

UC Irvine

UC Irvine Electronic Theses and Dissertations

Title

The Geography and Multiplexity of Personal Networks

Permalink

<https://escholarship.org/uc/item/29x128b7>

Author

Smith, Emily Jessica

Publication Date

2018

Peer reviewed|Thesis/dissertation

UNIVERSITY OF CALIFORNIA,
IRVINE

The Geography and Multiplexity of Personal Networks

DISSERTATION

submitted in partial satisfaction of the requirements
for the degree of

DOCTOR OF PHILOSOPHY

in Sociology

by

Emily J. Smith

Dissertation Committee:
Professor Carter T. Butts, Chair
Professor Katherine Faust
Professor John R. Hipp

2018

DEDICATION

To my undergraduate advisor, Amy Orr, for introducing me to sociology and encouraging me to take a risk and pursue a subject I was – and continue to be – genuinely passionate about.

And

To Edwin Vargas, whose unconditional support motivates me daily to be the best I can be.

TABLE OF CONTENTS

	Page
LIST OF FIGURES	v
LIST OF TABLES	vi
ACKNOWLEDGMENTS	vii
CURRICULUM VITAE	viii
ABSTRACT OF THE DISSERTATION	xi
1 Introduction	1
1.1 Background	1
1.2 Chapter Outline	2
2 Precision in Self and Proxy Reports of Residential Location	4
2.1 Motivation	4
2.2 Data	5
2.2.1 Procedure	6
2.2.2 Measures	7
2.2.3 Point in Polygon Estimation	8
2.3 Determinants of Precision	10
2.3.1 Respondent Factors	10
2.3.2 Institutional Factors	12
2.3.3 Study Design Factors	13
2.3.4 Prior Information	13
2.4 Measures	14
2.4.1 Precision	14
2.4.2 Ego Measures	14
2.4.3 Alter Measures	16
2.4.4 Geographic Measures	16
2.5 Analytic Strategy	16
2.6 Results	17
2.6.1 Descriptives	17
2.6.2 Models Predicting Imprecision of Ego Location Estimate	18
2.6.3 Models Predicting Imprecision of Alter Location Estimates	21

2.7	Discussion	21
3	The Multiplexity and Locality of Potential Job Lead Ties	25
3.1	Introduction	25
3.1.1	The Job Search Process and Social Ties	26
3.1.2	Briding the Gap	28
3.2	Data and Methods	29
3.2.1	Data	29
3.2.2	Measures	31
3.2.3	Analytic Strategy	32
3.3	Results	32
3.3.1	Descriptives	32
3.3.2	Latent Class Analysis: Categorizing Job Lead Ties	34
3.3.3	Predicting Multiplexity	37
3.4	Discussion	38
4	Spatial and Social Embeddedness of Emergency Contact Ties	42
4.1	Introduction	42
4.2	Data and Methods	44
4.2.1	Data	44
4.2.2	Measures	46
4.2.3	Analytic Strategy	46
4.3	Results	47
4.3.1	How Local Are Emergency Contact Ties?	47
4.3.2	How Socially Embedded Are Emergency Contact Ties?	51
4.4	Discussion	54
5	Conclusion	56
5.1	Contributions	56
5.2	Limitations	57
5.3	Future Research	58
	Bibliography	59

LIST OF FIGURES

	Page
2.1 Newport Beach polygon with sampled points, point estimate, centroid, and RMSD	10
2.2 RMSD versus SRPA for egos and alter locations	15
2.3 Histograms of RMSD and SRPA of ego locations	17
2.4 Coefficients for education effects	20
3.1 Job search process	27
3.2 Sample distribution of job lead tie multiplexity score	33
3.3 Sample distribution of job lead tie multiplexity score: urban vs. rural	33
3.4 Sample distribution of job lead ties by strength	34
3.5 Sample distribution of job lead ties by distance and locality	35
4.1 Degree distribution of emergency contact alters	47
4.2 Distribution of distances to non-household emergency contact alters	48
4.3 Emergency contact alters by locality	49
4.4 Emergency contact ties by strength	52
4.5 Emergency contact ties by strength	52
4.6 Emergency contact ties by strength and locality	53

LIST OF TABLES

	Page
2.1 Prediction of Imprecision of Ego Location Estimate	19
2.2 Models Predicting Imprecision of Alter Location Estimate	22
3.1 LCA results for rural sample, n = 4854	37
3.2 LCA results for urban sample, n = 1174	37
3.3 OLS model predicting job lead tie multiplexity	39
4.1 Model predicting log(distance) to emergency contact ties	50
4.2 Model predicting odds of strong emergency contact tie	54

ACKNOWLEDGMENTS

I would first and foremost like to thank my advisor, Carter Butts, for his years of dedication to helping me become the strongest researcher and academic that I can be. I truly could not have accomplished what I have if it was not for your consistent support.

I would also like to thank my other thesis committee members: Katie Faust and John Hipp. Your mentorship throughout my years as a graduate student have been invaluable.

In addition, I would like to thank the members, past and present, of the NCASD Lab for the insightful conversations and collaboration.

And last but not least, thank you to my family for their unending support and encouragement as I have pursued my academic goals.

In terms of funding, I would like to acknowledge the following organizations and grants for supporting my work: National Science Foundation (BCS-0927027, OIA-1028394, IIS-1526736, SES-1260798); Office of Naval Research (N00014-08-1-1015).

CURRICULUM VITAE

Emily J. Smith

EDUCATION

Doctor of Philosophy in Sociology	2018
University of California, Irvine	<i>Irvine, CA</i>
Master of Science in Statistics	2017
University of California, Irvine	<i>Irvine, CA</i>
Master of Arts in Sociology	2015
University of California, Irvine	<i>Irvine, CA</i>
Bachelor of Science in Mathematics	2012
Linfield College	<i>McMinnville, OR</i>
Bachelor of Science in Sociology	2012
Linfield College	<i>McMinnville, OR</i>

RESEARCH EXPERIENCE

Graduate Student Researcher	2012–2018
University of California, Irvine	<i>Irvine, California</i>
Student Researcher	2011
Arizona State University	<i>Tempe, AZ</i>

TEACHING EXPERIENCE

Instructor: Introduction to Data Analysis with R	2017–2018
University of California, Irvine	<i>Irvine, CA</i>
Instructor: Introduction to R	2016–2017
University of California, Irvine	<i>Irvine, CA</i>
Teaching Assistant: ChaMP Summer Program	2015
University of California, Irvine	<i>Irvine, CA</i>
Workshop Assistant: Sunbelt Conferences	2013–2018

REFEREED JOURNAL PUBLICATIONS

- Social fabric and fear of crime: Considering spatial location and time of day** 2017
Social Networks
- The built environment, spatial scale, and social networks: Do land uses matter for personal structure?** 2017
Environment and Planning B: Urban Analytics and City Science
- The Relationship of Age to Personal Network Size, Relational Multiplexity, and Proximity to Alters in the Western United States.** 2015
The Journals of Gerontology, Series B: Psychological Sciences and Social Sciences
- Networks, Space, and Residents' Perceptions of Cohesion** 2014
American Journal of Community Psychology
- Estimating Clique Composition and Size Distributions from Sampled Network Data** 2014
Proceedings of First IEEE International Workshop on Network Science For Communication Networks

CONFERENCE PUBLICATIONS

- Geography and Multiplexity of Job Lead Ties** Aug 2017
American Sociological Association Annual Meeting
- Spatial and Social Embeddedness of Emergency Contact Ties** Jul 2017
1st North American Social Networks Conference
- Predicting Respondent Precision of Geographic Location** Aug 2016
American Sociological Association Annual Meeting
- Predicting Respondent Precision of Geographic Location** April 2016
36th International Sunbelt Network Conference
- The Relationship of Age to Personal Network Size, Relational Multiplexity, and Proximity to Alters in the Western United States** Aug 2016
American Sociological Association Annual Meeting

The Relationship of Age to Personal Network Size, Relational Multiplexity, and Proximity to Alters in the Western United States

34th International Sunbelt Network Conference

April 2016

Understanding the Job Lead Ties of Los Angeles Residents

American Sociological Association Annual Meeting

Aug 2016

Understanding the Job Lead Ties of Los Angeles Residents

33rd International Sunbelt Network Conference

April 2016

ABSTRACT OF THE DISSERTATION

The Geography and Multiplexity of Personal Networks

By

Emily J. Smith

Doctor of Philosophy in Sociology

University of California, Irvine, 2018

Professor Carter T. Butts, Chair

While the importance of geographic space in structuring social interactions is well established, much has remained unknown to researchers due in large part to the lack of available data on personal networks across space. Utilizing data from the American Social Fabric Study (Butts et al., 2014), a spatially stratified egocentric network study, this thesis aims to answer various methodological and substantive questions relating to networks and geography that have not been studied previously. The first paper is a methodological piece exploring what factors are related to the precision of geographic location of survey respondents (egos) as well as their social ties on six relations (alters). The second study examines an oft-studied social relation, that of the job seeking social tie, from a new perspective. In particular, we focus on the pool of *potential* job lead ties (rather than those that were mobilized) and study the relation in terms of multiplexity (i.e., the overlap of social relations) and locality to ego. The final paper examines another vital social relation: the ties that ego would seek to notify in the event of a disaster or emergency affecting his or her area. Given that this type of phenomenon is inherently spatial, we seek to understand the extent to which these ties are both spatially and socially embedded.

Chapter 1

Introduction

1.1 Background

Humans are social beings, which we can conceptualize as being embedded in a complex *social fabric* that consists of both their personal social networks along with their locations in social and geographic space (Butts et al., 2012). Despite technological advances that better allow for communication over long distances, physical space is one of the most fundamental determinants of interaction (Spiro et al., 2016; Takhteyev et al., 2012; Kulshrestha et al., 2012; Backstrom et al., 2010). While the distribution of individuals across space influences network structures (Preciado et al., 2012; Daraganova et al., 2012; McPherson et al., 2001), it is also true that the embeddedness of social networks in space influences the types of social relations formed and their dynamics, such as tie retention (Habinek et al., 2015), gang violence (Papachristos et al., 2013), and youth co-offending (Schaefer, 2012).

Space structures interaction opportunities, and thus where one lives is closely related to where one is tied to. Exploiting this property, researchers have used *spatial interaction functions* to describe the probability of a tie between two randomly selected individuals

as a function of the distance between them (Butts et al., 2012). These propinquity-based models have been successfully used to predict social phenomena such as crime rates (Hipp et al., 2013) and regional identification (Almquist and Butts, 2015). While these methods are useful in describing spatial phenomena, large-scale network data across space had not been collected until recently, thus leaving a sizable gap in our understanding of interpersonal networks across space.

Open questions remain, for instance, regarding the measurement of geographic location and expectations regarding precision of location estimates of respondents and their social ties. Further, our knowledge of how the structures of certain types of functional relations (such as emergency contacts or job informants) vary across geographic space remains limited. This thesis uses data from the American Social Fabric Study (ASFS), a spatially stratified ego-centric network study of adults across the western United States, in three papers examining, broadly, the relationship between interpersonal networks, geographic space, and social processes. Further, it highlights my own contributions to the ASFS dataset itself; specifically with respect to point localization of respondents and their social ties.

1.2 Chapter Outline

The first paper is a methodological piece examining factors related to precision of self and proxy reports of residential location. Though precision in self-report data has long been of interest to researchers, no study to our knowledge attempts to characterize self reporting of geographic location. By understanding what is associated with precision of estimates, I hope to inform researchers working with self-reported geographic data at both the survey design and analysis stages in terms of expected data quality. The second and third papers delve more deeply into two of the network relations reported in the ASFS: job lead ties and emergency contacts, respectively. In the third chapter I characterize potential job lead

ties (i.e., ties that one *would* use to gain information about job opportunities) in terms of what other social relations are incident on the dyad (tie multiplexity). Further, I explore the spatial element of the job lead relationship by exploring both the spatial proximity of these ties as well as how they vary in rural versus urban settings. In the fourth chapter, I examine the emergency contact relation in terms of embeddedness in both geographic and social space in an attempt to characterize where information might travel in the event of a disaster or emergency. To summarize, the three chapters will address the following questions:

1. What factors are associated with the precision of the estimated geographic location of respondents and those to whom they have social ties on any of six different relations?
2. How do we characterize an individual's set of potential job lead ties, especially in terms of multiplexity and locality, and to what extent do we see differences in rural versus urban settings?
3. How do we characterize an individual's set of potential emergency contact ties, in terms of embeddedness in social and geographic space?

Chapter 2

Precision in Self and Proxy Reports of Residential Location

2.1 Motivation

In a survey where respondents are asked to report the residential locations of themselves or others, the resulting measurement will have some degree of uncertainty. The extent of such geographical uncertainty can arise from multiple factors, such as obfuscation of the data by the researcher, use of widely recognized (hence reportable) areal units, or other issues related to informant accuracy. In this paper we focus on the precision of estimated residential locations of respondents and their personal network neighbors from a large-scale social network study. The precision of such estimates can have important implications for both subsequent analyses and research conclusions. For instance, if one is using geographic location to make inferences about some social process, it is necessary to be aware of the uncertainty surrounding location measurement. Likewise, understanding what factors are associated with precision could be of use to researchers at both the study design and analysis

stages (e.g. to correct for or reduce measurement error).

While accuracy and precision in survey reporting is a general problem that has long been of interest to researchers (Bernard et al., 1984; Mabe and West, 1982; Prince et al., 2008; Sherry et al., 2007), no study to our knowledge has examined precision of individuals’ residential locations obtained from self-report data. Similarly, we are aware of no prior work that examines this question for individuals in respondents’ personal networks (aka *alters*). In this paper, we use data from the American Social Fabric Study (Butts et al., 2014), a large-scale, spatially stratified egocentric network study of adults in the western United States, to explore the factors associated with the precision of the estimated geographic location of respondents (egos) and those to whom they have social ties on any of six different relations (alters).

2.2 Data

The American Social Fabric Study (ASFS) is a spatially stratified egocentric network sample of non-institutionalized adults in the Western United States (Butts et al., 2014). The ASFS consists of four surveys with a common instrument, administration, and recruitment mechanism: the Twin Communities Network Study (TCNS), Los Angeles Network Study (LA), Southern California Regional Network Study (SCR), and the Western United States Network Study (Western). The four studies differ in sampling design and/or target population. For the TCNS, respondents were solicited from select tracts in Irvine, CA and Santa Ana, CA, with respondents sampled at random within each tract set. For the LA component, respondents were solicited at random from the Los Angeles city boundary as of the 2000 Census. For the SCR, respondents were solicited at random from the 2000 Census tracts of several counties in Southern California, with the probability of a respondent being solicited from any given tract being proportional to the tract’s land area. And finally, for the West-

ern component, respondents were solicited at random from 2010 Census block groups within the study area consisting of the states of Washington, Oregon, California, Idaho, Nevada, Wyoming, Utah, Arizona, Colorado, and New Mexico. The probability of a respondent being solicited from any given block group was proportional to the block group's land area. This areal – rather than population – based design gives the ASFS substantial coverage of rural areas, and provides a high level of spatial variation, making it well-suited to examining questions of spatial precision. The inclusion of several name generator items on the ASFS survey (questions in which alters are elicited from respondents on several different social relations (Marsden, 1990), from which we elicit residential location of each alter) allows us to investigate precision for proxy reports of alter locations as well.

