ICIS 2010 Proceedings

International Conference on Information Systems (ICIS)

2010

THE INFORMATION CONTENT OF ECONOMIC NETWORKS: EVIDENCE FROM ONLINE CHARITABLE GIVING

Lauren Rhue
New York University, lrhue@stern.nyu.edu

Arun Sundararajan
New York University, lrhue@stern.nyu.edu

THE INFORMATION CONTENT OF ECONOMIC NETWORKS: EVIDENCE FROM ONLINE CHARITABLE GIVING

Research-in-Progress

Lauren Rhue
New York University
lrhue@stern.nyu.edu

Arun Sundararajan
New York University
lrhue@stern.nyu.edu

Abstract

We measure the "information content" of online economic networks – sets of connected entities where links are created by realizations of shared prior outcomes. We conjecture that such electronic networks contain information about similarity in latent preferences across actors that are not captured by observable product or consumer features. We provide a methodology for measuring this information content in a rigorous and outcome-driven manner using matched-sample estimation techniques to mimic the optimal use of all observable non-network data. Using detailed transaction-level data about 257,851 contributions to 95,684 charitable projects by 99,720 donors on an leading online giving web site, we show that co-donors in an economic network have an 80-fold higher overlap in future choice than a random benchmark, the network outperforms even matched sample alternatives based on sophisticated feature-based predictive models 5-fold to 23-fold, and this inferred overlap in latent preferences persists with local network traversal.

Keywords: social networks, Web 2.0, econometrics, propensity score matching, electronic commerce

Introduction and Related Literature

A growing fraction of interpersonal and commercial interactions are conducted online. A by-product of these interactions is an explosion in the availability of information about the connections between people, between products and across other interacting entities. Some of these *networks* – sets of entities or *nodes* connected by *links* – form an integral part of the electronic interactions being conducted (for example, like those on social networking web sites). Others, like the co-purchase network on Amazon.com, are constructed by capturing, filtering, analyzing and/or summarizing the data trails left by these electronic interactions.

We report on a project that aims to analyze and measure the "information content" of *economic networks* – sets of connected entities where links are created by realizations of shared economic outcomes between entities (Dhar et al., 2009; Jackson 2008). Our research-in-progress paper describes our methodology for measuring information content in a rigorous and outcome-driven manner, using matched sample techniques to mimic the optimal use of *all non-network data* as an alternative to the economic network and a control for the robust quantifying of the true information content of the network. Using detailed transaction-level data from an online charitable giving web site, we construct person-to-person economic networks and measure their information content. Our first empirical results show that the use of the economic networks can increase the odds of predicting future shared choices almost 100-fold relative to a random pairing, and 5-fold to 23-fold over control groups that are chosen using matched sample techniques and a very detailed individual transaction profile.

Our work will add to a growing literature in information systems about the value and impact of social and economic networks. Over the last decade, links based on such economic networks have been made visible to consumers at a wide variety of web sites ranging from Amazon.com and YouTube to the Social Sciences Research Network. The specific economic network we will use in our paper is constructed by linking donors who have contributed to one or more shared "causes" (specifically, projects which share a requesting teacher, more on this later). In prior work about the value of economic networks, Dhar et, al. (2009) provide evidence that future demand trends can be predicted using shared economic links between books on Amazon.com; Oestreicher-Singer and Sundararajan (2009) show that position in this economic network is related to the distribution of demand between popular and niche products. Similarly, Oh et al. (2009) provide some evidence that outcome-based economic links between videos on YouTube can form the basis for the spread of user-generated content. We build on these findings about the impact of economic networks by providing a new calibration of their information content that is benchmarked with predictions based on matched samples.

While the economic network that forms the focus of our investigation are fairly widespread, they are distinct from *social networks* that link individuals with some form of a social relationship. The explosion of social media of various kinds has led to a variety of these networks being created and made visible. While a growth in interest of IS scholars in social networks has paralleled their commercial use, a primary focus of this inquiry is on how electronic social networks serve as a conduit for information flows, how the structure of these networks alters flows of various kinds (Bampo et al, 2008, . A comprehensive survey of the social networks in IS literature is beyond the scope of our document, so we mention a couple of papers that seem especially pertinent to ours. Hill et al. (2006) provided early evidence that social ties are informative by demonstrating that using friendship links to make predictions yields more accurate results than predictions based purely on individual features. Aral et al. (2009) provides a methodology to distinguish the influence that is mediated by social networks from the node homophily that underlies both the creation of the network and the correlation in choices between socially linked consumers. The latter paper is similar to ours in its incorporation of matched-sample techniques into a network setting, but its research question is quite different, its use of matching is more traditional and the techniques in ours paper are developed specifically to address our question of information content in networks.

