Detecting Social Network Effects on Willingness to Pay for Environmental Improvements using Egocentric Network Measures

John Janmaat*, Solomon Geleta†, John Loomis‡

October 27, 2019

Abstract

Since people care about each other, an individual's willingness to pay to protect an environmental good or service will reflect their concern for others who would also be impacted by a change in the good or service. Through the extended social network, an individual's willingness to pay will reflect the impacts on people who they do not immediately know. If this effect is not considered, willingness to pay estimates can be biased. However, extended social networks are difficult to measure. We therefore explored the potential for egocentric social networks to help explain variations in willingness to pay. Given the conventional way of describing social networks, we demonstrate that egocentric social network measures should not be related to willingness to pay if there is no relationship between the social network measure and the willingness to pay for a change in the environmental good or service. When the social network measures are increasing in the willingness to pay for an environmental improvement, then a regression of willingness to pay on these social network measures will show a positive relationship. Empirically, we find such a relationship in the results of a choice experiment conducted in the central Okanagan of British Columbia. However, we also find that a measure of peoples assessment of the benefits of development relative to the environmental impacts was a more effective predictor. This may be a consequence of how the respondent's egocentric networks were measured. Alternative approaches to measuring the egocentric social network may be necessary.

*john.janmaat@ubc.ca, Department of Economics, Philosophy and Political Science, University of British Columbia, Kelowna, Canada

†sog092@usask.edu, Department of Agricultural and Resource Economics, Colorado State University, Fort Collins, Colorado, USA

‡john.loomis@colostate.edu, Department of Agricultural and Resource Economics, Colorado State University, Fort Collins, Colorado, USA

Keywords : Social network effects, egocentric, sociocentric, prosocial, ego-alter, random utility, network centrality.

1 Introduction

We are social animals, and how much we care about the environment is likely influenced by those we associate with. However, most nonmarket valuation studies in economics do not consider people's social ties. This is partly due to the fact that the development of economic models demonstrating the importance of social networks for individual willingness to pay is relatively recent [Neilson and Wichmann, 2014]. These results have been developed based on a completely mapped social network. In most situations of interest, we cannot practically map the whole social network, limiting the applicability of the theoretical results. Researchers in a number of disciplines have been finding that egocentric social network measures add explanatory power to statistical models. Herein we similarly explore the use of egocentric social network measures as additional regressors for analyzing the results of a choice experiment. We find that these measures do make a significant contribution, and argue that incorporating such measures may reduce the bias that occurs if social network effects are ignored all together.

Formally, a social network is defined as "a group of people [,each of] who are connected to some or all of the others following a random or particular pattern in network system [He et al., 2014]." In network language, this translates into links between nodes in a graph [Watts and Strogatz, 1998, Watts and Witham, 2012]. Nodes show the position of different actors in a network and links measure how involved an actor is, how connected the actor is [Diani and McAdam, 2003] and represent a flow of potential information and interaction between those actors [Burt, 1984, Carrasco et al., 2008]. Network positions could facilitate or constrain actors [Diani and McAdam, 2003], and this can be related to centrality measures such as degree and betweenness centrality [Freeman et al., 1979]. The degree centrality of an actor is the number of other actors they are directly connected to, while betweenness centrality measures the frequency with which an actor lies on the shortest path between two other actors. High betweenness centrality may increase the potential to locate and gain resources from the network and/or exert direct influence on actions and priorities of others in the network [Watts and Witham, 2012, Scott, 2011, Jackson, 2009, Carrington et al., 2005, Carrasco and Miller, 2009]. Social network maps provide a way to visualize the routes for information and resource flow, and centrality measures provide a quantitative measure reflecting the different positions actors have in the network.

In this paper, we investigate the effects of egocentric network centrality measures on individuals marginal rates of substitution between attributes in a choice experiment and by extension the effects on the willingness to pay for the underlying environmental services reflected by the attributes. We begin by exploring the way that egocentric social network measures may relate to willingness to pay. We demonstrate that for a relationship to exist between egocentric social network measures and willingness to pay, either the whole social network is not considered or the egocentric structure is itself a function of willingness to pay. The choice experiment was administered in the central Okanagan region, including the city of Kelowna, in British Columbia, Canada. The experiment offered participants several opportunities to choose between a status quo and two alternatives. We analyze the survey results using both fixed effects multinomal logit and mixed logit models. We find that the network measures have signs consistent with there being a relationship between the structure of the network and the willingness to pay. However, significance of the estimated parameters depends on the presence of other covariates, principally one that measures the respondent's evaluation of the trade-off between development and environmental damage.

The remainder of the paper is organized as follows. Section 2 reviews the literature on social network theories and choice experiments. Section 3 uses the model of Neilson and Wichmann [2014] to examine how egocentric social network measures might be related to WTP estimates. Section 4 describes the survey methodology and the structure of the network data. Section 5 describes the econometric model. Section 6 presents the results. Section 7 offers some discussion of the results, and the paper is concluded with a brief conclusion.

2 Social Network Theories and Choice Modeling

The importance of social networks has long been recognized, and Social Network Analysis is a well established field [Prell, 2012]. Humans are social animals, and it follows that people recognize each other as part of a network system [Diani, 2003, Diani and McAdam, 2003, He et al., 2014, Golub and Jackson, 2010]. Social network analysis provides mechanisms for both characterizing the position of individuals within a network and describing the form of the entire network. [Watts and Witham, 2012, Emirbayer and Goodwin, 1994].

While economics has long recognized concern for others (e.g. altruism and bequest motives in relation to nonmarket valuation), explicit inclusion of social network metrics in economic analyses is relatively recent [Easley et al., 2012]. Some of the notable studies include: the importance of information obtained from social networks in influencing preferences and choice decisions [Manski, 1993, Banerjee, 1992, Samuelson, 2004]; interdependence of choices and social interaction [Brock and Durlauf, 2002]. [Stopher, 1980]; peer influence on product attribute preference [Narayan et al., 2011]; word-of-mouth effect in alternative fuel vehicle choices [Struben and Sterman, 2008]; social network effects on intercity travel mode choice behavior [Dugundji and Walker, 2005]; residential location choice using simulated network data with varying degrees of distributions and clustering parameters [Páez et al., 2008]; the effects of vehicle, contextual, and social network attributes on the latent demand for electric cars using fixed and random mixed logit models [Rasouli and Timmermans, 2016].

Applications of social network analysis to natural resources and non-market valuation are also emerging. Some of these studies that incorporate social interactions into public goods include Elliott and Golub [2019], Crona and Bodin [2006], Ostrom [2015], Pretty and Ward [2001], Prell et al. [2009], Bouma et al. [2008]. Key results include the increase in the likelihood of collective action, opinion formation, cooperation, resource sharing and social influence. For instance, Bramoullé and Kranton [2007] show how social networks can influence some individuals to contribute and others to free-ride. Neilson and Wichmann [2014] show how aggregate willingness to pay is related to network centrality and that people are willing to pay more in aggregate for a public good when that public good provides more benefit to people more central to the society. Watts and Witham [2012] demonstrates how willingness to pay or willingness to accept amounts are influenced by the position of actors (opinion leaders) in a network. Newton [2010] show the effect of coalition behavior on local public goods provision. Granovetter [2005] demonstrates how overcoming the free-rider problem is more likely in groups with a dense and cohesive social network.

The overarching result of this emerging work is that social interactions generate variations in overall valuations of non-market goods, and that treating individuals as independent units of analysis, as traditionally done in economic science, overlooks this fact. Valuing public goods therefore should

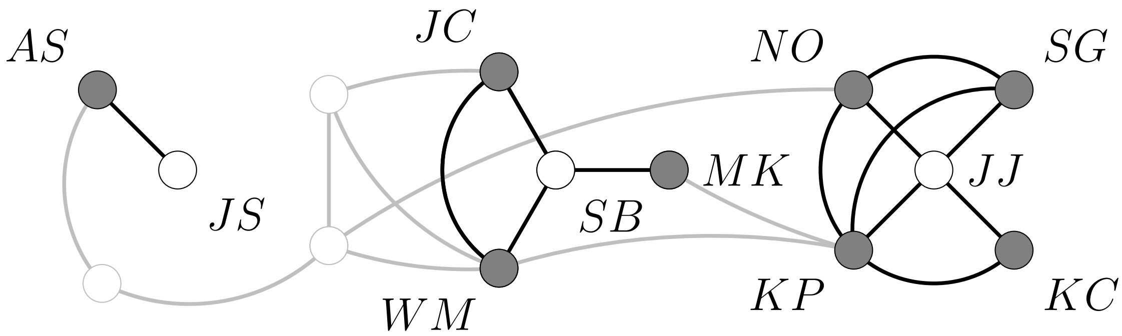

Figure 1: A social network. Three egos, JS, SB and JJ are identified, together with their egocentric networks.

be analyzed through both individual characteristics and social structural attributes [Carrasco et al., 2008, Neilson and Wichmann, 2014].

3 Modeling Social Network Effects

Figure 1 shows a hypothetical social network with 14 members. Many social network analysis focus on the structure of the entire network, how dense the linkages are, whether there are isolated sub-networks, whether there are key nodes through which many paths between nodes pass, etc. For surveys of the public, it is generally not possible to map the entire social network, making network level analyses impossible. However, surveys can gather information about the respondent's network, their 'egocentric' network. Below, we introduce egocentric network centrality measures into regressions predicting responses to our choice experiment. Before doing so, we identify one possible explanation for the existence of a positive relationship between egocentric network measures and choices [Borgatti et al., 2013, Hanneman and Riddle, 2005].

Networks can be directed or undirected. In directed networks, a tie goes from one node to another, and may not be matched by a tie in the other direction. In an undirected network, all ties between nodes are reciprocal. Most of the analysis of ego networks uses undirected graphs, which we will do as well, by making any reported directional ties reciprocal [Hanneman and Riddle, 2005].

Three egos are identified in Figure 1, JS, SB and JJ. We use these below to illustrate several common network centrality measures, two of which we will introduce into our regressions. JS has an ego network with only one alter, AS. SB has three alters, MK, WM and JC. Two of these alters are linked together, while the other is isolated from anyone else in SB's ego network. JS has four alters, SG, NO, KP and KC. Every alter in JJ's network has at least one tie to another alter. These three examples illustrate that ego networks can differ in both size and structure. The example also illustrates the challenge of using egocentric networks when the entire network may matter. The network has ties between alters across ego networks, and ties from alters to network members that do not have a direct tie to one of the three selected egos. These additional members and the links between them may be important to the behavior of the egos we have identified, through their links to the alters, but our measures of the ego networks do not register information about the alters' networks beyond links between alters who are connected to the ego.

