Research Report – UCD-ITS-RR-15-27

Social Networks and Travel Behavior: An Investigation into the Role of Social Influence in the Transportation Mode Choices of Students

December 2015

Susan Pike

Institute of Transportation Studies ◦ University of California, Davis
1605 Tilia Street ◦ Davis, California 95616
PHONE (530) 752-6548 ◦ FAX (530) 752-6572
www.its.ucdavis.edu

Social Networks and Travel Behavior: An Investigation into the Role of Social Influence in the Transportation Mode Choices of Students

By

SUSAN CAROL PIKE
B.S. (University of California, Davis) 2006
M.S. (University of California, Davis) 2012

DISSERTATION

Submitted in partial satisfaction of the requirements for the degree of

DOCTOR OF PHILOSOPHY

in

Transportation Technology and Policy

in the

OFFICE OF GRADUATE STUDIES

of the

UNIVERSITY OF CALIFORNIA

DAVIS

Approved:

____

Mark Lubell, Chair

____

Susan Handy

____

Robert Hijmans

Committee in Charge

2015

Social Networks and Travel Behavior: An Empirical Investigation into the Influence of Ego-Networks on the Transportation Mode Choices of Students

ABSTRACT

The research presented here explores the role of social networks in transportation mode choice. Social networks serve as a foundation for multiple social processes, such as cooperation, resource sharing and social influence. The focus of this study is social influence, whereby the knowledge, behaviors and/or opinions of one individual affect those of another. Through an online survey of students at the University of California, Davis, I collected information about students' transportation decisions for campus travel and their social networks. For each participant as an ego, I gathered information about their ego-network including up to five alters or contacts. The ego-networks, representing a subset of the social network of each respondent, were analyzed for similarities in transportation mode choice. Each respondent's mode choice was also compared to the mode use of geographic neighbors. Individuals tend to make transportation choices similar to both those to whom they are socially connected and those with whom they share geographic proximity. The first paper in this study presents evidence that although similarities in transportation mode choice occur both for socially connected individuals as well as geographic neighbors, social and neighborhood effects are not the same. To account for the possibility that similar commute characteristics simultaneously affect socially connected individuals, in the second paper presented here, I use an instrumental variables approach and find evidence of social influence even when accounting for the commute characteristics of both the alters and ego in ego-networks. In the third paper, I demonstrate that social influence does not affect all individuals equally; in particular social influence related to biking has a smaller effect on those with longer or

very short commute distances. This study improves our understanding of social influences in travel behavior and the results may inform sustainable transportation programs and policies.

INTRODUCTION

Social network theory recognizes that decisions are made in a social context, and social relationships may directly affect the costs and benefits of different choices. Recent research in the field of network science has demonstrated that social networks profoundly influence individual behavior ranging from political decisions (Klofstad, McClurg, and Rolfe 2009) to diet and exercise (Fowler and Christakis 2007). As in other behavioral research, how social networks affect behavior is becoming a central topic in travel behavior research and has important implications for transportation planning and the overall design of sustainable transportation policies

Many travel demand management programs already incorporate various forms of social influence into their design. The UC Davis goClub periodically asks members to invite their friends and colleagues to sign up and pledge to use an alternative to driving alone for their commutes to campus. May is Bike Month, an annual month-long drive to increase bicycle commuting, utilizes social tools such as “challenge a friend” and “share your accomplishments” to instigate friendly competition among participants as they log bike miles during the event. Private operators also use social mechanisms to increase ridership; Amtrak offers a socially based discount. The “Take 5” campaign allows riders to purchase up to five $5 tickets with the purchase of one full-fare ticket to encourage groups to travel together by train along the California-based Capitol Corridor line. All of these programs incorporate social mechanisms to increase the use of alternative modes of transportation. Programs like these will benefit from a deeper understanding of how social influence affects travel behavior.

This project explores the influence of social networks on transportation mode choice and addresses several fine points within this area of inquiry. First, the effects of social influence likely depend on how social networks are defined. Network connections may be explicitly social or defined by other relationships such as geographic neighborhoods; different types of relations may exert different types of social influence. How networks are defined also has implications for what is actually measured. The measurement of social influence is further complicated by the possibility that socially connected individuals make the same choices because of social influence, but also because they share environmental characteristics that impact their commute. Finally, the effect of social influence may be impacted by commute characteristics. For example, in the case of social influence on biking it is likely that individuals who have very long (or very short) commute distances are less affected by social influence. For these individuals, the differences between modes are more extreme, and the convenience of walking or driving is likely to take precedent over social influence.

BACKGROUND

Research in network analysis aims to define the structure of networks, identify relationships among members and determine the effects of networks on behavior and other outcomes (Wasserman and Faust 1994). There are numerous ways to study social networks including whole networks – whereby a network is defined by some group of interest, such as all students in a classroom. Another approach is a relational network; individuals are connected if they are both linked to an event or have an association with a group or place, for example a public meeting. Another approach to studying networks, and the primary means of analysis used here, is ego-network analysis. In this approach a sample of individuals is selected. The sampled individuals are *egos*, who report on their own personal network of contacts or *alters*. Ego-network based research may focus on the egos, and/or the ego-networks as the unit of analysis.

This project brings a social network perspective to the study of transportation mode choice and investigates the role of social influence among other factors important to transportation mode choice. To study factors affecting transportation mode choice, researchers traditionally use discrete choice models with a utility maximization framework, where dependent variables are a set of distinct outcomes (Ben-Akiva and Lerman 1985). Factors which have previously been identified as important in travel behavior include socio-demographics, trip and mode characteristics, the built environment (Ewing and Cervero 2010), attitudes (Mokhtarian and Salomon 1997) and information and communication technology (Haraldsson 2003).

Social Networks and Travel Behavior

Social networks are important to travel behavior in at least three ways; social networks contribute to trip generation, second, social network connections provide access to transportation resources, and third social networks act as pathways for social influence. As people make trips to participate in social activities, social networks generate travel (Carrasco and Miller 2006 and 2009). The types of social ties affect travel behavior when social interactions are sometimes replaced by communication over the internet (Mok et al. 2010), and the frequency of social interactions may be dependent on network structure and composition and with whom activities take place (Carrasco et al. 2008).

Transportation resources may be acquired through social networks. In elderly populations, those with active social networks and to some extent, those living in retirement homes are more likely to use ride-sharing (Silvis and Niemeier 2009). Lovejoy and Handy (2011) found that some Mexican immigrants have more access to rides and vehicles than others and that although the type of ties may matter, geographic and temporal factors are also relevant, as well as having either a car or ability to drive. Expanded networks (beyond closer personal networks) may provide more transportation resources (Lovejoy and Handy 2011).

Social Influence and Transportation Mode Choice

There is a growing literature exploring social influence related to transportation and travel behavior. Wilton, Páez, and Scott (2011) explore social influences on telecommuting through qualitative interviews. Social factors including interactions with co-workers at work (which is beneficial or annoying) and a culture around telecommuting are relevant in the choice to telecommute (Wilton, Páez, and Scott 2011). In a spatially autoregressive logit mode choice model in New York City, using the 40 nearest neighbors, Goetzke (2008) finds neighborhood network effects influence the use of transit. Residential districts and socioeconomic groups exhibit interdependence of decision making in transportation mode choice (Dugundji and Walker 2005). Cultural context affects bicycle ridership (Goetzke and Rave 2010) and has been characterized as a social network affect, though it is broadly defined. Perceptions of and attitudes towards new technologies such as electric vehicles are related to social influence through social network exposure to these new technologies (Axsen and Kurani 2011).

With few exceptions (such as Wilton et al. 2011 and Scott et al. 2012) the majority of studies exploring social influences in transportation mode choice represent social networks as neighborhoods and other broadly defined social groups. There are, however, many examples of explicit social interactions in studies of activity-based travel (for example Carrasco et al. 2006, Larsen et al. 2008). In this study, social influence is defined as the mode use of contacts with explicitly social, as opposed to geographic relationships, therefore the measured effects are classified as social, rather than neighborhood or cultural.

RESEARCH QUESTIONS AND EXPECTED OUTCOMES

Social networks serve as a foundation for multiple social processes, such as cooperation, resource sharing and social influence. The focus of this study is social influence, whereby the knowledge, behaviors and/or opinions of one individual affect those of another. In the three papers summarized below I investigate several details of this research area.

Paper 1: Social Reference Groups and Transportation Mode Choice

There are many ways that social networks may be defined, and many ways researchers may measure the presence of connections between individuals. In particular, connections may be explicitly social, or they may be based on shared socio-economic status, or geographic nearness. I test the hypothesis that the way researchers define social connections has implications for the measurement of the effects of social influence between connected individuals.

Paper 2: Geography and Social Influence

I find that the more contacts or alters within an ego's network using a particular mode of transportation (regardless of which), the more likely it is that the ego also uses that mode. This may be due to social influence, but may also be due to other factors. Individuals may make the same choices as those to whom they are socially connected because they are influenced by their social contacts or because they face similar commute environments. Shared environmental characteristics may include similar levels of access to transit, similar commute distance or quality of bicycle infrastructure. I hypothesize that even when taking shared environment into account, ego mode choice is influenced by the transportation choices of their alters.

Paper 3: External Impacts on the Effect of Social Influence

Factors which can be considered constraints on individual mode choice, such as costs or travel time, impact the extent to which social influence has an effect. Individuals with more options in their choice set are likely to be more affected by social influence than those with few transportation mode options. The focus of this paper is the interaction between commute distance and social influence. The primary hypothesis is individuals with very short or very long commute distances are less affected by social influence than individuals with a medium commute distance; those with very short or long commute distances respond more to factors related to commute distance, than factors related to social influence.

REFERENCES

Axelrod, Lawrence J., and Darrin R. Lehman. 1993. "Responding to Environmental Concerns: What Factors Guide Individual Action?" *Journal of Environmental Psychology* 13 (2): 149–159.

Axsen, John and Kenneth S. Kurani. 2011. Interpersonal influence within car buyers' social networks: applying five perspectives to plug-in hybrid vehicle drivers *Environment andPlanning A*. 2011 Volume: 340, Issue: 1, Pages: 1-26

Ben-Akiva, Moshe, and Lerman. 1985. *Discrete Choice Analysis*. MIT Press.

Carrasco, Juan Antonio, Bernie Hogan, Barry Wellman, and Eric J. Miller. 2008. "Collecting Social Network Data to Study Social Activity-travel Behavior: An Egocentric Approach." *Environment and Planning. B, Planning & Design* 35 (6): 961.

Carrasco, Juan Antonio, and Eric J. Miller. 2006. "Exploring the Propensity to Perform Social Activities: a Social Network Approach." *Transportation* 33 (5): 463–480.

Carrasco, Juan-Antonio, and Eric J. Miller. 2009. "The Social Dimension in Action: A Multilevel, Personal Networks Model of Social Activity Frequency Between Individuals." *Transportation Research Part A: Policy and Practice* 43 (1): 90–104.

Carrell, Scott E., Richard L. Fullerton, and James E. West. 2008. "Does Your Cohort Matter? Measuring Peer Effects in College Achievement". National Bureau of Economic Research. http://www.nber.org/papers/w14032.

Dugundji, Elenna R., and Joan L. Walker. 2005. "Discrete Choice with Social and Spatial Network Interdependencies: An Empirical Example Using Mixed Generalized Extreme Value Models with Field and Panel Effects." *Transportation Research Record: Journal of the Transportation Research Board* 1921 (1): 70–78.

Ewing, Reid, and Robert Cervero. 2010. "Travel and the Built Environment." *Journal of the American Planning Association* 76 (3) (June 21): 265–294. doi:10.1080/01944361003766766.

Fowler, James H. and Christakis, Nicholas A. 2007. The Spread of Obesity in a Large Social Network over 32 Years. *New England Journal of Medicine* 2007; 357:370-9.

Fowler, J. H, and N. A Christakis. 2008. “Dynamic Spread of Happiness in a Large Social Network: Longitudinal Analysis over 20 Years in the Framingham Heart Study.” *BMJ* 337 (dec04 2) (December 4): a2338–a2338. doi:10.1136/bmj.a2338.

Goetzke, F., and T. Rave. 2010. “Bicycle Use in Germany: Explaining Differences Between Municipalities with Social Network Effects.” *Urban Studies* 48 (2) (June 7): 427–437. doi:10.1177/0042098009360681.

Goetzke, Frank. 2008. “Network Effects in Public Transit Use: Evidence from a Spatially Autoregressive Mode Choice Model for New York.” *Urban Studies* 45 (2): 407–417.

Haraldsson, Mattias. 2003. *VTI notat 48A 200:.Travel Behavior and Land Use in the information society: Literature Review and Problem Discussion.* Swedish National Road and Transport Research Institute.

Klofstad, C. A. 2007. “Talk Leads to Recruitment: How Discussions About Politics and Current Events Increase Civic Participation.” *Political Research Quarterly* 60 (2) (June 1): 180–191. doi:10.1177/1065912907301708.

———. 2009. “Civic Talk and Civic Participation: The Moderating Effect of Individual Predispositions.” *American Politics Research* 37 (5) (August 14): 856–878. doi:10.1177/1532673X09333960.

Larsen, Jonas, Urry, John and Axhausen, Kay. 2008. Coordinating Face to Face Meetings in Mobile Network Societies. *Information, Communication and Society*. Vol. 11, No. 5, August 2008 pp. 640-658.

Lovejoy, Kristin, and Susan Handy. 2011. “Social Networks as a Source of Private-vehicle Transportation: The Practice of Getting Rides and Borrowing Vehicles Among Mexican Immigrants in California.” *Transportation Research Part A: Policy and Practice* 45 (4) (May): 248–257. doi:10.1016/j.tra.2011.01.007.

Mok, D., B. Wellman, and J. Carrasco. 2010. “Does Distance Matter in the Age of the Internet?” *Urban Studies* 47 (13) (November 16): 2747–2783. doi:10.1177/0042098010377363.

Mokhtarian, Patricia; & Ilan Salomon. Modeling the Desire to Telecommute: the Importance of Attitude Factors in Behavioral Models. Transportation Research Part A. Volume 31, Issue 1, January 1997, Pages 35–50.

Sacerdote, Bruce. 2001. “Peer Effects with Random Assignment: Results for Dartmouth Roommates.” *The Quarterly Journal of Economics* 116 (2): 681–704.

Silvis, Julia, and Debbie Niemeier. 2009. “Social Network and Dwelling Characteristics That Influence Ridesharing Behavior of Seniors.” *Transportation Research Record: Journal of the Transportation Research Board* 2118 (-1) (December 1): 47–54. doi:10.3141/2118-07.

Wasserman, Stanley, and Katherine Faust. *Social Network Analysis: Methods and Applications.* Cambridge University Press, Cambridge, 1994.

Wilton, Robert D., Antonio Páez, and Darren M. Scott. 2011. “Why Do You Care What Other People Think? A Qualitative Investigation of Social Influence and Telecommuting.” *Transportation Research Part A: Policy and Practice* 45 (4) (May): 269–282. doi:10.1016/j.tra.2011.01.002.

Travel Mode Choice, Social and Spatial Reference Groups;
A Comparison of Two Formulations

ABSTRACT

This paper investigates social influence on mode choice using two methods for defining reference groups; the respondents' social contacts, and respondents' geographic neighbors. The use of social network analysis builds on traditional models of travel behavior that rely on individualistic assumptions about decision making rather than the social context in which travel behavior takes place. Mode choice is explored using traditional socio-economic, attitudinal and trip characteristic variables. In order to address social influence, ego-centric social network factors, consisting of the behaviors of social contacts (alters), are incorporated into models to investigate whether alter behaviors influence ego transportation mode choice. Second, using spatially defined reference groups, neighborhood mode use variables are considered for their potential influence on mode choice. Findings show that for some modes, the choices of the ego-network and spatial neighborhood have similar effects while for other modes the effects on mode choice are different. Neighborhood characteristics such as access to transit conceivably account for the greatest portion of the correlation in mode choice between the respondent and their neighbors. In ego-networks, shared environment likely accounts for some of the observed similarities in behavior but social processes, including social influence, are thought to contribute to similarities in behavior as well. Findings suggest that ego-network processes related to mode choice are dissimilar from those of spatial neighborhoods.

INTRODUCTION

Recent research in the field of network science has demonstrated that social networks profoundly influence individual behavior ranging from political decisions (Klofstad et al. 2009) to diet and exercise (Fowler and Chirstakis 2008). Social networks also provide pathways for social influence, where the travel choices of one person affect the choices of individuals to whom they are socially connected. As in other behavioral research, how social network processes affect travel behavior is becoming a central topic in transportation research. Understanding social influences in travel behavior informs broader questions related to how social network-based policies and programs may be utilized to affect behavior changes for congestion relief and transportation demand management.

Social networks act as avenues for the diffusion of information; for example social groups may discuss alternative modes, changes or improvements to transit service, or new infrastructure. Social influence may also occur through normalization of behaviors, such as the use of a particular mode of transportation or by reinforcing and reaffirming behaviors. To find new and innovative solutions to reduce transportation emissions and support the use of alternative modes it is essential to better understand the mechanisms that influence individual travel behavior. Social influences may be a powerful tool that could be incorporated into campus-wide, local or regional policies in order to promote the use of sustainable means of transportation.

A growing body of research addresses social influences and social networks in transportation (Larsen et al. 2008, Wilton et al. 2011, Páez and Scott 2007 and others). How social effects and the groups within which social influence occurs are defined are important questions for consideration. I investigate two methods for defining social reference groups; the ego-networks made up of respondents' social contacts, and the networks of respondents' spatial or geographic neighbors. Both the ego-network and the spatial reference group have the potential to influence or to be related to mode choice. The results presented here highlight the differences between spatially defined reference group effects and socially defined reference group effects. Exploring the means by which social influence may be identified and how social reference groups are defined, this paper contributes to the study of these processes, as well as how they may be incorporated into transportation policy and programs.

BACKGROUND: SOCIAL NETWORKS AND TRAVEL BEHAVIOR

Travel behavior researchers have developed a strong understanding of factors that contribute to individual transportation choices. However, these models have left a portion of the influences unexplained and have generally looked at individual travel behavior as an atomized choice made without respect to the influences of social relationships. Travel behavior research typically utilizes a utility maximization framework and predominantly relies on trip characteristics and individual socio-demographics to explain and predict travel behavior. Social network theory recognizes that decisions are made in a social context, and social relationships may directly affect the costs and benefits of different choices, such as transportation mode; for example, by making it easier to find information, or through the establishment of behavioral norms. Thus, transportation research that ignores social networks is likely to miss a number of important variables; the nature of and the extent to which social network variables impact transportation behavior is a matter for empirical research like that presented here.

Social networks are involved in many aspects of transportation including trip generation; individuals make trips in order to spend time with others in their social networks. Thus, social networks influence daily activity patterns in ways that may be used to predict travel behavior and trip generation (Han et al. 2011). At the same time the frequency of social interactions, and the associated amount of travel may be dependent on network structure and composition as well as with whom activities take place (Carrasco et al. 2008). Information and Communication Technologies (ICT) coupled with increased mobility influence social travel through the continual coordination of and last-minute changes to plans (Larsen et al. 2008). ICT could also reduce social travel in some cases when social interactions are sometimes replaced by communication over the internet (Mok et al. 2010).

