Power Imbalance and Rating Systems

Bogdan State
Department of Sociology
Stanford University
bstate@stanford.edu

Bruno Abrahao
Department of Sociology
Stanford University
abrahao@cs.stanford.edu

Karen S. Cook
Department of Sociology
Stanford University
kcook@stanford.edu

Abstract

Ratings are critical to the function and success of services in the emerging sharing economy. They are a means through which users develop trust in one another and in the services themselves. Ratings are designed to give users a proxy for the expected quality and risk of potential online transactions. We expect online ratings to reflect an objective measure of quality, but such evaluations in fact may be systematically distorted by many, complex social-psychological processes. Decoupling these subjective factors from rating systems to correct for biases and to provide neutral assessments of risk and quality has proved extremely challenging. We focus on one of the most prevalent factors in virtually every form of social exchange. Differences in resource ownership affect the balance of power in interpersonal interactions, likely impacting online ratings. We demonstrate how power imbalance affects mutual ratings using a massive dataset from CouchSurng.org, an international online hospitality exchange network. Our methodology employs a deductive approach to knowledge discovery. Through a series of observational experiments, we find support for a sociological theory dating back to the 1960s, Power-Dependence Theory (PD), as a possible explanation. PD predicts that power-imbalanced relationships induce user behavior that attempts to balance power. We find support for status-giving as a likely mechanism driving the asymmetry of ratings between power-unequal users. Our findings underscore the need for ratings systems to account for the tendency of mutual ratings between users that hold differential resources to be asymmetrical, especially under conditions of resource scarcity.

1 Introduction

The Internet is changing how we exchange. A new wave of more than ten thousand online companies is now emerging under the banner of the sharing economy (Stein, 2015). Platforms such as AirBnB, Uber or TaskRabbit facilitate the trading of services by creating and maintaining trust between strangers. To encourage trust and prevent bad outcomes, these services invariably employ some form of rating system with the premise that the ratings associated with an entity are reliable indicators of the quality and risk involved in potential online transactions.

To provide for social exchange on a global scale these services face daunting obstacles. As exchanges in the sharing economy are conducted through interpersonal interactions governed by complex social norms, the ratings obtained through such systems are vulnerable to normative distortions, with which the large marketplaces of the Internet, such as Amazon, Alibaba and Expedia, need not contend (Zervas, Proserpio and Byers, 2015).

In the sharing economy trust is essential for the completion of transactions and negative interactions can have costly or even catastrophic consequences (Nuzzi, 2014). An extensive body of algorithmic research has focused on developing methods to combat the abuse of ratings systems (e.g.: Massa and Avesani, 2007). Similarly, human-computer interaction research has dealt with the construction of ratings systems in a way that accurately captures user intent (e.g. Lampe and Garrett, 2007).

As we will show, in social systems such as those created by the sharing economy, information collected about entities is also vulnerable to distortions due to underlying interpersonal processes, an area which has received comparatively little investigation (but see Teng, Lauterbach and Adamic, 2010). Because they are driven by offline interactions, these distortions are likely to occur even when a rating system is free from abuse and misunderstanding. For example, the magnitude of a rating might be influenced by the raters number and quality of prior experiences, their expertise, their position in the exchange, their mood, or by the number of outside options available, to list only a few of the potential social-psychological factors involved. Given that rating systems seek to extract objective assessments it is important to understand and account for interpersonal processes that affect such evaluations.

We focus on understanding how one of the most prevalent factors in virtually all forms of social exchange influences ratings in social networks: the fact that actors hold different resource endowments. For instance, an Uber driver controls the car (the "resource"), whereas the passenger is temporarily dependent on the driver to get to their destination. Similarly, a host using the paid hospitality service AirBnB controls the space they make available to their guest. Such differences in resource ownership and control affect the balance of power in interactions. We show that they may also affect the ratings generated by these interactions.

Social Exchange Theory offers an elegant, parsimonious definition of power, a basic social concept whose very nature has been theorized in many different ways over time. Social Exchange Theory proposes that power and dependence are intertwined, relational concepts (Emerson, 1962). An actor has power over another inasmuch as the other is dependent on him or her for valued resources for which alternative sources are scarce. This understanding is at odds with the early, positional definitions of power, which conceptualized it as a static quantity, inherent in an individuals ability to control resources or impose their will on others (e.g. Marx, 1965; Weber, 1978; Parsons, 1969).

We aim to understand how this power imbalance impacts user ratings. However, in data mining we are often faced with an inherent limitation in the ability to pick up signals reflecting internalized beliefs and induced behaviors. Social-psychological factors are extremely challenging to detect, measure, and quantify. We address this limitation by identifying a dataset that allows us to capture the relevant factors that come into play in social exchange. We analyze a large dataset from CouchSurfing.org, a service facilitating interactions between over 5 million users worldwide, who are seeking hospitality from, or willing to provide hosting to, other members of the site. CouchSurfing has been the object of a number of previous studies (Teng, Lauterbach and Adamic, 2010; Lauterbach et al., 2009; Bialski and Batorski, 2010), but, to our knowledge, no study has investigated rating dynamics from an exchange-theoretic perspective using data from CouchSurfing or other sharing economy platforms.

Studying power imbalance requires researchers to identify the resources of interest in our case, living space offered on CouchSurfing and owned by the hosts , and the exchange of mutual evaluations friendship and trust ratings. On CouchSurfing "couches" or spare beds are the primary resource of interest to a traveler (henceforth "surfer") in search of a place to stay in a particular city for a determined time period. Not only does opening one's house represent a less-than-universal act of generosity, but even the most committed host's ability to provide others with hospitality is limited. More interestingly, monetary exchange is prohibited on Couchsurfing. Thus, we argue that the intrinsic value of hospitality leads to a high valuation of the host's resources by the surfer, who depends on hosts to find accommodation, typically making for a power imbalance in the relationship.

Our main observation exposes the tendency of guests (CouchSurfers) to give higher ratings to their hosts than vice-versa. Likewise, we observe that guests are more likely to rate their partner than hosts are. This, we believe, is not a phenomenon that is idiosyncratic to CouchSurfing. To understand this mechanism and to generalize our findings, we turn to the social sciences to provide a cogent explanation for the rating imbalances we observe in the data. Power-Dependence Theory (PD), a sociological theory in the Social Exchange tradition, emerges as a potential candidate. Given the richness of its vocabulary, we use Power-Dependence to guide our exploration of the social-psychological mechanisms underpinning the observations (for examples of this approach see Danescu-Niculescu-Mizil et al., 2012). Furthermore, due to the ubiquity of resource imbalances in both online and offline social interactions, Power Dependence Theory can be powerful in the study of social networks from a computational perspective, particularly in the sharing economy. Indeed, Kleinberg and Tardos (2008) argued using a game-theoretic approach that power-balancing behaviors fundamentally influence social network formation and structure. Without excluding the possibility of multiple other interpersonal processes being at work in a real-world setting, we find increasingly strong support for PD, through a series of comprehensive observational studies.