The three matrices below, known as adjacency matrices, provide a representation of the ego network for the three individuals, JS, SB and JJ. The ego is located at the top left of the matrix. When there is a tie between two different people in the network, that entry in the matrix is set to 1. Where there is no tie, then the entry is 0. The matrix is symmetric, reflecting that this is an undirected network, and we are assuming that the network ties are all of equal strength.

$$\begin{array}{c} JS \\ AS \end{array} \begin{bmatrix} 0 & 1 \\ 1 & 0 \end{bmatrix} \qquad \begin{array}{c} SB \\ MK \\ JC \\ WM \end{array} \begin{bmatrix} 0 & 1 & 1 & 1 \\ 1 & 0 & 0 & 0 \\ 1 & 0 & 0 & 1 \\ 1 & 0 & 1 & 0 \end{bmatrix} \qquad \begin{array}{c} JJ \\ SG \\ NO \\ KP \\ KC \end{array} \begin{bmatrix} 0 & 1 & 1 & 1 & 1 \\ 1 & 0 & 1 & 1 & 0 \\ 1 & 1 & 0 & 1 & 0 \\ 1 & 1 & 1 & 0 & 1 \\ 1 & 0 & 0 & 1 & 0 \end{bmatrix}$$

For the statistical analysis to follow, we will use two numerical measures that represent the structure of the ego networks, degree centrality and transitivity, also know as the clustering coefficient. In the exposition here we also include closeness centrality, which is one additional example among several other network centrality measures that could be examined. We generated our network measures using Gephi [Bastian et al., 2009]. The definitions that follow are based on the calculations implemented in Gephi, which uses algorithms set out in Brandes [2001] and Latapy [2008] for the centrality measures we use.

Degree centrality simply measures the number of connections that the ego has [Freeman, 1978, O'Malley et al., 2012].

$$\text{Degree centrality}_i = C_D(i) = \sum_{j \in X_i} \theta_{ij} \tag{1}$$

where X_i is the set of vertices in the network that are joined to i by a single edge, and includes i itself, with $\theta_{ij} = 0$ when $i = j$ and $\theta_{ij} = 1$ when there is a connection between i and j. We define X to be the extended network for agent i, which included all the agents that can be reached along paths that include agent i. For regression purposes, we scale the degree centrality and the other centrality measures to lie between zero and one, by dividing the individual's degree by five, the maximum number of alters that could be named. This does introduce the potential for a truncation issue, when the network is larger. We assume that such situations are sufficiently infrequent that we can ignore them.

Closeness centrality measures the shortest paths linking the ego to their alters, either directly or indirectly [Opsahl et al., 2010]. Following the notation used in Brandes [2001] we define closeness centrality as:

$$\text{Closeness centrality}_i = C_C(i) = \frac{1}{\sum_{j \in X_i} d(i,j)} \tag{2}$$

where $d(i,j)$ is the length of the shortest path from i to j, with $d(i,i) = 0$. Notice that when X_i only contains agent i herself, closeness centrality is undefined. We use the approach implemented in Gephi, namely setting the closeness centrality equal to zero when the ego has not specified any alters. Such conventions are not uncommon for social network measures, and need to be considered when interpreting results. Notice also that $C_C(i) = 1/C_D(i)$ for ego networks where the ego is directly linked to all the alters. As for degree centrality, we normalize this value to lie between zero and one.

Clustering coefficient or transitivity is defined as

$$\text{Clustering coefficient}_i = C_T(i) = \frac{2\sum_{j,k \in X_i; j<k} I(j,k)}{C_D(i)[C_D(i) - 1]}$$

Table 1: Centrality measures for example ego networks.

Measure	Ego Network JS	SB	JJ
Degree centrality (unscaled)	1	3	4
Closeness centrality	$1/1 = 1$	$1/3 = 0.333$	$1/4 = 0.25$
Clustering coefficient	0[a]	$1/3 = 0.333$	$4/6 = 0.667$

[a] 0 by convention used in Gephi.

where $I(j,k)$ is an indicator that equals 1 if there is an edge joining j and k and zero otherwise [see Watts and Strogatz, 1998]. When there are C_D vertices, then there are $C_D(C_D-1)/2$ possible edges, and the summation provides the number of edges that are actually present. Perfect transitivity implies everyone in a network is connected , i.e., a network whose components are all cliques [Opsahl et al., 2010]. Notice that C_T is undefined when $C_D = 0$ or $C_D = 1$. The clustering coefficient can also be interpreted as the number of closed triangles connecting triplets of vertices in the network as a share of the total possible triangles connecting triplets. If the network only contains one or two vertices, then no triangles are possible. Again following the convention adopted by Gephi, for networks where C_T is undefined, we set the value to zero.

We aim to examine whether or not characteristics of the self reported network of a survey respondent impact on the respondents willingness to pay for environmental goods and services. To do so we propose that individuals care about each others utility. Adapting the formulation used by Neilson and Wichmann [2014] we begin by assuming that the individual's utility can be written in quasilinear form as:

$$V_i(x_i, g) = x_i + v_i(g) \tag{3}$$

where x_i is individual i's consumption of a private good and g a public good. If we assume that individual i cares about the utility of her friends, then we can write the public good contribution to her utility as

$$v_i(g) = (1-\lambda_i)u_i(g) + \lambda_i \sum_j a_{ij} v_j(g) \tag{4}$$

where $u_i(g)$is ego i's private utility from the public good and a_{ij} is a measure of how much i cares about the utility of j. By convention we assume that $\sum_j a_{ij} = 1$ and $a_{ii} = 0$. The parameter $0 \leq \lambda_i < 1$ measures the amount of i's utility that comes from the utility of i's friends. Neilson and

Wichmann assume that the entire network is known, and thereby express equation 4 in matrix form as

$$\mathbf{v} = (\mathbf{I} - \mathbf{\Lambda})\mathbf{u} + \mathbf{\Lambda}\mathbf{A}\mathbf{v}$$

which can be solved for $\mathbf{v}$ to yield

$$\mathbf{v} = (\mathbf{I} - \mathbf{\Lambda}\mathbf{A})^{-1}(\mathbf{I} - \mathbf{\Lambda}) = \mathbf{W}\mathbf{u}$$

The model set out in 3 can be used to determine the compensating variation for a change in the level of the public good from g^0 to g^1.

$$V_i(m_i, g^0) = V_i(m_i - C_i, g^1)$$

Applying the definition of V_i and we have

$$m_i + \sum_j w_{ij}u_j(g^0) = m_i - C_i + \sum_j w_{ij}u_j(g^1)$$

which simplifies to

$$C_i = \sum_j w_{ij}\left[u_j(g^1) - u_j(g^0)\right]$$

Our problem is to explore under what conditions would we expect centrality measures on egocentric networks to be related to willingness to pay. The simplest possible implementation of the model developed by Neilson and Wichmann is to set the benefit from a change in the public good equal to a constant, Δu. Neilson and Wichmann demonstrate that the row sums of $\mathbf{W}$ are equal to one. This result has a couple of interesting implications for our simplest possible case. First, the compensating variation will be equal to this constant utility, $C_i = \Delta u$. Second, since the value of λ_i does not impact on the row totals in $\mathbf{W}$, then when the change in utility is constant, differences in the values of λ_i do not impact on the value of C_i. Consequently, there should be no relationship between willingness to pay and any of the measures of centrality for the ego networks if utility is constant, or by extension there is no relationship between the structure of the network and utility.

We consider two possible sources of variation in the willingness to pay. One source focuses on $\mathbf{W}$. If agents have something akin to bounded rationality, then perhaps they only consider the utility of the people that the agent actually knows. In this case, the willingness to pay would be

$$C_i = \Delta u \sum_{j \in X_i} w_{ij}$$

where X_i is the set of alters connected to agent i. When this is the case, if $\sum_{j \in X_i} w_{ij}$ is increasing in the measure of X_i, then we would expect to see a positive relationship between the stated willingness to pay for the proposed change in the public good and measures of centrality for the ego networks that are increasing in the number of alters, measures such as degree centrality. Similarly, if $\sum_{j \in X_i} w_{ij}$ is increasing in the number of connections between alters in X_i, then we would also expect to see a positive relationship between C_i and measures of network density such as transitivity. Whether or not the restricted sum of the weights is increasing depends on what is happening to $\sum_{j \in X \setminus X_i} w_{ij}$. If the alters to ego i are connected, then this latter sum can increase, and since $\sum_{j \in X} w_{ij} = 1$, the sum of the weights on i's egocentric network can fall.

A second source of variation is to make the number of alters an agent has be a function of the utility that the agent gains from a change in the level of the public good. Since we measure our ego networks by asking people who they have spoken with in the last six months about environmental matters, it is reasonable to suppose that people who value the change in the public good more may also talk about it more. In this case, we would have that

$$C_i = \sum_j w_{ij} \Delta u(\# X_j)$$

where $\# X_j$ is the count of the number of alters in X_j. In this situation, $\Delta u(\# X_i)$ is higher, this contributes to an increase in C_i All else equal, when the ego network is more dense, then $\Delta u(\# X_j)$ will also be larger, and thereby contribute more to C_i.

To examine these two sources of variation in C_i,we turn to a random network simulation. The simulation involves the following steps:

1. Specify the utility for a set of network vertices. This is either constant and equal to one, or drawn from a distribution centered on one.

2. Specify the degree for each vertex. The degree is either drawn from a uniform distribution over the range from zero to five, or translated to a number between zero and five from the

variable utility. Therefore, either the network structure is unrelated to utility - people have the same utility, and talking to others about environmental issues is a random action - or the connections reflect increased utility from improvement in the public good.

3. Construct the network. The **iGraph** package [Csardi and Nepusz, 2006] in R [R Core Team, 2019] provides a method to construct a random network where each vertex has a specified degree.

4. Calculate the weighting matrix $\mathbf{W}$ for network. Calculate the compensating variation for each vertex, as per the two approaches set out above.

5. Sample without replacement a set of vertices, the egos, from the network and calculate a number of centrality measures for each ego network.

6. Regress the compensating variation on the calculated ego network centrality measures.

7. Repeat steps 1 through 6 a large number of times.

Figure 2 shows the simulation results for regressions on degree centrality and transitivity, for the truncated network case and the increasing utility case. Regressions for the truncated network case are mostly negative. With random networks, increasing both the size and density of the ego network on average reduces the sum of the utility weights for the members of the ego network. In contrast, with heterogeneity in the utility received from an increase in the public good and with people who experience a larger increase in utility having a larger network - people who care more talk more - impacts are positive. The slope estimate is positive for both degree centrality and transitivity (clustering) centrality.