In addition to affecting trip generation, social networks also provide transportation resources. In some elderly populations those with active social networks, and to some extent those living in retirement homes may be more likely to use ride-sharing (Silvis and Niemeier 2009). Other populations, such as immigrants depend on each other for transportation resources and although the type of social ties matter, geographical and temporal factors are also relevant, as well as having either a car or ability to drive (Lovejoy and Handy 2011). Expanded networks (beyond close personal networks) may provide more resources (Lovejoy and Handy 2011), highlighting the importance of weak ties (Granovetter 1973). There is also evidence that transportation mode choices affect social interactions; high levels of automobile use can limit social interactions (Farber and Páez 2009).

Reference Groups and Mode Choice

Recent work may be categorized by the type of reference group used in analysis. In some cases researchers consider an explicit social reference group – a network made up of individuals and their reported social connections. The connections in these networks are self-defined (by network members). In other cases a general or spatial reference group, such as neighbors, or peers (with respect to socio-demographics) has been used to estimate social influence in mode choice.

The connections in these assumed networks are exogenously defined (by the researcher). In this paper I aim to further the understanding of the ways reference group definition changes what is being studied and which types of reference groups are more suited to the study of social influence.

Looking at explicit social reference groups or social networks, Wilton et al. (2011) find factors such as learning and *validation* (a form of social reinforcement) from peers and co-workers about experiences have an effect on the choice to telecommute. Social factors also include interactions with co-workers at work (that can be either be beneficial *or* distracting) and a culture around telecommuting in some instances (Wilton et al. 2011) and Scott et al. (2012) also find that social effects may play an important role in the decision to telecommute, and the characteristics of relationships affect the relevance of social influence. Páez and Scott (2007) simulate a panel study and show the first-wave behaviors within social networks affect the behaviors of individuals in the second wave. In all of these cases the social reference group is defined as a group of networked individuals; observed, reported or simulated ties are present between individuals.

Spatial reference groups are also likely to have an influence on mode choice. In a spatially autoregressive logit mode choice model in New York City, using the 40 nearest neighbors, Goetzke (2008) finds neighborhood network effects influence the use of transit (Goetzke2008). Dugundji and Walker (2005) consider residential district, socioeconomic group and postal code in discrete choice models with social interdependence of decision making and find social influence occurs to some extent (Dugandki and Walker 2005. Further, differences in the mode share of bicycling among German cities can be attributed to a city-level cultural component characterized as a social network effect, though it is broadly defined (Goetzke and Rave 2010).

DATA AND METHODS

Surveys

In coordination with an annual Campus Travel Survey (CTS) at the University of California, Davis in the 2012-13 academic year a Social Networks and Travel Survey (SNTS) was administered to a sample of students. The CTS was sent to 28,838 members of the UC Davis Campus in October 2012, and resulted in 3,982 useable responses; a response rate of 13.8% (Driller, Brigitte 2013). At the end of the survey students were presented with an option to participate in the SNTS at a later point in time, and were asked to provide an email address to which the survey invitation could be sent. Of the 3,171 students who participated in the CTS, 56%, a total of 1,789 individuals, indicated an interest in the SNTS. The UC Davis Office of the President (UCOP) conducted its biennial Cost of Attendance Survey (COAS) during the same time as the Social Networks and Travel Survey; March 2013. We were required to exclude any COAS participants from our sample so from the 1,789 students interested in our survey 1,642 were sent invitations to participate. Ultimately 962 students completed the survey (an initial response rate of 59%) and 692 provided enough information to be included in analysis (22% of the initial 3,171 students who participated in the CTS).

The survey aimed to capture both the key variables of interest to this research, as well as variables known to be important factors in travel behavior such as socio-demographics, trip and mode characteristics, the built environment (Ewing and Cervero 2010 and Mokhtarian and Cao 2008), and attitudes (Mokhtarian and Salomon 1997). Some of this information was collected in the CTS and linked to SNTS responses. In the SNTS respondents were first asked what transportation modes are available to them, what mode of transportation they usually use for travel to campus, and why they consider some modes unavailable. Next, the survey asked about the importance (on a 5 point scale) of 18 factors, including social factors in the choice of their usual mode. The survey also included a section designed to elicit a social reference group for each participant.

Reference Groups

Ego-networks

In the "name generator" respondents were asked to identify contacts within their social networks. While there are numerous ways to study social networks (see Wasserman and Faust 1994), the primary method used here is ego-network analysis. In this approach a sample of individuals is selected. The sampled individuals are the *egos*, who are connected to their own personal network (which may be defined in multiple ways) of contacts or *alters*. In our survey, the name generator asked respondents to think about their social circle, including "people with whom you live, work or attend class, socialize or participate in activities etc. or people you speak with over the phone or internet." Spaces were provided for the ego to name up to five alters, with whom they had different types of interactions over the past six months.

Three versions of the name generator were included in the survey with one version randomly assigned to each respondent. In all three versions, the social circle is defined in the same way, but respondents were asked to list the names of different types of contacts in order to investigate the effects of name-generator wording on ego-network characteristics. (For a discussion of possible effects see Campbell and Lee 1991), Bernard et al. 1990 , and Klofstad, McClurg, and Rolfe 2009). The first name generator requested the names of "any five people who have been in your social circle over the past six months." The second version requested the names of "the five contacts you have had the most frequent regular interaction with over the past six months." The third asked for "five people in your social circle, with whom you spoke about transportation in the past six months."

Once alters were named, respondents were asked about their relationships with each alter including how long they've known each other, and how close they are. They were also asked the usual commute mode of transportation for each alter and where each alter lives in relation to the ego. The egos provided all information about alters and their ego-network. At the end of the survey, contact information for each alter was requested for a follow up snowball survey to gather self-reported information from each alter. There was very limited response to the snowball survey (and limited contact information provided by the egos). However, there are 90 alters who participated in either the snowball survey or, by chance, the CTS and reported their own usual mode of transportation. These reports were compared to the ego-reported usual mode for these alters. Egos correctly identified alter mode about 80% of the time. For this reason, and because the ego may be as influenced by what he/she thinks alters are doing as much as by what the alters are actually doing, ego accounts of alter behaviors are presumed to be correct. Egos also provided information about relationships among their alters. This information was used to calculate the ego-network density; the number of observed/reported ties out of the total possible number of ties. In a network of n individuals there are $2(n-1)$ possible ties. In this study most ego-networks include 6 individuals (one ego and up to five alters) and have a total of 10 possible ties.

Neighborhood Reference Groups

For each respondent, spatial reference groups were also identified. Respondents gave the cross streets for an intersection near their home address and these intersections were geocoded. As with the name generator – and in alignment with the overall theme of identifying a suitable means to define reference groups – an exploratory approach was utilized to determine the appropriate geographic scale for a spatial reference group. Selecting neighbors using an arbitrary distance may not pick up on spatial patterns that are scale dependent. If the geographic neighborhoods are too small, they may include too few individuals and poorly represent the spatial context. If the neighborhoods are too large, each individual's neighborhood will be equal to (or very similar to) that of every other person, and at the extreme, would include the entire sample. In order to address these issues related to neighborhood definition, 100 distinct neighborhood sizes were considered.

Each neighborhood was generated using a circle with radius = d with the respondent's residential cross streets at the center. The radii lengths, or distances from the respondent's cross streets ranged from 250 feet (2-3 houses in each direction) to 25,000 feet (about 5 miles). For each neighborhood size, neighborhood networks were exogenously defined using all Campus Travel Survey participants within the given distance of the respondent as neighborhood/network members. Mode choices for neighbors were identified and the percentage of neighbors using each mode of transportation were computed. As the distance used to define a neighborhood is increased the percentage of neighbors using each mode becomes closer to the population share of those using that mode. There is not a clearly identifiable distance that corresponds to a best definition of neighborhood, so models were estimated using each neighborhood size. Results for a small and a medium neighborhood are presented here.

Reference group variables are incorporated into models while controlling for variables typically considered in mode choice analysis. Analysis such as that presented here is faced with the challenge that there are multiple explanations for correlations in behavior within reference groups. There is likely social influence within ego-networks and/or within neighborhoods. At the same time, individuals within the same neighborhood face a similar choice context; for example similar commute distance. This is also true for *some* individuals within the same ego-network, though to a lesser extent, since ego-networks are made up of individuals from varying neighborhoods, different towns, etc. It is also possible that neighbors or members of the same ego-network share characteristics that drew them towards their neighborhood, or social circle. These same characteristics that caused individuals to be friends or neighbors may also predispose them to make similar transportation choices.

The present paper focuses on methods for defining reference groups; either by ego-network or by neighborhood. Related work that aims to address the challenges outlined above includes discrete choice models, (Brock and Durlauf 2001), the linear in means model (Manski 1993) and the spatial autoregressive model (Fotheringham et al. 2000). Lee (2007) and Bramoullé et al. (2009) provide some discussion of how these two models are used in social contexts as well as how they are functionally related. The linear in means model has been used in linear (An 2011, Bramoullé et al. 2009) and discrete applications (Brock and Durlauf 2001). Lee (2007) provides a detailed example of the estimation of peer effects, addressing endogeneity, using the spatial auto regressive model.

ANALYSIS AND OUTCOMES

Ego-networks are defined as the set of alters the ego names in the survey and the relationships between them. To analyze whether social influence affects mode choice after controlling for factors typically used in travel behavior research, the behaviors of the alters are used as explanatory variables in model estimation of the mode choice of the ego. Since the differences in name generators could affect characteristics of the ego-networks, ego-network properties are compared with respect to the name generator questions (Table 1).

There are 692 respondents who reported the name(s) of at least one contact. Roughly equal numbers of respondents saw each of the three name generator questions. All of the statistics about the contacts are given as percentages. The alternative is to use counts; however, both methods can distort lower numbers since 1 out of 1 would yield 100%, just as 5 out of 5 would yield 100%, but as a straight count 1 is quite different than 5. Since most (613 out of 692, or 88%) respondents have five contacts, and because percentages reflect the overall make-up of the ego networks the proportion of alters using each mode is used. Ego-network proportions are also more directly comparable with neighborhood mode use, for which percentages are used.

Few network characteristics differ by name generator. Namely, the transportation discussion generator yielded lower numbers of contacts than the other two. Those who saw the any-five-contacts generator listed more roommates and fewer contacts in a nearby town than those who saw the other two name generators. Those asked to name the contacts with whom they interact most frequently have the most “very close” contacts. The transportation based name generator produced the fewest contacts known for two to five years but the most contacts known for five years or more while the any-five-contacts name generator has the most contacts known for two to five years.

TABLE 1 Ego-network Characteristics By Name Generator

Ego Network Characteristic[1]	Any Five Contacts	Frequent Interactions	Discuss Transportation
Number of contacts named (p < 0.000)	4.73 (N = 231)	4.87 (N = 245)	4.45 (N = 222)
Ego-network Density (p = 0.306)	0.440 (N = 231)	0.427 (N = 245)	0.432 (N = 222)
Geographic Nearness	(N = 231)	(N = 245)	(N =222)
Roommates (p = 0.105)	35%	34%	30%
In Same Neighborhood (p = 0.311)	19%	16%	19%
In Same Town (p = 0.341)	29%	26%	26%
In Nearby Town (p = 0.005	6%	12%	10%
In Same State (p = 0.129)	6%	8%	9%
In Another State (p = 0.301)	1%	2%	3%
In Another Country (p = 0.201)	0%	1%	0%
Closeness in Relationship	(N = 231)	(N= 245)	(N = 222)
Not Close (p = 0.137)	2%	4%	2%
Somewhat Close (p = 0.287)	8%	10%	11%
Moderately Close (p = 0.798)	22%	21%	21%
Considerably Close (p = 0.086)	29%	23%	28%
Very Close (p = 0.252)	37%	41%	36%
Duration of Relationship	(N = 231)	(N = 245)	(N = 222)
Less than one Month (p = 0.617)	1%	0%	1%
One to Six Months (p = 0.765)	10%	10%	11%
Six Months to One Year (p = 0.550)	17%	16%	18%
One to Two Years (p = 0.682)	22%	22%	20%
Two to Five Years (p = 0.015)	32%	30%	24%
More than Five Years (p = 0.040)	16%	21%	22%

[1]p-values are shown for ANOVA in comparisons of means and for chi-squared tests for categorical variables

Other characteristics of the ego-networks were examined, but no other properties exhibited significant differences with respect to the name generator. The only exception is the frequency of interactions with alters. Those asked the frequent interactions name generator have the most contacts with whom they interact every day (about 50%, on average), and the fewest contacts with whom they interact less than once a month (about 0.2% on average). This result is not surprising, as the formulation of the question addressed frequency of interactions. While there is some variation in ego-network properties according to name generator it is not considered extensive enough to require nor warrant separate analysis for each group. Future work in this project will explore differences in ego-network properties and how these properties relate to the conformity of mode use within ego-networks.

Both the CTS and the SNTS surveys present mode choice as a choice among nine modes of transportation. Since very few students commute by modes other than bike, bus or driving alone, the analysis presented here focuses on the respondents using these three modes. This accounts for 633, or 91% of the respondents who named at least one contact in the SNTS. The dataset was also reduced to include only those respondents that live in Davis, since there is limited information about neighbors for those respondents who live outside of Davis. Table 2 presents the mode use of the remaining 576 respondents, and summarizes the mean percent of ego-network and neighborhood use for each mode.

TABLE 2 Mean Percent Reference Group Mode Use And Respondent Mode Choice

	Mean percent of ego-network alters using each mode			Mean percent of neighbors using each mode, d = 1,250 ft.			Mean percent of neighbors using each mode, d = 2,250 ft.		
Ego Mode[1, 2, 3]	Bike (p<.001)	Drive (p=.699)	Bus (p<.001)	Bike (p<.001)	Drive (p=.122)	Bus (p<.001)	Bike (p<.001)	Drive (p=.003)	Bus (p<.001)
Bike (N = 390) 52.2%	47%	21%	16%	63%	7%	22%	63%	7%	23%
Drive (N = 37) 13.7%	23%	40%	16%	44%	14%	32%	50%	13%	31%
Bus (N = 149) 25.8%	24%	20%	45%	46%	9%	36%	52%	8%	32%

[1] Each column represents the comparison between average mode use for each mode, according to the ego's mode
[2] p-values are shown for ANOVA in comparisons of means
[3] Mode use by ego-network/neighbors does not add up to 100% across rows because only relevant modes are shown

Respondents tend to use the mode used by the highest percentage of their reference group, whether the reference group is the ego-network or the neighborhood (highlighted cells, on the diagonals). Mode use within ego-networks is somewhat more spread out among the three modes than it is in neighborhoods, but overall percentages of the ego-network and the neighborhoods are fairly similar, with a few exceptions. Notably, the percentage of neighborhoods that drive is less than 10% for both those who choose bike or bus as their mode of transportation, but the ego-network percentage of drivers is roughly 20%. This indicates slightly less correlation in behavior within ego-networks than neighborhoods. This outcome highlights the importance of studying these different formulations of social reference groups. For the social networks, the correlations in behavior are likely linked to social processes, however at the neighborhood level it is probable that similarities in behavior are related to neighborhood characteristics such as infrastructure and land use. It is also interesting that bikers have fewer bikers in their ego-networks relative to their neighborhoods, but for both drive and bus the ego-network proportions are higher than the proportions in the neighborhood. The neighborhood percentages probably reflect the behaviors for individuals in the City of Davis, and although spatial variation occurs, there is more variation between Davis and other locations than between neighborhoods within Davis. The City of Davis is known for high levels of bicycling; with flat topography, a mild climate and extensive infrastructure; more than 50% of students bicycle to campus on an average day (Popovich 2014).

Descriptive Statistics

The remainder of this paper investigates how these two definitions of reference group relate to mode choice when considered alongside other factors typically important in travel behavior, such as trip and individual characteristics. Table 3 shows a selection of variables considered and/or included in model estimations. Respondent age differs among mode choices, as well as the mean distance travelled to campus. Both males and females tend to choose bike as their usual mode more than bus or drive however, more males bike than females, and almost 30% of females choose to ride the bus. Very few of those who drive alone report that “The cost of owning a car or other vehicle,” is more than moderately important in their decision to drive alone. This factor is more important for those who bike or bus.

TABLE 3 Respondent Characteristics With Respect To Mode Choice

Characteristic[1]	Bike		Drive Alone		Bus	
Mean age (p = 0.011) N = 575	22.11	N = 390	24.27	N = 37	20.92	N = 148
Mean Distance to Campus (miles) (p < .001) N = 549	1.72	N = 359	2.43	N = 28	2.11	N = 142
Gender (p = 0.001)						
Females N = 398	248	63%	31	8%	117	30%
Males N = 169	126	79%	6	4%	28	18%
Importance of "The cost of owning a car or other vehicle" in mode choice (p < .001)						
Not Important	66	17%	11	31%	20	14%
Slightly Important	37	10%	9	25%	17	11%
Moderately Important	76	20%	9	25%	25	17%
Considerably Important	83	22%	6	17%	48	32%
Extremely Important	120	31%	1	3%	38	26%
Importance of "Commuting at the times I prefer" in mode choice (p = 0.521)						
Not Important	12	3%	0	0%	1	1%
Slightly Important	10	3%	1	3%	8	5%
Moderately Important	39	10%	3	8%	16	11%
Considerably Important	107	28%	9	24%	39	26%
Extremely Important	215	56%	24	65%	85	57%
Familiarity with UC Davis Transportation and Parking Services GoClub Program (p = 0.196)						
It's new to me	177	46%	16	43%	60	42%
I've heard of it, but never used it	138	36%	15	41%	66	46%
I've used it	68	18%	6	16%	16	11%

[1]p-values are shown for ANOVA in comparisons of means and for chi-squared tests for categorical variables

Model Results

Model 1 uses ego-network mode use, among other factors, to predict mode choice. I estimated models using spatial reference groups for 30 different neighborhood radii; too many to present here. The selection of the two neighborhood sizes was somewhat arbitrary, however some considerations were: at too small of a distance, there may not be sufficient numbers of neighbors to reliably calculate the neighborhood proportion of mode use. At too large of a distance all 'neighborhoods' begin to look the same, as the neighborhood captures more and more of the city-wide variation in mode use, and the neighborhood mode shares become close to the overall shares of the city. In this dataset this happens at a neighborhood size above about 2 miles.

Two neighborhood sizes that are within the range of values that have sufficient numbers of neighbors, conceptually reasonable values of neighborhood radius; and are also not so large that variation is lost, are about ¼ mile radius and ½ mile radius. These two distances also are different from each other such that the ½ mile radius has 4 times the area of the ¼ mile radius neighborhood size. Model 2 uses mode use in a small neighborhood (1,250 ft radius) and Model 3 in a medium-sized neighborhood (2,250 ft radius). The base alternative in each model is bike, with coefficients estimated for the alternatives bus and drive.