1.1 Power Dependence Theory

Richard Emerson (1972a), the main proponent of Power-Dependence Theory put forth the hypothesis that a fundamental tension emerges in situations in which one actor is more dependent on an exchange partner than vice-versa. To resolve this tension, Emerson hypothesized that actors engage in behaviors that help move the relationship to a more balanced state,as perceived by the parties involved. Several processes may move a relationship closer to power balance. For example, the low-power actor may reduce her dependence by withdrawing from the relationship. The dependent party may also seek other, alternative sources for the resource provided by the high-power actor. Emerson likewise cited the case of less relevance to our particular context of low-power actors forming coalitions against high-power ones.

More interesting to us is status-giving, the fourth power-balancing mechanism proposed by Emerson. The low-power actor's dependence in the relationship may decrease if the dependent actor provides their partner with status, understood here as esteem (Emerson, 1972b) and reverence. If we assume that the high-power partner values the status they receive, then something valuable is flowing both ways in the exchange relation, reducing the power imbalance. Even this assumption may be unnecessary: the status giving behavior is likely to occur if the low-power actor has at least the expectation of the high power actor valuing the status given. We study rating asymmetries through the lens of status-giving, a power-balancing mechanism, which exposes the tendency of mutual ratings between users who hold differential resources to be asymmetrical, especially under conditions of resource scarcity.

1.2 Power balance through status-giving

Status and power are often conflated conceptually, given their frequent interrelation in social settings (Willer et al., 2012; Galinsky et al., 2008). In spite of this frequent coincidence (as assumed in Danescu-Niculescu-Mizil et al., 2012), there is no necessary relationship between the two concepts (Willer et al., 2012). Status Characteristics Theory proposes that status-giving may occur in power-imbalanced situations. It has been shown in experiments that status beliefs formed in the wake of a resource imbalanced encounter tend to be more favorable to resource advantaged individuals (Ridgeway et al., 1998). But rather than expressing a universal relationship, power seems to transform into status only when certain conditions are met. In particular, how power

is used is an important mediator of the relationship between power and status. Our case study concerns a particular type of power use generosity that has been shown to lead to status giving under laboratory conditions (Willer et al., 2012).

Operationalizing status. Status relations are a fundamental element of informal social hierarchies. At its core, status is an abstract, concept that data can capture only imperfectly. Importantly, status should not be equated with reputation alone. While reputation is a reported measure of the overall opinion about an entity, status is a more complex concept. It can be given, but not traded thus, it cannot be conceptualized as a commodity. Depending on context, status may connote esteem (Weber, 1978), competence (Ridgeway et al., 1998), technical knowledge (Blau, 1964), or cultural facility (Peterson, 1997). Furthermore, an individual may be considered high status (for instance, on account of their taste and manners) and have a low reputation (as a result of behavior judged unacceptable).

Because we are interested in how status flows from one individual to another in interpersonal relationships, the richness of CouchSurfings data presents a promising opportunity for measuring individual status-giving interactions. We focus on the asymmetry of users ratings toward one another as a likely proxy for this flow of status. Regardless of the nature of the rating (e.g. CouchSurfing measures of friendship and trust), if status is flowing from one user to the other, then one user is expected to hold the other in higher esteem than is the case in reverse. This discrepancy in the latent factors that lead to the generation of ratings is likely to be reflected in the observed imbalance between ratings. Asymmetries in friendship nominations have been a standard tool in the analysis of status hierarchies (e.g. Moody, 2001) in social networks. Rather than the typical binary nomination, our data include ordered evaluations of tie strength. But status is expected to flow not only in situations where ratings elicit opinions on friendship, but in any situation where the rating demands one user to give an opinion of another. For this reason we include in our analysis a second set of ratings, which are explicitly concerned with trustworthiness. Unlike friendship nominations, in our dataset trustworthiness ratings are not disclosed to the interaction partner, making them a stronger measure of *internalized*, rather than performative status giving. We use asymmetry in trustworthiness ratings to gauge for the deeper status exchange process that goes beyond the norms of gratitude.

Our analysis also looks at a third measure of status-giving, namely the asymmetry in the likelihood of giving a rating at all. Because giving ratings to others is in itself a costly act, having a high opinion of ones partner is expected to make one more likely to bear the cognitive cost of rating and thus impact the likelihood of observing a rating. This measure is relevant in the context of CouchSurfing, where surfers are traveling, often have only intermittent Internet access, and are thus in a situation where giving ratings is even costlier than it is for hosts who are likely to be at their place of residence.

1.3 Mediating Factors

Internalization. Power balancing mechanisms are not necessarily employed rationally, nor do they have to be consciously utilitarian. In situations of perceived unfairness, such as in the ultimatum game (Osborne and Rubinstein, 1994), the low-power actor is likely to reduce her dependence by withdrawing from the relationship, even though maintaining an unfavorable relationship may be the most rational strategy. Ultimatum game research shows that such strategies are employed even in one-time interactions, as sharing economy exchanges could be conceptualized.

Two main views of status are possible with respect to rationality. In the first, status-giving is a conscious, "performative" act in which the dependent actor engages, and which is not necessarily consistent with their beliefs. This view places status closer to hierarchical relations in formal organizations. A subordinate will (or is expected to) address a superior respectfully regardless of their internal beliefs. Power Dependence describes status as an internalized belief first and foremost, which may be manifested externally fully or even dampened by circumstances.

Our data offer a unique opportunity to examine the degree to which status- giving is internalized. Because CouchSurfing collects both public (friendship) and private (trust) interpersonal ratings, we can use the difference in the effect between the public and private ratings asymmetry to investigate whether status-giving is a deliberate or internalized process in this context. If status is given conscientiously as a calculative act, then we would expect the effect to be stronger in the public ratings. Conversely, if status-giving is based on internalized beliefs (i.e., gratitude), then we expect the effect to be visible primarily in private ratings.