The simulation results suggest hypotheses that we can test with the choice experiment results. If the ego states a willingness to pay based on a truncated network, then we expect the regression coefficients for both degree centrality and transitivity to be negative. If egos who value improvements in the public good talk to more people about this - have a larger ego network - then we expect the regression coefficients for degree centrality and transitivity to be positive. No relationship between willingness to pay and the centrality measures is consistent with equal utility for an increase in the public good across network members and/or no systematic relationship between network form and heterogeneity in the utility from an increase in the public good.

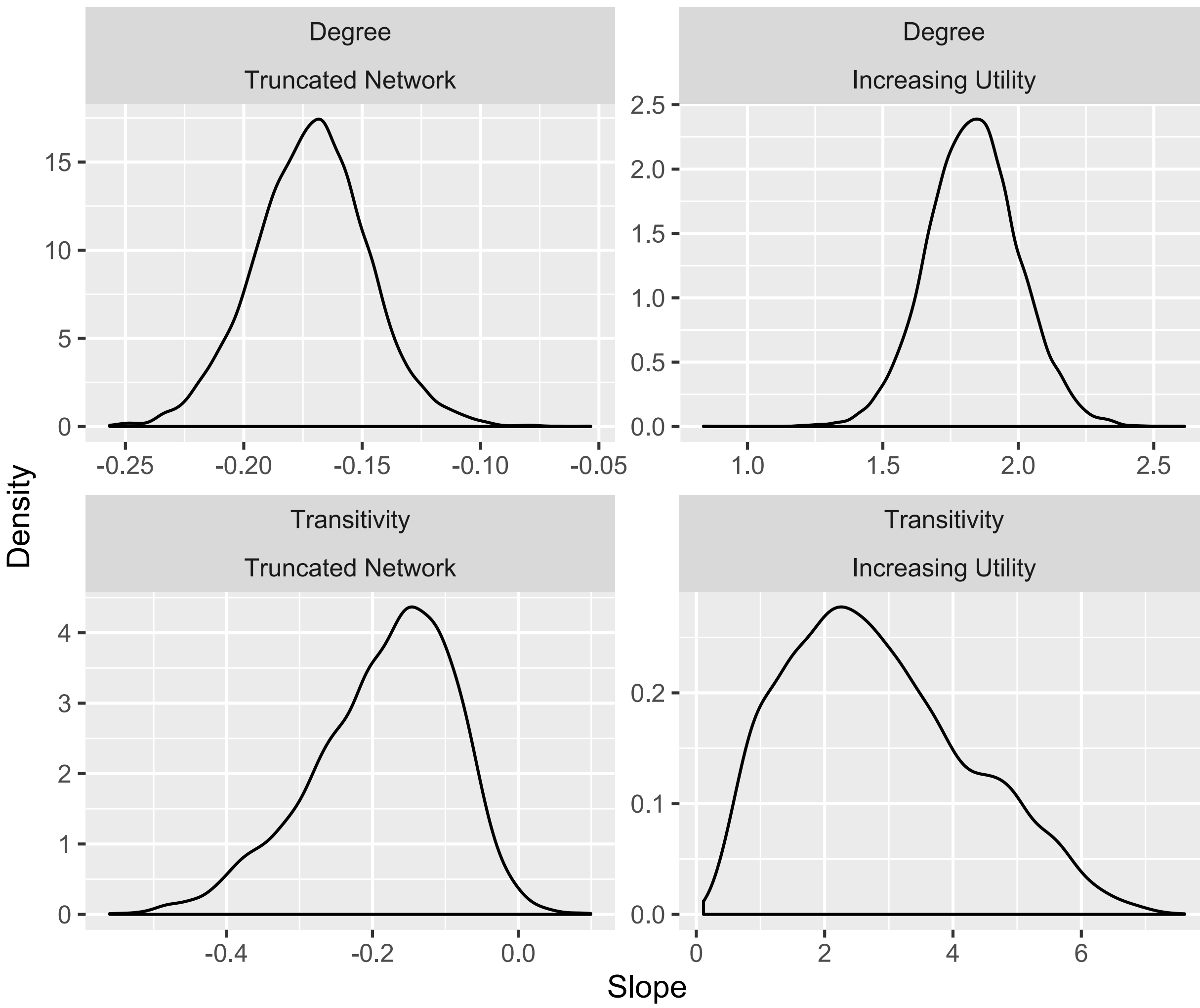

Figure 2: Simulation results.

4 Data

4.1 Choice Experiment Survey

The choice experiment and the egocentric network measurement instrument were administered online.The online survey was deemed to be an appropriate and cost-effective tool on the conviction that the use of computers and Internet among RDCO residents is high. According to Paul Budde Communication Pty Ltd [2016], Canada has one of the highest broadband penetration rates among the OECD nations and estimates that broadband availability, particularly in rural and regional areas close to 98%. [Dillman et al., 2014] also suggests online surveys as one of the tailored mixed design methods. We tested our survey with sample participants and adjusted the wording and organization based on responses. We also executed some laptop-based face-to-face and over the phone interviews at interviewee request. In these cases, each respondent was asked to complete the survey in the presence of an interviewer at their residence or over the phone with the help of employed student.

To generate the sample, a list of Central Okanagan addresses was harvested from an online source that permitted reverse postal code lookup (canada411.com). After cleaning the list, a random sample of addresses was selected and an invitation letter sent to the address. Addresses were individualized, using names that were part of the address list, under the expectation that personalized invitations will generate a higher response rate. The choice task was seen as a household decision, as the payment vehicle was applied at the level of a residence. However, some demographic questions (e.g. education, age) collected information on the individual respondent.

The choice experiment offered participants alternatives described using four attributes (Table 2). These attributes reflected aspects of the local environment that the Regional District of Central Okanagan (RDCO) had identified as important indicators of the state of the local environment. They reflect a locally scarce resource that people directly use, water; two aspects of local ecological health, natural terrestrial habitat and aquatic species population numbers; and one socio-cultural environmental measure, rural character. In consultation with local experts, including RDCO staff, we established a status quo forecast of how these attributes are likely to have changed thirty years in the future. We then generated two levels of improvement for each attribute. These improvement levels were chosen to be reasonably attainable, in order to enhance the credibility of the survey.

Table 2: Attribute levels and descriptions, as provided to participants through online survey.

Attribute	Levels	Description
Share of Total Water Use from Groundwater	10%, 15%, 20%	**GROUND WATER USE**: Groundwater is an important source of water we use and supports springs and wetlands that are important natural habitats. Increasing ground water use threatens these habitats. Most surface sources in the Okanagan are fully allocated. Increasing the share of total use from groundwater also reflects an increase in overall water use and in the amount of groundwater used.
Count of Spawning Kokanee Salmon	40,000, 50,000, 60,000	**AQUATIC HABITAT HEALTH**: Aquatic health is threatened by shoreline development, invasive species pollution, etc. Species such as Kokanee salmon are sensitive to the overall health of the aquatic environment. Since the 70's, spawning Kokanee numbers in Okanagan lake have fallen from over 1 million to about 40,000.
Sensitive Ecosystem Area Lost	100 km^2, 50 km^2, 0 km^2	**NATURAL HABITATS**: Natural habitats, such as wetlands, forests, natural grasslands, etc. provide a range of environmental goods and services. These environmental resources and services are largely lost if the land is developed.
Population Density in Rural Areas	60 ppl / km^2, 70 ppl / km^2, 75 ppl / km^2,	**RURAL CHARACTER**: Rural areas have a unique character that reflects their history and close ties to the land. Increasing development and population growth in rural areas increases traffic, reduces the amount of open space, increases conflict between farmers and non-farm residents, and impacts natural habitats.
Special Property Levy	\$0.00, \$20 / \$400,000 \$40 / \$400,000	**COST**: The tools that can reduce groundwater use, improve aquatic health, limit loss of natural habitats and reduce population growth in rural areas do have a cost. However they are paid, they will leave your household with less money to spend each year.

We used parcel levy as the payment vehicle. This was chosen as parcel levies have been used locally to finance a variety of projects. For example, the H_2O aquatic center in Kelowna was funded in part by a levy on Kelowna properties of about \$0.21 per \$1,000 value for twenty years, and ongoing activities to control the invasive aquatic plant, Eurasian milfoil in valley lakes is paid for by a \$0.057 per \$1,000 collected by the Okanagan Basin Water Board. Each alternative to the status quo involved a levy of either \$0.05 per \$1,000 or \$0.10 per \$1,000. These values scale to \$20 and \$40 for the average RDCO residential property, which at the time of the survey was selling for around \$400,000. The levy would take effect immediately and continue for thirty years.

Figure 3 provides an example of a choice card used in the experiment. Each respondent was presented, in sequence, with six such choice cards. Graphics were included to make it easier for the participant to recognize the relative size of the changes. In addition to offering the respondents the three way choice, we also asked them to choose between the two alternatives, if the status quo was not offered. For the analysis to follow, we only use results for the three choice situation.

A name generator was used to identify up to five members, by initials, of each participant's egocentric network, with the specific instructions "During the last six months, who have you talked with about how growth and development are impacting the environment in the Central Okanagan?" See Carrasco et al. [2008] for a similar survey. As noted by Campbell and Lee [1991] and Marsden [2005], name generators that specify a constrained time and a specific type of interaction tend to result in smaller networks. As our interest was specific to concern for the future state of the environment, requiring a specific type of relationship and relatively recent contact would be appropriate for our purposes. Respondents were asked to rate the strength of the relationship as very close, close or acquaintance, and asked to provide, as best they could, information about gender, income, education, political orientation, physical distance between residences, frequency of contact, length of relationship, and purpose (e.g. children are friends, common faith community, work, sports team, etc.). The name generator included a relationship table, where survey participants indicated if their named alters knew each other. They were able to specify directed connections, as very close, close, or stranger, between each alter-alter pair. For the analysis reported below, all connections were assumed to be undirected. The collected data will enable us in future to explore further network features [e.g. homophily, see Macias and Williams, 2016] on willingness to pay.

ATTRIBUTE	STATUS QUO	OPTION 1	OPTION 2
Share:	**20%**	**10%**	**20%**
GROUND WATER USE (Share of total water use from groundwater)			
Returns:	**40,000**	**50,000**	**60,000**
AQUATIC HABITAT HEALTH (Count of spawning Kokanee Salmon)			
Loss:	**100 km^2**	**50 km^2**	**100 km^2**
NATURAL HABITATS (Sensitive Ecosystem Area Lost)			
Density:	**75 ppl/km^2**	**75 ppl/km^2**	**60 ppl/km^2**
RURAL CHARACTER (population density in rural areas)			
Levy:	**$0.00**	**$0.05 per $1000**	**$0.10 per $1000**
COST (Special Property Levy)	$0.00	$20 on $400,000 house	$40 on $400,000 house
Which option would you prefer?	○	○	○
Your preferred option without status quo		○	○

Figure 3: Example choice card.