2.2.1 Procedure

Potential participants were identified by random sampling of non-institutionalized adult residents at the tract or block group level; names and addresses were obtained by a survey firm, and were not available to the researchers. Sampled individuals were mailed a letter (in both English and Spanish translations) inviting them to participate and providing them with a unique identification number they could use to log into an online survey. Participants were offered a \$2 incentive up front, as well as a \$10 completion incentive. Potential respondents were also sent postcard reminders one week after the initial mailing, follow-up letters one month after the initial mailing, and postcard reminders six months after the initial mailing. Data collection took place between 2010 (beginning with the TCNS) and 2013, with the LA, SCR, and Western samples being conducted from approximately 4/2012 to 1/2013. The overall response rate for all surveys was 19.3%.

2.2.2 Measures

Data collection for the ASFS was performed via a self-administered, web-based survey instrument hosted at UCI, available in both English and Spanish language versions. The survey instrument was identical for all four ASFS samples, and was structured as follows. The instrument begins with basic demographic questions, followed by a series of network generator questions and questions about ego’s neighborhood. Respondents were then asked to provide their and their alters’ places of primary residence. Finally, respondents were asked to provide complete egocentric network information (i.e., ties among alters) for two of the network generator items.

Network data was collected via egocentric name generator questions (Marsden, 1990). The six relations are as follows: (1) *Core discussion* indicates those with whom ego has discussed important matters in the previous six months; (2) *Social activities* indicates those with whom ego engages in social activities; (3) *Emergency contact* indicates those whom ego would seek to immediately inform in the event of an emergency; (4) *Neighborhood safety* indicates those with whom ego would discuss issues of safety in his or her neighborhood; (5) *Job leads* indicates those whom ego would contact to seek information about job opportunities; and (6) *Kin*, which is an aggregate of ego’s spouses or partners, parents, children, and siblings (separately elicited). There was no upper limit placed on the number of alters the respondent could name for each relation, and it was explicitly stated that the respondent could nominate the same alter for more than one relation.

With respect to the geographic data, respondents were asked to provide their and their alters’ places of primary residence. Initial elicitation requested addresses, with as much detail as the respondent was able to provide. Raw inputs were validated using the Google Maps API (with subjects being given the opportunity to validate and update initially parsed address information), and saved such that names and address(es) could not be linked. Further, this

data was automatically obfuscated so that geographic data were mapped to aerial units no smaller than predetermined clusters of two or more U.S. Census blocks (or a Landscan cell for alters outside of the U.S.). In the final dataset, each individual (egos and alters) is localized to a spatial polygon of varying degrees of precision (these can range from a block cluster or Landscan cell to an entire country), depending on the available information.

2.2.3 Point in Polygon Estimation

While each individual in the ASFS is localized to a spatial polygon, there are some instances where we want to localize each individual to a point (e.g., to estimate distance between ego and alter). The technique for estimating these points is presented here, which will directly relate to one notion of precision, discussed below.

Suppose person x is localized to polygon y . Knowing no other information other than that they reside somewhere in this polygon, we might be motivated to use the polygon centroid as a point estimate. However, we know the population of each Census block within the polygon, and thus have an idea about how the population is distributed, which we can use to achieve a better point estimate. Here, we treat each target individual x as exchangeable with everyone else in polygon y . In other words, each person in polygon y has an equal chance of being person x . Under this model, the probability that person x resides in some block z within polygon y is proportional to the population of z . This allows us to considerably refine location estimates for large but sparse polygons (common in the study area), without making more extensive assumptions regarding covariate availability or accuracy in either the survey or Census reference data.

To estimate a point in polygon y using this information, we do the following. For each individual in polygon y , we sample 1000 latitude/longitude coordinates from within the polygon. We first determine how many points to sample from each block by taking 1000

random samples, with the draws weighted proportionally to the population in each block (in other words, blocks with larger populations have a higher probability of being sampled). After determining how many points to sample from each block, we project each block onto the x - y plane, and draw points from within the bounding box of the block polygon until the predetermined number of points have been drawn from within the block. We repeat this for each block, subsequently transforming the points back into the original latitude/longitude coordinates. After sampling points for every individual in a given polygon, we average across all points to provide a point estimate of the specific location within the polygon. We also calculate three uncertainty measurements for that estimate: root mean squared distance (RMSD), median distance, and 95% quantile distance.

Figure 2.1 displays an example of the result of this process using the polygon of Newport Beach, CA. There were 37 individuals in the ASFS localized to this polygon, and thus 37,000 latitude/longitude points were sampled. Displayed are the polygon with the sampled points, along with the point estimate and its RMSD (in green) and the centroid (in red) for comparison. The RMSD for the point estimate is 3.61 km for the calculated point estimate, and is 3.88 km for the centroid. This suggests that our technique leads to a more precise estimate of the location of an individual in this polygon. It also follows that one source of potential locational uncertainty is the shape and size of the polygon to which an individual has been localized, as well as the concentration of the population distribution within it. These features, in turn, depend on institutional (i.e., the availability of well-recognized areal units such as cities or counties that are mnemonically available to respondents) and geographical (i.e., differences in population distribution) factors, as well as the behavior of respondents in providing more or less detailed address information. Our analyses will examine how these different factors combine to yield higher or lower precision estimates in practice.

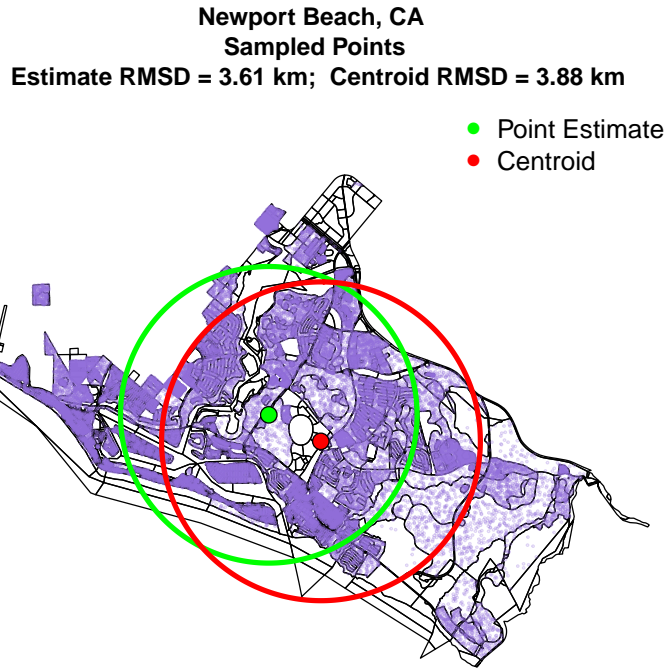


Figure 2.1: Newport Beach polygon with sampled points, point estimate, centroid, and RMSD

2.3 Determinants of Precision

When a respondent is asked to provide their place of primary residence in a survey, there are many different factors that could play a role in the precision of such an estimate. We consider four potential influences on precision: respondent-related, institutional, study design, and prior information.

2.3.1 Respondent Factors

The geographic information that a respondent is able and/or willing to provide in the survey will undoubtedly influence the precision of our location estimate. Researchers of cognitive geography have studied the different ways in which people think about and describe geographic space, which helps inform what we might expect to see vis a vis geographic precision (Montello and Friendschuh, 2005). Cognitive mapping is the process by which an individual

“acquires, stores, recalls, and decodes information about the relative locations and attributes of the phenomena in his everyday spatial environment” (Downs and Stea, 1973, pp.7). These cognitive maps consist of both knowledge of places as well as knowledge of spatial relationships (Kitchin, 1994). Golledge and Timmermans (1990) find that these maps develop with both age and education. With respect to wayfinding (the ability to learn and remember a route), familiarity with the environment is an influential factor (Gärling et al., 1983; Seidel, 1982). Finally, while some studies have found that males are better than females vis a vis spatial orientation tasks (Saarinen et al., 1988), the gender difference results have largely been mixed (Coluccia and Louse, 2004). Research has also suggested that cognitive effort in retrieval is related to the regularity or salience of events (Bound et al., 2001; Menon, 1994). With respect to the problem at hand, this factor would be particularly relevant in the retrieval of alter locations, as we might assume that addresses or locations more salient to the respondent would be more readily retrieved and thus more precise. While this set of literature is not directly related to survey response vis a vis geographic location, it adds insight into the types of cognitive processes involved when a respondent is asked to provide their place of primary residence.

In addition to what a respondent *can* tell us about their place of residence, it is also important to consider what a respondent *will* tell us. For instance, questions that are sensitive in nature may impact the types of answers the respondent provides (Bound et al., 2001). Questions related to sexual practices, eating disorders, drug use, or even income may elicit answers that distort the truth. While we are not aware of results examining this issue in the specific context of questions about location, we might assume that some individuals might be more or less inclined to report their place of residence in a survey. In addition to privacy effects, a respondent might also determine what they feel is an appropriate (i.e., conversationally sufficient) level of detail to provide about their location or that of their alter(s). For example, if an alter lives in a distant state, the respondent may feel it is appropriate to simply provide the city or the state, rather than a street address. Finally, respondent fatigue may also play

a role in how precise of an answer they are willing to provide (Hart et al., 2005; Hess et al., 2012).

While we are not aware of prior studies examining these effects for location reporting, the above literature above provides the basis for some hypotheses regarding possible predictors of precision. For instance, from a cognitive standpoint, because knowledge of one’s spatial environment tends to develop with age and education we might expect these variables to be positively associated with precision. Familiarity with one’s environment may be proxied by residential tenure, and thus we might expect the longer one lives in a location, the more precise of a location estimate we can obtain. Relatedly, we might expect alters who live farther away from ego or in another country to be associated with less precise estimates. Finally, we might expect that the location of an alter may be more salient if ego and alter are tied on strong relations (e.g., core discussion), and thus stronger ties may be associated with greater precision.

2.3.2 Institutional Factors

The precision of the estimate of geographic location is not only related to informant accuracy, but also institutional factors. Certain types of geographies are more common when humans talk about or describe geographic space, which we can conceptualize as a spatial “vocabulary,” which we refer to generically in terms of de-facto *reportable locations*. For instance, when asked for a place of residence, an individual is much more likely to provide the name of a town rather than a FIPS code corresponding to a Census tract, much less latitude/longitude coordinates. Each reportable location, such as a city or ZIP code, is mapped to an institutionalized polygonal marker that has political and functional borders. Such polygons vary in size, population distribution, etc. These properties are shaped by social factors that are external to the respondent (e.g., counties in the western U.S. tend to

be much larger than counties in the eastern U.S.).

2.3.3 Study Design Factors

Study design considerations can also impact precision. For instance, intentional obfuscation for purposes of privacy preservation puts an upper limit on achievable precision. In the context of the ASFS, locations were obfuscated to block clusters of two or more census blocks; thus, if a respondent provided an address, our estimate of their location is less precise than that, as we are by design unable to identify them below the cross-street level. On the other hand, sampling respondents from areal units can also provide a *lower* bound on spatial precision, as the respondent can (with rare exceptions due to recent moves) be immediately localized to their sampling unit. In the case of the ASFS, these units were US Census tracts or block groups (depending on survey wave), permitting a reasonably high degree of precision; note, however, that this only applies to estimates of ego locations, as alters are not localized by design.

2.3.4 Prior Information

Finally, the prior information we have with respect to the individual's polygon influences the precision of the estimate. The point in polygon estimation technique outlined above exploits the distribution of the population within the polygon as an informative prior, which can improve precision when population is unevenly distributed. For example, a polygon in which the population is densely clustered in a small region may lead to a more precise location estimate than a polygon of similar size with a more uniformly distributed population.

2.4 Measures

2.4.1 Precision

We use two measures of precision: root mean squared distance (RMSD) and square root of polygon area. RMSD, calculated using the sampled points and point estimate described above, is a measurement of how certain we are vis a vis an individual's location within a polygon given its population distribution. Polygon area is the area of the polygon (in square kilometers) to which an individual was localized, the square root of which provides a natural distance scale. We term this square root polygonal area (SRPA). The correlation between RMSD and SRPA is relatively high ($\rho = 0.85$), but there are some cases where the two are not in agreement. For instance, a large polygon with the population clustered in a small section of the polygon will have a much smaller RMSD than a polygon of similar area with a uniformly distributed population. Although RMSD provides a strong metric for applications such as point process analysis, SRPA may be a better indicator for areal unit studies. Figure 2.2 displays RMSD versus SRPA for both ego and alter locations, and highlights that while the two measures are similar, it is useful to study both when exploring spatial precision.

