IRLE
Institute for Research on Labor and Employment

IRLE WORKING PAPER
#141-13
February 2013

How Reputation Transfers Across Categorical Boundaries

Ming Leung and Weiyi Ng

Cite as: Ming Leung and Weiyi Ng. (2013). "How Reputation Transfers Across Categorical Boundaries." IRLE Working Paper No. 141-13. http://irle.berkeley.edu/workingpapers/141-13.pdf

Title:
How reputation transfers across categorical boundaries

Author:
Leung, Ming D., University of California, Berkeley
Ng, Weiyi, University of California, Berkeley

Publication Date:
02-14-2013

Series:
Working Paper Series

Publication Info:
Working Paper Series, Institute for Research on Labor and Employment, UC Berkeley

Permalink:
http://escholarship.org/uc/item/1r14400x

Keywords:
Categorization, Reputation, Online Markets

Abstract:
If categorical boundaries serve to assist audiences in lumping together like organizations and separating different ones, then does an organization's reputation also remain similarly bounded? We conceptualize reputation as an assessment of an actor's ability or quality, but in the circumscribed domain of an identified category. We hypothesize that reputation acts differentlydepending both on the relevance and focus (read: distance and spanning) of a marketparticipant's past experiences. We examine an online market for freelancing services. We find that better reputations benefit those sellers who have more relevant experiences with regard to a focal category. However, for those with poor reputations, this disadvantage is ameliorated as distance from the focal category is increased. For those sellers with better reputations, spanningacts detrimentally on their reputational advantage. Yet for those with poor reputations, spanning decreases their disadvantage vis-à-vis their better reputational alters. We also employ a novelmeasure of categorical distance using a text overlap measure.

How reputation transfers across categorical boundaries [1]

Ming D. Leung
Haas School of Business
University of California, Berkeley
2220 Piedmont Avenue
Berkeley, CA 94720
1-510-642-4952
mingdleung@haas.berkeley.edu

Weiyi Ng
Haas School of Business
University of California, Berkeley
2220 Piedmont Avenue
Berkeley, CA 94720
weiyi_ng@haas.berkeley.edu

DRAFT WORKING PAPER

FEBRUARY 15, 2013

[1] Helpful comments on earlier drafts from Huggy Rao, Mike Hannan, Jerker Denrell, Giacomo Negro and the Stanford Business School Macro Lunch participants improved this paper. We wish to thank Fabio Rosati, Ved Sinha, and James Lee of Elance for their generous help in allowing us to use their data. Of course, all errors remain our own.

ABSTRACT

If categorical boundaries serve to assist audiences in lumping together like organizations and separating different ones, then does an organization's reputation also remain similarly bounded? We conceptualize reputation as an assessment of an actor's ability or quality, but in the circumscribed domain of an identified category. We hypothesize that reputation acts differently depending both on the *relevance* and *focus* (read: *distance* and *spanning*) of a market participant's past experiences. We examine an online market for freelancing services. We find that better reputations benefit those sellers who have more relevant experiences with regard to a focal category. However, for those with poor reputations, this disadvantage is ameliorated as distance from the focal category is increased. For those sellers with better reputations, spanning acts detrimentally on their reputational advantage. Yet for those with poor reputations, spanning decreases their disadvantage vis-à-vis their better reputational alters. We also employ a novel measure of categorical distance using a text overlap measure.

KEYWORDS
Categorization, Reputation, Online Markets

INTRODUCTION

A fundamental challenge to buyers in markets is to gain insight as to the underlying ability of a potential transaction partner. Scholars who study the impact of classification systems on markets identify how demonstrating characteristics what lie within the socially recognizable confines of a particular category can demonstrate to naïve buyers a seller's abilities in that arena; thereby increasing their likelihood of securing additional work in that domain (Zuckerman et al 2003). In this sense categories act to develop expectations for what its members will possess. Those with demonstrated success in one arena can be expected to perform well again in that arena. Yet this also implies assumptions of inability in other, disparate, domains (Hsu 2006; Hannan et al 2007).