Our work is also related to a fairly extensive literature on collaborative filtering (CF) in electronic commerce that leverages prior overlaps in choice histories as the basis for product recommendations. (See, for example, Goldberg et al., 1992 and the literature which follows). A typical CF algorithm makes predictions for an individual by inferring his or her future tastes from their history of prior choices and a group of others with overlapping choices. In a sense, one might view a fraction of these CF methods as leveraging underlying economic networks between these individuals. Rather than constructing a specific recommendation method, our focus in this paper is in analyzing the information content of the building block of CF – the economic network – and thus our findings complement this literature while possibly providing conceptual insights that may be incorporated by it. For example, our results indicate that simply using the links in the economic network for prediction can outperform sophisticated

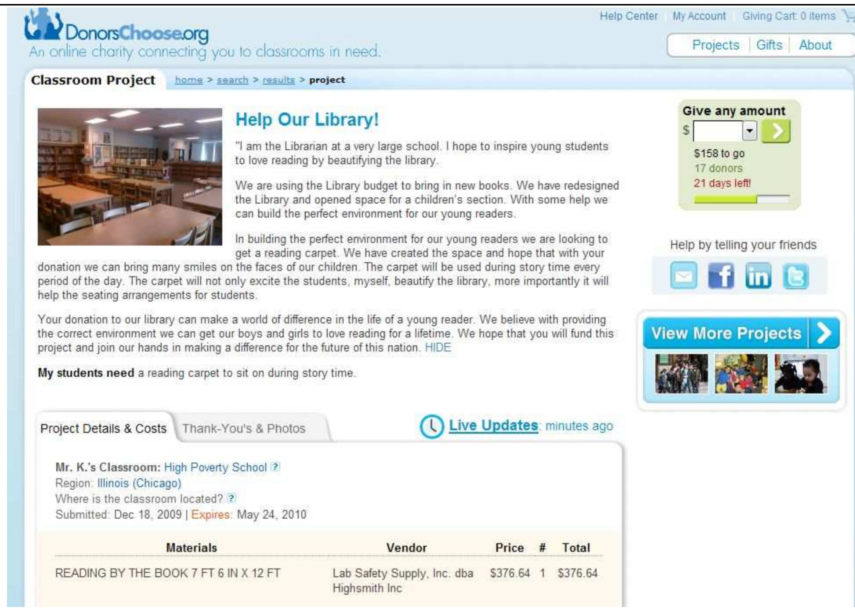

Figure 1: Illustrates a project posted on the DonorsChoose.org web site.

models that create "group of peers" based on prior overlaps in choice, and that second neighbors seem to have a preference overlap comparable to first neighbors.

Finally, our interest in measuring the information content in the context of networks that link donors engaged in *online giving* has its roots in findings which have suggested that when making charitable donations, people seem especially susceptible to peer influence. Prior work has shown that a person's decision to donate to charitable organizations is influenced by not just personal factors like beliefs, discretionary income and demographic characteristics, but social and interpersonal factors as well, such as the giving patterns among one's peer group. For example, Schervish and Havens (1997) find that beyond personal and experiential characteristics, participation in shared community organizations has the strongest correlation with fraction of income donated. Similarly, Andreoni and Scholz (1998) associate a two to three percent rise in individual contributions for each 10% increase in contributions from those in the same "social reference space". A recent experiment by Frey and Meier (2004) who study a university fundraiser suggests that a mere perception of giving by a peer group can influence charitable giving. This literature motivates our current work as well as its extensions discussed in our concluding section.

Data

Our data is obtained from a non-profit organization called DonorsChoose.org, which provides a platform that enables K-12 school teachers to post project listings requesting supplies, books and technological resources. Potential donors browse among the teachers' project listings and give money to specific projects based on their personal preferences. The company was founded in 2000 with a focus on New York City public schools, but has since expanded to include schools across the United States. This website differs from many other non-profit companies because their revenue model depends on effectively matching individuals to appropriate requests for funding (much like that of a for-profit retailer). Figure 1 provides a screenshot of a recent listing on the web site.