Anchoring. The expected internalization effect is net of that of anchoring, i.e. the process through which one user adapts their public ratings to the previously observed or expected ratings generated by their partner, due to norms of reciprocity. This effect has been shown to occur in CouchSurfing ratings (Teng, Lauterbach and Adamic, 2010). Corroborating our expectations with respect to both anchoring and internalization leads us to expect the ratings asymmetry to be stronger in private rather than public ratings.[1]

Reciprocity. Both gratitude and anchoring are manifestations of reciprocity, an overarching norm which dictates that individuals repay the others generosity. There is implied reciprocity in every relation of dependence, and status-giving itself can be conceptualized as a manifestation of the norm of reciprocity. Because reciprocity is a fundamental process of social interaction, it is impossible to decouple it from status-giving entirely, though it is possible to better specify the relationship between the two concepts. Reciprocity can be conceptualized as either gratitude or obligation (Goei and Booster, 2005). PD predicts that the guest will feel grateful, and not obligated, towards the host, and offers an explanation for why the guest is grateful, and reciprocity

[1] A consequence of status-giving not being an intentional behavior is that actors may not necessarily be aware of it. This is why even in-depth interviews may not capture the interpersonal dynamics at work in power-unequal situations.

is manifested in a particular way.

Status-giving behavior cannot be purely ascribed to reciprocity, however. Although PD predictions may seem intuitive, they are at odds with the core tenets of cognitive dissonance, an alternative explanation with respect to reciprocity. Contrary to PD's prediction, in interpersonal relationships, psychologists postulate that the receiver of resources is more likely to be held in higher esteem than vice-versa. The giver is likewise hypothesized to be more committed in the act of giving, whereas the receiver is expected to harbor more mixed feelings (Weinstein and Ryan, 2010; Horan and Booth-Butterfield, 2010). This could be due to the givers investment, financial, time-wise, and so on, in the act of giving as well as the fact that the receivers are perceived by the givers as having enough value to justify the sacrifice of their resources. Furthermore, if the receiver of hospitality tries to rationalize their position, they may also perceive the expectation of reciprocity as an obligation, or as the feeling of having been purchased. In the context of one-time interactions, rationally-construed obligations are not likely to be reciprocated (Goei and Booster, 2005). Thus, cognitive dissonance would lead us to expect status flowing in the opposite direction, from host to surfer.

Scarcity. Not all resources are created equal in the view of Power Dependence. Some resources are scarcer than others, due to their natural rarity, differences in demand or the lack of viable alternative sources. For instance, couches in a tourist hub may be more valuable than couches in the countryside. PD holds that the power imbalance is proportional to scarcity. As a consequence we expect the asymmetry measuring status-giving to be more pronounced in cases of scarcity.

Supporting PD. Power Dependence Theory predicts that in response to the power imbalance surfers may give status to their hosts. We find support for this hypothesis in our data (Section 3), which reveal that status-giving is reflected in the asymmetry of both private trust ratings and public friendship ratings, and in the commitment to providing ratings. While other mechanisms may be at work in this setting, we find support for PD using a comprehensive set of tests grounded in the main tenets of the theory. We control for other potential explanations through a number of experiments. Using role reversal (Section 3) we confirm that status-giving emerges from dependence and not from intrinsic personal characteristics, such as the position one has in the organization. We control for the effect of anchoring (Section 3.1), by performing a between-dyads analysis, where we only examine the first rater. Finally, in Section 4 we study the relationship between resource scarcity and status-giving on CouchSurfing.

2 Data and Methods

CouchSurfing users may play the roles of host, surfer, or both. The service makes no distinction between users that play different roles. The user account is universal and uniform regardless of the role the user intends to occupy. The roles are expressed by users through actions of social signaling, e.g., making a couch available or searching for a couch. Likewise, the ratings are assigned to users (not to their roles) and are accumulated over all interactions (regardless of role) in which the user participates. All users are presented with the same web interface, including the same rating interface and scale (potential differences in the psychometric scales for the host and surfer role are controlled for in Section 3). Thus, we draw our sample of surfers and hosts from the same user population, in which a large intersection between the two roles exists, i.e., many users have both surfed and hosted at least once.

We analyze a sample of anonymized dyads, defined as host-surfer relationships, with 1,446,690 hospitality interactions occurring between users across the world and facilitated by CouchSurng in the period 2003-2012. The data include tie strength and trust ratings. For the user to establish a tie with other users in the network, the service presents him or her with two mandatory rating tasks whose outcomes constitute our dataset: (1) the strength of the tie and (2) the level of trust between parties. Tie strength becomes visible to the other party, and we map its six ordered levels into integer values, namely "1: Acquaintance" "2: Couchsurng (CS) Friend," "3: Friend," "4: Good Friend," "5: Close Friend," and "6: Best Friend". Conversely, trust ratings are recorded, but never reported. This rating spans a set of five ordinal values: "1: Do not Trust," "2: Trust Somewhat," "3: Generally Trust," "4: Highly Trust," and "5: Would Trust with Life." Tie strength ratings are made available as soon as any unilateral rating is given. However, at the time the data were generated, the platform did not notify the rated party to visit a special page to find their ratings.

It is non-obvious to interpret the relative magnitude of the ratings, e.g., the difference between "Highly Trust" and "Trust with Life" may be much greater than the step from "Trust Somewhat" to "Generally trust." To avoid distorting the data, we preserve the order without interpreting the magnitude of the differences and, to the extent possible, we limit ourselves to non-parametric methods and tests to prevent misinterpretations of the ordered scale. We also observe that the rating system suffers from "ceiling" and "floor" problems. That is, a user who gives the highest (or the lowest) rating is already setting the status-giving behavior of the relationship. Fortunately, as our data show, extreme rating values are rare and the distribution of ratings concentrates on the middle of the rating scales, which minimizes this problem. The study of resource scarcity is conducted on a narrower sub-sample of hosting interactions initiated through a "CouchRequest." We impose the requirement that hosts included in this smaller sample live in cities with at least two other hosts, because we aim to compare hosts within the same city.

2.1 Independent Set Sample Construction

Non-independence between dyads is a major concern in the study we present in Section 3. A single host or surfer could generate multiple rating pairs and the ratings could be related in potentially problematic ways for our regression model. Independent sampling in networks is an open question, and there is no established principled way to accomplish this task. However, it can be addressed by extremely 'lossy' procedures, when we can afford to discard a poten-

tially large number of samples. Here, we rely on this technique by computing an approximate maximal independent edge set of the graph containing the 1,446,690 dyads. This results in a dataset consisting of 132,914 edges. In the third robustness check reported in Section 3, this dataset is used to check whether the rating inequality held between the ratings that were chronologically first in the dyad. Results from the same maximal independent set are reported in Section 3, though some observations were dropped due to missing values. Section 4 uses the same procedure to generate an independent set from an original dataset of 20,165 ratings between surfers and hosts, where a CouchRequest could be matched with the interaction, and where hosts had received at least one other CouchRequest prior to the interaction and lived in cities with at least two other hosts, so that z-scores could be computed. The resulting set contains 4,278 interactions in which each host and surfer is assured to have participated only once.