4.2 Survey Results

Invitations were mailed to 3,000 residential addresses, approximately 10% of which were invalid (moved, etc.). 550 invitees started the online survey, resulting in 468 completed questionnaires (17.3% response rate). Identifying the sample through internet phone book address harvesting was not ideal. For online phone books, listed people either have a land line, or have manually added address information to a mobile phone number. To the extent that younger, more technologically sophisticated and/or more mobile residents (renters, persons multiple residences, only one of which is in the Okanagan, etc.) do not have publicly available addresses, such people would be less likely to receive an invitation to participate in the survey. Sample demographic measures are consistent with this.

Figure 4 compares the Statistics Canada Community Profile for the Regional District of Central Okanagan age, education and income distributions with those reported by the survey participants. The sample is clearly over represented by the older half of the population. It is somewhat over represented by people who have continued their education beyond high school, and seriously under represents people who did not complete high school or only completed high school. The sample also has less representation by those with the lowest incomes, and somewhat higher representation among almost all the middle and upper income categories. The results therefore apply to an older, somewhat more educated and somewhat more affluent segment of the central Okanagan population, and extrapolation to the entire population needs to be done with care.

In the sample, 64% of respondents who answered the gender question chose male, with 118 of 468 not providing an answer. One person indicated that they were a member of a First Nation, with 115 of respondents not answering. Fully 91% of those who indicated home ownership owned their home. Among answering respondents, 81% were couples, with 25% having children living at home.

Almost 85% of respondents indicated an awareness of the environmental issues that were the subject of the survey, prior to completing the survey. When asked if on balance the benefits of development outweighed the negative impacts on the environment, only 10% saw the balance as neither negative or positive. The remaining 90% was almost equally divided between very negative, negative and somewhat negative, and the corollary positive opinions. Given the somewhat higher income and education levels, this balanced opinion is somewhat surprising. Very few

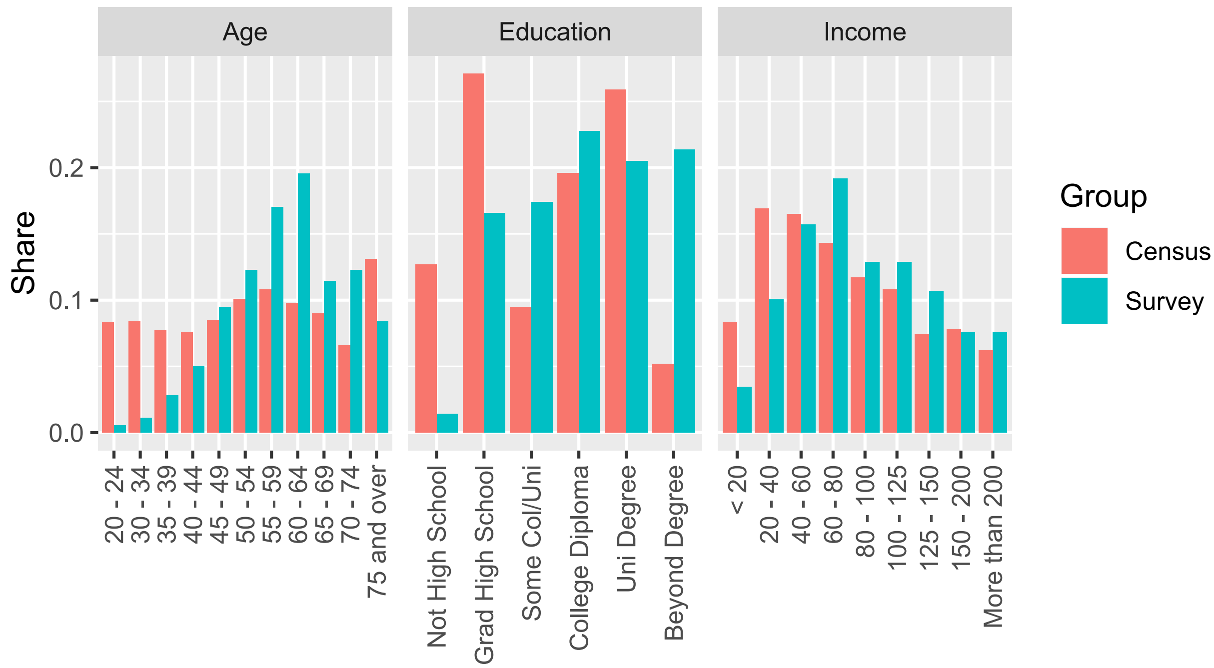

Figure 4: Comparison of sample and census distribution for age, education and income.

respondents indicated having any formal education in environmental or biological topics, nor being involved professionally with environmental issues. Overall, our sample is biased towards older, more educated and more affluent central Okanagan residents who consider themselves informed about environmental issues in the valley. However, it is not biased towards people who are strongly in favor of environmental protection or of development.

Figure 5 plots the share of the choice experiment responses that are not the status quo for each survey participant, against four different centrality measures. In all four cases, the likelihood that a participant chooses an alternative to the status quo is increasing in the network centrality parameter. The figure also illustrates differences in the way that centrality measures can be calculated. The iGraph package in R does not normalize the centrality measures. We chose to normalize by dividing by the maximum possible network size, five. The Gephi program divides closeness centrality by degree to normalize. For an ego network, any network with at least on member will have a closeness centrality measure equal to one, with the Gephi convention of setting undefined values to zero meaning all respondents who didn't specify any alters would be zero. This centrality measure is therefore a dummy indicating whether the participant named any alters. This highlights how the

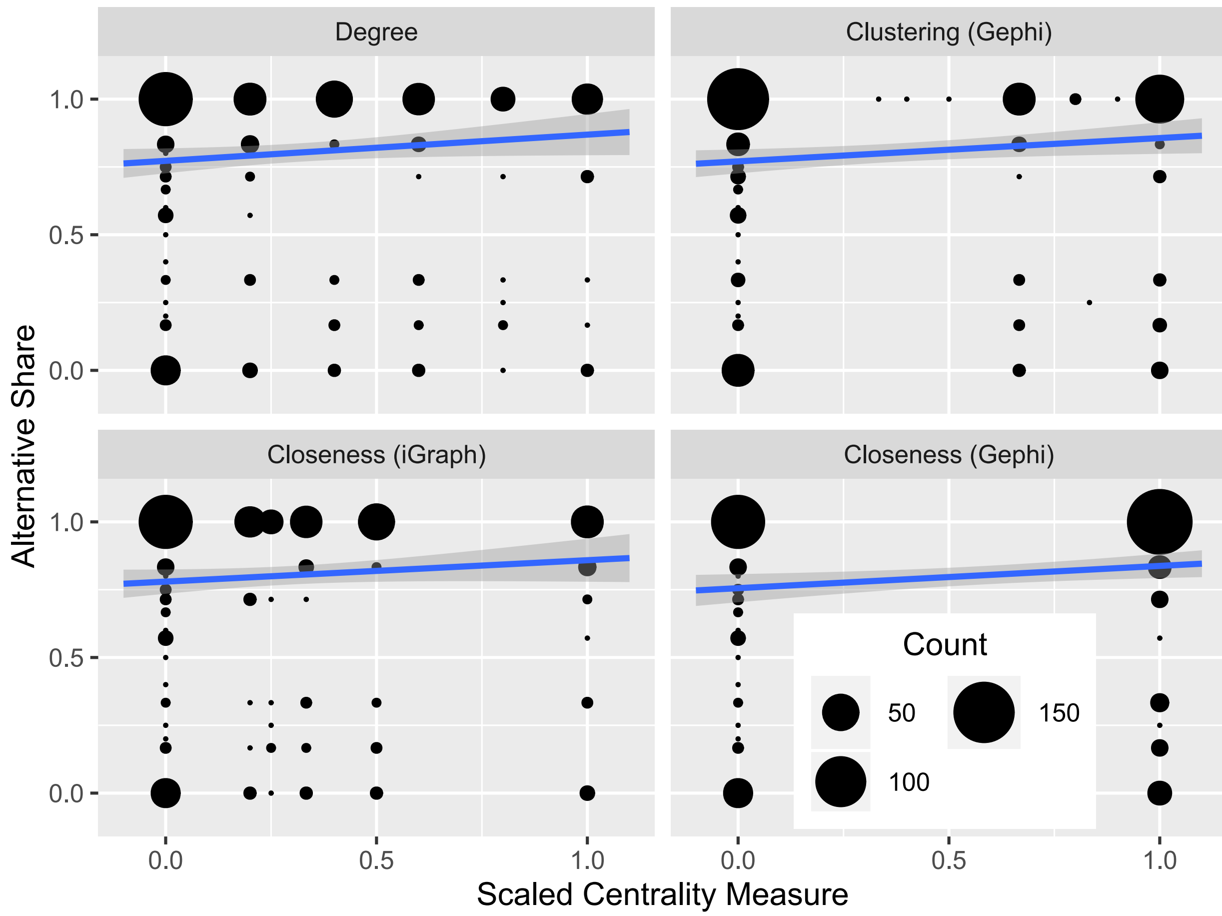

Figure 5: Frequency choosing alternative to status quo.

method of calculating and the way of measuring the networks will impact the relationships that are observed between network measures and other variables.

5 Econometric Model

Our survey offered participants scenarios that varied along four environmental quality dimensions and one cost dimension. Participants were also asked to provide demographic, knowledge and opinion information. We therefore have alternative specific variables that vary with each of the alternatives offered to the participant, and individual specific variables that are constant for each individual.

We follow the conventional random utility model [McFadden, 1980, 1978] where the utility individual i receives from choosing alternative $j \in [1, \ldots, J]$ in choice situation $k \in [1, \ldots, K]$ is

$$V_{ijk} = U_{ijk} + \epsilon_{ijk}$$

where U_{ijk} is the utility the individual receives that is a function of observable characteristics of the alternative and the choice situation. The error term ϵ_{ijk} captures unobserved factors contributing to the utility individual i receives, and is assumed drawn from an i.i.d. extreme value type 1 (Gumble) distribution.