Baseline Data

In what follows, a *donor* is a unique individual who has made a contribution at DonorsChoose.org, and a *project* is a request for resources made by an individual teacher at a specific point in time, with a specific timeline for donation, a description of what the funds are being requested for, and a target amount. Figure 2 illustrates one such project with a target amount of $376.74. Donors can contribute fractions of this target amount, and thus a project typically has multiple donors. We have been granted full-access to the donation histories of each donor who has contributed to a project on DonorsChoose.org, including their total number of donations, the details of each corresponding project and its attributes, as well as selected donor demographic data. These data are anonymized and provided to us for research in a set of archival files. For the analysis reported in this document, we have used a pilot subset comprising 95,684 projects, 99,720 donors, and 257,851 contributions

The donor variables constructed for our analysis are summarized in Table 1. Most of these are computed from individual projects the donor has contributed to, thus providing a very detailed profile of the donor's preferences for giving. This is appropriate given that our main question relates to predicting future overlaps in the giving of pairs of donors, and whether an economic network explains this better than this suite of variables.

Table 1: Summary of Key Donor Profile Variables

AvgDonation	The average dollar amount of donations made by the donor.
AvgMaterialPrice	The average cost of materials (project size) of projects donated to.
AvgPoverty	The average percentage of students who qualify for the free and reduced lunch program across the schools given to by the donor.
TeacherAffiliate_X	A set of variables indicating the number of donations made by the donor to teachers across a variety of affiliate programs like Teach for America.
Grade_X	Four variables measuring the number of donations made by the donor to projects in particular grades (K-2, 3-5, 6-8, 9-12).
Resource_X	Five variables indicating the number of donations made by the donor to projects requesting a particular type of resource (Books, Technology, Field Trips, Classroom Supplies, Visitors).
Discipline_X	Six variables indicating the number of donations made by the donor to a particular discipline/subject matter (Math and Science, Literacy and Writing, History, Special Needs, Music, Applied Learning).
DisciplineFocus_X	Twenty variables indicating the number of donations made by the donor to projects that focus on a particular sub-discipline.
Rural, Urban, Suburban	Three variables measuring the number of donations made by the donor to schools in rural, urban, or suburban areas.
State_X	Fifty-one variables indicating the number of donations to schools in each of the U.S. states and the District of Columbia.

We also have limited demographic information: specifically, self-reported gender and an indication of all states in which donor listed an address any time before 9/30/2009. In a subsequent section, we interact the grade, resource and discipline focus variables across donors. We have access to two "social networks": one from an ongoing "tell-a-friend" campaign and another created from a "donate-on-behalf-of your-friend" campaign during the holiday season. We discuss these briefly in our concluding section.

Construction of the Co-Donor Networks

We call the economic network we construct a *co-donor network* in which the nodes are individual donors and links are created between these donors based on shared contributions to projects. These links aim to capture similarity in *latent preferences* across the donors, based on our conjecture that donors who give to the same projects are likely to have a similarity in preferences that is not captured by either the attributes of the projects they contribute to, or the available donor attributes. Because not all projects are live simultaneously, but teachers tend to have many projects posted and tend to teach the same grade and subject at the same school over time, we have further conjectured that donors who give to the same teacher likely have similar latent preferences. We therefore first construct a bipartite donor-teacher graph (rather than the bipartite donor-project graph, which is extremely sparse) with donors on one side and teachers on the other, and with links from donors to the teachers whose projects they have contributed to. The co-donor network is then simply the incidence graph defined by the donor side of this bipartite graph – two donors have an edge in the co-donor network if they have contributed to the same teacher. This process is illustrated in Figure 2. We store this as a time-stamped multi-graph, which allows us to view the co-donor network that exists at any point in history.