3 Power and Status Giving

We delve into the relationship between power and status-giving by observing the behavior of Couchsurfers within and between dyadic relationships. We first test for the existence of status-giving for balancing unequal power relationships, as described by Emerson (1972b). We compare 1,446,690 pairs of ratings between hosts and surfers, recorded after each hospitality exchange mediated by the organization, collected over nine years, from the early days of the organization in 2003 until the spring of 2012.[2]

The observations of friendship strength show a weak effect of status-giving. In 242,339 (18.76%) rating pairs the surfer's friendship rating was greater than the host's and in 222,006 (17.19%) rating pairs the reverse was true. The corresponding figures for trust ratings exhibit a stronger signal of status-giving: 433,504 (35.38%) and 292,510 (23.87%), respectively. For every situation where friendship ratings indicate status flowing from host to surfer, there are 1.09 instances where the reverse happens. The same figure is 1.48 for trust ratings flowing in the predicted direction for every trust rating in the opposite direction.

Friendship Strength Ratings. Table 1 plots the value of friendship ratings traded between the two exchange partners, within the same dyad. Plotted on the diagonal are rating pairs of equal magnitude: above the diagonal we show instances where surfer-to-host ratings were higher than host-to-surfer ratings, and below the diagonal we count cases where host to surfer ratings were higher. We can compare frequency counts between cells symmetrical to the diagonal. There were, for instance, 1,910 cases where the surfer nominated the host as a friend, and the host responded with a counter nomination as merely an acquaintance. This count exceeds by 508 cases the 1,402 instances in the reverse pair of host to surfer ratings - an imbalance that occurs consistently for pairs of ratings involving low strength designations ranging from "acquaintance" to "best friend."

[2]An extended abstract that discusses preliminary results shown in Tables 1 and 2, using a more restricted dataset, has appeared in State, Abrahao and Cook, 2012.

S2H / H2S	**Acq.**	**CS**	**Friend**	**Good**	**Close**	**Best**	**NA**
Acq.	**987**	12,061	1,910	695	120	53	1,370
CS	9,211	**712,807**	122,450	54,515	7,246	2,810	47,282
Friend	1,402	108,460	**59,516**	24,464	4,120	1,206	9,640
Good	513	50,388	24,932	**39,952**	7,172	1,958	6,089
Close	90	6,988	4,278	7,176	**9,853**	1,559	1,708
Best	63	3,102	1,451	2,127	1,825	**4,233**	1,108
NA	1,845	59,640	13,640	6,988	1,760	846	**3,081**

Table 1: Friendship Strength Rating Inequality in Hospitality Exchanges mediated by CouchSurfing

$t = 14.74, dF = 2,734,165, p < 0.001. \chi^2 = 459264.7, dF = 25, p < 0.001.$

As expected, given the large sample size, both a χ^2 test and a t-test reveal a significant level of association between the matched ratings exchanged within dyads. To further guard against spuriousness, we also implemented a statistical test based on Monte Carlo approximation to the permutation distribution. We observed a share of $s = 16.75\%$ of all friendship rating pairs indicating status flowing from surfer to host (as opposed to 15.35% for the reverse situation, the rest corresponding to equal ratings or to missing values). We obtained an empirical distribution of this statistic by randomly reassigning the host and surfer roles in the dyad 1,000 times. Our results reveal an empirical p-value below 0.1% for the statistic: not once during our simulations did the estimated statistic under random reassignment of roles (mean 16.05%) exceed the value of 16.75%.[3]

Trust Ratings. Hosts and surfers exchange not only public evaluations, such as friendship ratings, but they also produce confidentially collected ratings of each other's perceived trustworthiness, which are unreported to the receiving party. Comparing the frequencies of anonymized public and unreported ratings helps determine to what extent status-giving persists when ratings are not shown to the recipient. In Table 2 we perform a bivariate analysis of the anonymized trust ratings exchanged between host and surfer, using the same conventions as before. In this case the results are much stronger: surfers say they trust hosts more than hosts declare they trust surfers in 433,504 cases, 48% higher than the 292,510 cases in which the reverse happens. As in the case of friendship ratings, statistical tests reveal the observed tendency of surfers to give status to hosts to be significantly different from the null hypothesis.

Missing data. We can observe another form of status-giving in the data in the Not Available (NA) columns of Tables 1 and 2. The host is not only more likely to rate the surfer lower on the friendship and trust scales, but they are also more likely not to give any rating at all. In 176,500 cases hosts did not award any ratings to their partners, a value 25% higher than the 140,556 instances in which surfers neglected to give any ratings. This is consistent with the status-giving mechanism. Given that rating other users on CouchSurfing

[3]Given the large size of our sample, we also performed a more conservative test: estimating the values based on samples of 10,000 interactions each yielded rejections of the null hypothesis of no status- giving from surfer to host in 96% out of 1,000 simulations.

S2H / H2S	**Not**	**Some**	**General**	**Highly**	**Life**	**NA**
Not	**41**	440	1,348	1,009	147	587
Some	493	**22,637**	76,659	54,313	5,641	15,821
General	1,200	64,433	**269,048**	236,744	26,066	45,109
Highly	654	30,326	154,470	**198,448**	31,137	27,036
Life	86	2,942	14,231	23,675	**9,229**	3,987
NA	693	16,818	53,241	42,703	5,640	**9,638**

Table 2: Trust Rating Inequality in Hospitality Exchanges mediated by CouchSurfing.

$t = 155.29, dF = 2,662,258, p < 0.001. X2 = 40612.85, dF = 16, p < 0.001.$

requires at least a few minutes to answer the 9 mandatory questions on the form used on the website, we would expect surfers to be more likely than hosts to spend the time required to rate their partner. Missing data for both friendship strength and trust ratings proves to be more prevalent in host to surfer ratings.

