	(0.007)	(0.007)**
blue share		-0.371			-0.275	
		[0.331]			[0.342]	
		(0.345)			(0.335)	
network*share			-0.058			-0.044
			[0.039]			[0.040]
			(0.038)			(0.038)
Vietnam Controls	Y	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y	Y
Observations	4,522	4,522	4,522	3,777	3,777	3,777
R-squared	[0.038]	[0.038]	[0.037]	[0.038]	[0.038]	[0.037]
	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)	(0.037)
KP F-stat	[1173.8]	[18.4]	[32.8]	[258.3]	[7.8]	[10]
	(2094.7)	(16.6)	(29.3)	(294.4)	(7.4)	(9.6)

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A18: Complexity Downgrading and Networks

	1	2	3	4	5	6
VARIABLES	2SLS	2SLS	2SLS	2SLS	2SLS	2SLS
log network	-0.002	-0.010	-0.022	-0.017	-0.009	-0.020
	[0.007]***	[0.008]	[0.007]***	[0.007]**	[0.009]	[0.008]***
	(0.006)***	(0.007)	(0.008)***	(0.006)***	(0.007)	(0.007)***
blue share		0.497			0.487	
		[0.325]			[0.340]	
		(0.377)			(0.378)	
network*share			0.074			0.070
			[0.037]**			[0.039]*
			(0.042)*			(0.042)*
Vietnam Controls	Y	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y	Y
Observations	4,522	4,522	4,522	3,777	3,777	3,777
R-squared	[0.039]	[0.038]	[0.038]	[0.039]	[0.038]	[0.038]
	(0.04)	(0.039)	(0.039)	(0.04)	(0.039)	(0.039)
KP F-stat	[1173.8]	[18.4]	[32.8]	[258.3]	[7.8]	[10]
	(2094.7)	(16.6)	(29.3)	(294.4)	(7.4)	(9.6)

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A19: OLS: Blue Collar Work and Networks

	1	2	3	4	5	6
VARIABLES	OLS	OLS	OLS	OLS	OLS	OLS
log network	0.005	0.017	-0.005	0.007	0.020	-0.003
	[0.008]	[0.008**]	[0.007]	[0.008]	[0.009]**	[0.007]
	(0.006)	(0.007)**	(0.006)	(0.006)	(0.07)***	(0.007)
blue share		0.944			0.960	
		[0.190]***			[0.200]***	
		(0.246)			(0.254)***	
network*share			0.145			0.147
			[0.024***]			[0.025]***
			(0.034)***			(0.035)***
Vietnam Controls	Y	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y	Y
Observations	4,522	4,522	4,522	3,777	3,777	3,777
R-squared	0.17	0.178	0.18	0.188	0.198	0.2

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A20: OLS: Professional Work and Networks

	1	2	3	4	5	6
VARIABLES	OLS	OLS	OLS	OLS	OLS	OLS
log network	0.007	0.004	0.010	0.007	0.003	0.009
	[0.004]*	[0.004]	[0.004]**	[0.004]	[0.004]	[0.004]***
	(0.003)	(0.004)	(0.003)***	(0.003)*	(0.04)	(0.003)***
blue share		-0.248			-0.263	
		[0.109]**			[0.115]***	
		(0.093)***			(0.095)***	
network*share			0.033			-0.034
			[0.016]**			[0.017]**
			(0.013)**			(0.013)**
Vietnam Controls	Y	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y	Y
Observations	4,522	4,522	4,522	3,777	3,777	3,777
R-squared	0.133	0.134	0.134	0.151	0.152	0.152

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A21: OLS: Cognitive Work and Networks

	1	2	3	4	5	6
VARIABLES	OLS	OLS	OLS	OLS	OLS	OLS
log network	0.027	0.027	0.026	0.028	0.028	0.027
	[0.004]***	[0.004]***	[0.004]***	[0.004]***	[0.004]***	[0.004]***
	(0.005)***	(0.005)***	(0.005)***	(0.006)***	(0.06)***	(0.005)***
blue share		0.003			-0.009	
		[0.123]			[0.129]	
		(0.096)			(0.099)	
network*share			0.010			0.010
			[0.018]			[0.019]
			(0.013)			(0.0134)
Vietnam Controls	Y	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y	Y
Observations	4,522	4,522	4,522	3,777	3,777	3,777
R-squared	0.244	0.244	0.244	0.266	0.266	0.266

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A22: OLS: Manual Work and Networks

	1	2	3	4	5	6
VARIABLES	OLS	OLS	OLS	OLS	OLS	OLS
log network	-0.003	-0.001	-0.009	=0.002	0.000	-0.008
	[0.003]	[0.003]	[0.003]***	[0.003]	[0.003]	[0.003]**
	(0.002)	(0.002)	(0.003)***	(0.003)	(0.003)	(0.003)**
blue share		0.300			-0.313	
		[0.072]***			[0.075]***	
		(0.090)***			(0.093)***	
network*share			0.047			0.048
			[0.010]***			[0.010]***
			(0.013)***			(0.014)***
Vietnam Controls	Y	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y	Y
Observations	4,522	4,522	4,522	3,777	3,777	3,777
R-squared	0.216	0.22	0.221	0.233	0.238	0.239

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Table A23: OLS: Complex Work and Networks

	1	2	3	4	5	6
VARIABLES	OLS	OLS	OLS	OLS	OLS	OLS
log network	0.011	0.009	0.016	0.011	0.009	0.016
	[0.003]***	[0.003]***	[0.003]***	[0.004]***	[0.004]**	[0.003]***
	(0.003)***	(0.003)***	(0.003)***	(0.002)***	(0.002)***	(0.003)***
blue share		-0.282			-0.298	
		[0.080]***			[0.083]***	
		(0.075)***			(0.077)***	
network*share			-0.040			-0.041
			[0.011]***			[0.012]***
			(0.010)***			(0.011)***
Vietnam Controls	Y	Y	Y	Y	Y	Y
Individual Controls	Y	Y	Y	Y	Y	Y
CZone Arrival FE	Y	Y	Y	Y	Y	Y
Observations	4,522	4,522	4,522	3,777	3,777	3,777
R-squared	0.206	0.209	0.209	0.223	0.226	0.226

Notes: Standard errors displayed under regression coefficients. Statistics pertaining to standard errors clustered by commuting zone of arrival are shown in square parentheses. Statistics pertaining to standard errors clustered by commuting zone of destination are shown in curly parentheses. *** p<0.01, ** p<0.05, * p<0.1.