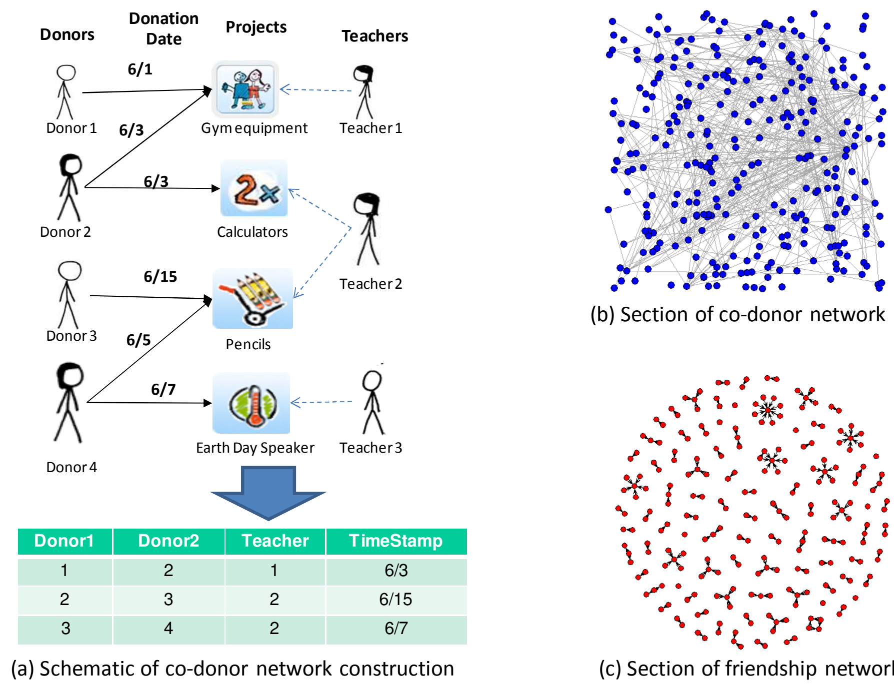

Figure 2: Network construction and visualizations of a segment of co-donor and social networks

Measuring Information Content

How informative is this co-donor network about future giving? Rather than imposing an exogenous measure of the informativeness of the edges in the co-donor network, we rely on an *outcome-based* measure. If the co-donor network is informative, the predicted overlap in future giving between current co-donors should be higher than the overlap in giving between donors and their predicted "matches". Briefly, we partition our pilot data in time; we construct a co-donor network using the earlier time-slice of data; we build predictive model of future shared giving using prior shared giving and all available donor and project data, and we contrast the empirical overlap in future giving based on being linked by the co-donor network with three benchmarks: a random control and two "matched" control groups that are matched using the predictive models. We will conclude that the co-donor network has higher information content if its predicted overlap is higher than that of the controls.

Now for the details. We partition our data into two subsets, reflecting all donations made prior to June 30, 2008 (the *baseline* set), and all donations made subsequent to July 1, 2008 (the *test* set). We construct the co-donor network based on all donations contained in the baseline set (but not the test set). We then randomly choose a set A of 500 donors from the co-donor network whom we label the "targets". For each "target" donor a in A, we identify the set of "first neighbor" donors $D_{FN}(a)$ who have a direct co-donor link to the donor a, and the set of second neighbor donors $D_{SN}(a)$ who have a direct co-donor link to a donor in $D_{FN}(a)$. These are two sets of economic neighbors based on the co-donor network. Denote the set of all first neighbors of the targets as,

$$\mathrm{N}(A) = \bigcup_{a \in A} \left[D_{FN}(a) \right],$$

and the set of all non-target and non-first-neighbor donors as

$$\mathrm{M}(A) = \Delta \setminus \left\{ A \cup N(A) \right\}.$$

If the co-donor network is informative about future shared giving, then there should be an empirical overlap in the test set between the donations made by a and its economic neighbors that is "above average". However, what should this overlap be benchmarked against? What is a good "average" to use? Our first benchmark is, for each a, a *randomly chosen* set of donors $D_R(a)$, chosen from the set M(A). This random benchmark is potentially a strawman because the overlap in shared future giving or the information contained in this co-donor link might simply reflect observable donor-specific and project-specific characteristics or features. Next, we therefore use *all the observable information we have* to construct two more benchmark sets $D_{MI}(a)$ and $D_{MP}(a)$ of donor "matches" for each target in A. The procedure we use is summarized below

1. For each a, we assign a "overlap incidence" of 1 to the elements of $D_{FN}(a)$, and an "overlap incidence" of zero to the elements of M(A). For readers familiar with matched sample estimation and propensity score matching, think of the overlap incidence as analogous to a binary-valued "treatment".

2. For each a, we then use logistic regression to estimate a propensity score model Π where the dependent variable is the overlap incidence, and the covariates or independent variables are *all the variables* listed in our Data section. We use $\Pi(x,a)$ to denote the score predicted by this model for donor x to be a co-donor of a, which reflects that every target has a propensity scoring model that is individually estimated[1].