Robustness under Role Reversal We have established so far the existence of a ratings asymmetry between hosts and surfers, which suggests the expected surfer-to-host status flow. An important question concerns the source of this asymmetry. We distinguish between two potential explanations. One readily available hypothesis states that hosts and surfers are selected from two different sets of individuals, the intrinsic characteristics of which make hosts more likely to receive and surfers to give status in any circumstance. For instance, perhaps the asymmetries we observe are due to hosts being older than surfers, or better read, of a different social class. According to this hypothesis it is the individual characteristics enumerated above as well as many others we may have overlooked that generate our results. Or, perhaps differences in psychometric scales account for the observed ratings asymmetry. Perhaps users who tend to be hosts assign different meanings to ratings like generally trust than surfers. The alternative explanation is that hosts are more likely to receive status not because of who they are, but because of the nature of the hosting role. Following Power-Dependence Theory we contend that the asymmetries we observe in trust and friendship ratings are driven by the role-specific power differentials between host and surfer.

A principled test of these two alternative explanations would require posing a counter-factual question: what would the rating process look like absent the power imbalance? Assuming individual characteristics completely accounted for differences in status-giving, we would expect the rating process to be independent of the resource exchange: just as much status should flow from one individual to another regardless of their roles in the hospitality exchange. The starkest contrast between the two theories would arguably involve a comparison in which the roles are reversed where the prior surfer is now the host and prior host the surfer.

Theoretical and empirical considerations prevent us from pursuing either test, however. Once two individuals have interacted as host and surfer, they are no longer strangers to each other. Obtaining ratings from a *de novo* interaction for the same dyad, in which no resource exchange takes place, or where roles are reversed is unfortunately a logical impossibility, while a randomized experiment assigning individuals to the hosting and surfing roles by chance would be unfeasible. Despite these difficulties, a series of tests for the influence of individual differences is nonetheless possible. We present four such tests.

S2H / H2S	**Acq.**	**CS**	**Friend**	**Good**	**Close**	**Best**	**NA**
Acq.	**0**	7	1	0	0	0	0
CS	2	**199**	50	15	4	1	4
Friend	0	25	**19**	8	0	0	0
Good	0	7	7	**12**	4	0	0
Close	0	2	1	0	**2**	0	0
Best	0	0	0	1	0	**2**	0
NA	0	2	2	2	0	0	**0**

Table 3: Friendship. Role-reversed Host and Surfer.

"Role-reversal" defined as having performed only the opposite role (at least three times) prior to the focal experience (from which the rating is drawn).
t = 2.87, df = 741.5, p = 0.004. $\chi^2 = 265.4, df = 25, p < 0.001$.

S2H / H2S	**Not**	**Some**	**General**	**Highly**	**Life**	**NA**
Not	**0**	0	0	0	0	0
Some	0	**3**	19	18	1	2
General	1	18	**92**	60	9	6
Highly	1	4	43	**66**	6	2
Life	0	2	0	5	**2**	0
NA	0	2	10	5	1	**1**

Table 4: Trust. Role-reversed Host and Surfer.

χ^2 = 29.03, dF = 12, p = 0.004. t = 3.06, dF = 725.354, p = 0.002

Role-Reversed Users. Given the size of our dataset, it is possible to find those interactions where both host and surfer play their current roles for the first time, after only having played the complementary role during prior exchanges with each other. We set a cutoff of three minimum prior experiences with the other role: our analysis looks at those encounters between a first-time host who has surfed at least three times before, and a first-time surfer who has been a host at least three times previously. We chose the threshold of three prior encounters as the highest threshold for which a reasonably large set of observations could be gathered, although the findings are qualitatively the same for lower thresholds.

Due to the rarity of such encounters, these observations of role reversal represent a small subset of all interactions, consisting of only 379 instances out of millions, shown in Tables 3 and 4. Nonetheless, the sample is large enough to allow us to draw the same conclusion as before: status flows from the current surfer to the current host. In our sample 90 friendship rating pairs are higher from surfer to host, compared to 45 that are higher from host to surfer. A similar dynamic holds for trust, which flows in the hypothesized direction in 113

S2H H2S	Acq.	CS	Friend	Good	Close	Best	NA
Acq.	**180**	2,305	402	147	23	3	69
CS	1,674	**134,853**	29,620	11,254	1,493	486	2,292
Friend	258	24,337	**16,724**	6,246	1,029	242	479
Good	102	10,478	6,521	**8,907**	1,532	359	285
Close	9	1,331	1,081	1,553	**2,016**	276	55
Best	14	634	351	455	331	**601**	32
NA	63	2,038	486	219	36	8	**55**

Table 5: Friendship. Experienced Host and Surfer.

"Experienced:" having both surfed and hosted at least once prior to the focal experience (from which the rating is drawn). Each interaction involves only individuals who had experience with both hosting and surfing before their current encounter. t = 2.9, df = 741.5, p = 0.004. $\chi^2 = 265.4, df = 25, p < 0.001$.

S2H H2S	Not	Some	General	Highly	Life	NA
Not	**15**	130	312	203	28	65
Some	132	**6,092**	18,998	11,607	1,149	1,749
General	303	15,971	**60,744**	46,489	4,776	4,459
Highly	132	6,651	31,511	**35,034**	5,053	1,964
Life	23	711	3,008	4,280	**1,182**	244
NA	61	1,795	5,148	3,095	335	**494**

Table 6: Trust. Experienced Host and Surfer.

χ^2 = 8986.1, df = 25, p $<$ 0.001. t = 63.0, df = 527,978, p $<$ 0.001.

cases compared to 74 in the opposite direction. Even with this very small sample, the difference is statistically significant, as revealed by a t-test, a chi-square test and the empirical, simulation-based hypothesis test described previously. We observe here that status flows to a greater extent from the low power surfer to the high power host, even in the rare occasion when both interaction partners are repeat surfers and hosts who are switching roles for the first time. We conclude that the roles of surfer and host and their associated resource constraints are most likely to explain status-giving, rather than the personal characteristics of individuals who either host or surf in these encounters.

Experienced Users. Individuals who have played both roles are expected to be more alike in their unobserved characteristics. A simple restriction stipulates that both the current surfer and the current host have engaged in hosting and surfing at least once before. Were the differences in rating behavior explained by individual characteristics that condition individuals' access to resources, we would expect such differences to disappear with this control.

We obtained 273,943 interactions (out of 1,446,690 total host-surfer interactions mediated by CouchSurfing) when this specification both host and surfer having hosted and surfed at least once before was imposed on the data. We show the results in Tables 5 and 6. There were 55,417 cases when the surfer gave higher friendship ratings to the host, compared to 49,128 cases when the host gave higher friendship ratings to the surfer. The analogous figures for trust ratings were 88,745 and 62,722.