Yet another prominent mechanism by which market participants can usefully infer seller ability is through their reputation. A reputation is a publically recognized indication of skill or ability and is derived from quality assessments of previous experiences. Reputation therefore acts as a proxy for indiscernible underlying quality. Following the literature on signaling (Spence 1973), whereby social actors rely on proxies or past observations to serve as a basis for judgment of a candidate's future performance, a market participant's reputation can act to bolster or undermine their future opportunities (Rao 1994; Fombrun and Shanley 1990; Roberts and Dowling 2002). For example, Rao (1994) has shown that greater reputation in the form of winning certification contests increases the survival odds of car manufacturers. Thus, reputation serves as an indicator of the fitness of a producer. In online markets, reputation has been shown to garner a higher price as well as a greater likelihood of transaction closure (Resnick and Zeckhauser 2002; Resnick et al 2006).

If categorical boundaries serve to assist audiences in lumping together like organizations and separating different ones, then does an organization's reputation also remain similarly bounded? In other words, if an organization is particularly adept at addressing a specific categorical niche, does this enhanced reputation bolster or undermine their opportunities in another one? Current work on categorization or reputation has not formally addressed this issue. However, because reputation has been identified as an intangible resource (Rao 1994; Fombrun 1990; Roberts and Dowling 2002) it should also be considered to have a diffuse quality. Fombrun (1990: 72) describes reputation as a measure of, "the firm's overall appeal." If this is so, an organization's positive accolades in one domain should be recognizable and positively influence evaluation of ability in another.

However, the principle of allocation, originating with the organizational ecology literature, (Dobrev, Kim, and Hannan 2001; Hsu 2006) may suggest the opposite effect. This theory states that actors have a fixed ability to concentrate their efforts. As the breadth of an actor's focus widens, they necessarily sacrifice in-depth knowledge in a particular area. So, if a social actor has accrued substantial experience in a particular niche, then they necessarily would have forsaken developing expertise in another. If a better reputation is at least partially conceived from extensive focus in an area, then this may suggest great reputation in one area may actually be a detriment to future opportunities in another. How can we reconcile these two, possibly opposing predictions?

We follow Jensen and his colleagues' (2012) lead by conceptualizing reputation as a role-specific attribute. That is we focus on reputation, though as a socially bounded measure of quality. Specifically, we conceptualize reputation as a measure of the expected quality of future actions of a particular actor, but for a particular socially defined set of behaviors. In this case, we leverage the work on categories, which act to circumscribe actions into recognized groupings, in a market context. We therefore conceptualize reputation as an assessment of an actor's ability or quality, but in the circumscribed domain of an identified category.

Taking this as a starting point, we hypothesize that reputation acts differently depending both on the *relevance* and *focus* (read: *distance* and *spanning*) of a market participant's past experiences. In particular, we first demonstrate the uncontroversial finding that better reputations benefit those sellers who have more relevant experiences with regard to a focal category. However, for those with poor reputations, this disadvantage is ameliorated as distance from the focal category is increased – suggesting an asymmetric effect of reputation on appeal. A similar counterintuitive effect is revealed for reputation and spanning. For those sellers with better reputations, spanning acts detrimentally on their reputational advantage, with increased spanning leading to a reduced (yet still positive) effect of a good reputation. Yet for those who are relegated to a poor reputational position, spanning decreases their disadvantage vis-à-vis their better reputational alters.

This study is notable for several reasons. First, we demonstrate how a particular, ostensibly objective, measure of reputation, can be viewed differently depending on the prevailing classificatory structure that market participants are beholden to. Unless a social actor's ability fits neatly within the confines of how the category structure has been defined, they risk devaluation due to their inability to adequately serve a recognized niche. Yet this disadvantage is asymmetric between those of good versus bad reputations. We also relax the assumption implied to date by suggesting that distinctive categories can be more or less related to one another (Kovacs and Hannan 2012). We also contribute to understanding as to how people can move beyond where they begin (Zuckerman et al 2003). In particular point out that forces towards specialism, at least in an employee's early career, are driven by lack of other sufficient signals to underlying quality. The intuition being that alternative signals such as reputation or credentials should alleviate concerns of generalist behavior being coded as lacking any particular skills. This paper extends on this conjecture by examining just how ones reputation acts to overcome generalist versus specialist positional outcomes.