3. We use Π to construct two matched sample sets for each target a, a *pooled matched set* $D_{MP}(a)$ and an *individualized matched set with interactions* $D_{MI}(a)$

***$D_{MP}(a)$*:** This is the set of donors in M(A).with the highest estimated propensity of being a co-donor of a. In a sense, these are the donors whose choices have not overlapped with those of a, but which our data tells us should have overlapped. That is, for each x in $D_{MP}(a)$, $\Pi(x,a) \geq \max_{y \notin D_{MP}(a),\, y \in \mathrm{M}(a)} \Pi(y,a)$. For balanced comparison purposes, the size of $D_{MP}(a)$ is chosen to be exactly the same as the size of the set of first neighbors $D_{FN}(a)$.

***$D_{MI}(a)$*:** This set of donors is chosen to account for the fact that donors may sometimes contribute to the same projects because of a common overlap between their preferences for specific project attributes. To account for this, the following procedure is used:

- Each donor d in $D_{FN}(a)$ is matched with the donor m in M(A) with the closest propensity score (that is, such that $\|\Pi(d,a)-\Pi(m,a)\|$ is minimized), and the matching is performed with replacement. The latter donor m is "assigned" the putative target a. Denote the set of all such "best matches" as C.
- A *second* propensity score model **Ω** is estimated, again using logistic regression but with additional variables. All the first neighbors of all the targets $\bigcup_{a \in A} D_{FN}(a)$ are assigned the "overlap value" of 1, and all the elements of C are assigned the overlap value of zero. The covariates used for this model include all the variables used in the first model as well as a new set of *interaction variables*, where each element of C is interacted with its putative target a. We interact the *Grade_X*, *Resource_X* and *Discipline_X* variables across targets and existing/candidate co-donors. For any arbitrary donor y, denote the propensity to be a co-donor of a predicted by this new model as **Ω**(y,a).
- For each a in A, the set $D_{MI}(a)$ is constructed by choosing the elements y in M(A) with the highest propensity scores **Ω**(y,a). That is, the set of most likely co-donors of a determined by applying the propensity score model **Ω** on all alternative candidate co-donors in M(A).

Results, Conclusions and Ongoing Work

The results we report provide strong evidence of high information content of the co-donor network. Our dependent (outcome) variable in this analysis is the overlap between the future project donation choices of each of the targets a in A and each donor in these five candidate sets $D_{FN}(a)$, $D_{SN}(a)$, $D_R(a)$, $D_{MP}(a)$ and $D_{MI}(a)$ described above. For any donor x in any of these sets $D_i(a)$, there is a unique level of overlap between the donations of x and the donations of a during the test interval July 1, 2008 through September 30, 2009. We use two dependent variables to record this overlap – a *binary indicator* which indicates whether there was any overlap at all (and which is used in both the T-tests and the logistic regressions below), and a *count* of the number of overlapping projects. Once these dependent variables have been recorded, we can simply pool all the elements of $D_i(a)$, treating each as a *distinct* observation. Denote the pooled set as $\mathbf{D_i}$. For example, $\mathbf{D}_{\mathrm{FN}} = \bigcup_{a \in A} D_{FN}(i)$.

[1] Further details about our matching estimators including the exact equations are available on request.

We perform all of our analysis as pair-wise comparisons between a candidate *economic network* group [$\mathbf{D_{FN}}$ or $\mathbf{D_{SN}}$] and *control* group [$\mathbf{D_R}$, $\mathbf{D_{MP}}$ or $\mathbf{D_{MI}}$]. That is, for each comparison (regression), we have one economic network group and one control group. We define the independent binary variable *EconomicNetwork* to indicate membership of the observation in $\mathbf{D_{FN}}$ or $\mathbf{D_{SN}}$ and distinguish between the two groups. For example, when we compare overlap with first neighbors to the pooled matched set, we use the observations in the sets $\mathbf{D_{FN}}$ and $\mathbf{D_{MP}}$, the variable *EconomicNetwork = 1* for each observation in $\mathbf{D_{FN}}$, the variable *EconomicNetwork = 0* for each observation in $\mathbf{D_{MP}}$, and the observations in $\mathbf{D_{SN}}$, $\mathbf{D_R}$ and $\mathbf{D_{MI}}$ are not used.

Our first set of results assesses whether the average outcome values (donation overlap) across the groups is sufficiently different to warrant further analysis. Two-sided T-tests indicate that each of the economic network groups has a significantly different mean outcome from each of the control groups.