S2H H2S	Acq.	CS	Friend	Good	Close	Best	NA
Acq.	**6**	62	6	2	0	0	0
CS	38	**2,516**	427	230	33	8	58
Friend	6	315	**189**	89	11	4	11
Good	1	147	76	**142**	27	10	5
Close	0	25	12	19	**39**	3	1
Best	0	6	5	4	2	**12**	0
NA	2	73	20	7	1	1	**3**

Table 7: Host and Surfer's friendship rating, where host has only been a surfer, and surfer has only been a host.

χ^2 = 1802, df = 36, p $<$ 0.001. t = 4.6, df = 9,099, p $<$ 0.001.

S2H H2S	Not	Some	General	Highly	Life	NA
Not	**1**	0	7	1	0	0
Some	3	**57**	249	187	13	23
General	8	224	**964**	834	90	86
Highly	3	92	490	**739**	115	41
Life	1	13	43	70	**27**	6
NA	0	38	124	83	6	**16**

Table 8: Host and Surfer's Trust rating, where host has only been a surfer, and surfer has only been a host.

χ^2 = 196.94, df = 25, p $<$ 0.001. t = 9.7, df = 8865.6, p $<$ 0.001.

Experienced Users Playing Current Role for the First Time. As a complement to the "Role-reversal" robustness check, we examined those interactions where both host and surfer played their current roles for the first time, after having played solely the complementary role during prior exchanges with other users. For instance, current host H has only surfed before, and current surfer S has only hosted before. This is again a more restricted subset of all interactions, only 4,654 instances. We show the bivariate distribution of ratings in Tables 7 and 8. As plots show, we draw the same conclusion as before: status flows from the current surfer to the current host, whether measured through friendship or trust ratings. Friend ratings given by surfers are higher than hosts' in 912 cases, compared to 656 cases when the reverse happens; for trust ratings the figures are 1,496 and 947.

Erasing anchoring effects. Some of the ratings presented in Tables 1 and 2 may be affected by anchoring effects (Bialski and Batorski, 2010), i.e., the first actor to give a rating in the dyad influences the rating of the second. To examine how the rating exchange sequence unfolds and to eliminate the possible effects of anchoring, we complement the within-dyad bivariate analyses presented previously with a between-dyads comparison. We compile a sample containing only one rating from each dyad. Specifically, we consider the rating given in a dyad by the party who initiated the interaction (i.e., the rating process) chronologically. Surfers were first to give ratings in 58% of cases, whereas hosts were first in 42% of the 993,802 cases for which the initiator

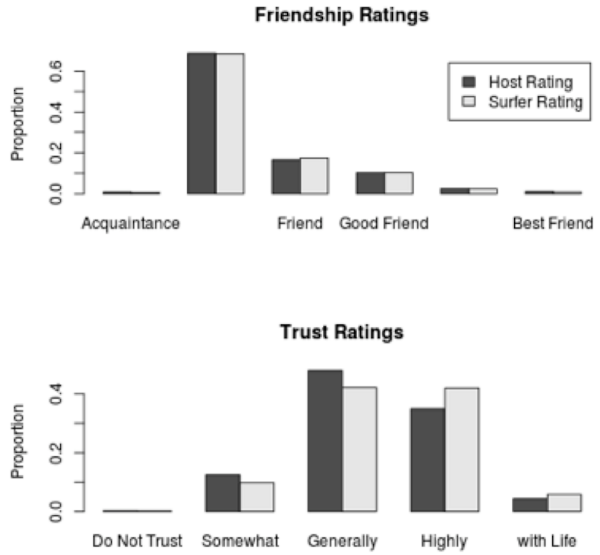

Figure 1: Ratings distribution for who is first to rate.

Covariate	Value	S.E.	T-statistic
A. Friendship Rating			
Surfer was First Rater	0.059	0.012	4.901
B. Trust Rating			
Surfer was First Rater	0.335	0.010	32.040

Table 9: Ordered logistic regressions. Response: Rating Given by First Rater (Chronologically).

Model A: Resid. dev.: 251897.05. AIC: 251909.05.
Model B: Resid. dev.: 300622.59. AIC: 300632.59.
All coef. significant at .001 level.
Sample size 993,802 interactions for which first rater (chronologically) could be accurately determined.

of the interaction could be determined accurately. Based on Power-Dependence Theory we predict that surfers are likely to hold hosts in higher esteem than is the case in reverse, which leads us to expect first ratings given by surfers to be higher than first ratings given by hosts.

The results confirm our suppositions. When surfers rate first and thus "set the tone" for the rating exchange, they will give higher ratings than when hosts first avail themselves of the opportunity to initiate the rating process. For trust ratings this conclusion is readily apparent from visual inspection (Figure 1). When comparing the distribution of hosts' and surfers' trust ratings, our findings show a rightward-shift for surfers, who give higher ratings. Indeed, this conclusion is confirmed by a one-sided Mann-Whitney U Test (Lehman and D'Abrera, 2006), which shows the shift to the right is statistically significant ($p < .001$). The conclusion holds for the trust ratings and for friendship ratings in an ordinal logistic regression (Table 9).

Covariate	Value	S.E.	T-statistic
A. Friendship Rating Inequality			
Surfer was First Rater	0.167	0.011	14.51
B. Trust Rating Inequality			
Surfer was First Rater	0.553	0.011	52.34
C. Friendship Rating Inequality, controlling for Trust			
Surfer was First Rater	0.052	0.012	4.42
Trust Rating Inequality (reference: equal ratings)			
. . . first rating > second	0.465	0.014	32.90
. . . second rating > first	-0.465	0.014	-32.95
D. Trust Rating Inequality, controlling for Friendship			
Surfer was First Rater	0.538	0.011	50.70
Friendship Rating Inequality (reference: equal ratings)			
. . . first rating > second	0.526	0.014	37.03
. . . second rating > first	-0.546	0.014	-38.26

Table 10: Ordered logistic regressions. Response: Rating given by first rater (randomly-assigned in each dyad) exceeds rating given by second rater.

Model A: Dev.: 229965.54. AIC: 229971.54. N = 129,051.
Model B: Dev.: 268944.31. AIC: 268950.31. N = 125,228.
Model C: Dev.: 220207.84. AIC: 220217.84. N = 125,091.
Model D: Dev.: 265031.81. AIC: 265041.81. N = 125,091.
All coef. significant at .001 level.
Each CouchSurfing user participates in at most one rating dyad.