In each choice situation k, the individual is assumed to choose that alternative which maximized utility. If J_k indexes the alternatives available in choice situation k, then the choice made will solve

$$\max_{j \in J_k} [U_{ijk} + \epsilon_{ijk}]$$

The probability that individual i chooses alternative j is

$$\begin{aligned} \Pr(j) &= \prod_{l \in J_k \setminus i} \Pr(U_{ijk} + \epsilon_{ijk} > U_{ilk} + \epsilon_{ilk}) \\ &= \prod_{l \in J_k \setminus i} \Pr(U_{ijk} - U_{ilk} > \epsilon_{ilk} - \epsilon_{ijk}) \end{aligned}$$

the probability that alternative j provides greater utility to i than any of the other alternatives in J_k. Given the assumptions on the distribution of ϵ_{ijk}, this probability can be expressed as

$$\Pr(j) = \frac{U_{ijk}}{\sum_{l \in J_k} U_{ilk}}$$

It is common to assume that the utility is a linear function

$$U_{ijk} = \mathbf{x}'_{ijk}\beta + \epsilon_{ijk}$$

where $\mathbf{x}_{ijk}$ is a $K \times 1$ vector describing the characteristics of the choice situation and alternative that the individual is evaluating. The parameter vector β measures the influence of the elements of $\mathbf{x}_{ijk}$ on the likelihood that j is chosen.

In addition to the impact of alternative and situation specific variables on the choice probability, we are also often interested in the effect of individual specific characteristics. This would redefine the utility function as

$$U_{ijk} = \mathbf{x}'_{ijk}\beta + \mathbf{z}'_i\gamma + \epsilon_{ijk}$$

where $\mathbf{z}_i$ is a vector of characteristics describing individual i and γ measures the impact of these individual characteristics on the utility individual i receives from choice j in situation k.

While we expect individual specific characteristics to affect the utility individuals receive from their choices, estimating these effects is challenging. For a fixed effects model, we add a dummy variable λ_i for each individual as

$$U_{ijk} = \mathbf{x}'_{ijk}\beta + \mathbf{z}'_i\gamma + \lambda_i + \epsilon_{ijk}$$

Since the variables in $\mathbf{z}_i$ are constant for each individual, they are perfectly colinear with each other and with λ_i. We can therefore add at most one individual specific variable to a regression model like this, and that is only when there is no fixed effect. However, interactions can be estimated, as

$$U_{ijk} = \mathbf{x}'_{ijk}\beta + \mathbf{x}'_{ijk} \odot \mathbf{z}'_i\gamma + \lambda_i + \epsilon_{ijk}$$

where $\odot$ indicates a Hadamard product (element by element). Typically many elements of γ are assumed to be zero, and restricted as such in the estimation.

An alternative method of including individual specific effects follows from a random parameters model. Here we assume that each individual has a unique vector of responses to the attribute and choice situation characteristics. Often the assumption is that $\beta_i \sim N(\beta, \mathbf{\Sigma})$. Relying on the fact that $\mathbf{\Sigma}$ can be factored as $\mathbf{L}\mathbf{L}'$ by a Cholesky decomposition, we can write

$$\beta_i = \beta + \mathbf{L}\eta,\ \eta \sim N(\mathbf{0}, \mathbf{I})$$

and individual specific characteristics can be included as

$$\beta_i = \beta + \mathbf{z}'_i\gamma + \mathbf{L}\eta$$

with utility now defined as

$$U_{ijk} = \mathbf{x}'_{ijk}(\beta + \mathbf{z}'_i\gamma + \mathbf{L}\eta) + \lambda_i + \epsilon_{ijk}$$

which requires integrating out the influence of η before estimation of the logistic regression. This integration is typically simulated, through a maximum simulated likelihood approach.

For policy purposes, we are often interested in the marginal rate of substitution between variables, in particular the marginal rate of substitution between attribute levels and income. The marginal rate of substitution is defined as

$$MRS_{xy} = -\left.\frac{dx}{dy}\right|_{\overline{U}} = -\left.\frac{\partial U/\partial x}{\partial U/\partial y}\right|_{\overline{U}}$$

with the amount of income an individual would give up to get a unit increase in a choice experiment attribute being

$$WTP_A = MRS_{CA} = -\left.\frac{\partial U/\partial x_C}{\partial U/\partial x_A}\right|_{\overline{U}}$$

where the subscripts C and A indicate cost and the attribute level respectively. When there are no interactions, then the willingness to pay resolves to

$$WTP_A = -\frac{\beta_A}{\beta_C}$$

However, if individual specific effects are included, either as interactions or as shifters for the means of the random parameters, then the willingness to may need to include these interactions. For example, if income is included as an interaction with the cost variable, then

$$WTP_A = -\frac{\beta_A}{\beta_C + \beta_{IC}\overline{x}_I}$$

if we are interested in the willingness to pay at the mean. With individual specific variables included in this way, we can also explore how willingness to pay changes across the respondents, and could also adjust either the mean income, or the distribution of incomes, to reflect a population different from that who responded to the survey invitation.

Hypothesis testing for ratios of parameter estimates is problematic. If the parameter estimates are assumed to follow a normal distribution, then their ratio follows a Cauchy distribution, and does not have defined moments. This presents particular problems if there is a significant probability that the parameter is close to zero - low significance of the parameter estimate - which translates into significant probability that large values can be realized. The standard error of the ratio calculated by the delta method [Greene, 2012] may therefore be inaccurate. Bootstrap methods have been proposed as alternatives [Krinsky and Robb, 1986, Hole, 2007], but are subject to the same issue as bootstrap realizations can generate values for the cost parameter close to zero [Carson and Czajkowski, 2019]. Different approaches to estimation have also been proposed. One suggestion is to estimate the entire system in willingness to pay space by restricting the parameter on the cost variable to equal one, and estimating a scaling parameter that applies to the attributes being priced [Scarpa et al., 2008]. Another is to define the cost parameter as a function, such as an exponential, that cannot be zero [Carson and Czajkowski, 2019].

6 Results

6.1 Regression Results

Results for five fixed effects multinomial regression models are shown in 3. The first model only includes the attribute levels and an intercept term for the alternatives. All models were also run with one alternative specific constant for each alternative. We have opted to estimate the models restricting the alternative specific constants to be equal for two reasons. First, when we consider the random parameters models, almost all of the models estimated with the alternative specific constants restricted to be equal score a lower AIC value than the models where this restriction is relaxed. Second, there is no theoretical reason to expect the alternative specific constants to differ. We acknowledge that an experimental design effect may be present, where subjects assume a ranking where left, status quo, is worst and right, the second alternative, is best. While the fixed effects models suggest that this may be the case, for ease of comparison with the random parameters results, we report only fixed effects model results where the alternative specific constants are restricted to be equal. Qualitatively, models with two alternative specific constants and as appropriate interaction terms with those constants are similar, and imposing this restriction does not impact on the interpretation of the results.

Table 3: Fixed effects multinomial logit regression results.

	#1	#2	#3	#4	#5
Groundwater Use	-0.014	-0.015	-0.015	-0.015	-0.015
	(0.008)	(0.008)	(0.008)	(0.008)	(0.008)
Aquatic Health	0.018^{***}	0.018^{***}	0.018^{***}	0.018^{***}	0.018^{***}
	(0.005)	(0.005)	(0.005)	(0.005)	(0.005)
Habitat Loss	-0.006^{***}	-0.006^{***}	-0.006^{***}	-0.006^{***}	-0.006^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
Rural Character	-0.011	-0.012	-0.012	-0.012	-0.012
	(0.006)	(0.006)	(0.006)	(0.006)	(0.006)
Levy	-0.760	-6.472^{**}	-6.415^{**}	-6.441^{**}	-6.432^{**}
	(1.579)	(1.968)	(1.993)	(1.993)	(1.994)
Alternatives Constant	0.216	0.014	-0.186	-0.220	-0.230
(ASC)	(0.120)	(0.132)	(0.396)	(0.396)	(0.396)
Levy × Income		0.056^{***}	0.055^{***}	0.055^{***}	0.055^{***}
		(0.011)	(0.012)	(0.012)	(0.012)
ASC × Degree				0.048	
				(0.038)	
ASC × Transitivity		0.494^{***}			0.303^{*}
		(0.133)			(0.138)
ASC × Age			0.008	0.007	0.007
			(0.005)	(0.005)	(0.005)
ASC × Education			0.155^{***}	0.145^{**}	0.136^{**}
			(0.044)	(0.045)	(0.045)
ASC × Devel. Good			-0.206^{***}	-0.196^{***}	-0.192^{***}
			(0.038)	(0.039)	(0.039)
Log Likelihood	-1865.241	-1844.967	-1830.295	-1829.515	-1827.882
χ^2 ($df = NA$)	71.655^{***}	112.202^{***}	141.548^{***}	143.107^{***}	146.373^{***}
pseudo R^2	0.019	0.030	0.037	0.038	0.038
Num. obs.	1823	1823	1823	1823	1823
AIC	3742.482	3705.935	3680.589	3681.030	3677.764

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

In all cases the parameters estimated for the attribute levels have the expected signs. Increasing groundwater use, increasing habitat loss, and increasing the population density in rural areas all are negative, indicating that utility falls when these measures increase. Increasing aquatic health, measured as one thousand additional spawning salmon each year has a positive sign, suggesting that this increases utility. Finally, the property levy has a negative sign, consistent with people feeling worse off if the levy they pay is greater, all else equal.

Absent any covariates (Model #1), the disutilty from an increase in the levy is far from significant. This lack of significance on the levy would render estimates of the willingness to pay for improvements to not be significantly different from zero.

In Model #2 we add two covariates, an interaction term between the level of the levy and the respondent's income, and an interaction term between the alternatives constant and the egocentric network transitivity measure. Both are highly significant. The transitivity estimate is positive, consistent with the simulation results described above. This is therefore consistent with those who gain more utility from the environmental improvements being more likely to speak with others about issues around development and the environment in the Okanagan. We introduce the interaction between levy and income to allow for a change in the sensitivity to the size of the levy as a function of income. This effect is strongly significant here, and for all the other models where it is present. The positive sign is consistent with the expectation that people with higher incomes are less sensitive to the size of the levy. Adding these two covariates leads to a large increase in the size of the parameter on the levy term, and with this increase it becomes strongly significant. Consequently, estimates of the willingness to pay for changes in the attribute levels are now likely to be significantly different from zero, at least where the parameter estimates are. Adding these two covariates does not substantially change the magnitudes nor significance of the four attribute levels, an effect that doesn't change with the other combinations of covariates.

In Models #3 through #5 we add two demographic variables, age and education, and one perception variable. The perception variable is based on the statement "Development and growth has led to large reductions in natural habitats. These changes have also brought jobs and improved the standard of living for Okanagan residents." Respondents were asked to indicate there evaluation of these changes by choosing on a Likert scale from "very bad (=0)" to "very good (=6)". Non responses and don't know were coded as neutral. Age is not significant in any models. Education is significant to at least the 1% level for all models. The interaction between the alternatives constant and the development perception is negative and strongly significant for all three models. If people who view the benefits of development as 'worth' the environmental cost, and see this as continuing to be the case, then we would expect them to be less willing to pay for actions that would make development more difficult.