TABLE 4 Multinomial Logit Models of Mode Choice With Reference Group Variables

Variables in Model Estimation[1, 2, 3]	Model 1 Ego-network		Model 2 d = 1250ft.		Model 3 d = 2250ft.	
	Drive	Bus	Drive	Bus	Drive	Bus
Constant	-5.66**	-2.89**	-5.94***	-1.86*	-3.71	-0.16
Male	-1.28*	-0.78**	-1.19*	-0.94***	-1.15*	-0.94***
Distance to Campus	0.67**	0.29*	0.72**	0.36**	0.70**	0.38**
Importance of "Cost of owning a car or other vehicle"	-0.52***	0.01	-0.58***	0.02	-0.56***	0.07
Importance of "Going other places before, during or after work"	0.85***	-0.10	0.86***	-0.22**	0.80***	-0.23**
Importance of "Using the same means of transportation every day"	0.13	0.30***	0.24	0.40***	0.22	0.43***
Agreement with "Feel safe biking" (reverse scale)	0.47**	0.53***	0.61***	0.54***	0.63***	0.54***
Number of sources of information about parking	-0.02	0.24**	-0.01	0.21**	0.00	0.23**
Familiarity with campus tire air repair stations (reverse scale)	-0.78*	-0.75***	-0.49	-0.48**	-0.45	-0.51**
Familiarity with in vehicle parking meter (reverse scale)	0.46	0.79***	0.42	0.73***	0.40	0.69***
Familiarity with UC Davis GoClub (reverse scale)	-0.20	-0.39*	0.17	-0.42**	0.21	-0.41**
Percent alters biking	**-2.40****	**-0.94**	**---**	**---**	**---**	**---**
Percent alters taking the bus	**0.38**	**3.66*****	**---**	**---**	**---**	**---**
Percent alters driving	**1.70**	**1.12**	**---**	**---**	**---**	**---**
Percent neighborhood biking	**---**	**---**	**-2.14****	**-1.99*****	**-4.80***	**-4.29*****
Percent neighborhood taking the bus	**---**	**---**	**-0.79**	**2.08*****	**-2.31**	**0.65**
Percent neighborhood driving	**---**	**---**	**-0.56**	**-0.55**	**-2.25**	**-2.03**
Model Diagnostics						
	Model 1		**Model 2**		**Model 3**	
Log-likelihood of full model estimation	-237.98		-266.12		-268.12	
Adjusted rho-squared (pseudo r-squared)	0.499		0.446		0.442	
Akaike Information Criterion	531.96		588.23		592.24	

[1] *, ** and *** indicate significance of parameter estimates to the .1, .05 and .01 levels
[2] Adjusted rho-squared (*adj-*ρ^2) indicates the proportion of variance explained by the model
[3] In final model estimations total sample is 483; 340 bike, 25 drive and 118 bus as usual mode of transportation

All three models perform fairly well, considering the adjusted rho-squared. However, the AIC for model 1, which uses the ego-network, is somewhat better than either of the models using neighborhoods. Socio-demographic and trip characteristics generally exhibit expected effects; gender and distance to campus are both important factors. Females are somewhat more likely to drive or take the bus than males, though the bike mode share is highest among all groups. Living further away from campus increases the likelihood that one takes transit or drives. The less important the cost of owning a car or other vehicle, the more likely individuals are to drive to campus; since the cost is less relevant, they are more willing to pay to drive. Further, the importance of going other places before, during, or after work increases the likelihood of driving, and the importance of using the same means of transportation every day increases the likelihood of taking the bus. Other variables included in model estimations include the familiarity with transportation resources on campus, level of information about parking, and feeling of safety with respect to biking, and all have expected effects on mode choice.

Both formulations of the references groups - ego-network and neighborhood – are relevant in mode choice, though not in exactly the same ways. One similarity across all models is the percentage of the reference group that drives does not have a significant effect on the likelihood the ego chooses to drive nor the likelihood the ego takes the bus.

In the ego-network model (Model 1), higher ego-network biking decreases the likelihood of choosing driving, but has no significant effect on the likelihood of taking the bus. At the same time, ego-network bus use increases the likelihood of taking the bus, but has no significant effect on the likelihood of driving.

In the neighborhood models, higher percentages of bikers in both sizes of neighborhood increase the likelihood of biking compared to *both* driving and taking the bus. In the small neighborhood higher percentages of neighbors taking the bus increases the likelihood of taking the bus, but this is not a significant effect in the medium neighborhood. The coefficient values are larger for the percentage of neighbors biking in the medium neighborhood. The medium size neighborhoods cover a larger geographic area and therefore likely have more people who live further from bus lines and/or more people who live closer to bike paths. These features would increase the neighborhood percent of bikers, and also increase the likelihood that any single neighborhood resident, i.e. the respondent, bikes.

Although there are similarities across all three models, similar results are thought to reflect different processes. In the neighborhood models, shared environment, that is spatial, geographic or neighborhood factors conceivably account for the greatest portion of the correlation in mode choice between the respondent and their neighbors. In the ego-network model, shared environment is thought to account for some portion of the correlation in behaviors but not as completely as in the neighborhood models. Social processes, including social influence, or even endogenous processes such as homophily and self-selection into relationships are surely reflected in these results as well. If spatial auto-correlation were driving the results in the ego-network model, the coefficients would be more similar to those in the spatial models, and the same effects on likelihood would be significant. They are not; the coefficient for the percentage of the ego-network that takes the bus is not significant in the ego-network model (though it is in both neighborhood models), suggesting that some processes occur in ego-networks that do not occur in neighborhoods. Though we do not tease out the extent to which social influence, relationship self-selection or other mechanisms operate, these results highlight the importance of reference group definition when studying social influence.

DISCUSSION AND CONCLUSIONS

The research presented here investigates relationships between social influence and travel behavior; considering the importance of ego-networks and neighbors. Social influence in transportation is becoming a topic of increasing attention, and it is important to explore various means of defining reference groups as techniques and methods are refined. The focus this analysis is how spatially defined reference groups (neighborhoods) and socially defined reference groups (ego-networks) differ both conceptually and in measured effects. There are also subtleties within each of these definitions; in this case, to identify ego-networks three versions of a name generator were utilized though few significant differences were found between them. Alternative means of defining ego-networks, however, may result in additional differences.

Further, there are many ways to define spatial reference groups. Here, neighbors within a given distance of the respondent were counted for 100 distances, ranging from 250 feet to about five miles. Although models were estimated for every distance, two were selected for presentation here; neighborhoods with radius 1,250 feet and with 2,250 feet. The coefficient estimate for neighborhood biking on the ego's choice to bike is much greater in the model with the larger neighborhood size. Another approach to defining neighborhoods in which social influence may be relevant are T-communities; "every household within a t-community is reachable from every other household by only using tertiary streets" (Grannis 1998 p. 1533). Though there are many potential ways to define both social and spatial reference groups, it was of primary interest here to compare social reference group *to* spatial reference group.

Much work remains to identify the most satisfactory means for defining reference group (and the best means surely differs between analytic contexts). However, conclusions of interest are drawn from the analysis presented here. First, whether defined socially or spatially, reference groups do have relevance in transportation mode choice. Although both types of reference group are relevant, and although within group behaviors tend to be correlated in both cases, the mechanisms operating in neighborhoods are not the same as those operating within ego-networks. For example the percentage of ego-network biking is important for the choice between biking and driving, whereas the percentage of spatial neighbors biking is important for both the choice between biking and taking the bus and the choice between biking and driving.

The results presented here are part of ongoing research exploring social influence in travel behavior. Future steps in this project include taking into account sources of endogeneity in the relationship between reference group and mode use. Future work will also identify how properties of ego-networks relate to ego and alter mode choices, as well as whether certain types of relationships are more influential than others. How reference group is defined is a key question for research aiming to understand social and spatial influences on travel behavior. As work in this area is furthered, policies seeking to improve the use of alternative modes can capitalize on this type of knowledge to implement socially-relevant programs.

REFERENCES

An, Weihua. *Models and Methods to Identify Peer Effects*. In The Sage Handbook of Social Network Analysis. Sage, London, 2011, pp. 515–532.

Bernard, H. Russell, Eugene C. Johnsen, Peter D. Killworth, Christopher McCarty, Gene A. Shelley, and Scott Robinson. Comparing Four Different Methods for Measuring Personal Social Networks. *Social Networks*. Vol. 12, No. 3), 1990, pp. 179–215.

Bramoullé, Yann, Habiba Djebbari, and Bernard Fortin. Identification of Peer Effects Through Social Networks. *Journal of Econometrics*. Vol. 150, 2009. No. 1, pp. 41–55.

Brock, William A., and Steven N. Durlauf. Discrete Choice with Social Interactions. *The Review of Economic Studies*. Vol. 68, No. 2, 2001, pp. 235–260.

Campbell, Karen E., and Barrett A. Lee. Name Generators in Surveys of Personal Networks. *Social Networks*. Vol. 13, No. 3, 1991, pp. 203–221.

Carrasco, Juan Antonio, Bernie Hogan, Barry Wellman, and Eric J. Miller. Collecting Social Network Data to Study Social Activity-travel Behavior: An Egocentric Approach. *Environment and Planning. B, Planning & Design*. Vol. 35, No. 6, 2008.

Driller, Brigette. *Results of the 2012-13 Campus Travel Survey*. Research Report UCD-ITS-RR-13-18. Institute of Transportation Studies, University of California, Davis, 2013.

Dugundji, Elenna R., and Joan L. Walker. Discrete Choice with Social and Spatial Network Interdependencies: An Empirical Example Using Mixed Generalized Extreme Value Models with Field and Panel Effects. *Transportation Research Record: Journal of the Transportation Research Board 1921*. Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 70–78.

Ewing, Reid, and Robert Cervero. Travel and the Built Environment. *Journal of the American Planning Association*. Vol. 76, No. 3, 2010, pp. 265–294.

Farber, Steven and Antonio Páez. My car, my friends, and me: a preliminary analysis of automobility and social activity participation. *Journal of Transport Geography*. Vol. 17, 2009, pp. 216–22.

Fotheringham, A. Stewart, Chris Brunsdon, and Martin Charlton. *Quantitative Geography: Perspectives on Spatial Data Analysis.* SAGE Publications, 2000.

Fowler, J. H, and N. A Christakis. Dynamic Spread of Happiness in a Large Social Network: Longitudinal Analysis over 20 Years in the Framingham Heart Study. *BMJ*. Vol. 337, 2008.

Goetzke, Frank. Network Effects in Public Transit Use: Evidence from a Spatially Autoregressive Mode Choice Model for New York. *Urban Studies*. Vol. 45, No. 2, 2008, pp. 407–417.

Goetzke, F., and T. Rave. Bicycle Use in Germany: Explaining Differences Between Municipalities with Social Network Effects. *Urban Studies*. Vol. 48, No. 2, 2010, pp. 427–437.

Grannis, Rick. The Importance of Trivial Streets: Residential Streets and Residential Segregation. *American Journal of Sociology*. Vol. 103, No. 6, 1998, pp. 1530-1564.

Granovetter, Mark S. The Strength of Weak Ties. *American Journal of Sociology*: 1973, pp. 1360–1380.

Han, Qi, Theo Arentze, Harry Timmermans, Davy Janssens, and Geert Wets. The Effects of Social Networks on Choice Set Dynamics: Results of Numerical Simulations Using an Agent-based Approach. *Transportation Research Part A: Policy and Practice*. Vol. 45, No. 4, 2011, pp. 310–322.

Klofstad, C. A., S. D. McClurg, and M. Rolfe. Measurement of Political Discussion Networks: A Comparison of Two 'Name Generator' Procedures. *Public Opinion Quarterly.* Vol. 73, No. 3, 2009, pp. 462–483.

Páez, Antonio, and Darren M Scott. Social Influence on Travel Behavior: a Simulation Example of the Decision to Telecommute. *Environment and Planning A*. Vol. 39, No. 3, 2007, pp. 647–665.

Larsen, Jonas, Urry, John and Kay Axhausen. Coordinating Face to Face Meetings in Mobile Network Societies. *Information, Communication and Society.* Vol. 11, 2008, pp. 640-658.

Lee, Lung-fei. Identification and Estimation of Econometric Models with Group Interactions, Contextual Factors and Fixed Effects. *Journal of Econometrics*. Vol. 140, 2007, pp. 333-374.

Lovejoy, Kristin, and Susan Handy. Social Networks as a Source of Private-vehicle Transportation: The Practice of Getting Rides and Borrowing Vehicles Among Mexican Immigrants in California. *Transportation Research Part A: Policy and Practice*. Vol. 45, No. 4, 2011, pp. 248–257.

Manski, Charles. Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies*. Vol. 60, 1993, pp. 531-542.

Mok, D., B. Wellman, and J. Carrasco. Does Distance Matter in the Age of the Internet? *Urban Studies.* Vol. 47, No. 13, 2010, pp. 2747–2783.

Mokhtarian, Patricia L., and Xinyu Cao. Examining the Impacts of Residential Self-selection on Travel Behavior: A Focus on Methodologies. *Transportation Research Part B: Methodological*. Vol. 42, No. 3, 2008, pp. 204-228.

Mokhtarian, Patricia; & Ilan Salomon. Modeling the Desire to Telecommute: the Importance of Attitude Factors in Behavioral Models. Transportation Research Part A. Volume 31, Issue 1, January 1997, Pages 35–50.

Páez, Antonio, and Darren M Scott. Social Influence on Travel Behavior: a Simulation Example of the Decision to Telecommute. *Environment and Planning A.* Vol. 39, 2007, pp. 647–665.

Scott, Darren M. Antonio Páez, Ivy Dam and Robert Wilton. Investigating the effects of social influence on the choice to telework. *Environment and Planning A.* Vol. 44, 2012 pp. 1016–1031.

Silvis, Julia, and Debbie Niemeier. Social Network and Dwelling Characteristics That Influence Ridesharing Behavior of Seniors. *Transportation Research Record: Journal of the Transportation Research Board 2118*. Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 47–54.

Wasserman, Stanley, and Katherine Faust. *Social Network Analysis: Methods and Applications.* Cambridge University Press, Cambridge, 1994.

Wilton, Robert D., Antonio Páez, and Darren M. Scott. Why Do You Care What Other People Think? A Qualitative Investigation of Social Influence and Telecommuting. *Transportation Research Part A: Policy and Practice*. Vol. 45, No. 4, 2011, pp. 269–282.

Geography and Social Networks in Transportation Mode Choice

ABSTRACT

In this paper we explore the effects of social influence on travel behavior and utilize an instrumental variables approach to address the potential endogeneity related to similarities in the choice environments of socially connected individuals. We expect individuals to use a mode of transportation that is used by others in their social network. At the same time, other factors influence transportation mode choice, and what's more, these factors may influence multiple members of a social network at the same time. We hypothesize that social influence is relevant to transportation mode choice, even when accounting for the effects of shared effects of these factors among members of a social network. We utilize survey data collected from a sample of university students, in Davis California. The survey collected information about respondents' social networks, the transportation mode choices of their social contacts, and geographic information for the respondent and other members of their social network. We estimate models using instruments based on the alters' distance to campus as well as bike use in the neighborhoods of the alters. Results provide evidence that social processes are important to travel behavior, even when accounting for similarities in behavior that may be attributed to similar choice environments.

INTRODUCTION

Social networks serve as a foundation for multiple social processes, such as cooperation, resource sharing and social influence. Evidence shows that social networks provide access to opportunities such as collaboration and co-authorship (for example Freeman 1984) and employment (Granovetter 1973), through network connections. In this study we focus on social influence, whereby the knowledge, behaviors and/or opinions of one individual affect those of others to whom they are socially connected. Social influence is linked to many behaviors, ranging from academic achievement (Sacerdote 2001) to civic engagement (Klofstad et al. 2009), and health issues such as obesity (Koehly and Loscalzo 2009).

Social influence also affects individual transportation decisions; such as when one person's mode choice affects the choices of individuals to whom they are socially connected. This occurs when social networks act as pathways for information sharing; individuals may seek information, for example about telecommuting, from their friends and colleagues who already telecommute (Wilton et al. 2011) Social influence also occurs when sustainable modes of transportation become the social norm, such as when using transit becomes the accepted or expected transportation mode within communities (Goetzke and Rave 2010), households, or friend groups. Social influence has thus begun to gain attention in studies of and programs related to travel behavior.

We investigate whether an individual's mode choice is affected by the mode choices of others in his social network, measured by similarities in mode choice among socially connected individuals. However, when socially connected individuals live in the same neighborhood, or face similar commute circumstances (though in different neighborhoods), it is possible similarities in commute environment, rather than social influence, causes similarities in behaviors. Though there are other potential sources of endogeneity, accounting for shared environment addresses the impacts of factors *external* to the social relationship. Furthermore, similar circumstances or shared environment is particularly relevant to social influence in transportation since we expect many geographic factors such as commute distance, land use characteristics, and the availability of infrastructure to be important in mode choice. Further, scholars have demonstrated that both geographic neighborhood mode use (Pike 2014, Dugandji et al. 2005, Goetzke 2008) *and* social influences (Scott et al. 2012, Wilton et al. 2011) are relevant in travel behavior. We measure the effect of social influence in transportation mode choice, while accounting for shared environment within social networks.

BACKGROUND: SOCIAL NETWORKS AND TRAVEL BEHAVIOR

Social networks have recently emerged as an important area of inquiry in transportation behavior and policy research. Transportation researchers have explored a number of topics related to social networks and transportation including how social networks provide access to transportation resources (Silvis and Niemeier 2009 and Lovejoy and Handy 2011) and how land use characteristics affect travel behavior and thereby social networks (Farber and Paez 2009). Travel behavior research related to social networks has been focused on two primary areas of inquiry; how social networks and communication with contacts affect trip generation, and how neighborhood effects and social influences affect transportation mode choice.

Social Networks, Communication and Travel Behavior

As individuals engage in face-to-face social activities, social networks contribute to trip generation (Carrasco et al. 2008 and Carrasco and Miller 2006 and 2009). Research into social networks and information and communication technologies finds that that increased mobility influences social travel through the continual coordination of and last-minute changes to plans (Larsen, Urry and Axhausen 2008). The types of social ties affects travel behavior when social interactions are sometimes replaced by communication over the internet (Mok, Wellman, and Carrasco 2010). The frequency of social interactions may be dependent on network structure and composition and with whom activities take place (Carrasco et al. 2008).

Social Influence and Mode Choice

Considering social influence in transportation mode choice, Wilton, Páez, and Scott (2011) find that interactions with co-workers at work (which can be beneficial *or* annoying) and a culture around telecommuting are among the social factors relevant in the decision to telecommute. Further, an individual's choice to use transit is influenced by neighborhood transit use (Goetzke 2008), and individuals connected through socioeconomic similarities or residing in the same zip-code zone exhibit interdependence in mode choice decision making (Dugundji and Walker 2005). Bicycling mode share in German cities has been attributed to a city-level cultural component characterized as a social network effect (Goetzke and Rave 2010).

With few exceptions (such as Wilton et al. 2011 and Scott et al. 2012) the majority of studies exploring social influence in transportation mode choice represent social networks as neighborhoods or other broadly defined social groups, though there are many examples of explicit social interactions in activity-based travel (for example Carrasco et al. 2006, Larsen et al. 2008). At the same time, despite the recognition of endogeneity as an issue in any study related to social influence (for example see Manski 1993), discussion of endogeneity in social influence related to transportation mode choice is limited (Walker et al. 2005 and Dugandji et al. 2005).