3.1 Friendship and Trust Rating Inequalities

Our findings suggest that asymmetries in trust ratings are a stronger indicator of status-giving than discrepancies in friendship ratings. In Tables 1 and 2, the frequency of friendship ratings where status flows from surfer to host is 18.76%, only 1.57% greater than the frequency of friendship ratings in which status flows in the opposite direction; the spread is much larger (11.51%) for trust ratings.

To better investigate the relationship between asymmetries in friendship and trust ratings in the status-giving process we devised the following experiment. Users in each dyad investigated in Section 3 (Tables 1 and 2) were ordered at random, and ratings were coded as 1 whenever the first (by random assignment) user's rating was higher than the second user's, as 0 when the ratings were equal, and as -1 when the first user's rating was lower than the second user's. The coding schema used expresses the idea of "status flow:" 1 indicates status flowing to the second user, and -1 indicates status flowing to the first.

Our primary hypothesis (already confirmed) holds that status should flow in the opposite direction (from surfer to host) of the flow of hosting (from host to surfer). This idea is expressed in Table 10, Models A and B, where we regress status flow, measured using friendship and trust ratings, on the direction of the exchange flow. The models use an independent set of the ratings' network edges (Boppana and Halldorsson, 1992; Hagberg, Schulte and Swart, 2008), where no two edges share a user in common (see Section 2 for details). Our prediction is again confirmed. In Model A there is a .167 increase in the log-odds ratio of status-flow - measured through friendship - towards the second user when

the second user hosted the first. Conversely, Model B shows a .553 log-odds increase when trust ratings are used as a measure of status flow. Model C estimates the effect of hospitality flow on the friendship rating inequality net of the variance explained by trust ratings. The model shows a robust effect for hospitality flow on the friendship ratings inequality, net of including this covariate. Model D leads us to the same conclusion with respect to trust ratings.

4 Scarcity

The preceding studies show the existence of a status-giving effect in which hosts are rated higher by surfers than vice-versa. Power-Dependence Theory posits a more complex relationship between power and status, however. The "scarcer" the host's hospitality resource, the higher the valuation we expect surfers to place on it, and, therefore, the more status we expect surfers to give their hosts (Emerson, 1962).

The scarcity of a hospitality resource may be conceived in two ways. One approach is to consider the host's own popularity and inclination to host, relative to other hosts. Here we expect that more popular hosts receive more status, whereas hosts who tend to accept many requests (compared to the prevailing local norm) would be given less status. The second possible interpretation of scarcity deals with the availability of alternatives. Power-Dependence Theory predicts that a surfer in a city that is in high demand but where hosts, as a rule, accept few requests (e.g., Paris) should value the hospitality they receive more than if they had surfed in a low demand, high acceptance city. Thus, we expect status given by the surfer to be directly correlated with the number of requests received by hosts in the city, but inversely correlated with the number of accepted requests.

To test this prediction we use a mechanism implemented by CouchSurfing starting with April 2010. From that point onward, surfers could issue hospitality requests to potential hosts through a standardized message, a CouchRequest. As a result, it is possible to observe whether or not a host agreed to a surfer's request and to compute each host's likelihood of accepting such a request. The focal measures in this study represent 4,278 post-interaction rating exchanges that could be matched with the pre-interaction CouchRequests sent by surfer to host, under the requirement that no user participate in more than one rating exchange. Our sample was constrained to include only those CouchRequests sent starting in July, 2010 as we computed indicators of each host's prior responses during the 90-day interval leading up to the moment when the request was sent. We only included in our sample those hosts living in a city with at least two other active hosts during the focal interval. These records were fully anonymized prior to our access to them. Results are presented in Table 11 using ordered logistic regression, with a response coded as 1 when the surfer's rating exceeds the host's rating, 0 when the ratings are equal and -1 when the host's rating exceeds the surfer's.

We first examine results for the friendship rating inequality in Table 11 (Model A). To test our hypotheses regarding the direct valuation of the host's hospitality resource, we include in the regressions the log-transformed prior number of requests received by the host during the 90 days before the request leading to the focal interaction. The magnitude of this effect (-0.004) is not statistically different from 0, however. We likewise include the log-transformed number of prior requests the host accepted, yielding an average effect on the log-odds of -0.013, for each unit increase in the log-transformed number of previously accepted requests. Our predictions regarding the availability of alternatives receive a test through the inclusion of the log-transformed total number of received and accepted requests in the host's city, by hosts other than the rated individual. Popular cities, where hosts receive a lot of requests report a positive effect (in the direction of surfer-to-host status-giving) of magnitude 0.074. The effect for cities where hosts tend to accept a lot of requests (-0.059) is in the expected direction, but is not statistically significant (t = 1.522).

Covariate	Value	S.E.	T-statistic
A. Friendship Ratings			
Host City Requests			
. . . Mean Received	0.074*	0.043	1.733
. . . Mean Accepted	-0.059	0.039	-1.522
Host Requests			
. . . Received	-0.004	0.006	-0.650
. . . Accepted	-0.013***	0.039	-3.020
B. Trust Ratings			
Host City Requests			
. . . Mean Received	-0.050	0.039	-1.280
. . . Mean Accepted	-0.033	0.034	-0.980
Host Requests			
. . . Received	-0.008	0.006	1.520
. . . Accepted	-0.004	0.004	-1.105

Table 11: Surfer-Host Rating Inequality (ordered logit).

DV coded as -1 when first rating in the dyad is lower than the second rating in the dyad, 0 when the ratings are equal, and 1 when the first rating exceeds the second rating. Each user participates in at most one dyad. All covariates express z-scores. City-level statistics exclude focal hosts requests.

Model A: Resid. Dev.: 7433.659, AIC: 7445.659. N = 4,278.

Model B: Resid. Dev.: 9166.157, AIC: 9178.157. N = 4,184.

*: $p < .10$, **: $p < .05$, ***: $p < .001$. Two-sided tests.

When considering the effect of scarcity on the trust rating asymmetry (Table 11, Model B), we note no statistically significant effects associated with the proposed measures. This result may be due to the model's lack of statistical power. The results are also suggestive of an alternative explanation. As previously discussed, we expect the status-giving to be mostly internalized by the individual through the mechanism of gratitude. But scarcity may engender a certain degree of calculation. Remembering how hard it was to get a couch in, say, London may lead the surfer to feel not only gratitude but also a slight degree of obligation to their host and thus lead them to consciously increase their friendship rating (which the host can see), without feeling any need to modify their

private trust rating that, while not revealed to the surfer, is nonetheless likely to be greater than the host's rating.