2.4.2 Ego Measures

The first set of variables we use to predict spatial precision are those relating to the respondent. These include variables such as gender, race, ethnicity, age, education, employment status, residential tenure, marital status, presence of children, and degree (number of alters). In addition, we include as a predictor the time to survey completion, as this could plausibly be associated with more precision (e.g., egos take their time answering the survey in detail) or less precision (e.g., respondent fatigue sets in). Because the data only includes the start

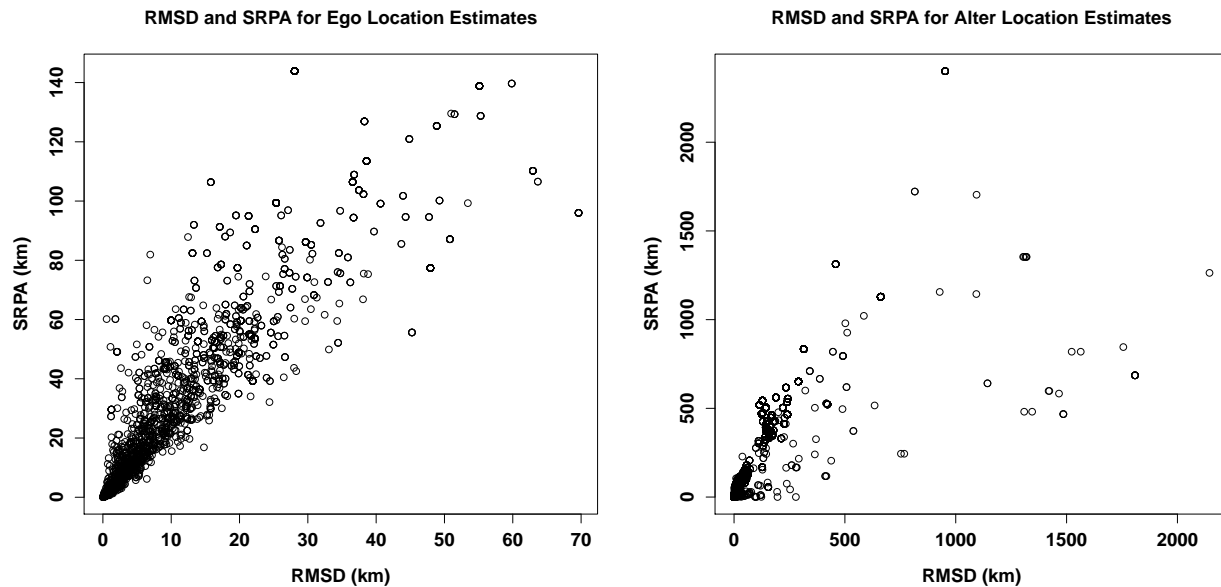


Figure 2.2: RMSD versus SRPA for egos and alter locations

and end dates and times, if the survey was completed over multiple distinct sessions it is impossible to determine exactly how much time was spent actually taking the survey. Thus, for the purposes of this analysis, we only include observations where the respondent took 300 minutes (5 hours) or less, under the assumption that any longer period of time was likely done in multiple sittings. Models were tested with different cutoff points and the results were not greatly changed. Finally, we include two measures as proxies for privacy or safety concerns. The first is an indicator variable for whether or not the respondent declined to state his or her income in the survey. The second is a composite measure of the level of safety the respondent feels in his or her neighborhood, made up of the average score from four scale items dealing with perceptions of neighborhood crime and safety. This score ranges from 1 to 5, with larger scores corresponding to greater perceptions of safety.

2.4.3 Alter Measures

The next set of variables are those relating to the alters. These include the available demographics (gender, race, and ethnicity), whether or not alters are reported to reside in ego's neighborhood, and estimated distance to alters. We also test for the type of relation, of which there were six: core discussion, social activities, emergency contact, neighborhood safety, job leads, and kin. It is to be noted that we exclude any alters living in ego's household for this analysis.

2.4.4 Geographic Measures

Finally, we include variables relating the geography of the polygon. These include population density (which can be thought of as a proxy for whether the area is urban or rural) and whether or not the polygon is in the United States (only applicable to alters).

2.5 Analytic Strategy

We fit four ordinary least squares models predicting $\ln(\text{RMSD})$ of ego location, $\ln(\text{SRPA})$ of ego location, $\ln(\text{RMSD})$ of alter location, and $\ln(\text{SRPA})$ of alter location. Using the survey package in R (Lumley et al., 2004), we correct the standard errors of the coefficients in the alter models for clustering within respondents (as more than one alter may be associated with a given respondent). Given the fact that little to no research has dealt with this topic, the analysis was largely of an exploratory nature and thus we tested various combinations of coefficients and chose the best fitting models as the ones with the lowest AICc scores.

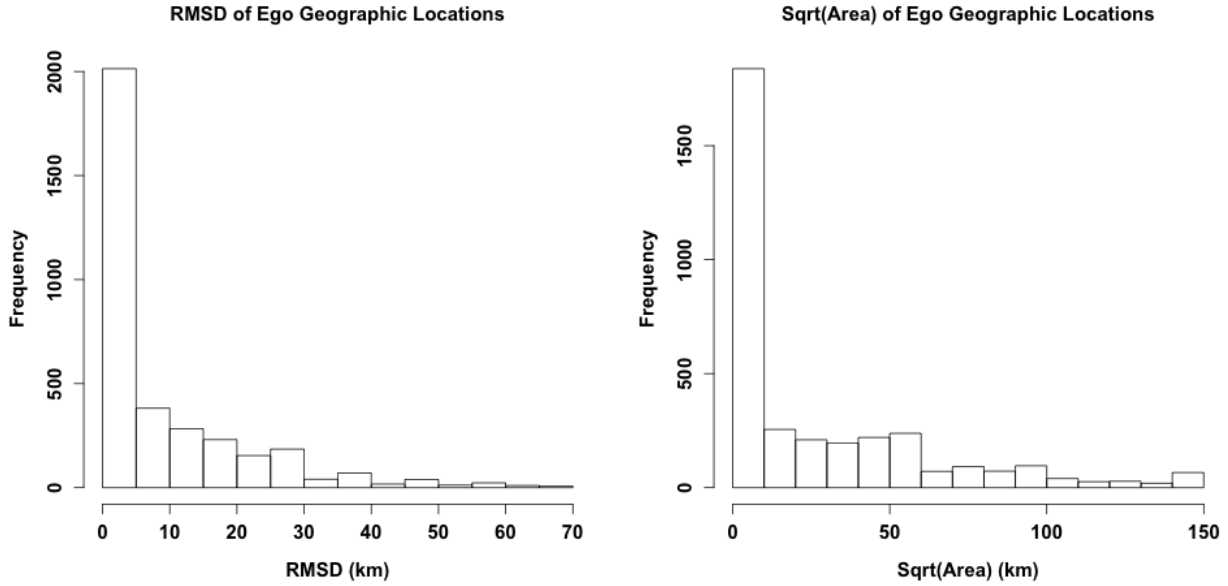


Figure 2.3: Histograms of RMSD and SRPA of ego locations

2.6 Results

2.6.1 Descriptives

Figure 2.3 displays histograms of the RMSD and SRPA of egos' geographic locations. The average RMSD is 8.80 km with a maximum of 69.63 km and the SRPA is 44.26 km with a maximum of 143.86 km. Thus, the distribution of both measures are heavily right-skewed. Not surprisingly, the estimate of alter locations are less precise on average, with a mean RMSD of 19.53 km and SRPA of 150.95 km (with maximum values of 2,144.61 km and 2,400.75 km, respectively).

Respondents are 56% male, 86% identify as white, and 28% identify as Hispanic. The average age is 54 years and the majority (34%) have completed some college. Respondents have on average 10 alters, the average time to complete the survey is 35 minutes, and about 5% of respondents declined to state their income. Finally, the average population density of ego's place of primary residence is 1,033 people per square kilometer.

About 47% of alters are male, 87% are white, and 10% are Hispanic. 22% of alters live in ego's neighborhood, and the average distance to an alter is 694 km. The average population density of alters is 1,063 people per square kilometer. About 53% of alters are kin of ego, 44% are core discussion partners, and 52% are emergency contacts.

2.6.2 Models Predicting Imprecision of Ego Location Estimate

Table 2.1 displays the coefficients for models predicting $\ln(\text{RMSD})$ and $\ln(\text{SRPA})$ of ego's location estimate. Since the dependent variable in each case reflects *imprecision* or uncertainty, a positive coefficient in either model corresponds to higher RMSD or SRPA (greater imprecision); negative coefficients on the other hand, indicate factors predictive of higher levels of precision.

We begin by examining factors significantly related to greater precision (lower uncertainty). The population density of ego's polygon is the strongest predictor of precision, with our estimate of ego's location being more precise when he or she resides in a place with higher population density (i.e., more urbanized areas). Greater time spent on the survey is also indicative of greater precision (i.e., the longer ego takes to complete the survey, the more precise his or her location estimate tends to be), as is neighborhood safety score (i.e., stronger perceptions of neighborhood safety are associated with more precise location estimates). Finally, greater residential tenure leads to greater precision vis a vis scale of ego's polygon (SRPA), but has no significant association with the RMSD of ego's estimated point location.

Turning to positive predictors of imprecision (i.e., lower precision), the strongest appears to be declining to state income. That is, if ego declined to state his or her income in the survey, the less precise his or her location estimates tend to be. Having a child is also associated with lower precision.

	ln(RMSD)	ln(SRPA)
Intercept	3.087 (0.184) ^{***}	4.335 (0.195) ^{***}
ln(Population Density)	-0.488 (0.005) ^{***}	-0.536 (0.005) ^{***}
ln(Survey Time)	-0.221 (0.029) ^{***}	-0.277 (0.031) ^{***}
Decline Income	0.414 (0.092) ^{***}	0.434 (0.098) ^{***}
Male	-0.062 (0.034)	-0.055 (0.036)
Hispanic	0.025 (0.021)	0.040 (0.022)
Age	0.001 (0.001)	0.001 (0.001)
Education		
Some College	0.263 (0.117) [*]	0.322 (0.124) ^{**}
Bachelor's	0.519 (0.133) ^{***}	0.655 (0.141) ^{***}
Graduate/Professional	0.735 (0.144) ^{***}	0.811 (0.152) ^{***}
White	-0.090 (0.100)	-0.072 (0.106)
Residential Tenure	-0.002 (0.002)	-0.004 (0.002) [*]
Has Child	0.054 (0.042)	0.103 (0.044) [*]
Neighborhood Safety Score	-0.069 (0.026) ^{**}	-0.064 (0.028) [*]
Education x White		
Some College x White	-0.191 (0.127)	-0.211 (0.134)
Bachelor's x White	-0.354 (0.143) [*]	-0.505 (0.151) ^{***}
Graduate x White	-0.604 (0.154) ^{***}	-0.658 (0.162) ^{***}
R ²	0.805	0.816
Adj. R ²	0.804	0.815
Num. obs.	2731	2731
RMSE	0.866	0.914

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 2.1: Prediction of Imprecision of Ego Location Estimate

A particularly interesting result is the positive relationship between education and imprecision. Egos with higher levels of education tend to have less precise location estimates compared to those with a high school education or less. Further, there is a significant interaction between ego being white and education. Figure 2.4 displays the coefficient estimates of education for those who identify as white and those who do not. For the RMSD model, there is no significant difference between the coefficient estimates between whites and non-whites for those who have completed some college or a bachelor's degree. Turning to those with a graduate degree, however, we see that whites with a graduate degree have significantly more precise location estimates than non-whites with a graduate degree. We observe a similar

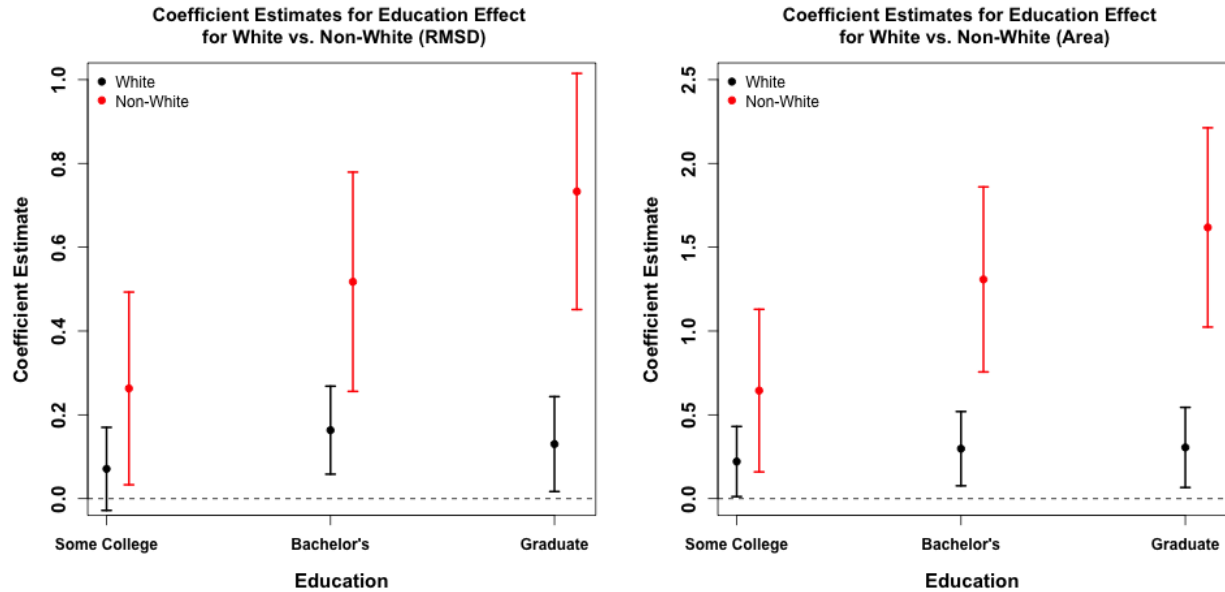


Figure 2.4: Coefficients for education effects

relationship in the SRPA model, though in this case the pattern holds for both those with a graduate degree and those with a bachelor's degree.

The positive relationship between education and imprecision could point to a privacy effect, whereby more highly educated individuals are more wary of sharing personal information over the Internet. This is supported by findings from the Pew Research Institute, which reports that those with a college or graduate education are more likely to take privacy precautions such as clearing cookie and browser history and encrypting email (Rainie et al., 2013) and those with at least some college education are more likely to believe they have the right to use the internet anonymously than those with a high school education or less (Madden and Rainie, 2015). Perhaps this, combined with an effect whereby those who are non-white (a more vulnerable population in the United States) are less likely to share personal information online, explain the interaction effects of the model (it should be noted that the Pew findings do not report any differences by race). The findings regarding declining to state income and the neighborhood safety score provide further support for this vulnerability and privacy hypothesis.

2.6.3 Models Predicting Imprecision of Alter Location Estimates

Table 2.2 displays the results from the models predicting imprecision of the estimate of alter location. In the SRPA model, as with the models predicting imprecision of ego’s location, population density is the strongest predictor of precision (i.e., more densely populated areas are associated with more precise location estimates). However, in the RMSD model, the strongest predictor of imprecision is alter being white, followed by population density. Survey time is once again a relatively strong predictor of precision, as is ego’s residential tenure. Finally, three relations (neighborhood safety, core discussion, and kinship) are all positive indicators of precision. That is, if alter is tied to ego on at least one of these relations, our estimate of alter location is more precise. This is intuitive given the nature of the relations in question, as two of the relations imply a relatively strong social tie (core discussion and kinship) while the other involves individuals with whom ego would discuss matters of community relevance (neighborhood safety).

The strongest positive predictor of imprecision is alter living outside of the U.S. Distance to alter is also positively associated with imprecision; that is, the farther alter lives from ego, the less our estimate of alter’s location tends to be. Our estimate also tends to be less precise when the alter is male, when ego is Hispanic, and if ego declined to state his or her income.