Model #4 includes the degree of the respondent's egocentric measure as an interaction covariate, and model #5 includes transitivity. Degree is not significant at the 5% level, although it does have a positive sign that is consistent with the hypothesis that respondents who earn more utility from environmental improvements are more likely to talk with people about environment and development issues. The transitivity measure is still positive, and significant at the 5% level. The somewhat smaller size and lower significance suggests that people who perceive the costs to the environment of development as being too large are also likely to talk with others about these issues.

Relative to a null model with only an intercept, these five models are all strongly significant. The AIC ranks Model #5 as best of these, and it is better than many, though not all, of the other specifications examined. Those that have a lower AIC relax the restriction that the alternative specific constants are equal. As noted above, we present results for these models as comparison to the random parameters models.

Estimation results for the random parameters models are shown in Table 4. The random parameters are the four attributes, the levy, and the alternatives constant. The mean of the levy parameter is allowed to be a function of the respondent's reported income level, while the mean of the alternatives constant is allowed to be a function of the respondent's age, education, attitude towards development, and their social network.

The parameters on the environmental attributes are very stable across the model specification. In Model #6, without any covariates, the mean of the levy parameter is not significantly different from zero. Including a relationship between the levy parameter and income substantially increases the size of the levy parameter estimate, making it significant at the 5% level for all four models where it is added. While the levy parameter is now significant, that is not the case for income effect on the levy parameter. The standard deviation estimates for the levy parameter do not change much when income is added as a predictor for the parameter value. Therefore, the effect of adding income is to right shift the distribution of the levy effect without substantially changing the shape

Table 4: Random parameters multinomial regression results.

	#6	#7	#8	#9	#10
Groundwater Use	-0.027^{*}	-0.027^{*}	-0.027^{*}	-0.026^{*}	-0.027^{*}
	(0.013)	(0.013)	(0.013)	(0.013)	(0.013)
Aquatic Health	0.029^{***}	0.029^{***}	0.029^{***}	0.029^{***}	0.029^{***}
	(0.007)	(0.007)	(0.007)	(0.007)	(0.007)
Habitat Loss	-0.010^{***}	-0.010^{***}	-0.010^{***}	-0.010^{***}	-0.010^{***}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Rural Character	-0.025^{*}	-0.025^{*}	-0.025^{*}	-0.024^{*}	-0.025^{*}
	(0.012)	(0.011)	(0.011)	(0.011)	(0.012)
Levy	-4.913	-12.681^{*}	-11.582^{*}	-12.432^{*}	-11.705^{*}
	(3.174)	(5.205)	(5.140)	(5.064)	(5.164)
Alternative Specific	4.555^{***}	3.274^{***}	-0.483	3.586	-0.277
Constant (ASC)	(0.738)	(0.787)	(2.374)	(2.505)	(2.520)
Levy $\times$ Income		0.075	0.070	0.077	0.068
		(0.040)	(0.040)	(0.040)	(0.040)
ASC $\times$ Degree				0.139	
				(0.242)	
ASC $\times$ Transitivity		2.118			1.063
		(1.110)			(0.834)
ASC $\times$ Age			0.721^{*}	0.613^{*}	0.696^{*}
			(0.289)	(0.296)	(0.338)
ASC $\times$ Education			0.076^{*}	0.011	0.060
			(0.035)	(0.031)	(0.035)
ASC $\times$ Devel. Good			-0.758^{**}	-0.763^{**}	-0.688^{**}
			(0.252)	(0.284)	(0.256)
Random Parameter Standard Deviations					
Groundwater Use	0.055^{*}	0.051	0.053	0.048	0.050
	(0.027)	(0.030)	(0.027)	(0.026)	(0.029)
Aquatic Health	0.028	0.030^{*}	0.029^{*}	0.030^{*}	0.031^{*}
	(0.015)	(0.013)	(0.013)	(0.013)	(0.013)
Habitat Loss	0.013^{***}	0.013^{***}	0.013^{***}	0.012^{***}	0.013^{***}
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Rural Character	0.098^{***}	0.095^{***}	0.099^{***}	0.096^{***}	0.099^{***}
	(0.018)	(0.019)	(0.018)	(0.018)	(0.018)
Levy	32.245^{***}	32.183^{***}	31.869^{***}	30.817^{***}	32.216^{***}
	(3.485)	(4.000)	(3.173)	(3.123)	(3.558)
ASC	5.990^{***}	5.936^{***}	6.001^{***}	6.464^{***}	5.894^{***}
	(0.811)	(0.931)	(0.680)	(0.856)	(0.780)
Log Likelihood	-1355.129	-1351.669	-1346.848	-1346.844	-1346.440
χ^2 ($df = NA$)	232.945	239.865	249.508	249.516	250.323
Pr($< \chi^2$)					
pseudo R^2	0.079	0.081	0.085	0.085	0.085
Num. obs.	1823	1823	1823	1823	1823
AIC	2734.259	2731.338	2725.695	2727.688	2726.881

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

of the distribution.

The egocentric network transitivity measure is added in Model #7. When included as a shifter for the alternatives constant, it is significant at the 10% level, but not quite attaining the 5% significance level. As for the fixed effects model, when other covariates are added, the significance of the network measure decreases. In this case, becoming insignificant even at the 10% level. The degree centrality measure (Model #8) is far from significant when the covariates are present, and didn't attain even 10% significance when added without the covariates.

The further three covariates are again Age, Education and the perspective on the development and environment trade-off. In contrast with the fixed effects model results, Age is now significant for all three models where it is included, and education is only significant in one. As for the fixed effects model, the respondent's perspective on development is a strongly significant shifter for the alternatives constant. It is also noteworthy that when these three covariates are present, the mean of the alternatives constant is zero.

The standard deviation estimates for the random parameters are strongly significant for habitat loss, rural character, the levy, and the alternatives constant. For most of the models, they are not significant for groundwater, and for aquatic health, they are significant at the 5% level, but not at the 1% level. Respondents are therefore fairly similar in how they respond to reductions in groundwater use and increases in salmon spawning returns. However, there is a large variety in the way that people respond to changes in the amount of habitat loss and the population density in rural areas. Using the standard deviation estimates, almost one quarter of respondents would see an increase in utility from more habitat loss, and almost half would similarly see an increase in utility from an increase in rural population density.

6.2 Willingness to Pay

Figure 6 shows the mean willingness to pay (WTP) for the four environmental attributes, with confidence intervals calculated using the Delta method, the parametric bootstrap method [Krinsky and Robb, 1986] and the nonparametric bootstrap method. The parametric bootstrap generates the distribution of WTP estimates from random samples of parameter values drawn from a multivariate normal distribution of the parameter space. The nonparametric bootstrap was generated by estimating the parameters from a resampling of the set of individuals participating in the choice

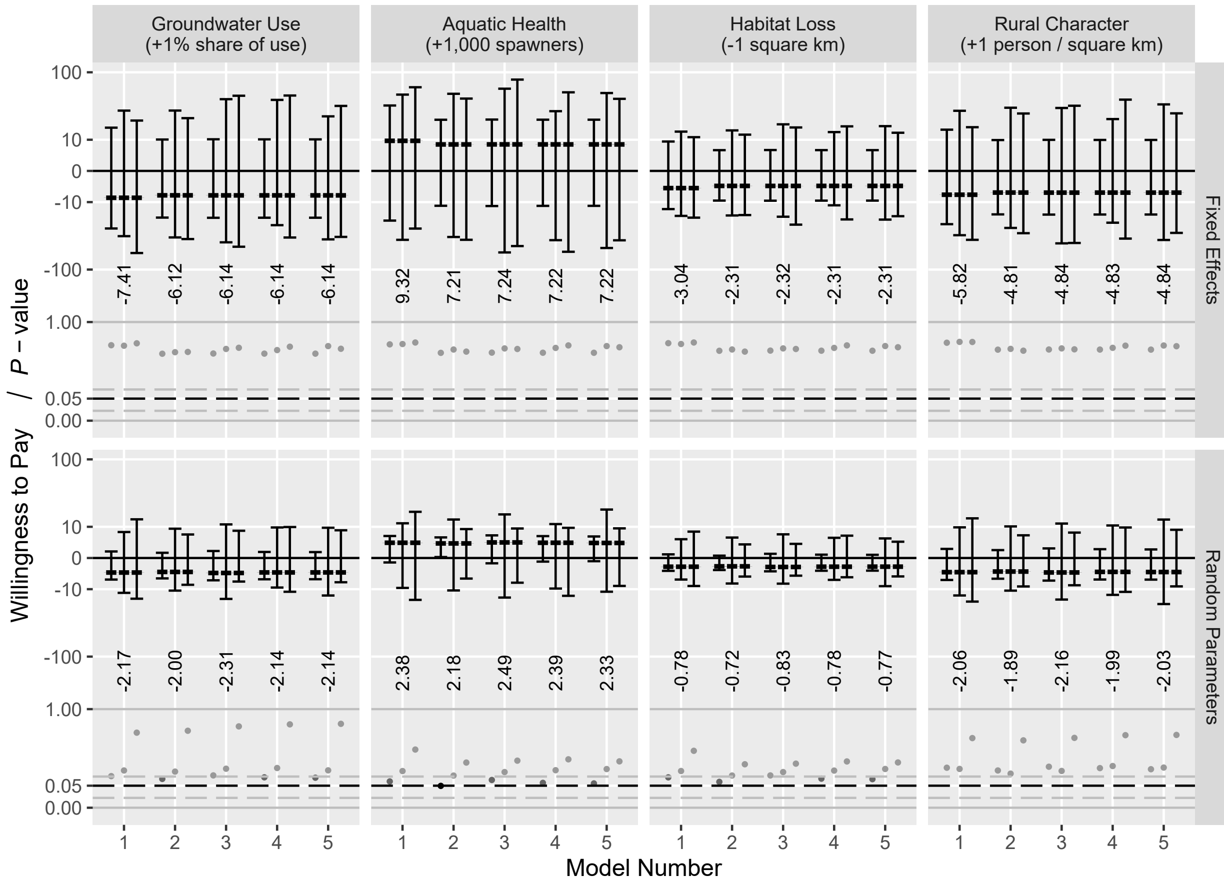

Figure 6: Willingness to pay mean values, confidence intervals, and P values, generated using the Delta method, the parametric (Krinsky Robb) and non-parametric bootstrap. WTP values and P values are square root scaled. For each model, the left confidence interval is generated using the Delta method, the central interval from the parametric bootstrap, and the right interval from the nonparametric bootstrap.

experiment, and using these parameter estimates to generate estimates of the WTP. P values for the Delta method are based on the z-score, while for the bootstrap methods are based on an ordering of the simulated WTP values and the interpolated position of the comparison value - zero - within this ordered list.