In this study, we define social influence in terms of the mode use of contacts with explicitly social relationships; thereby measuring social influence as opposed to neighborhood or cultural effects. Further, since we collected information about explicit social contacts who live in neighborhoods different than those of our focal individuals we can better address the issues of endogeneity. Another concern in this research is the possibility that there is joint decision-making related to mode choice between egos and alters when they live within the same household. To explore the possibility of inter-household joint decision making all of the analyses presented here were conducted including and excluding household members from the ego-networks of respondents.

SURVEY AND DATA

Campus Travel Survey

The setting for this project is the University of California, Davis (UC Davis), located in Davis, California. The community is known for high rates of bicycling among students and other residents, made possible by the extensive bicycle infrastructure throughout the city, the mild climate and flat topography. High rates of bicycling are exhibited by students; among those that commute to campus, 54% commute by bike; freshmen (77%) and doctoral students (61%) have the highest proportions of biking (Popovich 2014).

The data used in our analysis was collected through surveys of students at UC Davis who had participated in the 2013-14 annual campus wide transportation survey (CTS) and agreed to receive information about a social network and travel survey. Travel data as well as socio-demographic information and attitudes towards transportation and the environment were collected for survey participants in the CTS (the complete survey is contained as an appendix in Popovich 2014).

A total of 27,798 individuals were invited to participate in the CTS with a target response rate of 10.11% or 2,811 individuals. Ultimately 3,663 (13.2%) gave usable responses (Popovich 2014). Of these, 2,671 were students and were asked if they would like to receive information about a related survey involving social networks; 1,396 indicated they were interested and were invited to participate. There are 966 respondents to the social networks and travel survey; about 70% of those who indicated an interest, and one-third of the students who participated in the CTS.

Social Networks and Transportation Survey

The social networks survey focused on eliciting information about the ego-networks of respondents. Each respondent, as an ego, received a *name generator* question that asked them to list up to five social contacts or alters:

> In this question, think about all the people who have been in **your social circle** over the past six months; this includes **people with whom you live, work or attend class, socialize or participate in activities** etc. or people you speak with over the phone or internet.
>
> List the first names of the **five contacts** you have had the **most frequent regular interaction** with over the past six months.

Alternative specifications[1] of this question were used in the 2012-13 year of the survey. No substantial differences were found in the networks related to which version of the name generator was seen by the respondent (Pike 2014). Following the name generator, egos were asked a series of questions about their relationship with each of their alters, and the usual commute mode of each alter. To confirm that egos correctly reported about their alters we checked whether ego reports about alters matched self-reported information from alters. This was possible for a subset of alters (N = 149) from the previous year, 2012-13, of our survey who happened to have participated in the campus travel survey itself or who responded to a snowball survey that was sent to alters following the initial survey of egos[2]. For these alters, 90 provided their usual mode of transportation, and egos reported the correct mode for 70 of those, or 78%.

We collected geographic information for each ego by asking for the cross streets of the ego's home address. The ego also provided the cross streets or neighborhood of each alter's home

[1] In addition to the version used in the present paper, one alternative asked for five contacts with whom the respondent had discussed transportation, and the third version asked respondents to name "any five contacts".

[2] A snowball survey was administered to alters for which egos provided information; egos were also given the option to send a survey invitation to the alters they named. Due to very limited success of this effort, the data collected has been used primarily to confirm ego accuracy. The 2012-13 year's snowball survey was more successful than that of 2013-14 (though the 2013-14 data is used in all other analysis presented here).

location. This information was geocoded using ArcGIS software (www.esri.com) and “Google Maps” (maps.google.com). We identified the location (longitude, latitude) for each ego and alter. Geographic information was used to determine the geographic distances between members of each ego-network, and to identify neighborhood characteristics related to transportation, including the distance of each ego and alter to the UC Davis campus. Geographic computations were carried out using R (R Development Team 2013) and the package ‘rgeos’ (Bivand et al. 2015).

Neighborhood Mode Use

We are interested in distinguishing between the effects of social influence and neighborhood characteristics in transportation mode choice. One neighborhood characteristic we measure is the density of bicycling among an individual’s geographic neighbors. All of the approximately 3,300 individuals who participated in the CTS, or our social networks survey, or are alters whose usual commute mode and cross-streets were noted by an ego are potential neighbors. We calculate the density of bicycling among neighbors as the percentage of neighbors within a specified distance that bike. If there are 20 neighbors within 0.3 miles and 15 of them bike the neighborhood biking density is 0.75.We calculate the neighborhood biking density for each ego and each alter in our study using every 0.1-mile increment from 0.1 mile to 2.0 miles. Results presented here use neighborhoods with a 0.3 mile radius. Figure 1 shows the residential locations of alters and egos, respondents to the CTS who live in Davis (about 2600 individuals). Each point is colored by neighborhood biking density (proportions that do not occur are omitted), and points are transparent; darker points indicate more than respondent reported the same cross-streets.

FIGURE 1 Survey Respondents' Residential Location and Neighborhood (0.3 Mile Radius) Biking Density in the City of Davis

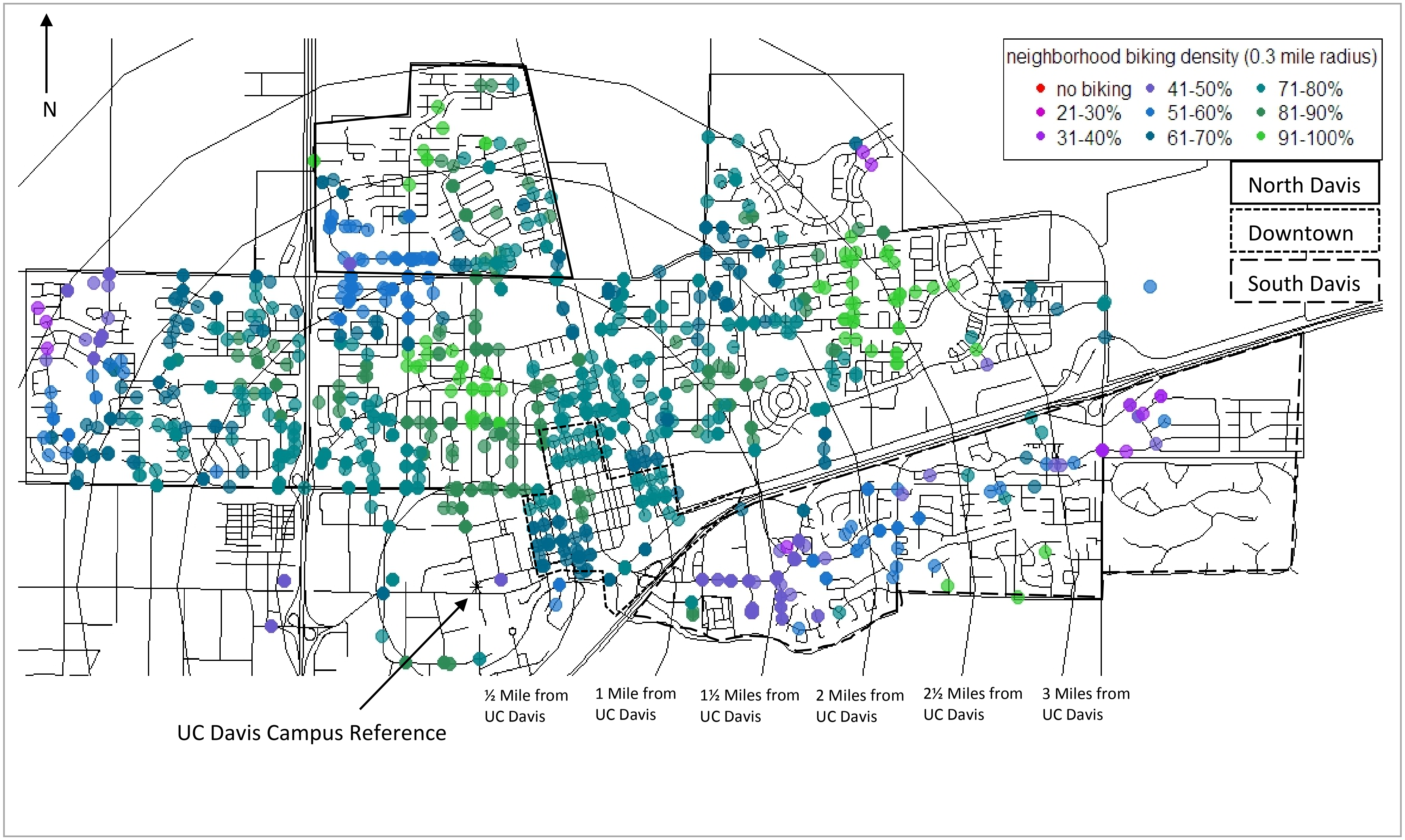


Figure 1 illustrates the geographic patterns of mode use throughout the study area, and highlights the non-random spatial distribution of transportation mode choice. Biking density is generally higher for those living closer to campus, although for those who live very close (within ½ mile, or just outside of ½ mile) to campus neighborhood biking density is somewhat lower because more individuals in these neighborhoods walk rather than bike. Among respondents to the CTS 20.8% of those who live downtown (outlined with a small-dashed line in Figure 1) report walking as their usual mode (Popovich 2014). Further from campus lower biking densities are attributed to higher rates of taking the bus and driving; from the CTS, North Davis (outlined with a solid line) and South Davis (outlined with a large-dashed line) have the highest proportions of individuals taking the bus, 37.4% and 32.5% respectively (Popovich 2014).

Some clusters of points that are adjacent or within a similar distance of campus have different neighborhood densities of bicycling. This could be due to differences in access to the bus, perceptions of bicycle infrastructure, or perceived barriers to bicycling. One barrier might be the freeway, which separates the southern portion of the City of Davis from the UC Davis campus and must be crossed by those respondents who live in this part of the city; South Davis has the highest proportion of CTS respondents that drive; 23.3% (Popovich 2014). The freeway is a barrier for bicycle commuting for high school students who live in South Davis because of the limited crossing locations and possibly the related increase in distance to school (Emond and Handy 2011).

As Figure 1 demonstrates, the density of bicycling among residents is related to geography, infrastructure and other features of the landscape. Failure to take this into account in our models might cause us to attribute social influence to effects better explained by shared environmental characteristics. In addition, these patterns provide a useful basis for our instrumental variables; the neighborhood biking density of an alter is a good indicator of the likelihood that alter bikes, but not directly related to whether that alter's ego bikes.

In our final models we have a sample of 388 individuals. Of the 966 who participated in our survey, 874 gave geographic information about their residential cross streets. We also restrict the sample to only those respondents who have at least one alter who lives in Davis, reported their own usual mode of transportation and indicated the transportation modes of their alters. We only include alters that live within 30 miles of the UC Davis campus, since we have more available data on the geographic characteristics for those that live in and around Davis[3]. The 388 individuals in our final sample are those that are not missing information on other variables used in final model estimations. In the analysis that follows, any additional changes in the sample are noted where applicable, and reduced samples used in our analysis are compared to the overall campus travel survey sample.

TWO STAGE RSIDUAL INCLUSION MODEL

The two Stage residual inclusion (2SRI) model used in our analysis is described as a straightforward process of estimating linear regression in the first stage, saving the residuals, and entering these into the second stage, along with the endogenous variable and all other exogenous explanatory variables (Wooldridge 2001; p 474, Rivers and Vuong 1988). The primary alternative to this model is the two-stage predictor substitution model, similar to two-stage least squares estimation, but used for modeling binary outcomes (also called the instrumental variable probit model). When compared to related or similar modeling approaches, the two-stage residual

[3] We explored limits on how far alters may live from campus ranging from 5 to 30 miles; there were no substantial differences in results related to the limit.

inclusion model is generally preferred (see Rivers and Vuong 1988, Terza et al. 2008 and Lewbel et al. 2012). Though there are a number of ways to write the model formula (see Rivers and Vuong 1988, Wooldridge 2001 p. 472), we present Wooldridge's notation (2001 p. 472-3). The first four equations show a binary probit model, with an endogenous regressor.

$$y_1^* = \boldsymbol{z_1}\boldsymbol{\delta_1} + \alpha_1 y_2 + u_1 \tag{1}$$

$$y_2 = \boldsymbol{z_1}\boldsymbol{\delta_{21}} + \boldsymbol{z_2}\boldsymbol{\delta_{22}} + v_2 \tag{2}$$

$$y_1 = 1[y_1^* > 0] \tag{3}$$

$$u_1 = \theta_1 v_2 + e_1 \tag{4}$$

The outcome of interest is y_1; whether the ego bikes or not. The endogenous variable is y_2; the proportion of alters that bike. The variables in $\boldsymbol{z}$ are exogenous regressors, including the instrumental variable(s); z_2. The proportion of alters biking; y_2 is estimated with respect to all of the exogenous regressors with error v_2. The error in the equation for y_1 has two parts. The portion e_1 is independent of the endogenous variable, and the portion $\theta_1 v_2$ represents the part that is correlated with the endogenous variable. In the second stage v_2 is included as a regressor and θ, α and $\boldsymbol{\delta}$ are estimated. We do not explicitly estimate ρ, the covariance of v_2 and u_1. The coefficient values in the second stage, θ, α and $\boldsymbol{\delta}$**,** are scaled by the factor $(1 - \rho^2)^{-1/2}$. The final reduced form of stage 2 is:

$$y_1^* = \boldsymbol{z_1}\boldsymbol{\delta_1} + \alpha_1 y_2 + \theta v_2 + e_1 \tag{5}$$

$$e_1 | \boldsymbol{z}, y_2, v_2 \sim Normal(0, 1 - \rho_1^2) \tag{6}$$

We are most concerned with the coefficient values in the second stage, θ, α and $\boldsymbol{\delta}$. The instrumented effect of the endogenous variable is given by α. The coefficient estimate α, and its significance make up the central test of whether social influence is relevant to mode choice, even when accounting for the potential shared environment of the egos and alters. The estimate θ is the coefficient on the stage-1 residual, and the components of $\boldsymbol{\delta}$ are the coefficients on the other exogenous variables in the model. See Wooldridge (2001) for further discussion of the model.

The 2SRI model has various diagnostic statistics associated with each stage. We also propose some additional caveats to these diagnostics. Instrumental variables approaches are concerned with the *validity* or weakness of the instrument; how well it does at predicting the endogenous variable. Validity is tested using the F-statistic of a Wald test. Instruments are weak if the F-statistic is less than 10 (Staiger and Stock 1997). The R^2 values of the stage-1 model are also used as an indicator of instrument validity, since R^2 tells us how much variation in the endogenous variable is explained by the stage-1 model; in which the primary explanatory variable is the instrument for the endogenous variable.

In the second stage of the model, the significance of the coefficient of the stage-1 residual serves as a test for the exogeneity of the suspected endogenous variable (Wooldridge 2001 p. 473). If the coefficient on the stage-1 residual is significant, the variable is endogenous, if not, the variable is exogenous. Recognizing that the stage-1 residual may also not be a significant predictor in the second stage when the instrument does not explain a large proportion of the variation in the endogenous variable, we suggest that the level of significance is an indicator of the validity of the instrument. In fact, in our results small changes to the instrument have an effect on whether the residual is a significant predictor, and the changes in this significance line up with changes in the stage 1 F statistic and R^2.

Instrumental Variables

In stage 1 of the 2SRI model, the instrumental variable is a predictor in the linear model of the proportion of alters that bike. Any instrumental variables approach requires identification of an instrument that is correlated with the endogenous regressor, in this case the proportion of alters that bike. The instrument must also not be correlated with variation in the outcome variable of interest; the transportation mode choice of the ego. To address the potential effects of shared environment the best candidate instruments are those that capture information about the alters' transportation environments, such as characteristics of their neighborhood, or their commutes.

The alters' neighborhood and commute characteristics affect the mode choices of the alters but we do not expect characteristics of the alters' neighborhoods to directly affect the mode choice of the ego. One may be concerned that, when an ego lives in the same neighborhood as one of his alters, they experience the same neighborhood characteristics. However, we calculate the effect of the alters' neighborhood characteristics on alters' mode choices for the group of alters collectively. Alters with neighborhood characteristics conducive to biking are more likely to bike, and alters with neighborhood characteristics more conducive to other modes are less likely to bike. Therefore, the *average* of a neighborhood characteristic across an ego's alters is different from that neighborhood characteristic for the ego, and is correlated with the proportion of alters that bike, but not whether the ego bikes or not. So, we use the instruments: the average distance to campus and the average neighborhood biking density among the alters in each ego's ego-network.

The alters' average distance to campus is potentially a good predictor of the proportion of alters that bike since most jobs in Davis are located at or near the university or downtown area (adjacent to the university) and because many of the alters are members of the university community themselves. The second instrument, average neighborhood biking density among alters, measures the percentage of neighbors within a specified distance of each alter's residential location that bike as their usual mode, and is also expected to be a good predictor of the proportion of alters that bike. It may be that neighbors influence one another, but can also be attributed to the similar commute characteristics of neighbors especially when, as with our sample, many commute to the same destination. There is likely a suite of neighborhood characteristics that contributes to higher levels of biking in some neighborhoods (see discussion of Figure 1).

Ego-Network Definitions

Related to the issue that if an ego and alter live in the same neighborhood is the issue of an ego and alter living in the same household. Egos and alters in the same household may make joint decisions about what modes of transportation they use. In order to determine how this affects results we conducted all analyses with ego-networks including and excluding household members. Any alter living within 0.1 miles of the ego was considered a household member[4]. The exclusion of particular alters from the ego-networks also changes the sample of egos, since some egos have only household members as alters. The sample of 428 egos in Table 1 includes all egos with alters that live within and around Davis but may be missing on other variables in our final models (the final model sample is 388). Analyses of ego-networks that include household members were conducted with the full sample of egos, even if they have only household members as alters. We also estimated models with a reduced sample of ego-networks including household members; only those egos/ego-networks with at least one alter outside of the ego's household (i.e. the egos in the sample when household members are excluded from the ego-networks).

[4]Presumably, the ego reported the same cross streets for them self and any household members they listed; we allow a 0.1 mile buffer. We also explored 0.5 mile and 0.25 mile as possible buffers, but found little difference.

When household members are excluded from the ego-networks, many of the ego-networks have only one or two alters (Table 1). There are over 100 egos that have only household members as alters. This is probably because they interact most frequently with these alters (the name generator question asked egos to list the contacts with whom they had the most frequent regular interaction) or because they don't know the residential locations of other alters. When household members are excluded from the ego-networks, there are on average, 2.14 alters in each ego-network. When household members are included in the networks there is on average one additional alter (3.14 alters) and there are 3.36 alters in each ego-network when the sample is reduced to exclude egos that have only household members as alters.