5 Conclusion

We have documented several ways in which status flows from dependent surfers to more powerful hosts in the context of hospitality exchanges mediated by CouchSurfing. We used the asymmetry of public friendship ratings and private trust ratings (in aggregated, anonymized data) as indicators of status- giving. Both sets of ratings appear to reflect status flowing from hosts to surfers. The ratings are influenced by a process that is non-malicious and based in genuine user intent. Nonetheless, this process is likely to bias conclusions that may result from taking the ratings at face value.

The large size of the dataset we investigated allowed us to perform a comprehensive analysis through robustness checks. We showed that even when the current surfer has a history of only being a host, and when the current host has the opposite history, of only being a surfer, status still flows from current surfer to current host. We see this finding as providing strong support for status-giving as a power-balancing mechanism.

Our findings also bolster to some extent the assumed relationship between resource scarcity and power-imbalance: hosts who accept more requests relative to other hosts in the same city receive less status as do hosts living in cities in which demand is higher and in which a request is less likely to be accepted as a result. The effects are small and do not seem to be reflected by the asymmetry in private trust ratings. This is suggestive of a potential relationship between scarcity and the occurrence of status-giving through calculated obligation rather than gratitude. We advance this only as a tentative explanation: scarcity effects are likely to be subtle and in need of deeper investigation.

In this study we have used social theory to illuminate fundamental structural features of social processes in exchanges mediated by the Internet to which we believe data mining cannot be indifferent. As a dividend of this approach, we provide the first test of Emersons theory on a scale not addressed in previous work. In fact, little has been done in more than fifty years to test some of the key tenets of Emerson's theory beyond controlled laboratory experiments. This, we believe, is the kind of fruitful intersection, where social science and computer science can work together to reveal new insights using large datasets not previously available for investigation of currently untested theories as well as new knowledge generation.

References

Bialski, P. and Batorski, D. 2010. From online familiarity to offline trust. In *Social Computing*, Taylor and Francis Group.

Blau, P. 1964. *Exchange and Power in Social Life.* New Brunswick, NJ: Transaction Publishers.

Boppana, R. and Hallddorsson, M. 1992. Approximating maximum independent set by excluding subgraphs. *BIT Numerical Mathematics*. 32: 180–196.

Danescu-Niculescu-Mizil, C., Pang, B., Lee, L., and Kleinberg, J. 2012. Echoes of power: Language effects and power differences in social interaction. In *Proc. of the ACM Int'l Conf on WWW*.

Emerson, R.M. 1962. Power-dependence relations. *Am. Sociol. Rev.*, 27(1):31–41.

Emerson, R.M. 1972a. Exchange theory, Part I: A psychological basis for social exchange. In *Sociol. Theories in Progress.*, Boston, MA: Houghton Mifflin.

Emerson, R.M. 1972b. Exchange Theory Part II: Exchange Relations and Network Structures. In *Sociol. Theories in Progress.*, Boston, MA.: Houghton Mifflin.

Galinsky, A.D., Magee, J.C., Gruenfeld, D.H., Whitson, J.A., and Liljenquist, K.A. 2008. Power reduces the press of the situation. *J. Pers. Soc. Psychol.*, 95(6).

Goei, R. and Booster, F.J. 2005. The roles of obligation and gratitude in explaining the effect of favors on compliance. *Communication Monographs*, 78(4):284–300.

Hagberg, A.A., Schult, D.A., and Swart, P. J. Exploring network structure, dynamics, and function using NetworkX. In *Proc. 7th Python in Science Conference.*, 11-15.

Horan, S. M. and Booth-Butterfield, M. 2010. Investing in affection: An investigation of affection exchange theory and relational qualities. *Communication Quarterly*, 58(4): 394–413.

Kleinberg, J. and Tardos, E. 2008. Balanced Outcomes in Social Exchange Networks. In *Proc. 40th ACM Symposium on Theory of Computing*.

Lampe, C. and Garrett, R. K. 2007. Its All News to Me: The Effect of Instruments on Ratings Provision. In *Proc. HICSS*. 180b. Washington, DC: IEEE Computer Society.

Lauterbach, D., Truong, H., Shah, T., and Adamic, L. 2009. Surfing a Web of Trust. In *Proc. IEEE ICCSE*. Vancouver, BC.

Lehmann, E. L. and D'Abrera, H. 2006. Nonparametrics: statistical methods based on ranks. New York: Springer.

Marx, K. and Engels, F. 1965. *The German Ideology*. London: Lawrence & Wishart.

Massa, P. and Avesani, P. 2007. Trust-aware recommender systems. In *Proc. of the ACM RecSys*. 17-24.

Moody, J. 2001. Race, School Integration, and Friendship Segregation in America. *Am. J. Sociol.*, 107(3): 679–716.

Nuzzi, O. 2014. More Bad News for Uber: Driver Arrested in Los Angeles Rape Case. *The Daily Beast*.

Osborne, M. J and Rubinstein, A. 1994. *A Course in Game Theory.* Cambridge, Mass.: The MIT Press.

Parsons, T. 1969. *Politics and Social Structure.* New York, New York: The Free Press.

Peterson, R. 1997. The Rise and Fall of Highbrow Snobbery as a Status Marker. *Poetics*, 25(2-3): 75-92.

Ridgeway, C.L., Boyle, E.H., Kuipers, K.J., and Robinson, D.T. 1998. How do status beliefs develop? *Am. Sociol. Rev.*, 63: 331.

State, B., Abrahao, B., and Cook, K. 2012. From Power to Status in Online Social Exchange. In *Proc. ACM WebSci* Evanston, IL.

Stein, J. 2015. Tales from the Sharing Economy. *Time* 185(4).

Teng, C., Lauterbach, D., and Adamic, L.. 2010. I rate you, you rate me? Should we do so publicly? In *Proc. of the 3rd USENIX Workshop on Online Social Networks*, Berkeley, CA.

Weber, M. 1978. *Economy and Society*. Berkeley, Calif: University of California Press.

Weinstein, R. and Ryan, R. 2010. When helping helps: Autonomous motivation for prosocial behavior and its influence on well-being for the helper and recipient. *J. Pers. Sociol. Psychol.*, 98(2): 222–244.

Willer, R., Youngreen, R., Troyer, L., and Lovaglia, M. J. 2012. How Do the Powerful Attain Status? The Roots of Legitimate Power. *Managerial and Decision Economics*. 33(5-6).

Zervas, G., Proserpio, D., and Byers, J. 2015. A First Look at Online Reputation on Airbnb, Where Every Stay is Above Average. *SSRN*. Available at `http://ssrn.com/abstract=2554500`.