2.7 Discussion

This paper examines the factors associated with the precision of our estimate of the geographic location of survey respondents and their alters. We fit models predicting precision using two different notions of geographic precision: RMSD and SRPA. Both measures of precision yield qualitatively similar results for precision of ego and alter location estimates.

	ln(RMSD)	ln(SRPA)
Intercept	3.169 (0.206) ^{***}	4.470 (0.204) ^{***}
ln(Population Density)	-0.323 (0.008) ^{***}	-0.362 (0.009) ^{***}
ln(Survey Time)	-0.215 (0.045) ^{***}	-0.222 (0.041) ^{***}
ln(Distance to Alter)	0.210 (0.011) ^{***}	0.210 (0.010) ^{***}
Alter Non-U.S.	1.286 (0.140) ^{***}	1.172 (0.125) ^{***}
Decline Income	0.475 (0.163) ^{**}	0.499 (0.166) ^{**}
Ego Male	0.003 (0.044)	0.015 (0.045)
Ego White	-0.110 (0.093)	-0.144 (0.089)
Ego Hispanic	0.062 (0.025) [*]	0.069 (0.028) [*]
Ego Age	0.004 (0.002)	0.003 (0.002)
Education		
Some College	0.004 (0.068)	-0.010 (0.066)
Bachelor's	-0.069 (0.069)	-0.059 (0.070)
Graduate/Professional	0.014 (0.075)	0.019 (0.075)
Ego Residential Tenure	-0.005 (0.002) [*]	-0.005 (0.002) [*]
Ego Has Children	0.004 (0.062)	-0.047 (0.059)
Alter Male	0.048 (0.019) [*]	0.051 (0.019) ^{**}
Alter White	-0.350 (0.058) ^{***}	-0.332 (0.052) ^{***}
Alter Neighborhood	-0.243 (0.049) ^{***}	-0.262 (0.052) ^{***}
Relation		
Neighborhood Safety	-0.091 (0.037) [*]	-0.104 (0.038) ^{**}
Core Discussion	-0.127 (0.029) ^{***}	-0.119 (0.029) ^{***}
Kin	-0.204 (0.035) ^{***}	-0.204 (0.036) ^{***}
Emergency Contact		-0.050 (0.034)
Deviance	41889.768	42791.474
Dispersion	1.910	1.951
Num. obs.	21937	21937

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 2.2: Models Predicting Imprecision of Alter Location Estimate

The population density of an individual's polygon was overall the strongest predictor of precision for both egos and alters, indicating that the geographical context of the region is often critical when considering precision of estimates. This is supported by the finding that our estimate of alter location is much less precise among alters who live outside of the U.S. Some of the latter doubtless arises from the difficulty in remembering and communicating address information from locales with very different local descriptors than those used in the

respondents' own context.

The ego models suggest an interaction between education and identifying as white, with highly educated non-whites having significantly less precise geographic reports than highly-educated whites. Indeed, we see no net effect of education for white respondents. This is compatible with the notion that members of vulnerable groups may elect to provide less information as they become more educated (and hence perhaps aware of privacy issues). This privacy hypothesis is further supported by the finding that those who decline to state their income tend to provide less precise estimates of their own location (as well as those of their alters). Having at least one child is also associated with lower precision, which could also be indicative of a privacy effect. Finally, the positive relationship between precision and perceptions of neighborhood safety lends further support this relationship between perceived vulnerability and imprecision.

We see that the demographics of ego have little association with precision of alter locations, with the exception of ego identifying as Hispanic which is negatively associated with precision. Rather, alter demographics (race and gender) as well as the relationship between alter and ego are significantly associated with precision. In particular, alters that are tied in stronger relationships, white, and closer are associated with more precise locations.

In sum, geographical context appears to be most important when predicting how precise an estimate of geographic location will be. Further, privacy and/or vulnerability effects are stronger with precision of one's own location, rather than that of one's alters. This is supported by the finding that respondents tend to provide more precise locations of their presumably stronger ties (discussion partners and kin). Thus, while respondents are in general more precise when reporting their own location, they show evidence of a tendency toward a concern about their own privacy rather than the privacy of their alters. One implication of these findings is that extra efforts to convey privacy-protection measures to subjects may be especially helpful in obtaining more precise location information from those in vulnerable

groups (particularly those in such groups with higher education levels). Another implication is that greater effort may be needed to obtain more precise location information for respondents in rural areas. Since such areas often offer relatively imprecise geographical vocabulary (e.g., counties are larger, roads more widely separated, etc.), measures that work well in urban settings may perform poorly there. Improved elicitation methods (particularly ones that can be easily automated) could assist in reducing this problem.

Finally, it should be noted that these results pertain to a particular geographical and cultural context, and method of elicitation. More research is needed to see which of the effects found here generalize, and which are idiosyncratic. Nevertheless, our findings provide a first look at this problem, and suggest patterns that can be investigated in a wider range of settings.

Chapter 3

The Multiplexity and Locality of Potential Job Lead Ties

3.1 Introduction

The job search process and the social ties that individuals use to find jobs has been the focus of a large body of research in not only the social networks literature, but across many subfields and disciplines. Granovetter’s (1973) classic “strength of weak ties” theory (SWT) argues ties that bridge different social groups can more effectively diffuse information across a large number of people, and that these bridging ties are necessarily weak ties.¹ Retrospective studies looking at the ties successfully used by job seekers to find employment tend to report that individuals often find jobs through weaker contacts, such as acquaintances, rather than stronger contacts, such as friends or family (Granovetter, 1995; Friedkin, 1980; Lin, 1981; Lin and Dumin, 1986; Greenberg and Fernandez, 2016).

¹This is the core assumption of SWT in its original form, that strong tie two-paths are always completed, and hence only weak ties can bridge (Granovetter, 1973, p1363); this is the ideal type case, however, and it is usually expected to hold only approximately.

While there is much merit in SWT in terms of understanding the processes of information diffusion, research in its role in the diffusion of job information is not without gaps. For instance, much of the data in support of SWT tends to come from already-employed individuals in professional, managerial, and technical positions (Langlois, 1977; Friedkin, 1980; Lin and Dumin, 1986; Granovetter, 1995; Greenberg and Fernandez, 2016), often in metropolitan areas (Green et al., 1999; Schmutte, 2015; Dustmann et al., 2016), a subpopulation whose job market opportunities differ greatly from unskilled or semi-skilled workers, unemployed persons, and many other groups in the general population. Likewise, most research (e.g., Green et al., 1999; Marmaros and Sacerdote, 2002; Belliveau, 2005; Greenberg and Fernandez, 2016) to date selects on the population of ties that happened to be used in a realized job search, omitting e.g. those ties available to persons not seeking employment. Some studies (e.g., Caliendo et al., 2011; Burke and Kraut, 2013; Schmutte, 2015) further select on ties that were used in *successful* job searches, omitting mobilized ties that did not immediately lead to employment. While there is merit in identifying such tie populations, they may also provide a severely biased view of ties *prospectively available* to individuals at risk of a job search. Since it is the latter population of ties that serve as the reservoir of *potential* social capital to which members of the workforce may turn in times of need, their omission leaves an important gap in our understanding of the role of social networks in the job search process.

3.1.1 The Job Search Process and Social Ties

Figure 3.1 outlines the job search process from an egocentric point of view. While most studies examining the role of social ties in finding jobs tend to select on those ties that were either used or successful, we focus on the subset of ties that ego would potentially use to obtain job information. This set of ties comprises the resource base which ego has for subsequent job searches. Characterizing this portfolio of alters that a given individual could use to advance

a job search is important, as different alters offer different advantages and disadvantages. Multiplexity and locality are the two characteristics of ties that are of interest here.

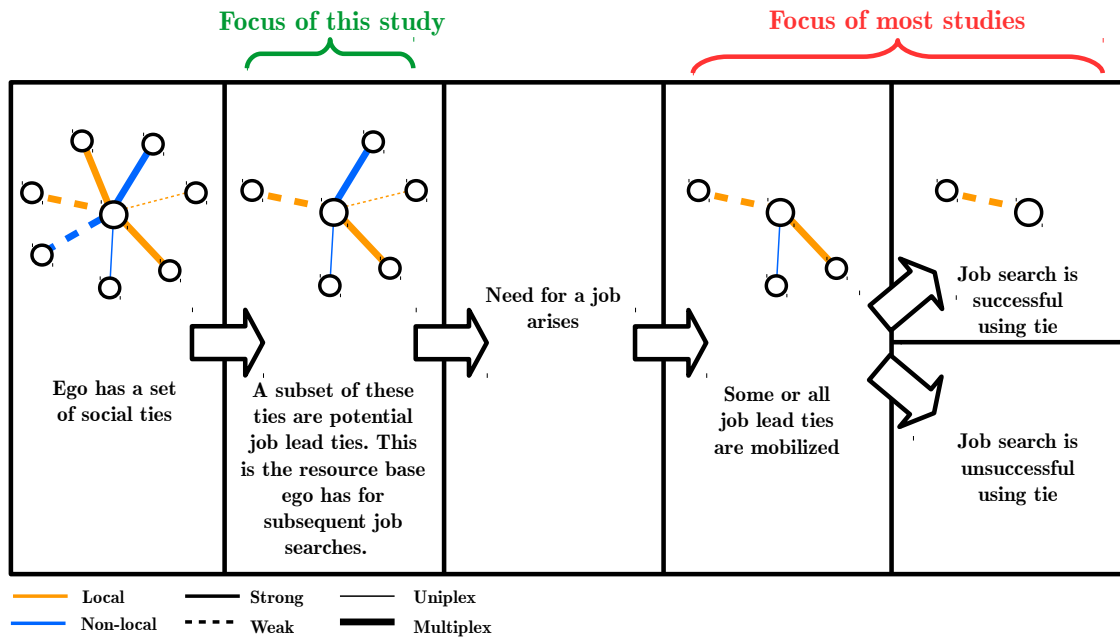


Figure 3.1: Job search process

Multiplex versus Non-Multiplex Ties

Multiplexity is the extent to which ties in one set of relationships overlap with ties in another set of relationships. Recalling Granovetter’s (1973) definition of tie strength as including interaction time, “reciprocal services,” and “emotional intensity” (i.e., an affective component distinct from other dimensions of the relationship), it is clear that multiplexity is deeply connected to tie strength per se. In the specific case of job leads, ties that coincide with many other types of interaction are likely to involve a greater degree of contact and an opportunity for more extensive exchange of “services” than those that do not. At an even more basic level, job lead ties that coincide with relationships broadly agreed upon to be strong ties (e.g., close kinship relations, confiding relations) are necessarily strong ties. The corresponding notion that lower levels of multiplexity are associated with weakness has had some support from prior studies (e.g., in his study of conversational ties in a kibbutz community,

Weimann (1983) suggests that weak ties are characterized by low multiplexity), although data allowing a direct investigation of this question has to date been limited.

Thus, to the extent that job lead ties are multiplex, ego has more contexts in which to interact with alter and seek his or her aid. Further, ego is easily able to mobilize these strong ties, and thus ego is more likely to access the information that alter may offer (Krackhardt et al., 2003). On the other hand, to the extent that the Granovetter (1973) condition holds (i.e., strong two-paths tend to be closed), strong-tie contacts are unlikely to bridge different parts of the social network which may contain new, non-redundant information.

Local versus Nonlocal Ties

Job leads offered by ties who are local to ego are more likely to be immediately useful to ego, as a local job is easier to take than a distant job. This could be due to factors such as unfamiliarity of housing markets (Roseman, 1971) or stress on the family (Munton, 1990). Alternately, economic conditions are spatially correlated, and thus distant alters are more likely to have access to opportunities that ego would not otherwise be able to discover. Furthermore, to the extent that urban and rural residents have systematically different patterns of job lead ties, the above assets and liabilities (with respect to both multiplexity and locality) will be distributed in a spatially unequal way.

3.1.2 Briding the Gap

This paper attempts to address gaps in our understanding of job lead ties by studying the ties prospectively identified as those to whom individuals *would* turn for information or assistance in a job search, were such a search required. Rather than focusing on specialized groups of workers (e.g., technical or managerial professionals), our study population consists

of the labor force of the western United States, including persons not currently employed.² Given this more representative population of job lead ties, we examine the incidence of these ties within the population (i.e., who sends them, and who receives them). We also attempt to characterize the embeddedness of job leads in dyads involving other relationships (i.e., multiplexity), particularly relationships conventionally viewed as “strong” ties (e.g., kinship or confiding relations). Finally, we characterize job lead ties in terms of their locality with respect to the residential location of ego.

Our analyses focus on the following questions: (1) how multiplex are job lead ties? (i.e., do they tend to be the specialized, narrow relationships often imagined in SWT theory, or do they tend to represent a single facet of a broader relationship?); (2) how local are job lead ties? (i.e., do they tend to be near or distant to ego’s residential location?); (3) to what extent do specific kinds of ties tend to co-occur with job lead ties (in particular, relational types conventionally viewed as “strong”)?; and, finally, (4) does job lead tie multiplexity vary systematically based on sender characteristics, or are similar patterns found throughout the population? To foreshadow our findings, a broader look at the population of prospective job lead ties reveals that these relations are considerably richer than often assumed, typically occurring as part of a complex of coincident relationships.

3.2 Data and Methods

3.2.1 Data

The data is from the American Social Fabric Study (ASFS), a spatially stratified egocentric network sample of non-institutionalized adults in the western United States (Butts et al.,

²Note that the labor force as conventionally defined (United States Department of Labor, Bureau of Labor Statistics (BLS), 2016) includes both employed persons and persons unemployed but looking for work; see also section 3.2.1.

2014). Respondents (egos) were recruited by mail and invited to take the survey online, where they answered demographic and geographic questions about themselves and their alters on six network relations. Data collection took place from approximately 4/2012-1/2013; participants were offered a small financial incentive up front, as well as completion incentives. The data consists four samples: (1) a random sample of residents in select tracts in Irvine, CA and Santa Ana, CA; (2) a population sample of adults in the city of Los Angeles; (3) a spatially stratified sample of adults in the Southern California region; and (4) a spatially stratified sample of adults in the western United States.

The survey contains six egocentric network generator questions, making it particularly useful for studying multiplex relationships. For each question, name generators were used, in which the data was uncensored and respondents could mention any number of names. The job leads question was worded as follows: *“When searching for a job, many people turn to friends, family, and/or associates for information regarding potential employment opportunities. Which of the following people would you seek to contact to obtain job leads or other information? Please check all that apply, and list other people (one-by-one) in the box below”* (Butts et al., 2014). The other five networks asked about are as follows: (1) with whom ego engages in social activities; (2) who ego would contact with information about an emergency or disaster; (3) with whom ego has discussed important matters within the previous six months; (4) who ego would contact to discuss issues of safety in his or her neighborhood; and (5) who ego considers a family member. We refer to these relations as social activities, emergency contact, core discussion, neighborhood safety, and kin, respectively.