TABLE 1 Ego-network Size; Including and Excluding Household Members

	Ego-Network Size:	1 alter	2 alters	3 alters	4 alters	5 alters
Household members excluded from ego-networks (N = 273; average size is 2.14 alters)	Count of Egos	108	77	44	30	14
	Percent of Egos	40%	28%	16%	11%	5%
Household members included in ego-networks (N = 428; average size is 3.14 alters)	Count of Egos	82	68	87	90	101
	Percent of Egos	19%	16%	20%	21%	24%
Household members included in ego-networks (N = 273; average size is 3.36 alters)	Count of Egos	34	47	50	70	72
	Percent of Egos	12%	17%	18%	26%	26%

RESULTS

Descriptive Statistics

Table 2 provides background information for a selection of characteristics on the individuals in our final sample, according to whether they bike as their usual mode. Commute distance is measured for the road-network distance which has the shortest commute time from an individual's home location to campus (Popovich 2014). Binary variables are 1 if the respondent has the noted characteristic, and 0 otherwise. Annual parking permits allow parking in designated areas on campus. The goClub is a UC Davis Transportation and Parking Services program. Members commit to using a mode of transportation other than driving for their commutes to campus and receive benefits depending on which mode of transportation they commit to. For example, those outside of Davis who commit to taking the train receive discounted tickets. Attitudes and preferences are based on five-point Likert-type scales. The 'importance of' survey questions asked respondents to indicate, "How important to you are the following factors when choosing your usual means of transportation to travel to work or school?" The 'agreement with' CTS questions asked respondents to indicate "To what extent do you agree or disagree with the following statements?"

TABLE 2 Selected Sample Characteristics by Bike as Ego's Usual Mode

	Not Bike 38% (N = 147)	Bike 62% (N = 241)	Full Sample (N = 388)
Demographics and Trip Characteristics – Continuous			
Distance to campus in miles (p < 0.000)	4.34	1.86	2.80
Age (p = 0.253)	22.32	22.85	22.65
Days per week commuting to campus (p < 0.000)	4.83	5.26	5.10
Attitudes and Preferences			
Importance of "The time it takes to make the trip" (p = 0.001)	4.37	4.08	4.19
Importance of "Environmental Impacts" (p = 0.16)	2.92	3.22	3.11
Importance of "Using a mode that is socially acceptable" (p = 0.401)	1.74	1.83	1.79
Agreement with "I like biking" (p < 0.000)	3.53	4.45	4.10
Agreement with "Travel time is wasted time" (p < 0.000)	3.44	2.93	3.12
Agreement with "I like using transit" (p = 0.27)	3.36	3.10	3.20
Demographics and Trip Characteristics – Binary			
Gender (1 = respondent is female) p = 0.001	119 (81%)	157 (65%)	276 (71%)
Undergraduate student (1 = undergraduate) p = 0.059	110 (75%)	157 (65%)	267 (69%)
Annual parking permit (1 = has permit) p < 0.001	15 (10%)	2 (< 1%)	17 (< 1%)
Driver's license (1 = has license) p = 0.818	133 (92%)	223 (93%)	356 (92%)

p-values: for ANOVA for continuous variables or chi-squared tests of distributions for categorical variables; between those that bike and those that do not

Ego and Alter Mode Use

In our sample, egos tend to use the same transportation mode as the majority of their alters. Table 3 shows the average proportion of alters within each ego-network that use the three most common modes of transportation (among egos), according to the usual mode choice of the ego. The sample in Table 3 includes all respondents to our survey for which we have ego and later mode use; a total of 646 respondents. Percentages do not add up to 100% since some alters use other modes such as carpooling or walking. The proportion of alters that bike is higher for egos that bike (49%) than egos that take the bus (26%) or drive (27%). At the same time, looking down the column for the 388 egos that bike we can see that on average 49% of their alters bike, 13% of their alters take the bus, and 20% of their alters drive. These patterns of ego-alter mode use are similar for each of these three modes[5], and highlight the tendencies for individuals to use the same mode of transportation as people with whom they share social connections. In order to rule out endogeneity related to shared or similar choice environments, we explore this outcome while accounting for spatial patterns in mode use.

TABLE 3 Proportion of Alters Using Each Mode, According to Ego's Mode

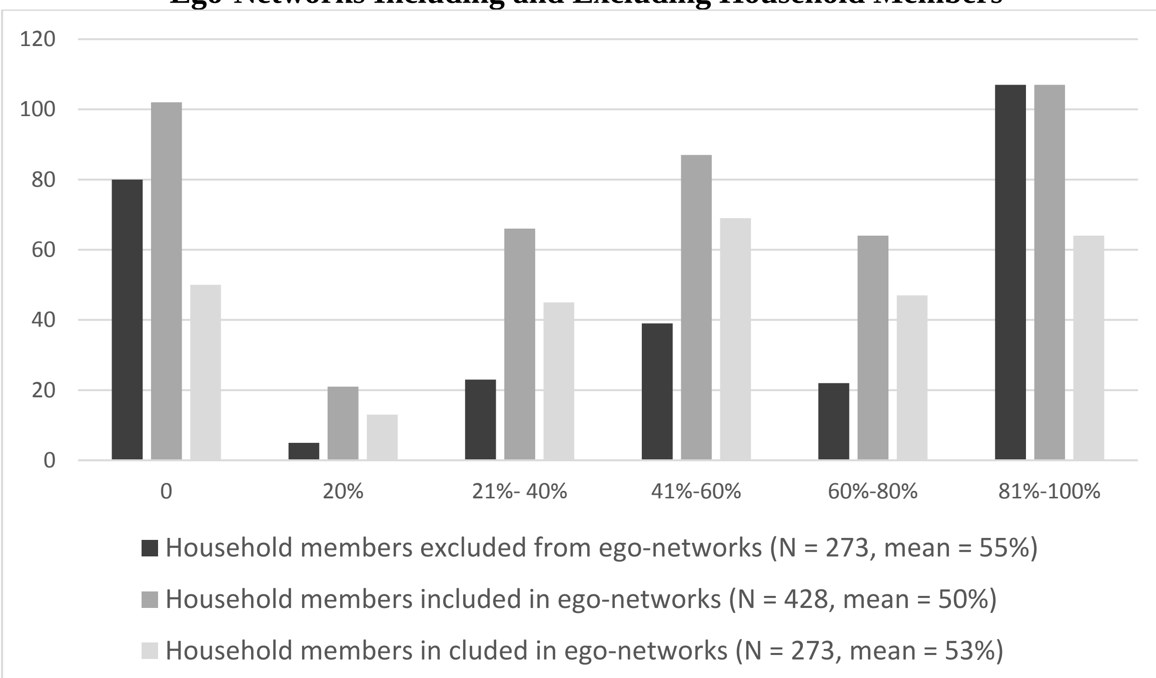
Ego Mode Choice[1]	Bike	Bus	Drive
Sample Size	388	162	96
Average proportion of alters that bike; $p < 0.000$	49%	26%	27%
Average proportion of alters that bus; $p < 0.000$	13%	41%	10%
Average proportion of alters that drive; $p = 0.041$	20%	20%	45%

[1]p-values are shown for ANOVA across each row (i.e. average percentage of alters biking, across ego mode choices)

[5]This pattern is similar for all of the modes included in the survey; however, since some modes are used by very few egos, we focus on these three most common modes.

The inclusion and exclusion of household members from ego-networks has a small effect on the distribution, across the ego-networks, of the proportion of alters that bike. Since there are more individuals in ego-networks that include household members, there is somewhat more variation in the proportion of alters that bike. Figure 2 shows the distribution of the proportion of alters that bike for the three formulations of ego-networks (excluding household members, including household members, and including household members but reducing the sample).

FIGURE 2 Proportion of Alters Biking for Ego-Networks Including and Excluding Household Members



When household members are excluded the proportion of alters that bike is moderately more concentrated at 0% and 100%, than when household members are included. This is because when household members are included there are more alters in each ego-network which leads to more variation. Nonetheless, the average proportion of alters that bike is similar whether household members are included (50%) or excluded (55%). For ego-networks including household members but with the reduced sample, the average proportion of alters that bike is also similar (53%). These patterns indicate egos' household members are not more or less likely to bike than non-household member alters.

Baseline Models

In our models we consider a number of other factors known to be important to transportation mode choice. These include attitudes, socio-demographics, and commute characteristics, such as the ego's distance to campus. Prior to estimating the two-staged models, we considered the ego's decision to bike or not with respect to variables and while a number of factors were considered, we retained variables that were significant predictors of the ego's choice to bike. These models, shown in Table 4, provide a baseline for comparison and include as predictors: age, gender, distance to campus, whether or not the ego is a member of the UC Davis goClub, has a parking permit, and how many days a week the ego commutes to campus.

The results of the baseline models are generally as expected. First, the older the ego, the higher the likelihood of biking. Though we might expect younger individuals to bike more, in our sample the older egos are graduate students who are somewhat more likely to bike than undergraduates, other than freshmen. Females are less likely to bike than males, as is the case for the overall student population at UC Davis (Popovich 2014). The farther the distance to campus, the less likely the ego is to bike. GoClub members are more likely to bike than others, however; it is probably the case that students who bike are more likely to join goClub; membership may not cause biking, but membership is associated with biking. The more days per week one commutes to campus, the more likely they are to bike; possibly because of costs associated with other modes.

TABLE 4 Baseline Binary Probit Models of the Ego's Decision to Bike[1]

	Excluding Household Members		Including Household Members		Including Household Members – Reduced Sample	
Variables in Model (Coefficients reflect effect on likelihood of biking)	N = 249; 156 Bike (63%)		N = 388; 241 Bike (62%)		N = 249; 156 Bike (63%)	
	Estimate (Std. Error)		Estimate (Std. Error)		Estimate (Std. Error)	
Intercept	-3.206***	(0.957)	-1.750*	(0.729)	-3.340***	(0.982)
Proportion of Alters Biking	0.554*	(0.217)	0.972***	(0.202)	1.080***	(0.275)
Age (years)	0.092**	(0.028)	0.053*	(0.023)	0.093**	(0.029)
Gender (1 = female)	-0.387.	(0.209)	-0.495**	(0.17)	-0.415.	(0.213)
Distance to Campus (miles)	-0.205**	(0.069)	-0.177***	(0.05)	-0.190**	(0.066)
Member of GoClub (1 = member)	0.297	(0.251)	0.446*	(0.212)	0.301	(0.256)
Annual Park Permit (1 = has permit)	-2.065**	(0.65)	-1.463**	(0.46)	-2.097**	(0.654)
Days Per Week Commute to Campus	0.385***	(0.111)	0.229**	(0.081)	0.354**	(0.112)
Model Diagnostics						
AIC:	267.29		417.39		259.93	
Log Likelihood	-125.65		-199.693		-120.97	
Rho-Squared – Market Share Base	0.2364		0.2243		0.2648	
Chi-squared/Likelihood Ratio Test (Proportion of Ego-Network Biking)	Statistic: 6.606 Pr(>Chisq): 0.010		Statistic: 23.701 Pr(>Chisq): < .001		Statistic: 15.97 Pr(>Chisq): < .001	

[1]Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Those with parking permits are less likely to bike and though we would expect those intending to drive to campus to purchase a parking permit it is also possible that possession of a parking permit reinforces the habit of driving. The model diagnostics, rho-squared and likelihood ratio test for the inclusion of the proportion of alters biking, indicate that these variables explain a reasonably good percent of the variation in the ego's choice to bike. Further, the proportion of alters that bike is significantly related to the likelihood that the ego bikes. As noted above, this coefficient is likely biased, since the proportion of alters that bike is almost certainly endogenous.

Two Stage Residual Inclusion Models

Distance to Campus Instrument

In our 2SRI models, the first stage is a linear model[6], predicting the proportion of alters that bike using the instrument and including all of the other exogenous regressors. The first stage models using the distance-based instrument for all three samples are shown in Table 5. The higher the average distance to campus among the alters, the lower the proportion of alters that bike. These models have small R^2 values, presumably because (as expected) few of the exogenous variables are related to the proportion of alters that bike. The F-statistic and the R^2 values suggest that the distance-based instrument is weak in all three models. Possibly because the average distance to campus is only a moderately good predictor of alter behavior; if the alters are not commuting to campus or downtown, their distance to campus wouldn't be related to their behavior. The stage 1 residuals are not normally distributed. This is likely because the proportion of alters that bike is not strictly continuous.

[6]Stage 1 models estimated with robust standard errors; iterated weighted linear regression: 'rlm' in MASS for R.

TABLE 5 Stage 1 Linear Models of Proportion of Alters Biking with Distance-Based Instrument[1]

	Household Members Excluded		Household Members Included		Household Members Included – Reduced Sample	
	N = 249; 156 Bike (63%)		N = 388; 241 Bike (62%)		N = 249; 156 Bike (63%)	
	Estimate (Std. Error)		Estimate (Std. Error)		Estimate (Std. Error)	
Intercept	0.194	(0.224)	0.083	(0.162)	0.051	(0.197)
Alters' Average Distance to Campus (miles)	-0.034***	(0.012)	-0.047***	(0.01)	-0.044***	(0.012)
Ego's Age (years)	0.010	(0.006)	0.009.	(0.005)	0.012.	(0.005)
Ego's Gender (1 = female)	-0.003	(0.06)	0.004	(0.042)	0.014	(0.052)
Ego's Distance to Campus (miles)	0.002	(0.005)	0.004	(0.004)	0.003	(0.004)
Ego Member of GoClub	0.047	(0.068)	0.003	(0.05)	0.027	(0.059)
Ego has Annual Parking Permit	-0.142	(0.126)	-0.196*	(0.094)	-0.152	(0.111)
Days Per Week Ego Commutes to Campus	0.039	(0.028)	0.062**	(0.019)	0.056.	(0.025)
Model Diagnostics						
R-squared	$R^2 = 0.055$		$R^2 = 0.097$		$R^2 = 0.077$	
Wald Test (Average Distance)	F = 8.48		F = 23.76		F = 13.04	

[1]Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The second stage is a binary probit model of the ego's decision to bike including the stage 1 residuals along with the proportion of alters biking, as explanatory variables. Results of the second stage, with the distance-based instrument are presented in Table 6. The coefficient values in the second stage are scaled by the factor $(1 - \rho^2)^{-1/2}$; the parameter ρ is the (un-estimated) covariance of the error in the first stage with the error in the second stage. The magnitudes and directions of the coefficient estimates are, nonetheless, similar to those in the baseline models.

The variables of interest are the proportion of alters that bike and the residual from the first stage. The proportion of alters that bike is significant only in the model excluding household members. Though the stage-1 residual is not significant in any of the three models, we do not consider this sufficient evidence to indicate exogeneity of the proportion of alters that bike. We suspect this outcome is more related to the validity of the distance-based instrument. The distance-based instrument may not be sufficiently related to the proportion of alters that bike to make definitive conclusions about social influence. If the instrument only weakly explains variation in the endogenous variable, there is still a high level of bias in the stage 2 coefficient estimate for the proportion of alters that bike, making it difficult to draw conclusions about social influence.

TABLE 6 Stage 2 Probit Model of Ego's Decision to Bike; Distance Based Instrument[1]

	Excluding Household Members		Including Household Members		Including Household Members – Reduced Sample	
Variables in Model (Coefficients reflect effect on likelihood of biking)	N = 249; 156 Bike (63%)		N = 388; 241 Bike (62%)		N = 249; 156 Bike (63%)	
	Estimate	(Std. Error)	Estimate	(Std. Error)	Estimate	(Std. Error)
Intercept	-3.64***	(1.036)	-1.728*	(0.755)	-3.353***	(1.008)
Proportion of Alters Biking	2.427.	(1.432)	0.786	(1.451)	1.194	(1.716)
Stage 1 Residual	-1.905	(1.433)	0.190	(1.459)	-0.117	(1.723)
Ego's Age (years)	0.081**	(0.029)	0.054*	(0.024)	0.092**	(0.032)
Ego's Gender (1 = female)	-0.397.	(0.21)	-0.494**	(0.17)	-0.417.	(0.214)
Ego's Distance to Campus	-0.210**	(0.073)	-0.180**	(0.057)	-0.188**	(0.07)
Ego Member of GoClub	0.185	(0.264)	0.448*	(0.212)	0.297	(0.261)
Ego has Annual Parking Permit	-1.811**	(0.687)	-1.500**	(0.54)	-2.082**	(0.695)
Days Per Week Ego Commutes to Campus	0.321**	(0.119)	0.240*	(0.117)	0.348*	(0.139)
Model Diagnostics						
AIC:	267.6		417.37		259.93	
Log Likelihood	-124.7999		-199.6843		-120.9627	
Rho-Squared – Market Share Base	0.242		0.224		0.265	
Likelihood Ratio Test (Stage1 Residual)	Statistic: 8.2962 Pr(>Chisq): 0.016		Statistic: 23.718 Pr(>Chisq): < 0.001		Statistic: 15.971 Pr(>Chisq): < 0.001	

[1]Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

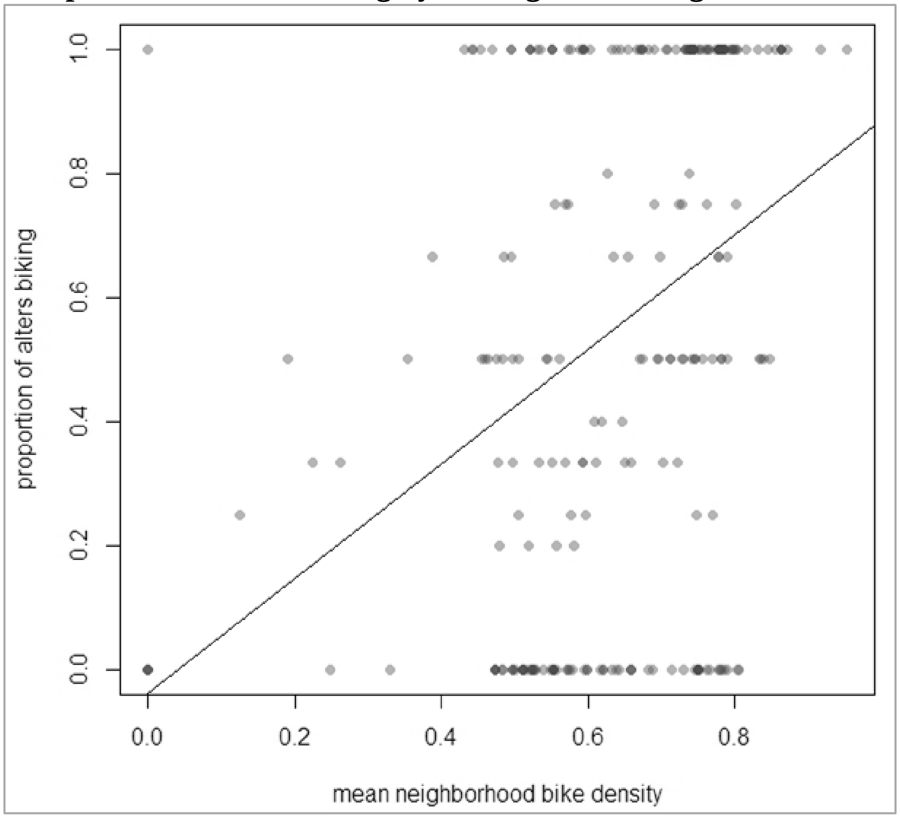
Neighborhood Biking Density Instrument

We examined multiple neighborhood sizes for the neighborhood biking density instrument, and completed a large set of model estimations. In general, as neighborhood size increases the differences in average neighborhood biking density decrease. The effect of average neighborhood biking density on the proportion of alters biking also decreases. The validity of the density-based instrument (measured by the stage 1 F-statistic and the R^2 value) is worst for neighborhood distances above about 1.0 mile. Results for the neighborhood biking density model for 0.3 mile neighborhoods are shown in Table 7[7]. The first stage coefficient estimates are not shown, but the F-statistic and R^2 value for the first stage are included in the model diagnostics. As with the distance-based instrument models, the residuals are not normally distributed; likely because the proportions of alters that bike are not strictly continuous.