While the parameter estimates for aquatic health and habitat loss are strongly significant, the willingness to pay values are not. When the size of the levy is interacted with income, the levy parameter becomes strongly significant in the fixed effects models, and significant at the usual level of 0.05 for the random parameters models. If the WTP was calculated using only the levy parameter, it would be significantly different from zero. However, as derived above, this is incorrect when there are interaction terms. The levy parameter and the parameter on the interaction between levy and income in the fixed effects model are strongly negatively correlated, with the result that the variance of their mean income weighted sum is not significantly different from zero, and the resultant WTP estimates are themselves not significantly different from zero. For the random parameters models, the parameter estimated for the interaction between levy and income is not significant, and consequently when combined with the levy parameter, the resultant combination again has a high variance, and consequently the WTP estimates are not significant.

Including the interaction between income and the size of the levy allows us to examine the impact of income on willingness to pay. That the interaction parameter is positive means that the disutility from an increase in the levy is diminishing with increasing income. As estimated, this disutility becomes zero for incomes not far beyond the sample mean income level. The implication is that the higher income participants in this experiment would be better off with a higher levy. To examine the robustness of this result, we also estimated the models using the logarithm of income. Results are no different.

Accepting that the results are not very precise, we can discuss the implied value of environmental improvements if the true value is the average estimated value. Using the random parameter estimates, which are smaller than the fixed effects estimates, the average respondent is willing to pay approximately \$2.20 to reduce the share of Okanagan water coming from groundwater, approximately \$2.40 to see an increase of 1,000 spawning Kokanee salmon, approximately \$0.75 for a one square kilometer reduction in the amount of natural habitat loss, and approximately \$2.00 to reduce the increase in rural population density by one person per square kilometer. Assuming 100,000 households, the annual value of these environmental improvements is \$220,000, \$240,000, \$75,000 and \$200,000 respectively. Of the four attributes described, aquatic health and habitat loss are probably the best understood by the participants. The limited or absent significance on the parameter estimates for groundwater share and rural character suggest that either people didn't particularly care about these attributes, or were confused by them.

Examining the standard deviation estimates for the random parameters provides a perspective on the variation in how the participants respond to changes in the attribute levels. Figure 7 shows the coefficient of variation, the ratio of the estimate of the parameter standard deviation to the estimate of the mean. The confidence intervals are calculated using the Delta method, although

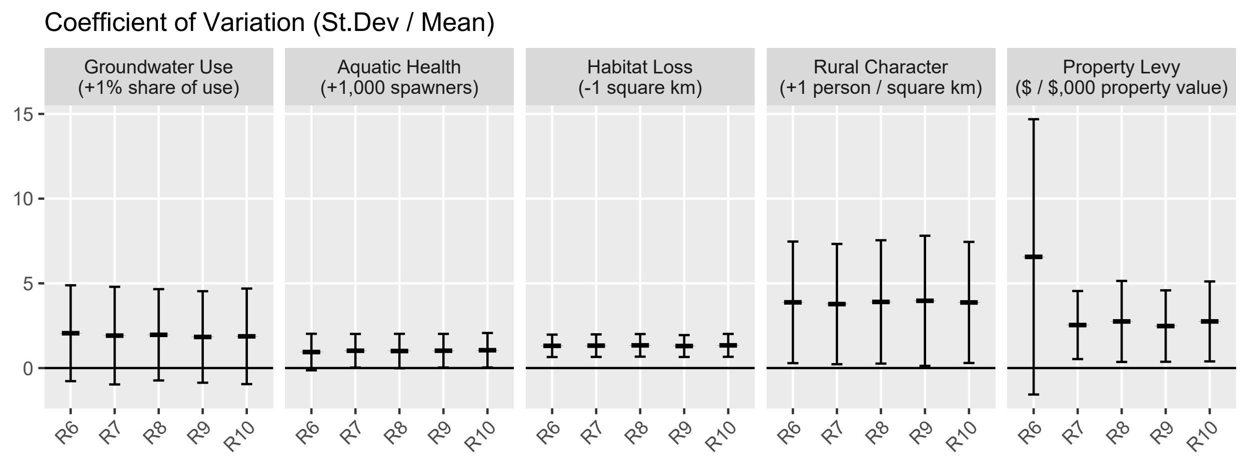

Figure 7: Coefficient of variation for random parameter estimates.

these estimates are also ratios of random variables and therefore face the same issues as for the willingness to pay.

The absolute variation is largest for rural character, and the ratio tends to be significant. The absolute value of the ratio is the smallest for aquatic health, and is only marginally significant. Using the estimated mean and standard deviation, 29.8% of the population is expected to have a positive response to an increase in the rate of groundwater use, 16.7% a negative response to an increase in salmon returns, 22.1% would have a positive response to an increase in habitat loss, and 39.8% a positive response to an increase in rural population density. The relatively weak significance for the standard deviation estimates for groundwater share and aquatic health suggest that there may not be that much variation in responses. However, for habitat loss and increasing rural population density, the estimates suggest that there is a large variation between people in the population in how they respond to changes in these attributes, and that a relatively large share of the population actually see their utility decreasing if these attributes are improved (as defined in the experiment).

7 Discussion

Our experimental design sought to test for the presence of a social network influence on the willingness to pay for environmental improvements. While egocentric social network measures do add significant explanatory power to the model when there are no other covariates in the estimated models, adding covariates, particularly a measure of the respondents perceptions about the relative benefits of development, reduce or eliminate the significance of the egocentric centrality measures.

Our measurement of the egocentric network is consistent with previous work using a name generator and framing the name generator with a time limit and topic of conversation [Campbell and Lee, 1991, Yousefi-Nooraie et al., 2019]. This means that the measured social network is not a network of connections reflecting concern for others, the type of network that Neilson and Wichmann [2014] examine in their analysis. The egocentric networks that we measured reflect communication about development and environmental issues in the Okanagan, communication that is likely more common among people who feel that the harm done by ongoing development in the Okanagan is large. The dependence of the measured social network on the way that the network was measured is a topic of ongoing research [Marsden, 2011, Shakya et al., 2017], and our results highlight the importance of the name generator. Mapping the networks of concern requires a name generator that identifying those alters whose utility changes most impact the ego's utility. The subsequent name interpreter would be used to explore the direction of and extent to which the subject of the choice experiment affects the utility of the alters. Further work such as this is needed to develop methods for capturing the potential bias identified by Neilson and Wichmann.

The use of egocentric social network measures in nonmarket valuation research is limited. The influence of peers is indirectly accounted for by including variables like membership in an environmental organization or participating in volunteer work. Some recent work, such as Axsen et al. [2013], Wall et al. [2014] and Kim et al. [2017] have included measures derived from egocentric network data in choice experiment regressions. The consistency of the effects is mixed, and the theoretical development varies. Kim et al. [2017] do develop a model reflecting social distance, and find that closer individuals are more influential on transportation choice decisions. That these studies incorporate egocentric network information in quite different ways suggests that there is scope for further development in this area.

In our survey, the measured egocentric social networks seem to capture concern for the environmental impact of development in the Okanagan. This is also captured by a perception measure we included, that asks respondents if they think the environmental impacts of development are 'worth it'. This variable was a strongly significant and stable predictor in many of the models examined. Our choice experiment would fit broadly with studies that estimate the willingness to pay for some form of environmental protection. Implicitly, that protection is being considered means that some activities or uses will be limited, uses which must generate some benefit to some portion of the population. We have not seen measures that directly measure people's relative assessment of the benefits and costs of not introducing protections. Recent examples would include Gundlach et al. [2018], who examined the willingness to pay for a car free city center, without asking directly about the perceived benefit of car access relative to the costs - noise, congestion, pollution - generated by car access. Similarly, Jin et al. [2018] examine the willingness to pay for cultivated land protection in China, without asking participants if feel that the benefits of loosing cultivated land are worth it. It is more common to include measures of education and environmental awareness, measures which often are found to have little explanatory power. Perhaps the weak predictive power comes from the fact that people can be educated and environmentally knowledgeable, and also view the benefits resulting from environmental damage being worth it. We suggest that including a measure of participant's perception of the benefits of not protecting the environmental resource in question can improve the precision of willingness to pay estimates, as demonstrated in our regression results.

This perception that the benefits of development are worth the environmental consequences may be particularly relevant in the Okanagan. The Regional District of Central Okanagan has been experiencing ongoing high rates of population growth, and the local economy is heavily dependent on real estate development. This may extend to the willingness to pay estimates as well. The willingness to pay to protect natural habitats is quite low, relative to the price of land in the Okanagan, and the willingness to pay to protect rural character is not significant. Both of these parameters also have highly significant standard deviations. Development in the Okanagan, a destination for retirees and amenity migrants, typically occurs in rural areas and areas that were here to for natural. Many retirees move to the Okanagan to build their dream home in the woods with a view of the lake, and many Okanagan developers and real estate agents work hard to sell this dream. Those who understand their connection to the development industry likely would not choose alternatives that in effect mean limiting development.

In contrast, there is a fairly high willingness to pay to enhance aquatic health, as measured by the spawning rate of Kokanee salmon. Salmon are an iconic species in British Columbia, central to the indigenous cultures, and seen by many in the settler culture as something that needs to be protected. In effect, if British Columbians are able to protect the salmon, then they can feel good about how they are protecting the environment. Salmon is something easy to understand, and restricted to riparian areas, where terrestrial needs, such as connectivity, are more complex and more directly impact economic activities.

8 Conclusion

Environmental goods and services are often public goods, and therefore changes in the level of these public goods affects many people. As social animals, humans are embedded in social networks, with an individuals well being connected, through this network, to the wellbeing of others. If ones goal is to measure each individual's private willingness to pay for environmental goods and/or services, then that part of the individual's willingness to pay which is due to their concern for others needs to be removed. Doing so requires knowing the complete social network that an individual is embedded in, and the utility of all others in that network, a generally impossible task.

We have examined the relationship between two measures of egocentric network centrality, degree centrality and transitivity, and willingness to pay for improvements in a set of environmental attributes using a choice experiment. We demonstrate that if the centrality measures are increasing in the utility that the respondent receives from an improvement in the environment, then these centrality measures are positively related to willingness to pay, a result we find in the choice experiment. However, we find that this effect is greatly diminished if we include a measure of the respondents belief that the impact of development on the environment was 'worth it'. As our tool for measuring the egocentric social network identified alters with whom the ego recently spoke about this trade-off, our network measure may have been measuring this belief. Our results suggest that egocentric network measures have potential to adjust some of the bias that network effects can introduce to willingness to pay measurement. More work is needed to identify the most useful approaches to measuring egocentric networks that will offset this bias.