This paper focuses specifically on job lead ties, with the edge itself being the unit of analysis. An edge here is defined as a job lead tie, where the sender is the potential job seeker (the respondent, or ego) and the receiver is the individual who would potentially provide information about a job lead (the alter). Because we are concerned only with individuals who would utilize job leads, we restrict the analysis to members of the labor force by omitting

respondents who identify as either (a) retired or (b) unemployed but not looking for work under the assumption that they will not be utilizing job leads and are not part of the labor force. After removing ineligible, there is a total of 6,028 job lead ties across 1,653 respondents. (unique ego/alter pairs).

3.2.2 Measures

We study multiplexity and locality of job lead ties. In order to study the multiplexity of job lead ties, we examine the other relations incident on each job lead tie (core discussion, social activities, emergency contact, neighborhood safety, and kin). We create a *multiplexity score* by summing the number of additional relations incident on each job lead dyad. Because there are five other network relations, this score ranges from 0 to 5, with 0 indicating the least multiplex (i.e., uniplex) job lead ties and 5 indicating the most multiplex. While multiplexity is one proxy for tie strength, we define a job lead tie that is also a kinship and/or core discussion tie as necessarily strong. We define a local job lead tie as a job lead tie being residing within 50 km of ego.

We include additional covariates in our analysis, some of which were collapsed for simplicity. All of the following variables are related to characteristics of the sender of the job lead tie. Race was simplified from fifteen categories to four: white, black, Asian, and other. Education was reduced to four categories: high school or less, Associate’s degree or some college, Bachelor’s degree, and graduate degree. We approximate yearly income by taking the midpoint of each income range in the survey. Other variables include gender, age, ethnicity (Hispanic or not Hispanic), employment status (unemployed and looking for work, employed part-time, and employed full-time), and job lead tie degree (number of job lead ties for a given respondent). The data also contains three demographic variables for the alters: gender, race, and ethnicity, operationalized in the same way that they were for the senders.

3.2.3 Analytic Strategy

The analysis will consist of three parts. First, descriptive figures highlight the important characteristics of job lead ties (multiplexity and locality). Second, a latent class analysis characterizes job lead ties in terms of other relations embedded in the dyad (multiplexity). Third, using OLS regression, we model job lead tie multiplexity as a function of geographic context and demographic variables.

3.3 Results

3.3.1 Descriptives

Multiplexity

Figure 3.2 displays the sample distribution of the job lead tie multiplexity score (i.e. the number of job lead ties in the data that have a multiplexity score of 0, 1, 2, etc.) in the ASFS. The shaded bars indicate multiplexity (i.e., a score greater than 0), and it is immediately clear that most of the job lead ties in the data are multiplex. Further, of those that are multiplex, the job lead ties do not tend to overlap with just one other social relation, but rather tend to be deeply embedded in multiple other relations. Figure 3.3 displays the multiplexity score distribution broken down by whether ego lives in a rural and urban area (a polygon was coded as urban if at least half the land area was classified as an urbanized area by the census). Urban areas have a significantly higher proportion of uniplex (multiplexity score of 0) ties than rural areas. In contrast, rural areas tend to have higher proportions of ties with high multiplexity scores (though this difference is not significant).

Figure 3.5 displays the distribution of job lead ties by strength, whereby a strong tie is

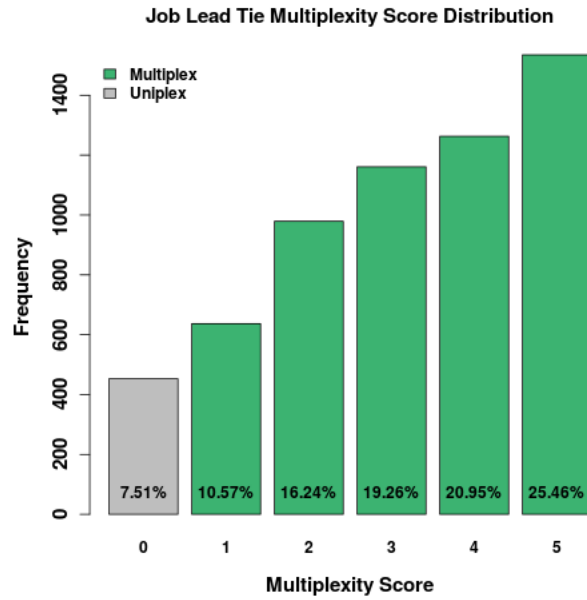


Figure 3.2: Sample distribution of job lead tie multiplexity score

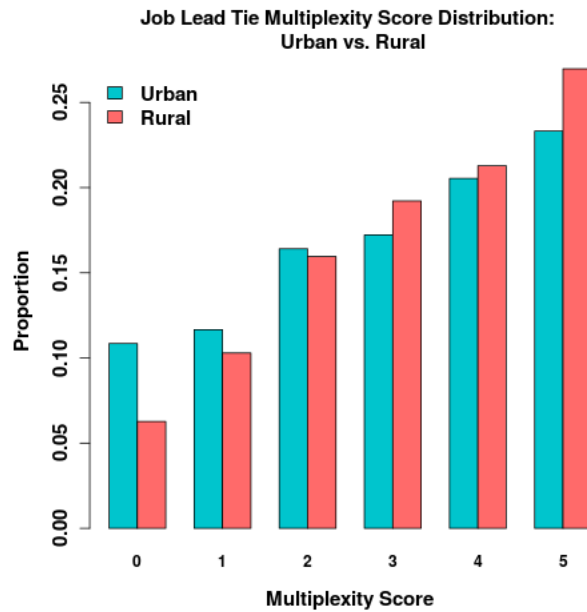


Figure 3.3: Sample distribution of job lead tie multiplexity score: urban vs. rural

defined as a job lead tie that is also a kinship and/or core discussion tie. Job lead ties are more often strong, with the majority being both a kinship and a core discussion tie.

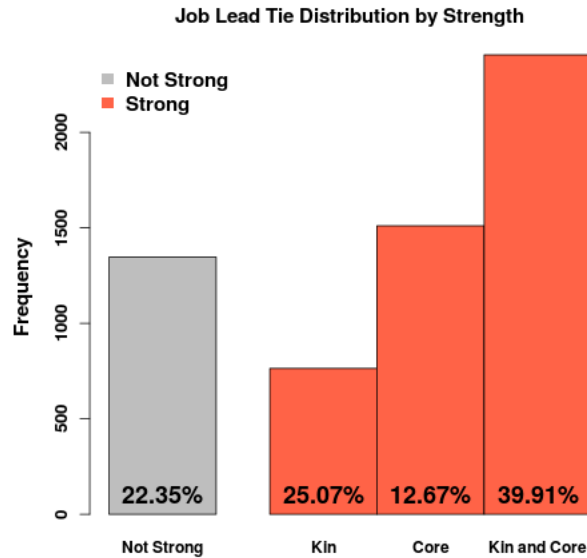


Figure 3.4: Sample distribution of job lead ties by strength

Locality

Figure 3.5 display descriptives of job lead ties with respect to locality to ego. The figure on the left suggests that while the median distance to job lead ties is relatively close to ego (42.29 km), these ties often extend far beyond ego’s local environment. Further, the majority of job lead ties are local to ego (i.e., within 50 km of ego’s residential location), with most of these local ties residing outside of ego’s household.

3.3.2 Latent Class Analysis: Categorizing Job Lead Ties

Latent Class Analysis (LCA) is a statistical technique used to analyze multivariate categorical data (Linzer and Lewis, 2011). The latent class model takes a set of observed (manifest) variables and stratifies them into a set of unobserved (latent) classes. In other words, given a sample of cases measured on several variables, we can use the LCA model to determine if there are a number of basic groups in which the cases fall. Once the model determines the

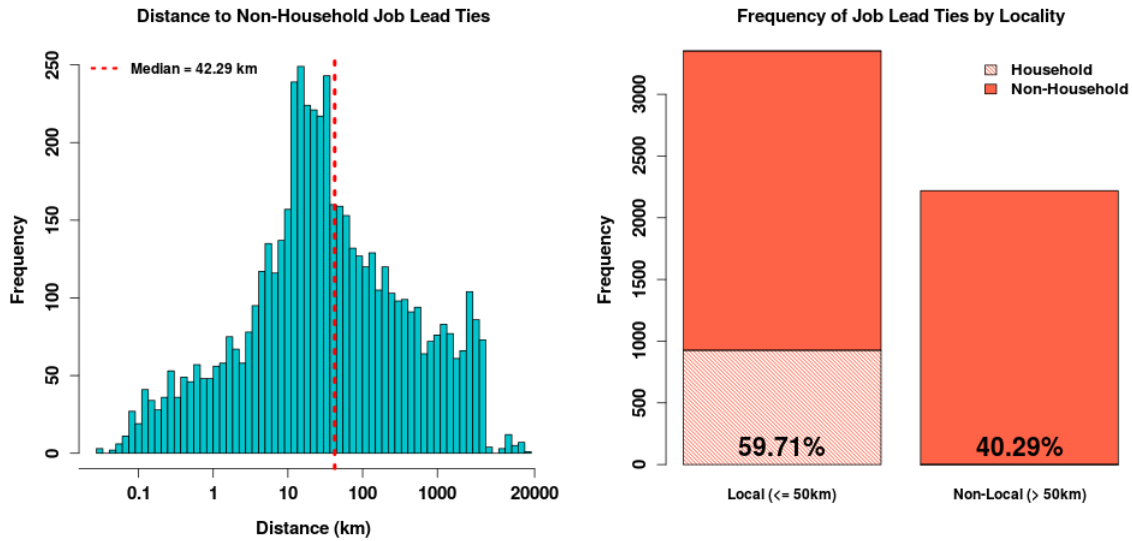


Figure 3.5: Sample distribution of job lead ties by distance and locality

latent classes, it produces conditional probabilities in regard to how each manifest variable will respond. That is, for a given member in a latent class, we can determine the probability that that member will belong to each of the manifest variables. Furthermore, conditional independence is assumed; within each latent class, each manifest variable is assumed to be statistically independent of every other variable.

Here, we use LCA to determine classes of relations in which job lead ties belong. In other words, given that ego is tied to an alter through a job lead tie, we can determine the probabilities that he or she will be tied to the alter on each of the other five network relations as well. This will be useful to explore trends in the multiplexity of job lead ties. All calculations were done using the `poLCA` package for **R**, “Polytomous Variable Latent Class Analysis” (Linzer and Lewis, 2007).

We fit two LCA models to the data, one for the rural egos and one for the urban egos, to examine whether relational patterns differ for the two samples. For each job lead tie, we take the other five relations (social activities, emergency contact, core discussion, neighborhood safety, and kin) and use LCA to determine if there are underlying groups of relational patterns

in which the job lead ties exist. It is important to note that all edges are observed with equal probability, and therefore reweighting is not necessary. In some applications, the number of latent classes is selected for theoretical reasons; however, because this is an exploratory analysis, the best model was chosen based on the AIC and Pearson's χ^2 goodness of fit test statistic. Ultimately, the model with three classes was chosen for both the rural and urban samples.

Tables 3.1 and 3.2 display the LCA results for the rural and urban samples, respectively. For each of the three classes, the tables present the conditional probabilities that a job lead tie is also each of the other five relations. The tables also show the estimated mixing proportions, i.e., the estimated proportion of the population of job lead ties that belongs to each of the three classes.

Beginning with the rural sample (Table 3.1), an estimated 39% of job lead ties belong to Class A. We call this the "Multiplex Class", as given a job lead tie belongs to this class, it has a high probability of also being each of the other five types of ties. We note that kin has the lowest probability of co-occurring with job lead ties in this class at 0.75. Class C is the smallest class, with an estimated 14% of ties belonging to it. Given that a tie belongs to Class C, it has a 0.32 probability of also being a social activities tie, followed by a 0.22 probability of also being a core discussion tie. We call this class the "Uniplex" class, as given a job lead tie belongs to it, it has a relatively low (or nearly 0) probability of being another type of tie. Finally, Class B is the largest class, encompassing approximately 47% of job lead ties. Given a job lead tie is in Class B, it has probability of 0.77 of also being a social activities tie and a probability of 0.66 of being an emergency contact tie. Ties in this class have a low probability of being a neighborhood safety tie and a relatively low probability of being a kinship tie. We therefore term this the "Social Friendship" class.

Turning to the urban sample (Table 3.2), the overall patterns and proportions of the three classes are relatively similar to that of the rural sample, though the job lead ties in the

uniplex class have a higher probability of being truly uniplex. This suggests that relational patterns of job lead ties do not vary considerably in urban and rural contexts, though we will explore this in more detail in the section below. Further, the LCA results suggest that these potential job lead ties do tend to be multiplex, despite prior research suggesting that utilized job lead ties are often weak.

	Class A	Class B	Class C
Kin	0.76	0.48	0.12
Social	0.95	0.77	0.32
Emergency	1.00	0.66	0.09
Neighborhood	0.92	0.26	0.00
Core	0.93	0.56	0.22
Est. Mixing Proportion	0.39	0.47	0.14

Table 3.1: LCA results for rural sample, n = 4854

	Class A	Class B	Class C
Kin	0.70	0.36	0.00
Social	0.91	0.68	0.19
Emergency	0.97	0.55	0.00
Neighborhood	0.89	0.14	0.00
Core	0.89	0.52	0.15
Est. Mixing Proportion	0.45	0.42	0.14

Table 3.2: LCA results for urban sample, n = 1174

3.3.3 Predicting Multiplexity

In order to examine what factors predict job lead tie multiplexity, we fit an OLS regression model to the data, with the multiplexity score of each tie as the response variable. Models were tested using combinations of variables and the model with the lowest AIC score was ultimately chosen. Table 3.3 displays the results of the model. We can immediately see that the baseline multiplexity score is very high, at 4.39 (recall that the maximum multiplexity score a job lead tie can have is 5). Job lead tie degree negatively predicts multiplexity, i.e., the more job lead ties an ego has, the less multiplex (or, more specialized) those ties tend to be.

The next set of predictors are related to the geography of ego and alter locations. First, as distance to alters increases, job lead tie multiplexity decreases. In other words, ties that are farther away tend to be more specialized. If we consider less multiplex ties to be weaker ties, then ties that span a greater spatial distance may also span a greater social distance as per Granovetter's (1973)'s argument regarding weak ties being more socially distant. Further, as population density of ego increases, multiplexity decreases. This can be considered an urban effect, as more densely populated areas tend to predict more specialized job lead ties.

Next we look at predictors related to ego, the first of which being ego's education level. The results indicate that individuals with either a bachelor's degree or a graduate/professional degree tend to have less multiplex job lead ties than individuals with a high school education or less. This fits with the narrative of prior research that suggests individuals in professional, technical, or managerial positions tend to report using weak ties to find jobs. White respondents also tend to have less multiplex job lead ties than those who do not identify as white.