Figure 3 shows the relationship between the average neighborhood biking density (for neighborhood radius 0.3 miles) and the proportion of alters that bike, for ego-networks excluding household members. Linear regression was estimated for this relationship and the parameter estimates are shown; the coefficient for neighborhood density is significant at 0.1%, and the intercept is not significant. All neighborhood sizes from 0.1 to 2.0 miles were tested[8], but 0.3 miles was ultimately selected as one of the best instruments based on the model diagnostics described above. In addition, as neighborhood size increases, the differences between neighborhoods decrease and consequently the effect of average neighborhood biking density on the proportion of alters biking becomes very small. Though average neighborhood biking density does predict the proportion of alters that bike, the absolute differences in neighborhood biking density are quite small. The average neighborhood biking density of alters ranges from just over 50% to about 70%, which is a narrower range than the neighborhood biking density for individuals (shown in Figure 1).

[7] The full set of all model estimations are available in online appendix.
[8] The full results, for all distances are available in online appendix.

FIGURE 3 Proportion of Alters Biking by Average Alter Neighborhood Biking Density



$$proportion\ of\ alters\ biking\ \approx\ -0.04 + 0.92 \times neighborhood\ bikiking\ density\ (R^2 =\ .13)$$

Models including household members have larger coefficients on the proportion of alters that bike than the models excluding household members, indicating a stronger relationship between the behaviors of household members and the ego, than other members of the ego-network. This might also be due to the reduced sizes of the ego-networks without household members.

In all three models the effect of social influence is significant, even when accounting for shared environment. However, the stage 1 residual is not estimated to be significantly different from 0 in the model excluding household members; we don't consider this sufficient evidence that the proportion of alters is exogenous, because this result is sensitive to the definition of the instrument. It is significant for some instrument values for the samples excluding household members and for more than half of the models for both samples that include household members.

TABLE 7 2SRI Model of Ego's Decision to Bike; Density-Based Instruments[1]

	Excluding Household Members		Including Household Members		Including Household Members – Reduced Sample	
Variables in Model (Coefficients reflect effect on likelihood of biking)	N = 249; 156 Bike (63%)		N = 388; 241 Bike (62%)		N = 249; 156 Bike (63%)	
	Estimate	(Std. Error)	Estimate	(Std. Error)	Estimate	(Std. Error)
Intercept	-3.326***	(0.972)	-1.794*	(0.733)	-3.308***	(0.989)
Proportion of Alters Biking	1.313*	(0.65)	2.148***	(0.57)	2.517***	(0.709)
Stage 1 Residual	-0.857	(0.685)	-1.328*	(0.595)	-1.673*	(0.746)
Ego's Age (years)	0.085**	(0.029)	0.042.	(0.023)	0.074*	(0.03)
Ego's Gender (1 = female)	-0.390.	(0.209)	-0.501**	(0.171)	-0.432*	(0.214)
Ego's Distance to Campus	-0.197**	(0.071)	-0.156**	(0.052)	-0.18*	(0.07)
Ego Member of GoClub	0.244	(0.254)	0.431*	(0.212)	0.229	(0.258)
Ego has Annual Parking Permit	-1.983**	(0.667)	-1.266**	(0.472)	-1.934**	(0.662)
Days Per Week Ego Commutes to Campus	0.356**	(0.113)	0.156.	(0.087)	0.281*	(0.117)
Model Diagnostics						
First Stage R-squared	R^2 = .145		R^2 = .177		R^2 = .169	
First Stage Wald Test	F = 37.65		F = 68.45		F = 50.65	
AIC:	267.76		412.57		255.22	
Log Likelihood	-124.879		-197.29		-118.61	
ρ^2 (Market Share Base)	0.24		0.24		0.28	
Likelihood Ratio Test (Stage1 Residual)	Statistic: 1.53 Pr(>Chisq): 0.216		Statistic: 30.534 Pr(>Chisq): < .001		Statistic: 20.68 Pr(>Chisq): < .001	

[1]Signif. Codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

The estimated coefficients are not substantially different from the base model, and although the magnitude of coefficient estimates are scaled (Wooldridge 2001 p. 474) coefficients generally have the expected signs. The stage-1 F-statistic and R^2 values suggest the density-based instrument contributes more explanatory power than the distance-based instrument. Neighborhood biking density is likely a proxy for a suite of geographic variables related to mode choice and is therefore a better instrument than solely distance. Our results suggest that shared neighborhood characteristics explain part of the similarity in behaviors within social groups, but that social influences are an important factor in transportation mode choice.

DISCUSSION AND CONCLUSIONS

Our findings provide evidence that social influence affects transportation mode choice, even after accounting for similarities in choice environments within social groups. We explored two different formulations of instrumental variables and found that neighborhood biking density may be preferred as it is likely a proxy for a number of geographic features, while distance captures the effect of just one neighborhood characteristic. The differences in the magnitude of the coefficients for the proportion of alters that bike when household members are included/excluded may be evidence that there are household level processes related to mode choice not captured even when addressing shared characteristics of alters' and egos' commute environments. Future work should further investigate interactions at the household level related to joint decision making about mode choice, which may be linked to joint decision making about residential location. It would also be helpful for survey efforts to collect data explicitly on household member and non-household member contacts.

We focus on endogeneity related to shared or similar environments since it is most relevant for transportation. Shared characteristics of the commute environment, such as the quality of infrastructure, access to particular modes and land-use features are expected to affect mode choice, whereas other sources of endogeneity such as self-selection into relationships, are more related to social processes. In addition to shared environment, self-selection into relationships and uncertainty about the direction of social influence are considered important endogeneity problems (Manski 1993). Accounting for endogeneity related to shared environment addresses the potential impacts of factors *external* to the social relationship. Nonetheless, future work should address the other two sources of endogeneity: self-selection into relationships and reflection. Snowball surveys, panel studies or whole-network social network analysis (see Wasserman and Faust 1994) may allow researchers to gather better information about the dynamic processes of relationship formation, bi-directional social influence and joint-decision making. Improved sampling and data collection would also allow for more continuous distributions of alter mode choices thereby improving model precision.

The outcomes of this study, along with the work of others (such as Scott et al. 2012), contribute to building evidence that social influence is relevant to transportation mode choice, as well as other aspects of travel behavior. Social influence has been incorporated into some programs aimed at increasing the use of sustainable transportation modes already. Taking advantage of the social networks of its members, the UC Davis goClub periodically asks members to invite their friends and colleagues to sign up and pledge to use an alternative to driving alone for their commutes to campus. May is Bike Month, an annual month-long drive to increase bicycle commuting, utilizes social tools such as 'challenge a friend' and 'share your accomplishments' to instigate friendly competition among participants as they log bike miles during the event. We encourage the simultaneous investigation of social influence effects and these types of sustainable transportation programs using experiments. Our results, along with the results of these types of studies enable a more informed use of social influence as a tool in sustainable transportation programs.

References

Axelrod, Lawrence J., and Darrin R. Lehman. 1993. "Responding to Environmental Concerns: What Factors Guide Individual Action?" *Journal of Environmental Psychology*. Vol. 13, 1993, pp. 149–159.

Bivand, Roger, Colin Rundel, Edzer Pebesma, Karl Ove Hufthammer. 2015. Package 'rgeos'; Interface to Geometry Engine - Open Source (GEOS). (https://r-forge.r-project.org/projects/rgeos/).

Carrasco, Juan Antonio, and Eric J. Miller. Exploring the Propensity to Perform Social Activities: a Social Network Approach. Transportation. Vol. 33, 2006, pp. 463–480.

Carrasco, Juan Antonio, Bernie Hogan, Barry Wellman, and Eric J. Miller. 2008. “Collecting Social Network Data to Study Social Activity-travel Behavior: An Egocentric Approach.” Environment and Planning. B, Planning & Design 35 (6): 961.

Carrasco, Juan-Antonio, and Eric J. Miller. The Social Dimension in Action: A Multilevel, Personal Networks Model of Social Activity Frequency Between Individuals. *Transportation Research Part A: Policy and Practice.* Vol. 43, 2009, pp. 90–104.

Carrell, Scott E., Richard L. Fullerton, and James E. West. Does Your Cohort Matter? Measuring Peer Effects in College Achievement. *Journal of Labor Economics*, University of Chicago Press. Vol. 27, 2009, pp. 439-464.

Dugundji, Elenna R., and Joan L. Walker. Discrete Choice with Social and Spatial Network Interdependencies: An Empirical Example Using Mixed Generalized Extreme Value Models with Field and Panel Effects. *Transportation Research Record: Journal of the Transportation Research Board 1921*. Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 70–78.

Emond, Catherine R. and Susan L. Handy (2011). Factors Associated with Bicycling to High School: Insights from Davis, CA. *Journal of Transport Geography* 20, 71 – 79.

Freeman, Linton C. The Impact of Computer Based Communication on the Social Structure of an Emerging Scientific Specialty. *Social Networks* Vol. 6, 1984, pp. 201–221.

Goetzke, Frank. Network Effects in Public Transit Use: Evidence from a Spatially Autoregressive Mode Choice Model for New York. *Urban Studies*. Vol. 45, No. 2, 2008, pp. 407–417.

Goetzke, F., and T. Rave. Bicycle Use in Germany: Explaining Differences Between Municipalities with Social Network Effects. *Urban Studies*. Vol. 48, No. 2, 2010, pp. 427–437.

Granovetter, Mark S. The Strength of Weak Ties. *American Journal of Sociology*: 1973, pp. 1360–1380.

Klofstad, C. A., S. D. McClurg, and M. Rolfe. Measurement of Political Discussion Networks: A Comparison of Two ‘Name Generator’ Procedures. *Public Opinion Quarterly*. Vol. 73, No. 3, 2009, pp. 462–483.

Koehly LM, Loscalzo A. Adolescent obesity and social networks. *Preventing Chronic Disease*. Vol. 6, 2009, pp 1-8.

Larsen, Jonas, Urry, John and Axhausen, Kay. Coordinating Face to Face Meetings in Mobile Network Societies. Information, Communication and Society. Vol. 11, No. 5, 2008 pp. 640-658.

Lewbel Arthur, Yingying Dong and Thomas Tao Yang. Comparing Features of Convenient Estimators for Binary choice Models with Endogenous Regressors. *Canadian Journal of Economics.* Vol. 45 No. 3, 2012, pp. 809-829.

Lovejoy, Kristin, and Susan Handy. Social Networks as a Source of Private-vehicle Transportation: The Practice of Getting Rides and Borrowing Vehicles Among Mexican Immigrants in California. *Transportation Research Part A: Policy and Practice*. Vol. 45, No. 4, 2011, pp. 248–257.

Manski, Charles. Identification of Endogenous Social Effects: The Reflection Problem. *Review of Economic Studies*. Vol. 60, 1993, pp. 531-542.

Mok, D., B. Wellman, and J. Carrasco. Does Distance Matter in the Age of the Internet? *Urban Studies.* Vol. 47, No. 13, 2010, pp. 2747–2783.

Pike, Susan (2014). Travel Mode Choice and Social and Spatial Reference Groups; Comparison of Two Formulations. *Transportation Research Record: Journal of the Transportation Research Board, No. 2412,* Transportation Research Board of the National Academies, Washington, D.C., 2014, pp. 75–81.

Popovich, Natalie (2014) Results of the 2013-14 Campus Travel Survey. Institute of Transportation Studies, University of California, Davis, Research Report UCD-ITS-RR-14-14

R Core Team (2013). R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria. URL http://www.R-project.org/.

Rivers, D., and Q. H. Vuong. Limited information estimators and exogeneity tests for simultaneous probit models. *Journal of Econometrics.* Vol. 39, 1988, pp. 347–366.

Sacerdote, Bruce. Peer Effects with Random Assignment: Results for Dartmouth Roommates. *The Quarterly Journal of Economics*, Vol. 116, 2001, pp. 681–704.

Scott, Darren M. Antonio Páez, Ivy Dam and Robert Wilton. Investigating the effects of social influence on the choice to telework. *Environment and Planning A*. Vol. 44, 2012 pp. 1016–1031.

Silvis, Julia, and Debbie Niemeier. Social Network and Dwelling Characteristics That Influence Ridesharing Behavior of Seniors. *Transportation Research Record: Journal of the Transportation Research Board 2118*. Transportation Research Board of the National Academies, Washington, D.C., 2009, pp. 47–54.

Staiger, Douglas and James H. Stock (1997) 2SRI Regression with Weak Instruments. *Econometrica.*Vol. 65, No. 3 (May, 1997), pp. 557-586.

Terza, Joseph V. Anirban Basu, and Paul J. Rathouz. Two-Stage Residual Inclusion Estimation: Addressing Endogeneity in Health Econometric Modeling. *Journal of Health Economics*. Vol. 27, No. 3, 2008, pp. 531–543.

Wasserman, Stanley, and Katherine Faust. *Social Network Analysis: Methods and Applications.* Cambridge University Press, Cambridge, 1994.

Wilton, Robert D., Antonio Páez, and Darren M. Scott. Why Do You Care What Other People Think? A Qualitative Investigation of Social Influence and Telecommuting. *Transportation Research Part A: Policy and Practice*. Vol. 45, No. 4, 2011, pp. 269–282.

Wooldridge, Jeffrey M. *Econometric Analysis of Cross Section and Panel Data.* The MIT Press, Cambridge, Massachusetts and London, England 2001.

External Impacts on the Effect of Social Influence in Transportation Mode Choice

ABSTRACT

Recent research suggests that an individual's transportation mode choice is partially influenced by the mode choice of other people in their social network. This paper advances this basic idea by testing the hypothesis that the extent of social influence is conditional on external context factors such as commute characteristics. Through an online survey of students at the University of California, Davis, we collected information about students' transportation decisions for campus travel and their social networks. For each participant as an *ego*, we gathered information about their *ego-network* including up to five *alters* or contacts. A series of complementary statistical models find the strength of social influence is lower for those with longer commute distances where biking is more costly than driving, and is also lower at distances where walking has higher utility than biking. Social influence is most important when the external commute characteristics entail relatively equal travel costs for different modes. As social influence and other social processes are evaluated as potential policy instruments, these heterogeneous effects should be taken into account.

INTRODUCTION

Social networks serve as a foundation for multiple social processes, such as cooperation, resource sharing and social influence. There is evidence that social networks are important in a broad range of outcomes including academic achievement (Sacerdote 2001 and Carrell et al. 2008), happiness (Fowler and Christakis 2008), and health issues such as obesity (Koehly LM, Loscalzo 2009). Social influence affects environmental actions (Axelrod and Lehman 1993), co-authorship and collaboration (for example Freeman 1984), and social networks provide access to resources such as employment opportunities, through network connections (Granovetter 1973).

The focus of this study is social influence, whereby the knowledge, actions and/or opinions of one individual affect those of another. For example, social networks act as avenues for the diffusion of information. Social groups may discuss their use of alternative modes, changes to transit routes or schedules, or new bicycle infrastructure; influencing shifts in the behavior or attitudes of individuals in the group. Social networks also provide pathways for the establishment of social norms where individuals conform to the behavior of those to whom they are socially connected.

Social Influence in Transportation Mode Choice

Social influence is important in transportation mode choices such as telecommuting (Wilton et al. 2011 and Scott et al. 2012). Neighborhood transit use is related to an individual's likelihood of using transit (Goetzke 2008) and social and cultural context affect bicycle ridership (Goetzke and Rave 2010). Perceptions of and attitudes towards new technologies such as electric vehicles are affected by social network exposure to these new technologies (Axsen and Kurani 2011). There is a growing literature exploring social influence in transportation mode choice and methods used to measure such effects (Dugandji and Walker 2005 and Walker et al. 2011, Paez and Scott 2007, Pike 2014, Manness et al. 2015).

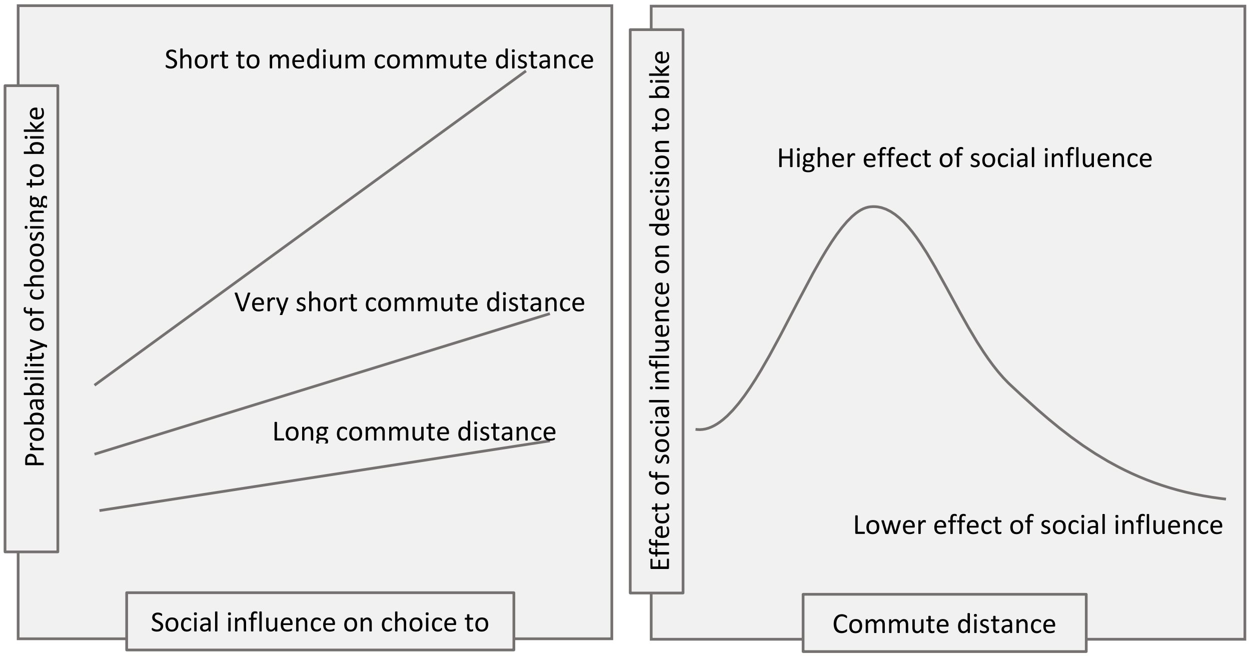
Other Important Factors

As travel behavior researchers explore new factors such as social influence it is important to consider how they related to and are impacted by more traditional factors, such as trip and mode characteristics as well as socio-demographics, attitudes (Mokhtarian and Salomon 1997), land use and the built environment (Ewing and Cervero 2010). There is growing evidence of social and spatial interdependencies in transportation decision making; however, these interdependencies have primarily been explored in relation to activity travel (for example Carrasco and Miller 2006 and 2009, Larsen et al. 2008) or to the impacts of land use characteristics on social interactions (Farber and Paez 2009). In this study we investigate how commute characteristics impact the effect of social influence.