We test these predictions in an online marketplace for services, www.elance.com. The online environment is particularly suited to testing effects of reputation, as website users are constantly evaluated from previous transactions. These evaluations are prominently displayed and stored by the category in which a job was completed.

CATEGORICAL DISTINCTIONS

The quickly growing literature on classification schemes in markets (Zuckerman, 1999; Rao, Monin, & Durand, 2003; Hannan, Polos, & Carroll, 2007) has identified audience derived categories as socially important boundaries circumscribing organizational action. A category is a socially agreed upon grouping of like-items attached to expectations for characteristics and/or behavior of those items. For example, Zuckerman's (1999) seminal study demonstrated that companies which recognized by finance analysts as worthy comparable organizations of an industry resulted in high stock prices. On the other hand, organizations which were unable or unwilling to display appropriate category specific behaviors were ignored by stock analysts and consequently discounted by the stock market. This study highlights two valuable components to this paradigm. First, category boundaries limit expectations to organizational action and second, external audiences (not the organization themselves) determine an organization's fitness in a category.

As social mechanisms, categories usefully lump and separate social actors into recognizable groupings (Zerbuval 1997). This implies that members of one category differ on some dimensions to those of another. In labor markets, these distinctions usefully partition jobs into recognizable distributions of skill. Take, for example, the market for film actors. As Zuckerman and colleagues (2003) identified, categorical distinctions based on the genre of film an actor has worked in necessarily suggests their expertise or abilities in a given niche, "To the extent that employers believe that labor-market categories represent distinct skills, experience in one category will be regarded as *prima facie* evidence that the candidate does not have the necessary skills to participate in another category," (Zuckerman et al, 2003: 1027). This is because is categorical distinctions are recognized, and accordingly, serve to bound related or identical sets of skills and exclude those unrelated.

To the extent categories usefully identify differentiated skills would suggests that those social actors who attempt to span several categories would suffer a discount. For example, Hsu (2006) has shown that movies which do not fit neatly into a recognizable film genre are less appealing. This is because attempts to span, or be a member of, several categories simultaneously is either seen as confusing to an audience or difficult to do well. In particular, the principle of allocation (suggests that if actors were to have a fixed amount of ability to focus on a particular niche, so attempts at focusing across several niches necessarily implies lower ability in any particular one.

Categorical digressions are perceived by audiences as a reflection of a candidate's skills or abilities *a priori* (Negro and Leung 2012). This is because audiences generally hold expectations as to what appropriate behaviors or characteristics are for candidates who purport to be classified in a particular area. For example, industrial beer brewers are not expected to be able to brew high quality microbrews. Despite the fact that in reality they could, beer enthusiasts refused to accept their legitimacy (Carroll and Swaminathan 2000). If evaluation of the abilities of a social actor were tainted by their categorical membership, then does more deliberate and direct evidence of their actual abilities transcend these preconceived notions of disadvantage?

REPUTATION

The literature on reputation conceptualizes the term as a perception of an organization's past actions and acts as a signal of their future abilities (Rao 1994; Fombrun 1996; Podolny and Phillips 1996). In essence it is a measure of the expected capability derived from past displays. It acts as a signal (Spence 1973) to would be buyers as to what they may expect in terms of quality. The absence of knowledge regarding an applicant's skill induces an audience to attend to informational cues to decipher a candidate's abilities. Reputation can be considered a cue that market participants use to consider a candidate's fitness.

A good reputation therefore should be a positive asset to a firm's future prospects. Rao (1994) conceptualizes reputation of nascent auto manufacturers as the legitimacy received from winning certification contests. He demonstrates how this improves the organizations chances of survival. Roberts and Dowling (2002) identify how a good reputation can help an organization sustain superior profits over time. In online environment, where uncertainty of a transaction partner may be high due to the remote nature of the market, good reputations helped e-bay sellers increase the price they can charge (Resnick et al 2006) and close auctions more easily (Resnick and Zeckhauser 2002).