Our results also indicate that characteristics of the alters predict multiplexity. In particular, male alters are associated with lower job lead tie multiplexity whereas white alters are associated with higher multiplexity. Unsurprisingly, alters who live in ego's household are associated with higher job lead tie multiplexity. Finally, if ego and alter are the same gender, the job lead tie tends to be less multiplex.

3.4 Discussion

The aim of this paper is to explore the complexities of the job seeking social tie, with an emphasis of embeddedness in social and geographic space. Reconceptualizing how we study job lead ties (i.e., focusing on potential ties, multiplexity, and geography) provides

	Estimate
Intercept	4.18 (0.23)***
Job Lead Tie Degree	-0.04 (0.01)***
Distance to Alter	-0.08 (0.01)***
Population Density: Ego	-0.04 (0.01)***
Ego Education	
Some College	-0.05 (0.11)
Bachelor's	-0.22 (0.11)*
Graduate	-0.42 (0.13)**
Ego Residential Tenure	-0.01 (0.00)
Ego White	-0.53 (0.13)***
Ego Married	-0.03 (0.03)
Ego Male	0.05 (0.07)
Ego Job Tenure	0.00 (0.03)
Alter Male	-0.22 (0.04)***
Alter White	0.44 (0.10)***
Alter Household	0.79 (0.14)***
Gender Homophily	-0.21 (0.05)***
Deviance	7579.29
Dispersion	1.81
Num. obs.	4194

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 3.3: OLS model predicting job lead tie multiplexity

deeper insight into the dynamics of this relation that has not previously been studied. For instance, given the breadth of literature on the effectiveness of weak ties in the diffusion of job information, we might have expected most job lead ties to exhibit little overlap with other types of network relations.

We find that job lead ties in our sample tend to be multiplex and strong, suggesting that job information is readily accessible to ego. Further, job lead ties of senders in rural areas tend to be more multiplex than those of senders in urban areas. While job lead ties are more often geographically local, they do extend far beyond ego's local environment. This has implications for mobility and usefulness of information, as distant job lead information may require the seeker to relocate, which may or may not be feasible.

The latent class analysis provided insight into what types of relations tend to be present along with job lead ties. In order to explore whether rural and urban ties display different relational patterns, we fit LCA models to each subset. Both samples display similar results; in particular, the majority of job lead ties tend to belong to classes that display high probabilities of multiplexity, though there is a smaller class in each model where the uniplex ties tend to be located. Overall, respondents in our sample view their available network of job lead ties as being embedded in other types of relations, which is perhaps indicative of relatively strong ties. Whether or not these ties are mobilized is another question, but understanding that individuals tend to perceive their job ties as being ones of more complex relationships draws into question the widespread use of weak ties for finding jobs. While this analysis suggests that relational patterns of job lead ties in the urban and rural samples do not differ markedly, we use OLS regression to explore the relationship between multiplexity, geography, and social space in more detail.

The OLS regression suggests that many different types of factors predict job lead tie multiplexity. Geographic context was a particularly strong predictor, with egos residing in more densely populated areas (a proxy for urban regions) tending to have less multiplex ties. Additionally, the farther an alter lives from ego, the less multiplex the tie tends to be. More highly educated respondents also predict less multiplex ties, as do white respondents. Finally, alter characteristics also predicted multiplexity. In particular, alters who are male or the same gender as ego are associated with lower multiplex job lead ties, whereas alters who are white or living in ego's household are associated with more multiplex ties. Taken as a whole, the OLS findings suggest that geography does indeed play an important role in job lead tie multiplexity, as do specific characteristics of both the sender and receiver of the tie. Thus, researchers interested in understanding the job search process and its relation to networks may want to do so in a more socially and geographically varied context than what has previously been studied.

Given the previously narrow focus on mobilized job lead ties belonging to professional, technical, and managerial workers in urban environments, we argue that expanding our focus to include other potential ties of a representative population provides insight into the complexities of job search networks that have not been able to be addressed previously. Through this analysis, we see that multiplexity is a pervasive attribute of job lead ties, which changes the way we think about this relationship; the traditional vision of job lead ties as very specialized or weak may not necessarily be the correct vision. Of course, we might not know if these ties are used or if they are even useful, but these results show that they are there, and they should continued to be studied in greater depth. Reconceptualizing how we study job lead networks can provide deeper insight into the dynamics of this network, which could ultimately have implications for policies and outcomes such as inequality in the labor market. Finally, this work emphasizes the importance of continuing to explore this phenomenon of multiplexity, not only for the job lead relationship, but for other relationships as well.

Chapter 4

Spatial and Social Embeddedness of Emergency Contact Ties

4.1 Introduction

Social networks serve as a crucial conduit for the dissemination of information, whether through face-to-face interaction (Granovetter, 1973; Centola and Macy, 2007) or online (Zinoviev, 2011; Bakshy et al., 2012; Guille et al., 2013). How information is diffused during emergencies or disasters is of interest to researchers (Toriumi et al., 2013; Fitzhugh et al., 2016) and to emergency managements officials and organizations as well (Manoj and Baker, 2007).

Classic studies have examined how information spreads through social contacts in the aftermath of a disaster. For example, Erickson et al. (1978) looked at the structure of interpersonal relations in an urban environment during an office building explosion, finding that increased status of workers was associated with the number and quality of contacts. They also found that housewives relied more on the media and family to propagate information

diffusion. Richardson et al. (1979) studied two communities of different sizes and examined the differences in the flow of information following a disaster, finding that while information flow was structured similarly in workplaces for the two communities, housewives' networks tended to be denser in the smaller town. More recent research in information transmission during and following disasters has focused on online environments, such as Twitter (Fitzhugh et al., 2016; Hui et al., 2012; Zhu et al., 2011), or through the media, e.g., television (Wei et al., 2010).

During a disaster, information diffuses quickly (Yoo et al., 2016), and this information often diffuses through pre-existing personal ties. Therefore, in order to understand where information will go when a disaster occurs, we must understand these conduits for information diffusion and where they are located. Disasters are a spatial phenomenon; however, we currently do not have a grasp on the geography of these ties due to the (previous) lack of available data. While some processes (e.g., information dissemination, resource mobilization) occur at the local level during a disaster, broader mobilizations occur as well if, for instance, local resources have been overwhelmed (Holguín-Veras et al., 2014). Thus, information that gets disseminated by an event may not necessarily remain local. For example, prior research on contacts used to discuss issues of neighborhood safety suggests that these networks are not spatially concentrated in the local neighborhood (Boessen et al., 2014).

Supposing a disaster affects ego's local environment, we can imagine scenarios where ego would push information to either local or distant contacts. Pushing individuals local to ego could serve a number of functions. First, this may mobilize those ties to take action to protect themselves. Second, these local individuals may be in a position to bring in immediate aid. By mobilizing local resources, we are better able to establish who needs aid and subsequently deliver that aid. Alternately, ego may push information about the disaster to ties that are farther away. If local resources are overwhelmed, it may be necessary to seek outside support. The Tennessee Valley coal ash disaster is a prime example of such mobilization,

whereby information diffusion via Twitter occurred beyond the local environment (Sutton, 2010). Distant communication may also offer a source of emotional support, especially if those ties are socially close to ego.

The goal of this chapter is to examine the social ties one would seek to notify in the event of an emergency or disaster, in an attempt to characterize the relation with respect to social and geographic space. Specifically, we ask the following questions: (a) To what extent are individuals pushing information locally versus non-locally? (b) To what extent are individuals pushing information to strong ties, such as kin or confidants? We might expect that in a disaster or emergency, individuals may tend to push information to those in their local environment (neighbors, community members, etc.). Alternatively, individuals might push information to individuals to which they are emotionally (but not necessarily geographically) close (kin, strong ties, etc.).

4.2 Data and Methods

4.2.1 Data

We use data from the American Social Fabric Study (ASFS), a spatially stratified egocentric network sample of non-institutionalized adults in the Western United States (Butts et al., 2014). Respondents (egos) answered demographic and geographic questions about themselves and their alters on six network relations. The data consists of four samples: (1) a random sample of residents in select tracts in Irvine, CA and Santa Ana, CA; (2) a population sample of adults in the city of Los Angeles; (3) a spatially stratified sample of adults in the Southern California region; and (4) a spatially stratified sample of adults in the western United States.

Network data was collected via six egocentric name generator questions (Marsden, 1990): (1) *Emergency contact* indicates whom ego would seek to immediately inform in the event of an emergency. This is the focal relation of the paper; (2) *Core discussion* indicates with whom ego has discussed important matters in the previous six months; (3) *Social activities* indicates with whom ego engages in social activities; (4) *Neighborhood safety* indicates with whom ego would discuss issues of safety in his or her neighborhood; (5) *Job leads* indicates whom ego would contact to seek information about job opportunities; and (6) *Kin*, which is an aggregate of ego’s spouses or partners, parents, children, and siblings. There was no upper limit placed on the number of alters the respondent could name for each relation, and it was explicitly stated that the respondent could nominate the same alter for more than one relation.

Respondents were then asked to provide the place of primary residence of themselves and their alters. This was validated using the Google Maps API and the data was automatically obfuscated so that geographic data were mapped to aerial units of clusters of two or more U.S. Census blocks (or a Landscan cell for alters outside of the U.S.). In the final dataset, each individual (egos and alters) is localized to a spatial polygon of varying degrees of precision (these can range from a block cluster or Landscan cell to an entire country). We also localize each individual to a point within their respective polygon, which is useful when examining, e.g., distance between two individuals.

ASFS respondents were asked the following question: *“Imagine that you were warned about an emergency, such as an earthquake, tornado, flash flood, fire, or toxic gas release affecting your area. Which of the following people would you see to immediately contact to pass on this information? Please check all that apply, and list other people(one-by-one) in the box below”*. Ego-alter pairs reported as falling under this relation are denoted as emergency contact ties. We note that here we examine an individuals’ set of *potential* emergency contacts (i.e., those they *would* contact in the event of an emergency). We argue that by studying the emergency

contact relation before it has necessarily been realized, we gain insight into where and to whom information might travel in the event of a crisis, emergency, or disaster.

4.2.2 Measures

The two dependent variables of interest are locality and strength. We split locality into four categories: ties in ego’s household (indicated in the ASFS survey), ties in ego’s neighborhood (indicated in the ASFS survey), local (less than 50km from ego), and non-local (greater than 50km from ego). Tie strength can be defined in many ways, but here we say an emergency contact tie is strong if that tie is also a kinship tie and/or a core discussion tie. We feel confident that an emergency contact tie that also falls under one of these two relations can be characterized as strong. However, it must be noted that all other ties cannot be necessarily denoted as “weak” (rather, we call them “not strong”).

4.2.3 Analytic Strategy

The analysis proceeds in two parts. First, we aim to answer the question “how local are emergency contact ties?” by exploring the embeddedness of emergency contact ties in geographic space. We present (a) distributions and frequencies of ties by locality, and (b) results of an OLS regression predicting distance to emergency contact ties. Next, we address the question “how socially embedded are emergency contact ties”, examining the embeddedness of these ties in socio-emotional space. We present (a) frequencies of ties by strength and (b) the results of a logistic regression predicting the odds of a strong emergency contact tie.

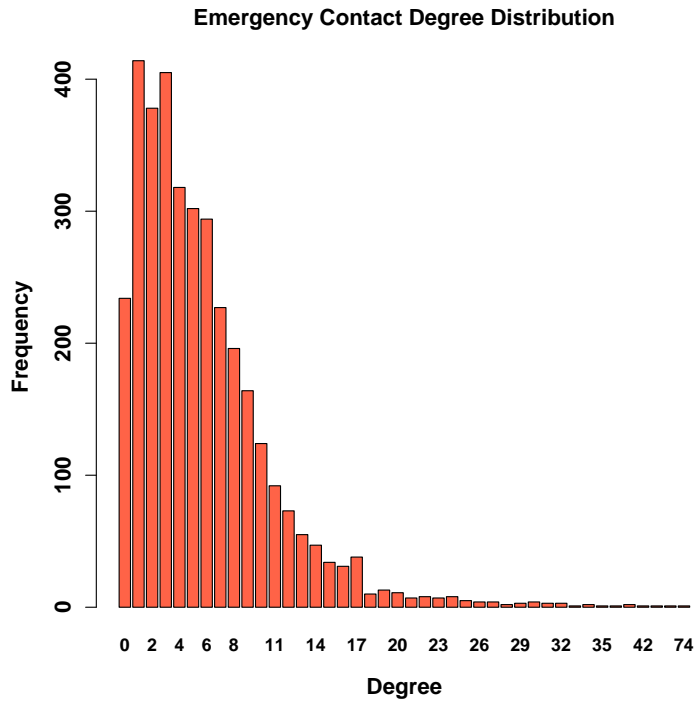


Figure 4.1: Degree distribution of emergency contact alters

4.3 Results

4.3.1 How Local Are Emergency Contact Ties?

Descriptives

Figure 4.1 displays the degree distribution of emergency contact ties in the ASFS. This is the distribution of the number of emergency contact ties per ego, which allows us to get a sense of how many emergency contact ties individuals tend to have. Characteristic of many social network degree distributions, it is right skewed (see, e.g., Girvan and Newman, 2002). The average emergency contact degree is 5.76, with a range from 0 to 74, indicating that while the egos in our sample have an average of just under 6 emergency contact ties, the range is quite wide.

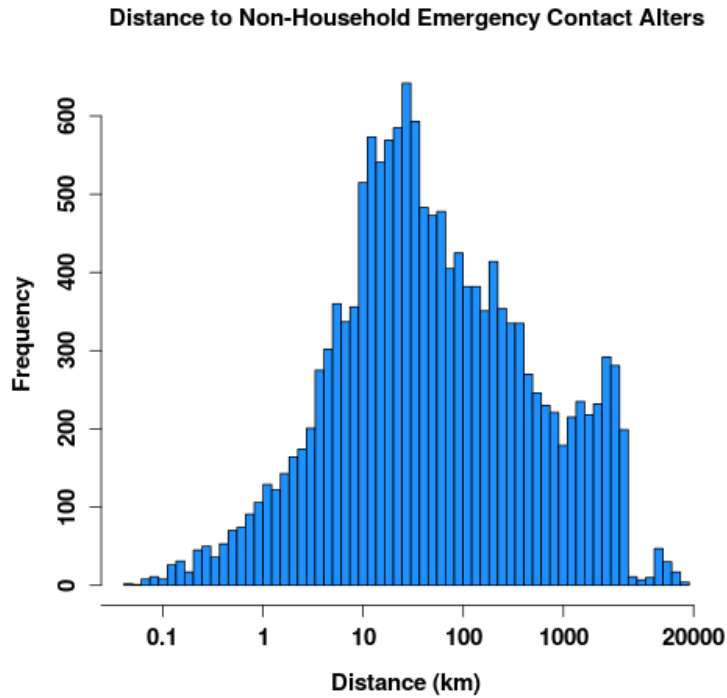


Figure 4.2: Distribution of distances to non-household emergency contact alters

Figure 4.2 displays the distribution of distances to non-household emergency contact alters (shown on log scale), allowing us to get a sense of how far away emergency contact ties tend to be. The median distance is 41km and the mean is 452km, suggesting that while the majority of ties tend to remain relatively close to ego, many are quite far.