Conceptual Model

Figure 1a presents a conceptual model of the conditional nature of social influence for a sample segmented based on commute distance. We expect the probability an individual chooses to bike to increase as social influence to bike increases. At the same time, we propose those with shorter commute distances are more affected by social influence (steeper curve) and those with longer commute distances are less affected (flatter curve). At very short commute distances we hypothesize that the effect is also small since the convenience of walking may outweigh the importance of social influence these individuals. Figure 1b illustrates how this relationship might change continuously with respect to commute distance. The exact shape of this curve is not known, though we expect the effect of social influence to bike to be low at very short commute distances and to have the greatest effect for those with short to medium commute distances. As commute distance increases the effect of social influence decreases and eventually tapers off.

FIGURE 1 Conceptual Model of the Impact of Distance on Social Influence Effects



Panel a: theoretical linear model of social influence; slope/effect varies by

Panel b: theoretical change in social influence with respect to commute

In the next section we describe our study location and sample, as well as the variables important to our analysis. In our results we present two approaches to test the impact of commute distance on the effect of social influence in the choice to bike as a usual commute mode.

DATA AND METHODS

<u>Survey and Sample</u>

Our study area is the University of California, Davis where extensive bicycle infrastructure throughout the city and the university campus, the mild climate and flat topography all contribute to high shares of bicycling within the city and the campus community. Survey data was collected at the UC Davis in coordination with the 2013-14 academic year's annual campus travel survey

(CTS), conducted in October of 2013. Travel data as well as socio-demographic information and attitudes towards transportation and the environment were collected in the CTS (the complete survey is contained as an appendix in Popovich 2014).

A total of 27,798 individuals were invited to participate in the CTS with a target response rate of 10.11% or 2,811 individuals. Ultimately 3,663 (13.2%) gave usable responses (Popovich 2014). Of these, 2,671 were students and were asked if they would like to receive information about a related survey involving social networks; 1,396 indicated they were interested and were invited to participate. There are 966 respondents to the social networks and travel survey; about 70% of those who indicated an interest, and one-third of the students who participated in the CTS.

The primary dependent variable in our analysis is transportation mode choice. We asked respondents "What mode of transportation do you usually use to travel to campus or your normal off-campus location for school or work?" A similar question was asked in the CTS, and for campus overall, among students that physically commute to campus, 54% commute by bike; with even higher proportions for freshmen (77%) and doctoral (61%) students (Popovich 2014).

The CTS collects cross-streets for respondents' home locations and this data was geocoded to determine the road-network distance to campus for each respondent. The road-network distance was estimated for the route to campus with the shortest commute time (Popovich 2014). The CTS does not ask for the residential location of students who live on campus; however, in the social networks survey we asked everyone, including students that live on campus, to list their nearest cross streets. For those on campus we computed the straight line distance from their home location to a central campus location. This is a realistic estimation of commute distance, since there is a dense network of campus bike paths and many possible walking or biking routes through campus.

The social networks survey asked participants to indicate the importance of various factors in their mode choice decisions, and gathered information about the ego-networks of respondents. Each respondent, as an ego, was asked to name up to five social contacts or alters, in the *name generator* that asks:

> In this question, think about all the people who have been in **your social circle** over the past six months; this includes **people with whom you live, work or attend class, socialize or participate in activities** etc. or people you speak with over the phone or internet.
>
> List the first names of the **five contacts** you have had the **most frequent regular interaction** with over the past six months.

Egos were also asked a series of questions about their alters, including each alter's usual commute mode. To confirm egos correctly reported about their alters we checked whether ego reports about alters matched self-reported information from alters. In our social networks survey for the 2012-13 academic year there were 149 alters (listed by egos) that either responded to a snowball survey or happened to be respondents to that year's CTS[9]. For these alters, 90 provided their usual mode of transportation, allowing us to cross-check the alters' self-reported modes against the ego-reported modes. Egos were correct mode for 70 of these alters, or 78%.

Because egos are reasonably accurate when indicating alter modes of transportation, we rely on the ego-reported typical mode of transportation for each alter. For each ego-network we compute the proportion of alters that uses each of nine modes of transportation. These are bike, walk, skate/board, motorcycle/scooter, drive alone, get a ride, carpool, take the bus or take the train. Alters are most likely to bike, drive or take the bus. Some egos did not provide the

[9] A snowball survey was administered to alters for which information could be obtained from egos; egos were also given the option to send a survey invitation to the alters they named. Due to very limited success of this effort, the data collected has been utilized primarily to validate ego reports. The 2012-13 year's snowball survey was more successful than that of 2013-14 (2013-14 data is used in all other analysis presented here).

transportation mode of every contact they named. The proportion of alters using each mode of transportation is computed with only the alters for which the ego noted a usual mode of transportation in the denominator.

Descriptive Statistics

Respondents are excluded from the sample if they do not live in Davis, or if they did not provide a usual mode of transportation for themselves and/or their alters. Respondents are also removed from the sample if they are missing geographic information. Davis is surrounded by agricultural lands and very few people live between five and ten miles from campus. This distorts the distribution of distances to campus since any individual living more than five miles from campus in effect lives at least ten miles from campus, so we restricted the sample to only those students who live in Davis. We also restricted our sample to those students who reported that they typically bike, take the bus or drive to campus. The numbers of students using other modes are too small for inclusion in models. There are 521 students, about 25% of the students who participated in the CTS, in our final model sample. Summary statistics of variables in our analysis, including socio-demographics, commute behaviors, preferences and attitudes are presented in Table 2.

Binary variables are equal to 1 if the respondent has the noted characteristic (0 otherwise). Annual parking permits allow parking in designated areas on campus. The goClub is a UC Davis Transportation and Parking Services program. Members commit to using a mode of transportation other than driving for their commutes to campus and receive benefits depending on which mode of transportation they commit to. For example, those outside of Davis who commit to taking the train receive discounted tickets. Social network variables are the proportion of alters using each mode of transportation. Attitudes and preferences are based on five-point Likert-type scales. The 'importance of' survey questions asked respondents to indicate, "How important to you are the following factors when choosing your usual means of transportation to travel to work or school?" The 'agreement with' CTS questions asked respondents to indicate "To what extent do you agree or disagree with the following statements?"

TABLE 1 Selected Sample Characteristics by Respondent Usual Mode

	Bike; 66% (N = 345)	Drive; 7% (N = 39)	Bus; 26% (N = 137)	Total (N = 521)
Demographics and Trip Characteristics – Binary				
Female (p = .002)	222 (67%)	32(82%)	109 (82%)	363 (72%)
Has annual parking permit (p < .001)	2 (1%)	17 (44%)	1 (1%)	20 (4%)
Member of GoClub (p = 0.383)	68 (20%)	6 (16%)	19 (15%)	93 (18%)
Undergraduate Student (p < .001)	225 (65%)	16 (41%)	126 (92%)	367 (70%)
Demographics and Trip Characteristics – Continuous				
Distance to Campus (p < .001)	1.65	2.24	2.16	1.83
Age (p = 0.001)	22.60	24.18	20.93	22.27
Years at UC Davis (p = 0.061)	1.89	2.74	2.13	2.02
Social Network				
Proportion of alters biking (p < 0.001)	52%	32%	28%	44%
Proportion of alters driving (p = 0.690)	21%	41%	20%	23%
Proportion of alters taking bus (p < 0.001)	14%	13%	44%	22%
Attitudes and Preferences				
Importance of: "Time it takes to make the trip" (p = 0.02)	4.10	4.62	4.26	4.18
"…cost of owning a car or other vehicle" (p = 0.431)	3.42	2.74	3.61	3.41
"Safety" (p < .001)	3.33	3.82	3.83	3.50
"Environmental impacts" (p = 0.347)	3.16	2.90	3.07	3.12
"Cost of transit" (p < .001)	3.05	1.97	3.79	3.17
"…physical exercise during my commute" (p < .001)	3.01	2.00	2.12	2.70
Agreement with: "I like biking" (p < .001)	4.39	3.28	3.55	4.09
"I need a car to do many of the things" (p < .001)	3.12	4.36	3.45	3.30
"I like transit" (p < .001)	3.08	2.38	3.79	3.21

p-values: ANOVA for continuous variables or chi-squared tests of distributions for categorical variables, across modes

In our sample, 72% are female and 70% are undergraduates. About 20% of the respondents are members of the goClub. The mean distance to campus is slightly less than two miles across the sample, and is somewhat shorter for those that bike than those that typically use either of the other two modes. The mean age of 22 is expected for a sample of university and graduate students, as is the average number of years at UC Davis; two. On average respondents have higher proportions of alters that bike than proportions of alters that drive or take the bus, however on average those that drive have higher proportions of alters that drive, and those that take the bus have higher proportions of alters that take the bus. Some of the attitudes and preferences are different across mode choices; drivers indicated higher levels of importance of “the time it takes to make the trip”, and lower levels of importance of “the cost of owning a car or other vehicle. Agreement with liking biking and liking transit is highest for bikers and bus riders, respectively.

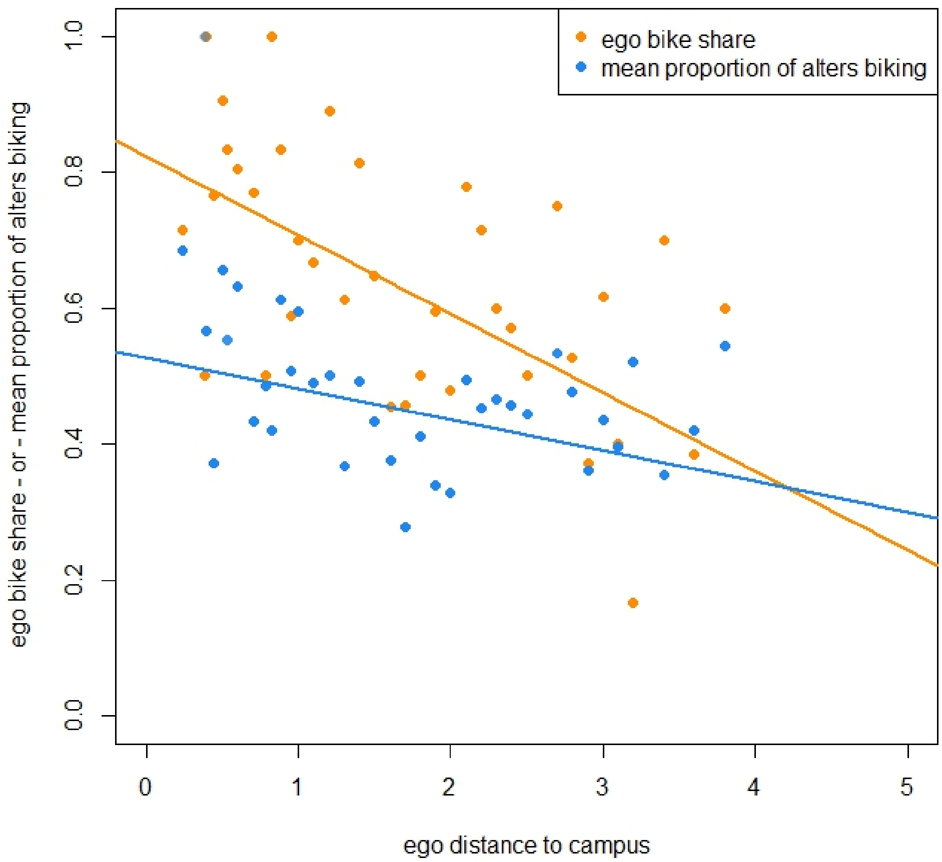
Commute Distance, Social Influence and Bike Share

Next we examine the relationships between the variables central to our analysis; bicycle mode share among egos, proportions of alters biking and commute distance. The ego bike share was computed by constructing commute distance bins based on rounding ego commute distance to campus to the nearest 10^{th} of a mile. Bins with fewer than 10 egos were combined. For each bin we computed the bike mode share among egos. The bike share is equal to the percentage of egos in the bin that commute by bike as their usual mode. For each bin that did not initially have at least 10 egos, after rounding distance to the nearest 10^{th} of a mile, we also computed the mean distance to campus. For bins that initially had more than 10 egos the mean distance to campus is simply the distance rounded to the nearest 10^{th} of a mile. In general, the higher the commute distance the lower the share of egos that bike (Figure 2).

In order to detect social influence, it is important to demonstrate that there is variance in the modes of alters even for egos that live very close to or far from campus. For example, if a student who lives far from campus and drives only has friends who drive, it would be impossible for their friends to exert social influence for a different mode choice. For each distance bin, we found the mean proportion of alters biking among the egos in that bin. Figure 2 presents the relationship between ego distance to campus and the mean proportion of alters biking for egos in each distance bin.

Since both the bike share of egos and the mean proportion of alters biking take on values between 0 and 1, the y-axis values are the same for both variables. Linear models were estimated for both relationships. Model results are plotted and the equations are presented below. Coefficients are significant at 0.1%, however in the model of the proportion of alters biking, the value is very small. There is a fairly high overall ego bike share as well as a moderate relationship between distance and bike share. Distance explains 40% of the variation in ego bike share. The proportion of alters that bike is relatively stable with respect to ego distance to campus; the coefficient is very small, and the model accounts for only 14% of the variation.

FIGURE 2 Share of Egos/Mean Proportion of Alters Biking by Ego Commute Distance

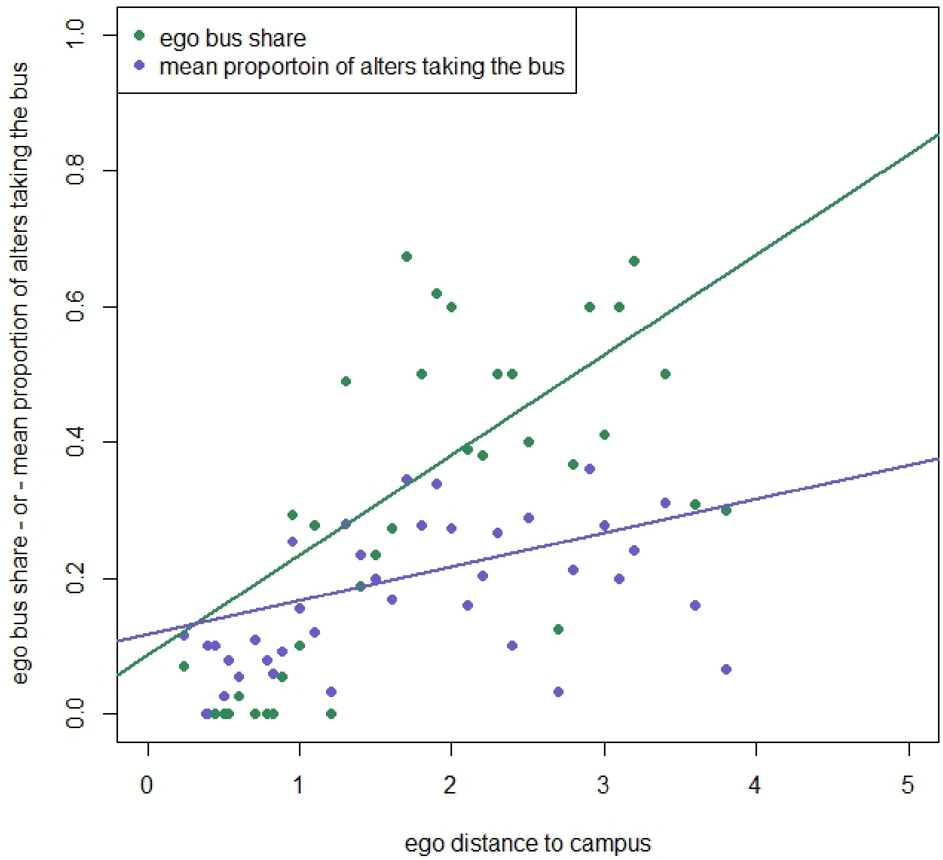


$$ego\ bike\ share\ \approx\ 0.82\ - 0.11 \times distance\ (R^2 =\ .40)$$

$$proportion\ of\ alters\ biking\ \approx\ 0.52\ - 0.04 \times distance\ (R^2 =\ .14)$$

A similar pattern occurs, though in the opposite direction, for the other two primary modes; bus and drive. Figure 3 shows ego bus share and the mean proportion of alters taking the bus for each commute distance bin. Ego bus share and the mean proportion of alters taking the bus were computed in the same ways for bus as they were for bike. As distance increases, ego bus share increases, as does (though to a lesser extent) the mean proportion of alters that take the bus. The linear models explain 34% of the variation in ego bus share and only 18% of the variation in the mean proportion of alters that take the bus. The coefficient in the model of the mean proportion of alters taking the bus is very small, though all coefficients are significant at 0.1%.

FIGURE 3 Share of Egos and Proportion of Alters Taking Bus by Ego Commute Distance



$$ego\ bus\ share\ \approx\ 0.09 + 0.14 \times distance\ (R^2 =\ .34)$$

$$proportion\ of\ alters\ taking\ bus\ \approx\ 0.\ 12 + 0.05 \times distance\ (R^2 =\ .18)$$

Figures 2 and 3 show ego mode choice changes relative to ego commute distance. Although the mean proportion of alters using each mode also changes with ego distance to campus, there is a much weaker relationship between ego commute distance and the mean proportion of alters that bike or take the bus. We use two approaches to explore the relationships between distance, social influence and transportation mode choice.

MODELING APPROACHES

<u>Two Stage Residual Inclusion Model</u>

We use a two-stage residual inclusion (2SRI) model, described as an instrumental variable or control function approach (for a detailed discussion of the model and its properties see Rivers and Vuong 1988 and Wooldridge 2001 chapter 15). We use the 2SRI model to account for endogeneity that may arise when an ego and his alters face similar commute characteristics. Egos who bike tend to have higher proportions of alters that bike, than egos who drive or take the bus (Table 1). While this is likely evidence of social influence, it may also be the result of other mechanisms. For example, an ego and his alters may make the same mode choice because they commute about the same distance, or have similar levels of access to transit. To reliably estimate the effect of social influence, while accounting for this potential endogeneity; the shared environment of an ego and his alters, we use the 2SRI model. Any instrumental variable approach requires the identification of an instrument that is correlated with the endogenous variable, but not correlated with the unobserved (or observed) variation in the outcome variable of interest – the choice of the ego to bike.

The alters' neighborhood and commute characteristics affect the mode choices of the alters but we do not expect characteristics of the alters' neighborhoods to directly affect the mode choice of the ego. One may be concerned that, when an ego lives in the same neighborhood as one of his alters, they experience the same neighborhood characteristics. However, we calculate the effect of the alters' neighborhood characteristics on alters' mode choices for the group of alters collectively. Alters with neighborhood characteristics conducive to biking are more likely to bike, and alters with neighborhood characteristics more conducive to other modes are less likely to bike.

Therefore, the *average* of a neighborhood characteristic across an ego's alters is different from that neighborhood characteristic for the ego, and is correlated with the proportion of alters that bike, but not whether the ego bikes or not.

Instrumental Variable

Our instrument is the alter neighborhood characteristic; neighborhood biking density. All of the approximately 3,300 individuals who participated in the CTS, or our social networks survey, or are alters whose usual commute mode and cross-streets were noted by an ego are potential neighbors. For each alter, we compute neighborhood biking density as the percentage of neighbors within 0.3 miles of the alter, that bike. If there are 20 neighbors within 0.3 miles of an alter and 15 of them bike that alter has a neighborhood biking density of 0.75. For each ego-network we calculate the average neighborhood biking density among the alters, who live in Davis. We exclude alters outside of Davis because we have limited neighborhood information for these alters; there may be only a few neighbors who participated in the surveys noted above, and these few neighbors are not likely to accurately represent neighborhood biking density. The use of the alters' geographic locations also excludes from the sample any ego that did not provide geographic information for their alters, and results in a total sample for these models of 397.