Table 2: Differences in Mean between the Economic Network and Control Groups

	Economic Network Group	
Control Group	First Neighbors (FN)	Second Neighbors (SN)
Matched - Pooled (MP)	*t-stat=8.311, p <0.001*	*t-stat=5,821, p <0.001*
Matched - Individual (MI)	*t-stat=4.179, p <0.001*	*t-stat=6.954, p <0.001*
Random Group (R)	*t-stat=13.29, p <0.001*	*t-stat*=7.726*, p <0.001*

What is perhaps equally interesting is that we were unable to establish any significant difference in mean between the two economic network groups (t-stat=0.1646, p=0.892). In other words, first and second neighbors in the co-donor network are indistinguishable on outcome.

Our second set of results compares the binary occurrence of co-donations across pairs of one economic network group and one control groups by estimating a logistic regression with the binary indicator outcome as the dependent variable. The *EconomicNetwork* variable is included as an independent variable, as are all the covariates described in prior sections. The results of these regressions are summarized in Table 3.

Table 3: Logistic Regression Results

	First Neighbors (FN)			Second Neighbors (SN)		
	Random (R)	Matched-Pooled (MP)	Matched-Individual (MI)	Random (R)	Matched-Pooled (MP)	Matched-Individual (MI)
Constant	-6.781*** (0.560)	-5.742*** (0.3796)	-4.492*** (0.321)	-7.272*** (0.5748)	-6.605*** (0.4543)	-4.725*** (0.351)
EconomicNetwork	4.373*** (0.5262)	3.165*** (0.329)	1.674*** (0.224)	4.86*** (0.5166)	4.003*** (0.362)	2.237*** (0.241)
AvgDonation	-0.001 (8E-4)	-0.001 (7.97 E-4)	-0.001 (7.92 E-4)	-0.004*** (0.001)	-0.003** (0.001)	-0.004*** (0.001)
AvgMaterialPrice	1.501 E-4* (6.79 E-5)	1.19 E-5 (6.1 E-5)	9.11E-5 (4.8 E -5)	6.39 E-5 (6.33 E-5)	5.25 E-5 (6.45E-5)	4.71 E-5 (5.87 E-5)
AvgPoverty	-0.021*** (0.003)	-0.019*** (0.003)	-0.014*** (0.003)	-0.009* (0.004)	-0.006 (0.004)	-0.008* (0.004)
Rural	-0.173 (0.122)	-0.130 (0.119)	0.075 (0.144)	0.096 (0.188)	0.071 (0.230)	0.0288 (0.153)
Urban	-0.238 (0.126)	-0.1867 (0.122)	0.038 (0.142)	0.136 (0.188)	0.065 (0.216)	0.059 (0.155)
Suburban	-0.305* (0.132)	-0.258* (0.126)	-0.066 (0.147)	0.094 (0.183)	0.070 (0.213)	0.048 (0.157)

Table 4 summarizes the results of our Poisson regressions with the same independent variables and using the *count* of the number of shared project donations as the dependent variable.

Table 4: Poisson Regression Results						
	First Neighbors (FN)			Second Neighbors (SN)		
	Random (R)	Matched-Pooled (MP)	Matched-Individual (MI)	Random (R)	Matched-Pooled (MP)	Matched-Individual (MI)
Constant	-6.762*** (0.553)	-4.538*** (0.247)	-3.141*** (0.229)	-7.268*** (0.573)	-5.568*** (0.342)	-2.956*** (0.260)
EconomicNetwork	4.33*** (0.52)	1.995*** (0.186)	0.968*** (0.152)	4.824*** (0.512)	2.911*** (0.233)	1.180*** (0.166)
AvgDonation	-1.44 E-3 (8.14 E-4)	-0.002* (8.0 E-4)	-0.003*** (0.001)	-0.004** (0.001)	-0.003* (0.001)	-0.005*** (0.001)
AvgMaterialPrice	1.42 E-4* (6.23 E-5)	6.384 E-5 (6.496 E-5)	1.812 E-5 (6.434 E-5)	6.624 E-5 (5.659 E-5)	2.561 E-5 (7.344 E-5)	-2.159 E-5 (8.073 E-5)
AvgPoverty	-2.13 E-2*** (3.38 E-3)	-0.020*** (0.003)	-0.020*** (0.003)	-0.007 (0.004)	-0.008* (0.004)	-0.017*** (0.003)
Rural	-0.17 (0.13)	-0.217* (0.084)	-0.227** (0.086)	0.210 (0.288)	-0.183 (0.196)	0.147 (0.205)
Urban	-0.22 (0.14)	-0.260** (0.086)	-0.308*** (0.086)	0.140 (0.267)	-0.009 (0.181)	0.118 (0.186)
Suburban	-0.28 (0.14)	-0.393*** (0.093)	-0.425*** (0.094)	0.245 (0.267)	-0.076 (0.180)	0.0816 (0.192)