Figure 4.3 shows the proportion of emergency contact ties by locality. The plot suggests that emergency contact ties tend to be geographically close to ego (in household, in neighborhood, or within 50km), though approximately one third of these ties are greater than 50km from ego. Along with Figure 4.2, we see that while the majority of potential disaster-related communication remains within ego’s local environment, information has the potential to traverse long distances as well.

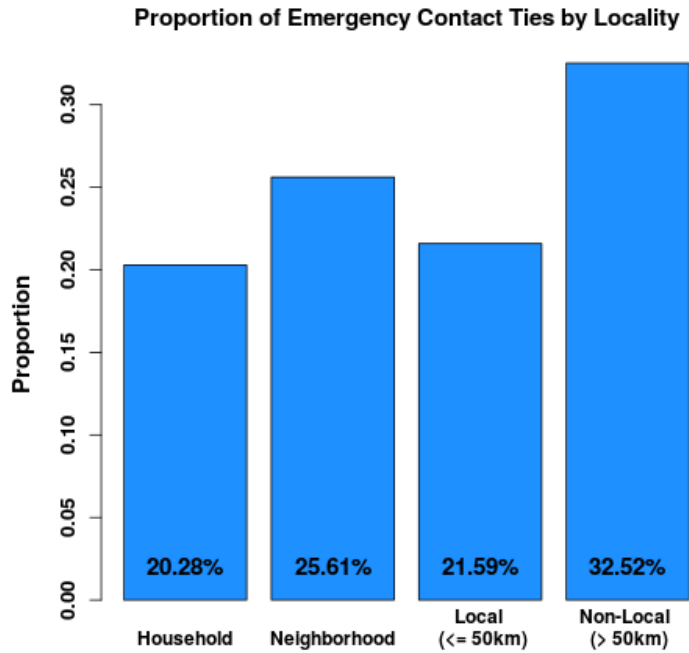


Figure 4.3: Emergency contact alters by locality

Predicting Distance to Alters

In order to further explore the spatial embeddedness of emergency contact ties, we fit an OLS model predicting $\log(\text{distance})$ to non-household alters as a function of ego, alter, and geographic variables. In other words, given that a tie in our sample is an emergency contact tie, what factors predict the distance to such a tie?

The results of the model fit are presented in Table 4.1. The strongest predictor of distance to emergency contact ties is ego holding a graduate/professional degree (compared to a high school education or less). The term is positive, indicating that individuals with a high degree tend to have emergency contact ties that are farther away. We note that there is a clear gradient trend for all levels of education (despite not being statistically significant), suggesting that those with more education are expected to push information to those who are farther away. With respect to other ego-related variables, males tend to have emergency

	Estimate
Intercept	4.199 (0.194) ^{***}
log(Ego Pop Density)	-0.187 (0.011) ^{***}
log(Alter Pop Density)	0.180 (0.012) ^{***}
Ego Education	
Some College	0.039 (0.097)
Bachelor's	0.096 (0.108)
Graduate/Professional	0.315 (0.118) ^{**}
Residential Tenure	-0.014 (0.003) ^{***}
Ego Male	0.148 (0.070) [*]
Ego White	-0.459 (0.134) ^{***}
Alter Male	0.010 (0.038)
Alter White	0.226 (0.093) [*]
Ego Neighborhood Belonging	-0.067 (0.013) ^{***}
Num. obs.	14112

^{***} $p < 0.001$, ^{**} $p < 0.01$, ^{*} $p < 0.05$

Table 4.1: Model predicting log(distance) to emergency contact ties

contact ties that are farther away, and white egos tend to have ties that are closer. The sole significant alter variable suggests that alters who are white tend to be more distant from ego.

Turning to variables related to geographic context, egos in areas with higher population density (i.e., more urban environments) tend to have ties that are closer. Conversely, alters in more urban environments tend to be farther away from ego. Taken together at equal density, the ego effect is stronger, suggesting that the net effect of urbanicity is to encourage shorter ties. Variables related to neighborhood attachment also display significant relationships with distance. The longer ego has lived in an area (residential tenure), the closer his or her emergency contact ties tend to be. Further, the stronger sense of belonging ego feels in his or her neighborhood, the closer these ties are as well.

In all, these results suggest that while many emergency contacts are geographically close to ego, a non-trivial amount do extend beyond ego's local environment. Geographic context in particular plays an important role in how far information could reach, shown through both population density as well as ego's attachment to his or her local environment. Further, the

education level of ego displays a clear association with distance, with more highly educated egos tending to have emergency contact ties that are farther away.

4.3.2 How Socially Embedded Are Emergency Contact Ties?

Descriptives

In this section we examine emergency contact ties in terms of tie strength. Figure 4.4 displays the proportion of emergency contact ties by strong versus not strong. We can see here that the vast majority (80%) of emergency contact ties are also a kin and/or core discussion tie, suggesting that most emergency contact ties in our sample are strong. Next we break the strong ties down by whether they are a kinship tie, core discussion tie, or both (Figure 4.5), and we see that strong emergency contact ties are most often both kinship and core discussion ties. Thus, most emergency contact ties are family members who are also confidants.

Figure 4.6 displays the proportion of emergency contact ties by strength and locality, allowing us to get a sense of whether there is a relationship between geographic and social space. Unsurprisingly, the majority of ties within ego's household are strong. Similarly, a higher proportion of non-local ties are strong as well. Because it takes more effort to maintain relations that are geographically distant, we argue it makes sense that ties upheld at greater distances are more likely to be strong ties. On the other hand, a greater proportion of ties within ego's neighborhood or local to ego (within 50km) are not strong. This suggests that there is indeed a relationship between geographic and social space with respect to emergency contact ties, which we will explore further in the regression analysis below.

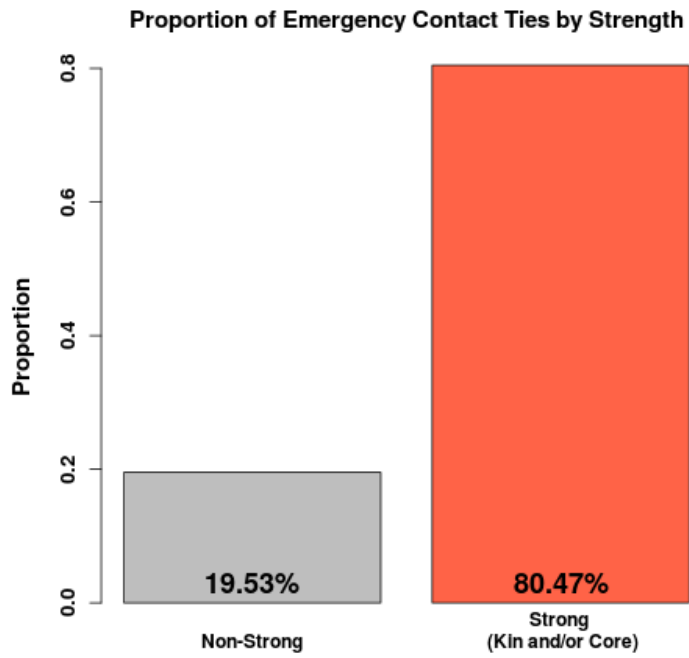


Figure 4.4: Emergency contact ties by strength

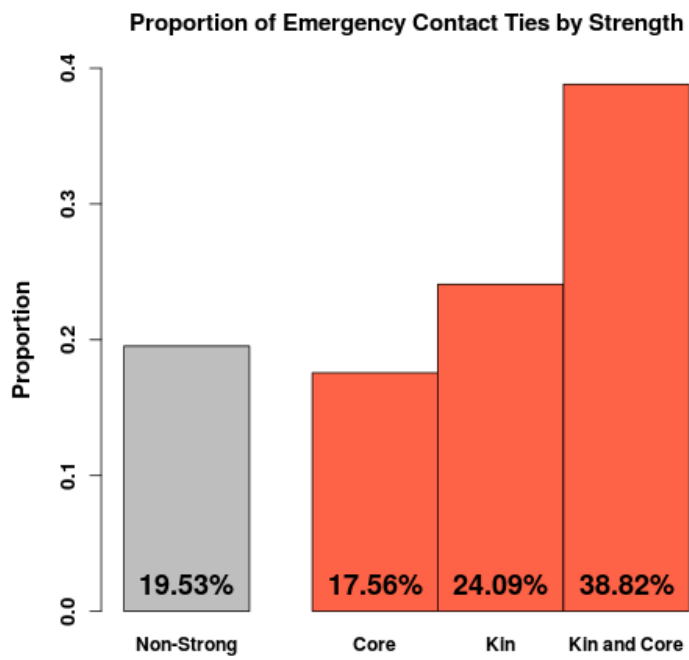


Figure 4.5: Emergency contact ties by strength

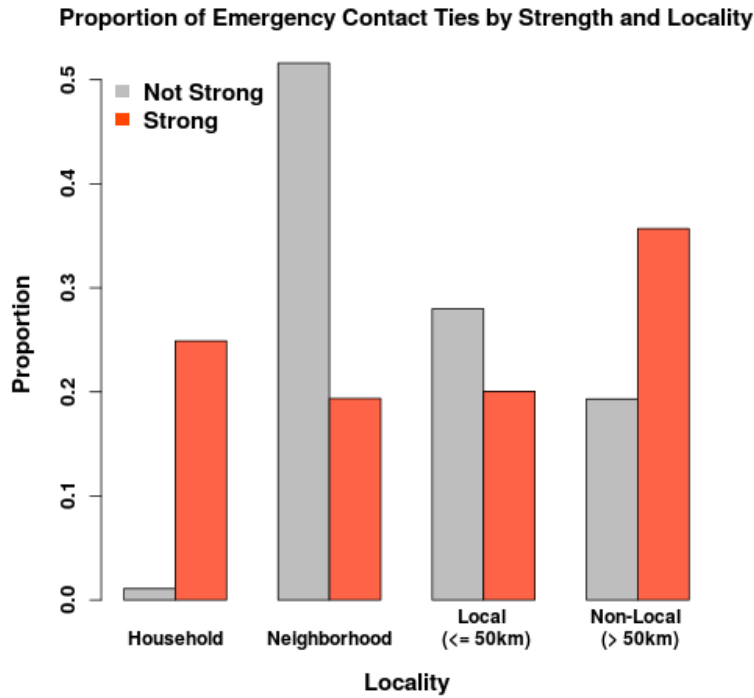


Figure 4.6: Emergency contact ties by strength and locality

Predicting Strong Ties

Next we fit a logistic regression model to explore the odds that an emergency contact tie is as a strong tie. In other words, given that a tie in our sample is an emergency contact tie, what predicts the odds of the tie being strong? The results are presented in Table 4.2. Beginning with geographic context, ties that are farther away tend to be stronger (note here we exclude ties in ego’s household), and alters living in more urbanized areas tend to be stronger emergency contact ties. Emergency contact alters living in ego’s neighborhood, on the other hand, are less likely to be strong.

Ego’s education is significantly associated with emergency contact tie strength. Egos with more education are less likely to have emergency contact ties that are strong. Thus, those with more education are more likely to push information to weaker ties. In terms of alter effects, male alters are less likely to be strong emergency contact ties, whereas white alters

	exp(Estimate)
Intercept	2.233 (0.223)***
log(Distance)	1.232 (0.017)***
log(Alter Pop Density)	1.091 (0.013)***
Ego Education	
Some College	0.731 (0.135)*
Bachelor's	0.655 (0.141)**
Graduate/Professional	0.589 (0.157)***
Ego Male	1.134 (0.090)
Ego White	0.701 (0.197)
Alter Male	0.740 (0.054)***
Alter White	1.273 (0.119)*
In Neighborhood	0.917 (0.011)***
Num. obs.	10677

*** $p < 0.001$, ** $p < 0.01$, * $p < 0.05$

Table 4.2: Model predicting odds of strong emergency contact tie

are more likely to be strong emergency contact ties.

Overall, these results suggest that emergency contact ties tend to be strong (i.e., these ties are often also a kinship and/or core discussion tie as well). However, geographic context plays an important role in tie strength, as alters that are farther away (or in ego's household) or in urban environments are more likely to be strong emergency contact ties. Characteristics related to ego and alter (gender, race, education) also play a role in predicting the probability of a strong emergency contact tie, though these results are subtle and warrant future investigation.

4.4 Discussion

This paper provides important insight into a vital social relation: the social ties one would seek to notify in the event of a disaster or emergency affecting their immediate area. This relation has implications for information diffusion and resource mobilization. In particular,

given that disasters are inherently spatial phenomena, adding a geographic layer of knowledge to the emergency contact relation is crucial to understanding where information may spread in the event of an emergency. Further, by understanding how and where different social groups or individuals in different geographic settings may spread information, we can better understand the social variation in information diffusion.

While the local environment plays an important role with respect to where information would travel in the event of an emergency or disaster, emergency contact ties often expand beyond where the disaster would occur. Thus, it is safe to assume that information could reach individuals far from the source of the disaster, which could have important implications for aid and resources to be delivered to affected regions. Further, ties that are close in a socio-emotional sense are prevalent among an individual's set of potential emergency contacts, even if these ties are not geographically near. Thus, when considering where information would go in the event of an emergency (and the resulting collective behavior or mobilization), we need to think in terms of both geographic and social space surrounding individuals and the disaster.

Chapter 5

Conclusion

The broad aim of this thesis is to provide insight into the complex relationship between interpersonal networks and geographic space. Despite the known importance of space in structuring social relations, much has remained unknown simply due to lack of available data. The ASFS is a first-of-its-kind dataset that allows for the study of multiple network relations on a large geographic scale. As a result, we are able to answer both methodological and substantive questions related to egocentric networks and geography. The following is a discussion of the major contributions of the thesis, as well as its limitations and avenues for future research.

5.1 Contributions

- Geographic context is the strongest predictor of location precision of both respondents and their social ties, with evidence that privacy and vulnerability effects are strong with respect to the precision of ego's own location (but not of their alters). This study lends insight into strategies researchers can use to potentially improve upon location

precision, such as better conveying privacy-protection measures for respondents, particularly among those in vulnerable populations, or greater efforts in obtaining higher precision in rural areas.