Acknowledgment

We are indebted to Dr.Stephen Davies, Stephan Kroll and Brett Buyere for helpful comments and suggestions. We are also grateful to the University of British Columbia, at Okanagan, Irving K. Barber School of Arts and Sciences for hosting the first author during field work. We thank Margaret Bakelaar and others at Regional District Central Okanagan for their helpful comments and suggestions during the early stage of this work. This project was conducted with generous support from Social Sciences and Humanities Research Council (SSHRC), Canada. All errors are our own.

References

Jonn Axsen, Caroline Orlebar, and Stephen Skippon. Social influence and consumer preference formation for pro-environmental technology: The case of a UK workplace electric-vehicle study. *Ecological Economics*, 95:96–107, 2013.

Abhijit V Banerjee. A simple model of herd behavior. *The Quarterly Journal of Economics*, pages 797–817, 1992.

Mathieu Bastian, Sebastien Heymann, Mathieu Jacomy, et al. Gephi: an open source software for exploring and manipulating networks. *ICWSM*, 8:361–362, 2009.

Stephen P Borgatti, Martin G Everett, and Jeffrey C Johnson. *Analyzing social networks*. SAGE Publications Limited, 2013.

Jetske Bouma, Erwin Bulte, and Daan van Soest. Trust and cooperation: Social capital and community resource management. *Journal of environmental economics and management*, 56(2): 155–166, 2008.

Yann Bramoullé and Rachel Kranton. Public goods in networks. *Journal of Economic Theory*, 135 (1):478–494, 2007.

Ulrik Brandes. A faster algorithm for betweenness centrality. *The Journal of Mathematical Sociology*, 25(2):163–177, 2001. doi: 10.1080/0022250X.2001.9990249. URL `https://doi.org/10.1080/0022250X.2001.9990249`.

William A Brock and Steven N Durlauf. A multinomial-choice model of neighborhood effects. *American Economic Review*, pages 298–303, 2002.

Ronald S Burt. Network items and the general social survey. *Social networks*, 6(4):293–339, 1984.

Karen E Campbell and Barrett A Lee. Name generators in surveys of personal networks. *Social networks*, 13(3):203–221, 1991.

Juan-Antonio Carrasco 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*, 43(1):90–104, 2009.

Juan Antonio Carrasco, 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*, 35(6):961, 2008.

Peter J Carrington, John Scott, and Stanley Wasserman. *Models and methods in social network analysis*, volume 28. Cambridge university press, 2005.

Richard T Carson and Mikołaj Czajkowski. A new baseline model for estimating willingness to pay from discrete choice models. *Journal of Environmental Economics and Management*, 95:57–61, 2019.

Beatrice Crona and Örjan Bodin. What you know is who you know? communication patterns among resource users as a prerequisite for co-management. *Ecology and society*, 11(2):7, 2006.

Gabor Csardi and Tamas Nepusz. The igraph software package for complex network research. *InterJournal*, Complex Systems:1695, 2006. URL `http://igraph.org`.

Mario Diani. Leaders or brokers? positions and influence in social movement networks. *Social movements and networks: Relational approaches to collective action*, pages 105–122, 2003.

Mario Diani and Doug McAdam. *Social movements and networks: Relational approaches to collective action: Relational approaches to collective action*. OUP Oxford, 2003.

Don A Dillman, Jolene D Smyth, and Leah Melani Christian. *Internet, phone, mail, and mixed-mode surveys: the tailored design method*. John Wiley & Sons, 2014.

Elenna Dugundji and Joan 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):70–78, 2005.

David Easley, Jon Kleinberg, et al. Networks, crowds, and markets: Reasoning about a highly connected world. *Significance*, 9:43–44, 2012.

Matthew Elliott and Benjamin Golub. A network approach to public goods. *Journal of Political Economy*, 127(2):730–776, 2019.

Mustafa Emirbayer and Jeff Goodwin. Network analysis, culture, and the problem of agency. *American journal of sociology*, pages 1411–1454, 1994.

Linton C Freeman. Centrality in social networks conceptual clarification. *Social networks*, 1(3): 215–239, 1978.

Linton C Freeman, Douglas Roeder, and Robert R Mulholland. Centrality in social networks: II. experimental results. *Social networks*, 2(2):119–141, 1979.

Benjamin Golub and Matthew O Jackson. Naive learning in social networks and the wisdom of crowds. *American Economic Journal: Microeconomics*, 2(1):112–149, 2010.

Mark Granovetter. The impact of social structure on economic outcomes. *The Journal of Economic Perspectives*, 19(1):33–50, 2005.

William H Greene. *Econometric Analysis*. Prentice Hall, Boston, 7th edition edition, 2012.

Amely Gundlach, Marius Ehrlinspiel, Svenja Kirsch, Alexander Koschker, and Julian Sagebiel. Investigating people's preferences for car-free city centers: A discrete choice experiment. *Transportation Research Part D: Transport and Environment*, 63:677–688, 2018.

Robert A Hanneman and Mark Riddle. Introduction to social network methods, 2005.

Lin He, Mingxian Wang, Wei Chen, and Guenter Conzelmann. Incorporating social impact on new product adoption in choice modeling: A case study in green vehicles. *Transportation Research Part D: Transport and Environment*, 32:421–434, 2014.

Arne Risa Hole. Estimating mixed logit models using maximum simulated likelihood. *Stata Journal*, 7(3):388–401, 2007.

Matthew O Jackson. Networks and economic behavior. *Annu. Rev. Econ.*, 1(1):489–511, 2009.

Jianjun Jin, Rui He, Wenyu Wang, and Haozhou Gong. Valuing cultivated land protection: A contingent valuation and choice experiment study in China. *Land use policy*, 74:214–219, 2018.

Jinhee Kim, Soora Rasouli, and Harry JP Timmermans. Investigating heterogeneity in social influence by social distance in car-sharing decisions under uncertainty: A regret-minimizing hybrid choice model framework based on sequential stated adaptation experiments. *Transportation Research Part C: Emerging Technologies*, 85:47–63, 2017.

Itzhak Krinsky and A Leslie Robb. On approximating the statistical properties of elasticities. *The Review of Economics and Statistics*, pages 715–719, 1986.

Matthieu Latapy. Main-memory triangle computations for very large (sparse (power-law)) graphs. *Theoretical computer science*, 407(1-3):458–473, 2008.

Thomas Macias and Kristin Williams. Know your neighbors, save the planet: Social capital and the widening wedge of pro-environmental outcomes. *Environment and Behavior*, 48(3):391–420, 2016.

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

Peter V Marsden. *Models and Methods in Social Network Analysis*, volume 8, chapter Recent Developments in Network Measurement, pages 8–30. Cambridge University Press, 2005.

Peter V. Marsden. *The SAGE handbook of social network analysis*, volume 25, chapter Survey methods for network data, pages 370–388. Sage Thousand Oaks, CA, 2011.

Daniel McFadden. Modeling the choice of residential location. *Transportation Research Record*, (673), 1978.

Daniel McFadden. Econometric models for probabilistic choice among products. *Journal of Business*, pages S13–S29, 1980.

Vishal Narayan, Vithala R Rao, and Carolyne Saunders. How peer influence affects attribute preferences: A bayesian updating mechanism. *Marketing Science*, 30(2):368–384, 2011.

William Neilson and Bruno Wichmann. Social networks and non-market valuations. *Journal of Environmental Economics and Management*, 67(2):155–170, 2014.

Jonathan Newton. Coalitional behaviour and the provision of public goods on networks. *University of Cambridge [mimeo]*, 2010.

A James O'Malley, Samuel Arbesman, Darby Miller Steiger, James H Fowler, and Nicholas A Christakis. Egocentric social network structure, health, and pro-social behaviors in a national panel study of americans. *PLoS One*, 7(5):e36250, 2012.

Tore Opsahl, Filip Agneessens, and John Skvoretz. Node centrality in weighted networks: Generalizing degree and shortest paths. *Social Networks*, 32(3):245–251, 2010.

Elinor Ostrom. *Governing the commons*. Cambridge university press, 2015.

Antonio Páez, Darren M Scott, and Erik Volz. A discrete-choice approach to modeling social influence on individual decision making. *Environment and planning. B, Planning & design*, 35(6): 1055, 2008.

Paul Paul Budde Communication Pty Ltd. Canada-broadband market-insights, statistics and forecasts, 2016. URL https://www.budde.com.au/Research/Canada-Broadband-Market-Insights-Statistics-a

Christina Prell. *Social network analysis: History, theory and methodology*. Sage, 2012.

Christina Prell, Klaus Hubacek, and Mark Reed. Stakeholder analysis and social network analysis in natural resource management. *Society and Natural Resources*, 22(6):501–518, 2009.

Jules Pretty and Hugh Ward. Social capital and the environment. *World development*, 29(2): 209–227, 2001.

R Core Team. *R: A Language and Environment for Statistical Computing*. R Foundation for Statistical Computing, Vienna, Austria, 2019. URL https://www.R-project.org/.

Soora Rasouli and Harry Timmermans. Influence of social networks on latent choice of electric cars: a mixed logit specification using experimental design data. *Networks and Spatial Economics*, 16 (1):99–130, 2016.

Larry Samuelson. Information-based relative consumption effects. *Econometrica*, 72(1):93–118, 2004.

Riccardo Scarpa, Mara Thiene, and Kenneth Train. Utility in willingness to pay space: a tool to address confounding random scale effects in destination choice to the Alps. *American Journal of Agricultural Economics*, 90(4):994–1010, 2008.

John Scott. Social network analysis: developments, advances, and prospects. *Social network analysis and mining*, 1(1):21–26, 2011.

Holly B Shakya, Nicholas A Christakis, and James H Fowler. An exploratory comparison of name generator content: Data from rural India. *Social networks*, 48:157–168, 2017.

Peter R Stopher. Captivity and choice in travel-behavior models. *Transportation engineering journal of the American Society of Civil Engineers*, 106(4):427–435, 1980.

Jeroen Struben and John D Sterman. Transition challenges for alternative fuel vehicle and transportation systems. *Environment and Planning B: Planning and Design*, 35(6):1070–1097, 2008.

Thomas A. Wall, Gregory S. Macfarlane, and Kari Edison Watkins. Exploring the use of egocentric online social network data to characterize individual air travel behavior. *Transportation Research Record*, 2400(1):78–86, 2014.

Duncan J Watts and Steven H Strogatz. Collective dynamics of 'small-world' networks. *nature*, 393 (6684):440–442, 1998.

Richard Watts and Aaron Witham. Social network analysis of sustainable transportation organizations. Technical report, 2012.

Reza Yousefi-Nooraie, Alexandra Marin, Robert Hanneman, Eleanor Pullenayegum, Lynne Lohfeld, and Maureen Dobbins. The relationship between the position of name generator questions and responsiveness in multiple name generator surveys. *Sociological Methods & Research*, 48(2): 243–262, 2019.