Space constraints preclude including all estimated coefficient values, but the main result is striking. The *EconomicNetwork* variable is positive and significant across all of the logistic regressions and Poisson regressions. Furthermore, while different controls are statistically significant across the different specifications, this is the *only* variable that is economically significant across all of the regressions, a significance that persists with both forward and backward stepwise regression. The economic significance of these coefficients is substantial, a fact made clear by computing the corresponding odds ratios. For example, relative to the average donor in the individually matched group, being a first neighbor in the increases the odds of making a future donation to the same project by over 500% (since exp(1.674) = 5.33). Corresponding transformations of other Tale 3 coefficients to odds ratios suggests that this amplification is up to 23-fold relative to the individually matched group, and close to 80-fold relative to the randomly chosen group[2]. The coefficients for second neighbors were slightly higher than those of the first neighbors, although we found no evidence of these being statistically different. In other words, the informativeness of the economic network is not restricted to just neighboring nodes, but rather, the overlap in preferences implied by the network *does not diminish much with network traversal.*

A central idea underlying our research is that not all online networks are created equal. They reflect different kinds of relationships between the entities they connect, these connections have differing levels of strength or closeness, and the connection between these entities may either be visible or hidden to the actors in any electronic commerce setting. Our ongoing work identifies three possible dimensions of an online network that affect its informativeness (1) the extent of the network's visibility to consumers; (2) whether the links in the network represent "social" ties or economic ones, and (3) the strength of these ties. Our work-in-progress is *contrasting* the information content of the *economic* network with two *social networks* between these donors, assessing how this informativeness varies with tie strength, and will run a series of field experiments at DonorsChoose.org towards assessing the *influence* of making these different links *visible* to potential donors.

[2] Space constraints preclude a more detailed analysis of our Poisson regression coefficients. We also found (not reported in the table) that in both cases, the *Discipline_Literacy* variable was significant and positive, which indicates that this specific disciplinary focus is an interest shared by people who have fairly cohesive preferences and consistency in the projects that they give to.

References

Andreoni, J., and Scholz, J. K. 1998. "An Econometric Analysis of Charitable Giving with Interdependent Preferences," Economic Inquiry (36:3) July, pp. 410-428.

Aral, S., Muchnik, L and Sundararajan, A. 2009. "Distinguishing Influence Based Contagion from. Homophily Driven Diffusion in Dynamic Networks," Proceedings of the National Academy of Science (106:51), December, pp. 21544-21549.

Bampo, D., Ewing, M., Mather, M., Stewart, D. and Wallace, M. 2008. "The Effects of the Social Structure of Digital Networks on Viral Marketing Performance. Information Systems Research (19:3), September, pp. 273–290.

Dhar, V., Oestreicher-Singer, G., Sundararajan, A. and Umyarov, A. 2009. "The Gestalt in Graphs: Predition Using Economic Networks," available at http://ssrn.com/abstract=1500834

Frey, B. S., and Meier, S. 2004. "Social Comparisons and Pro-Social Behavior:Testing "Conditional Cooperation" in a Field Experiment," American Economic Review (94:5), December, pp. 1717-1722.

Hill, S., Provost, F., and Volinsky, C. 2006. "Network-based Marketing: Identifying Likely Adopters via Consumer Networks," Statistical Science (21), pp. 256-276.

Jackson, M. O. 2008. *Social and Economic Networks.* Princeton University Press.

Oestreicher-Singer, G. and Sundararajan, A. 2010. "Recommendation Networks and the Long Tail of Electronic Commerce," available at http://ssrn.com/abstract=1324064

Oh, J., Susarla, A., and Tan, Y. 2009. "Examining the Diffusion of User-Generated Content in Online Social Networks," available at http://ssrn.com/abstract=1182631

Schervish, P. G., and Havens, J. J. 1997. "Social Participation and Charitable Giving: a Multivariate Analysis," Voluntas (8:3), September, pp. 235-260.