- Focusing on an understudied part of the job search process (i.e., an individual’s pool of potential job informants), we find that these ties tend to be multiplex and strong, suggesting that job information is accessible to ego. Further, while job lead ties tend to be local to ego, they are often located far away, information which may or may not be useful to ego given that relocation is costly and not always feasible.
- We find that emergency contacts often extend beyond ego’s local environment, suggesting that information about a local event may indeed reach far distances. Further, these emergency contacts tend to overwhelmingly be strong ties (i.e., kin and or confiding ties), and thus when considering where information might travel in the event of an emergency, it may be more useful to consider where strong contacts are located.

5.2 Limitations

- This thesis utilizes data that comes from a particular geographical and cultural context, the western United States. Thus, we cannot make broad generalizations about any of the findings at the population level, indicating that additional studies must be undertaken, perhaps at a larger scale.
- All three studies examine variables related characteristics of alters. However, due to concerns of respondent fatigue very few demographic questions were asked about alters (we only have alter’s gender, race, and Hispanic ethnicity), thus limiting the conclusions we can draw about alter effects.
- The emergency contact paper only examines the potential for information diffusion

(i.e., where information *could* go in the event of an emergency). Thus, in thinking about implications of the research, we are not able to make decisive conclusions about information diffusion in the event of disasters.

5.3 Future Research

- In further exploring the job lead relation, it would be useful to examine embeddedness of these ties in triads. We have information about the alter-alter relations on two ASFS relations (core discussion and neighborhood safety), and thus integrating whether or not these job lead ties are part of triads (particularly the “strong” core discussion triads) would allow us to comment more directly on Granovetter’s arguments regarding bridging ties.
- As another future avenue for the job leads research, it would be useful to research both the potential ties along with ones that were used (successfully or not). While we do not presently have the data to answer this question, this would perhaps give us the most complete picture of the relation to date, thus connecting potential with activated social capital.
- Because the education effect in the precision paper went the opposite direction as expected (i.e., those with more education tended to have *less* precision location estimates), the combination of this privacy effect along with vulnerability (as shown in Figure 2.4) warrants further research. In particular, I would be interested in doing a more rigorous examination of of factors associated with both privacy and vulnerability to examine what interactions might be at play that may be associated with precision of residential location.

Bibliography

- Almquist, Z. W. and Butts, C. T. (2015). Predicting regional self-identification from spatial network models. *Geographical analysis*, 47(1):50–72.
- Backstrom, L., Sun, E., and Marlow, C. (2010). Find me if you can: improving geographical prediction with social and spatial proximity. In *Proceedings of the 19th international conference on World wide web*, pages 61–70. ACM.
- Bakshy, E., Rosenn, I., Marlow, C., and Adamic, L. (2012). The role of social networks in information diffusion. In *Proceedings of the 21st international conference on World Wide Web*, pages 519–528. ACM.
- Belliveau, M. A. (2005). Blind ambition? the effects of social networks and institutional sex composition on the job search outcomes of elite coeducational and womens college graduates. *Organization Science*, 16(2):134–150.
- Bernard, H. R., Killworth, P., Kronenfeld, D., and Sailer, L. (1984). The problem of informant accuracy: The validity of retrospective data. *Annual review of anthropology*, 13(1):495–517.
- Boessen, A., Hipp, J. R., Smith, E. J., Butts, C. T., Nagle, N. N., and Almquist, Z. (2014). Networks, space, and residents? perception of cohesion. *American journal of community psychology*, 53(3-4):447–461.
- Bound, J., Brown, C., and Mathiowetz, N. (2001). Measurement error in survey data. *Handbook of econometrics*, 5:3705–3843.
- Burke, M. and Kraut, R. (2013). Using facebook after losing a job: Differential benefits of strong and weak ties. In *Proceedings of the 2013 conference on Computer supported cooperative work*, pages 1419–1430. ACM.
- Butts, C. T., Acton, R. M., Hipp, J. R., and Nagle, N. N. (2012). Geographical variability and network structure. *Social Networks*, 34(1):82–100.
- Butts, C. T., Hipp, J. R., Nagle, N. N., Boessen, A., Acton, R. M., Marcum, C. S., Almquist, Z. W., Smith, E. J., and Lickfett, J. (2014). The American Social Fabric Study.
- Caliendo, M., Schmidl, R., and Uhlendorff, A. (2011). Social networks, job search methods and reservation wages: Evidence for germany. *International Journal of Manpower*, 32(7):796–824.

- Centola, D. and Macy, M. (2007). Complex contagions and the weakness of long ties¹. *American journal of Sociology*, 113(3):702–734.
- Coluccia, E. and Louse, G. (2004). Gender differences in spatial orientation: A review. *Journal of environmental psychology*, 24(3):329–340.
- Daraganova, G., Pattison, P., Koskinen, J., Mitchell, B., Bill, A., Watts, M., and Baum, S. (2012). Networks and geography: Modelling community network structures as the outcome of both spatial and network processes. *Social networks*, 34(1):6–17.
- Downs, R. M. and Stea, D. (1973). *Image and environment: Cognitive mapping and spatial behavior*. Transaction Publishers.
- Dustmann, C., Glitz, A., Schönberg, U., and Brücker, H. (2016). Referral-based job search networks. *The Review of Economic Studies*, 83(2):514–546.
- Erickson, B. H., Nosanchuk, T., Mostacci, L., and Dalrymple, C. F. (1978). The flow of crisis information as a probe of work relations. *Canadian Journal of Sociology/Cahiers canadiens de sociologie*, pages 71–87.
- Fitzhugh, S. M., Gibson, C. B., Spiro, E. S., and Butts, C. T. (2016). Spatio-temporal filtering techniques for the detection of disaster-related communication. *Social Science Research*.
- Friedkin, N. (1980). A test of structural features of granovetter’s strength of weak ties theory. *Social Networks*, 2(4):411–422.
- Gärling, T., Lindberg, E., and Mäntylä, T. (1983). Orientation in buildings: Effects of familiarity, visual access, and orientation aids. *Journal of Applied Psychology*, 68(1):177.
- Girvan, M. and Newman, M. E. (2002). Community structure in social and biological networks. *Proceedings of the national academy of sciences*, 99(12):7821–7826.
- Golledge, R. G. and Timmermans, H. (1990). Applications of behavioural research on spatial problems i: cognition. *Progress in Human Geography*, 14(1):57–99.
- Granovetter, M. (1973). The strength of weak ties. *American journal of sociology*, 78(6):1.
- Granovetter, M. (1995). *Getting a job: A study of contacts and careers*. University of Chicago Press.
- Green, G. P., Tigges, L. M., and Diaz, D. (1999). Racial and ethnic differences in job-search strategies in atlanta, boston, and los angeles. *Social science quarterly*, pages 263–278.
- Greenberg, J. and Fernandez, R. M. (2016). The strength of weak ties in mba job search: A within-person test. *Sociological Science*, 3:296–316.
- Guille, A., Hacid, H., Favre, C., and Zighed, D. A. (2013). Information diffusion in online social networks: A survey. *ACM SIGMOD Record*, 42(2):17–28.

- Habinek, J., Martin, J. L., and Zablocki, B. D. (2015). Double-embeddedness: Spatial and relational contexts of tie persistence and re-formation. *Social Networks*, 42:27–41.
- Hart, T. C., Rennison, C. M., and Gibson, C. (2005). Revisiting respondent fatigue bias in the national crime victimization survey. *Journal of Quantitative Criminology*, 21(3):345–363.
- Hess, S., Hensher, D. A., and Daly, A. (2012). Not bored yet—revisiting respondent fatigue in stated choice experiments. *Transportation research part A: policy and practice*, 46(3):626–644.
- Hipp, J. R., Butts, C. T., Acton, R., Nagle, N. N., and Boessen, A. (2013). Extrapolative simulation of neighborhood networks based on population spatial distribution: Do they predict crime? *Social Networks*, 35(4):614–625.
- Holguín-Veras, J., Taniguchi, E., Jaller, M., Aros-Vera, F., Ferreira, F., and Thompson, R. G. (2014). The tohoku disasters: Chief lessons concerning the post disaster humanitarian logistics response and policy implications. *Transportation research part A: policy and practice*, 69:86–104.
- Hui, C., Tyshchuk, Y., Wallace, W. A., Magdon-Ismail, M., and Goldberg, M. (2012). Information cascades in social media in response to a crisis: a preliminary model and a case study. In *Proceedings of the 21st International Conference on World Wide Web*, pages 653–656. ACM.
- Kitchin, R. M. (1994). Cognitive maps: What are they and why study them? *Journal of environmental psychology*, 14(1):1–19.
- Krackhardt, D., Nohria, N., and Eccles, B. (2003). The strength of strong ties. *Networks in the knowledge economy*, 82.
- Kulshrestha, J., Kooti, F., Nikraves, A., and Gummadi, P. K. (2012). Geographic dissection of the twitter network. In *ICWSM*.
- Langlois, S. (1977). Les réseaux personnels et la diffusion des informations sur les emplois. *Recherches sociographiques*, 18(2):213–245.
- Lin, N. (1981). *Social resources and instrumental action*. State University of New York, Department of Sociology.
- Lin, N. and Dumin, M. (1986). Access to occupations through social ties. *Social networks*, 8(4):365–385.
- Linzer, D. A. and Lewis, J. (2007). polca: Polytomous variable latent class analysis. r package version 1.1.
- Linzer, D. A. and Lewis, J. B. (2011). polca: An r package for polytomous variable latent class analysis. *Journal of Statistical Software*, 42(10):1–29.

- Lumley, T. et al. (2004). Analysis of complex survey samples. *Journal of Statistical Software*, 9(1):1–19.
- Mabe, P. A. and West, S. G. (1982). Validity of self-evaluation of ability: A review and meta-analysis. *Journal of applied Psychology*, 67(3):280.
- Madden, M. and Rainie, L. (2015). Americans attitudes about privacy, security and surveillance. *Pew Research Center*, May, 20.
- Manoj, B. S. and Baker, A. H. (2007). Communication challenges in emergency response. *Communications of the ACM*, 50(3):51–53.
- Marmaros, D. and Sacerdote, B. (2002). Peer and social networks in job search. *European economic review*, 46(4-5):870–879.
- Marsden, P. V. (1990). Network data and measurement. *Annual review of sociology*, 16(1):435–463.
- McPherson, M., Smith-Lovin, L., and Cook, J. M. (2001). Birds of a feather: Homophily in social networks. *Annual review of sociology*, pages 415–444.
- Menon, G. (1994). Judgments of behavioral frequencies: Memory search and retrieval strategies. In *Autobiographical memory and the validity of retrospective reports*, pages 161–172. Springer.
- Montello, D. R. and Friendschuh, S. (2005). Cognition of geographic information. *A research agenda for geographic information science*, pages 61–91.
- Munton, A. G. (1990). Job relocation, stress and the family. *Journal of Organizational behavior*, 11(5):401–406.
- Papachristos, A. V., Hureau, D. M., and Braga, A. A. (2013). The corner and the crew: the influence of geography and social networks on gang violence. *American sociological review*, page 0003122413486800.
- Preciado, P., Snijders, T. A., Burk, W. J., Stattin, H., and Kerr, M. (2012). Does proximity matter? distance dependence of adolescent friendships. *Social networks*, 34(1):18–31.
- Prince, S. A., Adamo, K. B., Hamel, M. E., Hardt, J., Gorber, S. C., and Tremblay, M. (2008). A comparison of direct versus self-report measures for assessing physical activity in adults: a systematic review. *International Journal of Behavioral Nutrition and Physical Activity*, 5(1):56.
- Rainie, L., Kiesler, S., Kang, R., Madden, M., Duggan, M., Brown, S., and Dabbish, L. (2013). Anonymity, privacy, and security online. *Pew Research Center*.
- Richardson, R., Erickson, B. H., and Nosanchuk, T. A. (1979). Community size, network structure, and the flow of information. *Canadian Journal of Sociology/Cahiers canadiens de sociologie*, pages 379–392.

- Roseman, C. C. (1971). Migration as a spatial and temporal process. *Annals of the Association of American Geographers*, 61(3):589–598.
- Saarinen, T., MacCabe, C., and Morehouse, B. (1988). Sketch maps of the world as surrogates for world geographic knowledge. In *Skills in Geographical Education Symposium '88 Proceedings*, (Ed. R. Gerber, J. Lidmore) C, volume 2, pages 628–664.
- Schaefer, D. R. (2012). Youth co-offending networks: An investigation of social and spatial effects. *Social networks*, 34(1):141–149.
- Schmutte, I. M. (2015). Job referral networks and the determination of earnings in local labor markets. *Journal of Labor Economics*, 33(1):1–32.
- Seidel, A. (1982). Way-finding in public spaces: The dallas/fort worth, usa airport. In *20th International Congress of Applied Psychology*, pages 15–31.
- Sherry, B., Jefferds, M. E., and Grummer-Strawn, L. M. (2007). Accuracy of adolescent self-report of height and weight in assessing overweight status: a literature review. *Archives of pediatrics & adolescent medicine*, 161(12):1154–1161.
- Spiro, E. S., Almquist, Z. W., and Butts, C. T. (2016). The persistence of division geography, institutions, and online friendship ties. *Socius: Sociological Research for a Dynamic World*, 2:2378023116634340.
- Sutton, J. N. (2010). *Twittering Tennessee: Distributed networks and collaboration following a technological disaster*. Citeseer.
- Takhteyev, Y., Gruzd, A., and Wellman, B. (2012). Geography of twitter networks. *Social networks*, 34(1):73–81.
- Toriumi, F., Sakaki, T., Shinoda, K., Kazama, K., Kurihara, S., and Noda, I. (2013). Information sharing on twitter during the 2011 catastrophic earthquake. In *Proceedings of the 22nd International Conference on World Wide Web*, pages 1025–1028. ACM.
- United States Department of Labor, Bureau of Labor Statistics (BLS) (2016). Consumer Price Index-All Urban Consumers. <<https://www.bls.gov/bls/glossary.htm#L>>. Accessed (June 18, 2018).
- Wei, J., Zhao, D., Yang, F., Du, S., and Marinova, D. (2010). Timing crisis information release via television. *Disasters*, 34(4):1013–1030.
- Weimann, G. (1983). The strength of weak conversational ties in the flow of information and influence. *Social Networks*, 5(3):245–267.
- Yoo, E., Rand, W., Eftekhari, M., and Rabinovich, E. (2016). Evaluating information diffusion speed and its determinants in social media networks during humanitarian crises. *Journal of Operations Management*, 45:123–133.

- Zhu, J., Xiong, F., Piao, D., Liu, Y., and Zhang, Y. (2011). Statistically modeling the effectiveness of disaster information in social media. In *Global Humanitarian Technology Conference (GHTC), 2011 IEEE*, pages 431–436. IEEE.
- Zinoviev, D. (2011). Information diffusion in social networks. *Social Networking and Community Behavior Modeling: Qualitative and Quantitative Measures: Qualitative and Quantitative Measures*, page 146.