Sample Segments

We divide the sample of 397 respondents into three segments based on commute distance; those within 1½ miles of campus, those living 1½ to 3 miles of campus, and those living 3 to 5 miles from campus. The last segment is slightly broader in terms of distance, but has the fewest observations. Because segmenting the sample results in fairly small samples we utilize only the focal variables in these estimations; the instrumental variable, the proportion of alters that bike and the binary choice of the ego to bike or not bike.

The average neighborhood biking density of the alters is used to predict the proportion of alters that bike in the first stage of the 2SRI model. The residuals from the first stage are saved, and entered into the second stage as an explanatory variable along with the endogenous variable, and any other predictors. The stage-1 residual acts as a control variable for the proportion of alters biking. The coefficient estimate for the proportion of alters biking and its significance in this stage of the model make up the central test of whether social influence is relevant to mode choice, even when accounting for shared environment of the egos and alters.

Multinomial Logit Model

Following the 2SRI model, we present a multinomial logit (MNL) model (see Ben-Akiva and Lerman 1985) of the ego's mode choice. In the MNL model, we include the effect of distance on mode choice, the effect of social influence on mode choice and the effect of the interaction of these two terms. Interaction terms are used when the effect of one independent variable, x_1 is expected to vary with respect to another independent variable, x_2. The coefficient on the interaction term is negative if larger values of x_2 result in a smaller effect of x_1. The coefficient is positive if larger values of x_2 result in a larger effect of x_1.

We expect the coefficient on the interaction of commute distance and social influence to be negative for the mode choice of bike. That is, at longer commute distances, the effect of social influence on the choice to bike is smaller. Although we also hypothesize at very short commute distances the effect of social influence is small, the interaction term does not capture the potentially small effect of social influence for those who have very short commute distances. In the MNL model we do not require alter neighborhood information, so all alters are retained in the ego-networks, regardless of whether or not they live in Davis. We also include socio-demographic variables, attitudes and trip characteristics as explanatory variables.

RESULTS

Two-Stage Residual Inclusion Model

The 2SRI model results are presented in Table 2 and demonstrate how the effect of social influence varies with respect to distance from campus. Model estimations were carried out using the statistical software R (R Core Team 2013). Those that live closest to campus have a high overall share of biking, but the choice to bike is less affected by social influence than it is for those that live a medium distance from campus. Those who are farthest from campus have a similar response to social influence, but a lower share of biking. Social influence has the largest effect on the middle segment. These results also confirm social influence is relevant to transportation mode choice even when accounting for potential endogeneity related to the shared environment of the ego and alters.

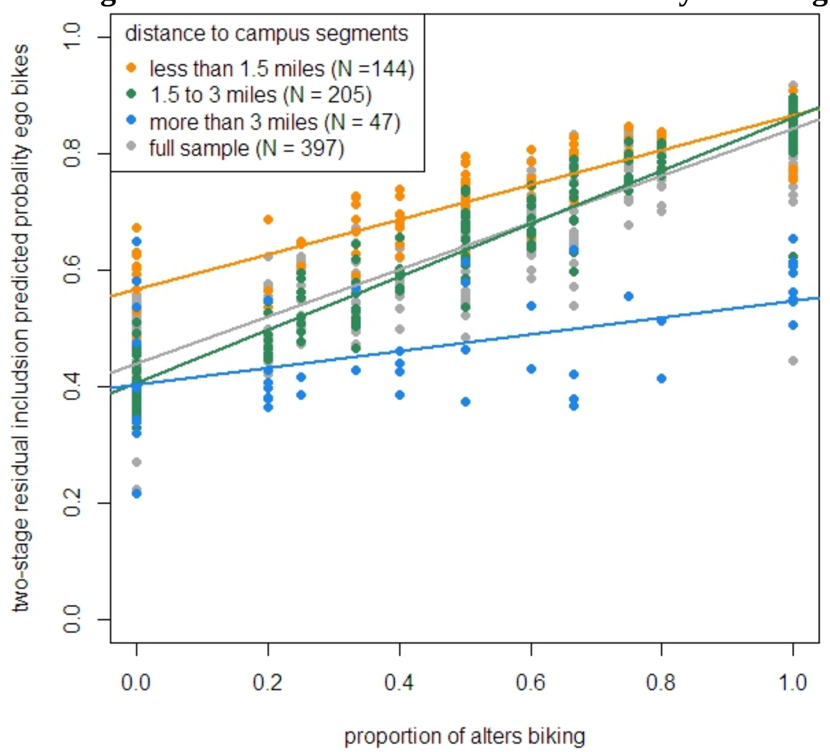
TABLE 2 Two-Stage Residual Inclusion Models of Ego's Choice to Bike

Segment	Bike share N (%)	Stage 2 Binary Probit Model Coefficients			F statistic (stage 1 instrument)
		Intercept	alters biking	stage 1 residual	
0 to 1½ miles (N = 144)	109 (76%)	-0.527 (p = .351)	**1.995 (p = .026)**	-1.211 (p = .203)	22.341
1½ to 3 miles (N = 205)	126 (61%)	**-0.896 (p = .058)**	**2.693 (p = .009)**	-1.477 (p = .156)	12.637
3 to 5 miles (N = 47)	22 (35%)	-0.910 (p =.150)	1.847 (p = .169)	-1.764 (p = .228)	8.877
Full sample (N = 397)	258 (61%)	**-0.875 (p = .002)**	**2.497 (p < .001)**	**-1.506 (p = .008)**	50.081

Though the stage 1 residual is not a significant predictor in any of the 2SRI models, we don't take this as evidence of exogeneity (see Wooldridge 2001 p. 473), since this test of fails for the full sample model. Although the F statistic, a criteria for determining the validity of an instrument (Staiger and Stock 1997), is less than 10 for the third model it may be partially due to the small sample size, and the results indicate that overall neighborhood biking density is a valid instrument for the proportion of alters that bike; as the F statistic is greater than 10 in the other three model estimations. The stage 1 residuals are not normally distributed, likely because the proportion of alters that bike is not strictly continuous.

Figure 5 illustrates how the effect of social influence is impacted by commute distance. The predicted probabilities, from the results of the 2SRI, according to the proportion of alters that bike, as well as a best-fit line for each segment are plotted. The effect of social influence is greatest (steepest curve) for those at medium distances from campus, and smallest for those at long distances from campus and those with the shortest commutes to campus.

FIGURE 4 Two-Stage Residual Inclusion Predicted Probability of Biking by Segments



Multinomial Logit Model Results

MNL model results are summarized in Table 3. The base alternative in this model is bike. The mode shares are somewhat imbalanced; the model for drive has 9 variables for only 39 cases. The ρ^2 value of 0.45 indicates that the model explains about half of the variance in ego mode choice compared to a model with constants only. The likelihood ratio test compares the model presented to a model without the interaction term (but with distance and the proportion of alters biking). The interaction term adds explanatory power.

The control variables have the expected signs (magnitudes are not directly comparable, however the odds ratios are shown). Those who indicate transit costs are important are more likely to take the bus, while those who indicated lower importance are more likely to drive. The higher the importance of safety the higher the likelihood of biking. Undergraduates are 13 times more likely to take the bus than graduate students. Those who need a car are more likely to drive than bike. Liking transit and liking biking have the expected effects.

TABLE 3 Multinomial Logistic Regression of Mode Choice (Bike, Drive, Bus)

Full Model Sample is 521	Drive		Bus	
Base Alternative Bike; N = 345 (66 %)	N = 39 (7%)		N = 137 (26%)	
Variables in Model (Coefficients reflect effect on likelihood of choosing drive or bus)	Estimate (Std. error)	Odds ratio	Estimate (Std. error)	Odds ratio
Proportion of alters biking	-2.278 (2.025)	0.10	-4.411*** (1.359)	0.01
Ego's distance to campus	0.666 (0.440)	1.95	0.622** (0.304)	1.86
Interaction: alters biking × distance	-0.202 (0.902)	0.82	0.838 (0.629)	2.31
Ego is undergrad (1 = undergraduate student)	-0.852* (0.479)	0.43	2.588*** (0.445)	13.3
Importance of "Safety"	0.502** (0.188)	1.65	0.091 (0.122)	1.10
Importance of "Transit costs"	-0.632*** (0.179)	0.53	0.389*** (0.108)	1.48
Agreement with "Need a car to do most things…"	0.852*** (0.243)	2.34	0.167 (0.122)	1.18
Agreement with "Like transit"	-0.274 (0.195)	0.76	0.722*** (0.150)	2.06
Agreement with "Like biking"	-0.947*** (0.215)	0.39	-1.199*** (0.172)	0.30
Intercept	-0.888 (1.906)	0.41	-2.985** (1.279)	0.05
Model Diagnostics				
Log likelihood: -233.89	Likelihood ratio test statistic: 20.44; Pr(>Chisq) < 0.001		Rho-squared (market share base): 0.45	

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Of primary interest are the effects of social influence and commute distance, and the interaction of these two variables. Higher proportions of alters biking decreases the likelihood of choosing either of the alternative modes (drive or bus). Those that live further from campus have a higher likelihood of choosing bus or drive. The coefficient on the interaction of distance and the proportion of alters biking is positive for the choice of bus. Related to biking, this term indicates that as distance increases, the effect of the proportion of alters that bike on the likelihood of choosing to bike decreases. The p-value on this coefficient is only 0.183, however the standard error is .629 and is smaller than the coefficient value of 0.838. The estimate likely indicates the true direction of the effect, even though it is not significant. Not so for the coefficient on the interaction term for the choice of drive; the p-value and standard error are very large and we expect this coefficient to be positive.

We also estimated two alternative specifications of this model; one with only the choices of bike and bus, because the drive share is very small, and one where we relaxed the restriction that respondents must live in Davis. This allows individuals with longer commute distances to be in the sample; however, commute distance is not continuously distributed since anyone who lives outside of Davis lives at least 10 miles from campus. The variables of primary interest are presented in Table 4[10]. In both alternative specifications the coefficients of interest are in the expected directions and are significant; except for distance to campus for the bus alternative in the model including respondents outside of Davis. There is a larger sample in the model without the distance restriction, and most of the added cases drive.

[10] Full model results are available in our online appendix.

TABLE 4 Select Terms from Alternative MNL Models of Mode Choice

Model Including Respondents Outside of Davis				
Model Sample is 567	Drive		Bus	
Base Alternative Bike; N = 347 (61 %)	N = 78 (14%)		N = 142 (25%)	
Variables in Model (Coefficients reflect effect on likelihood of choosing drive or bus)	Estimate (Std. error)	Odds ratio	Estimate (Std. error)	Odds ratio
Proportion of alters biking	-4.291*** (1.279)	0.014	-4.573*** (0.921)	0.01
Ego's distance to campus	0.427** (0.141)	1.532	0.189 (0.140)	1.20
Interaction: alters biking × distance	0.856** (0.400)	2.353	0.929** (0.385)	2.53
Model with Bike and Bus Only				
Model Sample is 482	Drive		Bus	
Base Alternative Bike; N = 345 (72%)	-		N = 137 (28%)	
Variables in Model (Coefficients reflect effect on likelihood of choosing drive or bus)	- -	-	Estimate (Std. error)	Odds ratio
Proportion of alters biking	- -	-	-4.761*** (1.386)	0.01
Ego's distance to campus	- -	-	0.529* (0.309)	1.70
Interaction: alters biking × distance	- -	-	1.071* (0.639)	2.92

Significance codes: 0 '***' 0.001 '**' 0.01 '*' 0.05 '.' 0.1 ' ' 1

Both the 2SRI and the MNL model results demonstrate that the effect of social influence is impacted by the distance an ego commutes to campus. Except at very short commute distances, where the effect of social influence is small; as commute distance increases, the effect of social influence to bike decreases.

DISCUSSION AND CONCLUSIONS

Our results support the hypothesis that the effect of social influence is impacted by external factors, such as commute characteristics. In 2SRI and MNL models we found that as commute distance increases the effect of social influence related to biking decreases, except for individuals with very short commute distances, for whom social influence also has a small effect. Our findings are somewhat limited by the study area – no one in our sample lives between 5 and 10 miles from campus, distances that might be particularly relevant to our research question. Additionally, the mode shares in our sample are somewhat imbalanced and this has an effect on results. Future work should explore this question in areas with a more continuous distribution of commute distances and balanced mode shares.

Future work should also investigate how other relevant factors in transportation mode choice interact with and impact the effects of social influence. Responses to social influences may vary by attitudes, land use characteristics, car ownership, gender or other factors. Further, we found that social influence has most effect on individuals with medium commute distances; it may be of interest to more precisely identify bounds within which social influence to bike is most important, or how these bounds might change depending on transportation mode.

A better understanding how social influence on transportation mode choice is impacted by external factors will allow sustainable transportation programs to more effectively utilize social influence as a tool. If social influence to bike is only relevant for individuals expected to have commute distances within a specific range, these are the individuals that should be targeted by programs aimed at increasing bicycle commuting.

REFERENCES

Axelrod, Lawrence J., and Darrin R. Lehman. 1993. “Responding to Environmental Concerns: What Factors Guide Individual Action?” Journal of Environmental Psychology. Vol. 13, 1993, pp. 149–159.

Axsen, John and Kenneth S. Kurani. Interpersonal influence within car buyers’ social networks: applying five perspectives to plug-in hybrid vehicle drivers. Environment and Planning A. Vol. 340, Issue 1, 2011, pp. 1-26.

Ben-Akiva, Moshe, and Lerman. 1985. Discrete Choice Analysis. MIT Press.

Carrasco, Juan Antonio, and Eric J. Miller. Exploring the Propensity to Perform Social Activities: a Social Network Approach. Transportation. Vol. 33, 2006, pp. 463–480.

Carrasco, Juan-Antonio, and Eric J. Miller. The Social Dimension in Action: A Multilevel, Personal Networks Model of Social Activity Frequency Between Individuals. Transportation Research Part A: Policy and Practice. Vol. 43, 2009, pp. 90–104.

Carrell, Scott E., Richard L. Fullerton, and James E. West. Does Your Cohort Matter? Measuring Peer Effects in College Achievement. Journal of Labor Economics, University of Chicago Press. Vol. 27, 2009, pp. 439-464.

Dugundji, E. R., and J. L. Walker. Discrete Choice with Social and Spatial Network Interdependencies: An Empirical Example Using Mixed Generalized Extreme Value Models with Field and Panel Effects. In Transportation Research Record: Journal of the Transportation Research Board, No. 1921, Transportation Research Board of the National Academies, Washington, D.C., 2005, pp. 70–78.

Ewing, Reid, and Robert Cervero. Travel and the Built Environment. Journal of the American Planning Association. Vol. 76, No. 3, 2010, pp. 265–294.

Farber, Steven and Antonio Páez. My car, my friends, and me: a preliminary analysis of automobility and social activity participation. Journal of Transport Geography. Vol. 17, 2009, pp. 216–22.

Fowler, J. H, and N. A Christakis. Dynamic Spread of Happiness in a Large Social Network: Longitudinal Analysis over 20 Years in the Framingham Heart Study. British Medical Journal. Vol. 337, 2008, pp. 1–9.

Freeman, Linton C. The Impact of Computer Based Communication on the Social Structure of an Emerging Scientific Specialty. Social Networks Vol. 6, 1984, pp. 201–221.

Goetzke, F, and T. Rave. Bicycle Use in Germany: Explaining Differences Between Municipalities with Social Network Effects. Urban Studies. Vol. 48, No. 2, 2010, pp. 427–437.

Goetzke, Frank. Network Effects in Public Transit Use: Evidence from a Spatially Autoregressive Mode Choice Model for New York. Urban Studies. Vol. 45, No. 2, 2008, pp. 407–417.

Granovetter, Mark S. The Strength of Weak Ties. American Journal of Sociology: Vol. 78, 1973, pp. 1360–1380.

Koehly LM, Loscalzo A. Adolescent obesity and social networks. Preventing Chronic Disease. Vol. 6, 2009, pp 1-8.

Larsen, Jonas, Urry, John and Axhausen, Kay. Coordinating Face to Face Meetings in Mobile Network Societies. Information, Communication and Society. Vol. 11, No. 5, 2008 pp. 640-658.

Maness Michael, Cinzia Cirillo and Elenna R Dugundji. Generalized behavioral framework for choice models of social influence: Behavioral and data concerns in travel behavior. Journal of Transport Geography. Vol. 46, 2015, pp 137–150.

Mokhtarian, Patricia; & Ilan Salomon. Modeling the Desire to Telecommute: the Importance of Attitude Factors in Behavioral Models. Transportation Research Part A. Volume 31, Issue 1, January 1997, Pages 35–50.

Páez, Antonio, and Darren M Scott. 2007. Social Influence on Travel Behavior: a Simulation Example of the Decision to Telecommute. Environment and Planning A, Vol. 39, 2007, pp. 647–665.

Pike, Susan (2014). Travel Mode Choice and Social and Spatial Reference Groups; Comparison of Two Formulations. Transportation Research Record: Journal of the Transportation Research Board, No. 2412, Transportation Research Board of the National Academies, Washington, D.C., 2014, pp. 75–81.

Popovich, Natalie. Results of the 2013-14 Campus Travel Survey. Institute of Transportation Studies, University of California, Davis, Research Report UCD-ITS-RR-14-14, 2014.

R Core Team. R: A language and environment for statistical computing. R Foundation for Statistical Computing, Vienna, Austria, 2013. URL http://www.R-project.org/.

Rivers, D., and Q. H. Vuong. Limited information estimators and exogeneity tests for simultaneous probit models. Journal of Econometrics. Vol. 39, 1988, pp. 347–366.

Sacerdote, Bruce. Peer Effects with Random Assignment: Results for Dartmouth Roommates. The Quarterly Journal of Economics, Vol. 116, 2001, pp. 681–704.

Scott, Darren M. Ivy Dam, Antonio Páez, and Robert Wilton. Investigating the effects of social influence on the choice to telework. Environment and Planning A. Vol. 44, 2012, pp. 1016–1031.

Staiger, Douglas and James H. Stock. 2SRI Regression with Weak Instruments. Econometrica.Vol. 65, No. 3, 1997, pp. 557-586.

Walker, Joan L., Emily Ehlers, Ipsita Banerjee, and Elenna R. Dugundji. Correcting for Endogeneity in Behavioral Choice Models with Social Influence Variables. Transportation Research Part A: Policy and Practice Vol. 45, No. 4, 2011, pp. 362–374.

Wilton, Robert D., Antonio Páez, and Darren M. Scott. Why Do You Care What Other People Think? A Qualitative Investigation of Social Influence and Telecommuting. Transportation Research Part A: Policy and Practice. Vol. 45, No. 4, 2011, pp. 269–282.

Wooldridge, Jeffrey M. Econometric Analysis of Cross Section and Panel Data. The MIT Press, Cambridge, Massachusetts and London, England 2001.