Reputations are constructed from several disparate domains and have been considered diffuse in nature. Scholars of reputation have generally agreed that reputation is best defined in terms of ‘overall assessments’ or even ‘generalized favorability’ (Lange et al 2011). Because an actor’s reputation is a perception of their quality, it should influence evaluations across diverse fields. For example, those with better reputations are more likely to associate with others who are considered to have a high reputation (Stuart et al 1999; Podolny and Phillips 1996). Better reputations also may induce potential buyers to pay more for good or services because of the expectation of higher quality (Shapiro 1983). Therefore, our first order prediction is:

Hypothesis 1: Greater overall reputation increases appeal.

REPUTATION AND CATEGORY RELATEDNESS

Yet the literature on categories suggests that not all previous experiences should be perceived as equally relevant. The literature to date suggests that exemplary performance in a particular arena suggests underlying ability in that area, but not necessarily in others (Zuckerman et al 2003). Yet, what this stream of literature has not taken into account is the relatedness between differently categorized experiences. Previous literature has either examined the oppositional nature of categories, thereby precluding any overlap in (identity) relatedness (Carroll and Swaminathan 2000; Rao et al 2003, 2005). Or assumed relatedness between similarly organized companies as reflected in the coverage overlap of analysts (Zuckerman 1999). Though some earlier work has implied that boundaries between disparate categories can erode through either repeated attempts at disparate combinations or high status adoption of oppositional identities (Rao et al 2005), only some nascent work has begun to examine how categories can be more or less related to one another, (Kovacs and Hannan 2012).

Therefore, we relax the assumption implied to date by suggesting that distinctive categories can be more or less related to one another. By related we refer to the similarities that may exist between them. For example, in their examination of movie actors and their ability to work between genres, Zuckerman and colleagues (2003) identified how the cleave between comedy and drama genres was particularly sharp, thereby noting that those actors who began their careers in comedy, were much less likely to be able to move into drama (or any other genre), but this effect did not necessarily hold for other movie genres. Presumably, this strong boundary between comedy and drama reflects beliefs by casting agents, and other movie industry insiders, as to the differences between the skills required to act in a comedy as differing markedly from those required to succeed in dramatic roles.

Relaxing this assumption then allows us to proceed in two directions. First, that past experiences should differentially affect a social actor's appeal depending on the distance/similarity of those past categorized experiences to the focal category. We hold an actor's reputation here constant in our theorizing and suggest that the strongest indicator of a social actor's underlying quality would be evidence of previous success in either the precise same arena (Zuckerman et al 2003) or similarly relevant domains. To the extent that the underlying characteristics of a category overlap more with another, the skill necessary to succeed in one will be more relevant to success in a more proximate one. Therefore, we predict:

> *Hypothesis 2a: Greater relevant experience (holding constant reputation) increases appeal.*

Second, we believe reputation should moderate this relationship, though in a less than straightforward manner. Reputation for previous performance in a categorized task acts as an indicator of quality in that task but should also serve to indicate potential for broader capabilities.

More specifically, the ability to perform a task adequately is a function of both task specific capabilities as well as more general facilities. Take for example Michael Jordan's move from basketball to baseball. While at face value we would likely conclude that very little of the skills which made him a good basketball player would translate directly to the baseball field, there were perhaps more generalizable skills that could have helped. So his ability to jump very high probably did little to enable him to hit a fastball well. Yet more general skills, such as the dedication and discipline required to become a skilled athlete, would likely be seen as translating to other sporting domains.

Yet past success at tasks highly dissimilar to the focal one should not weigh as heavily as those which are more similar. As the relevance of past experiences decreases, greater reputation in these areas serve as a weaker signal to the focal job. This is because what the social actor has accomplished in the past becomes less and less relevant. Therefore:

Hypothesis 2b: For those with higher reputations, the positive effect of a good reputation decreases as relevant experience decreases.

However, for those with a poor reputation to begin with, the lower appeal one has isn't further diminished by less relevant experiences because poorer performance in a domain which is less related. In fact, we suggest that audiences may interpret a poor reputation, though in an unrelated domain less harshly than in a related domain. Comparisons are more relevant with more related past experiences, thereby making salient the negative reputation. Yet ones poor reputation loses its relevance as distance to the focal category increases. This leads us to predict:

Hypothesis 2c: For those with lower reputations, the negative effect of a bad reputation decreases as relevant experience decreases.

REPUTATION AND CATEGORY FOCUS

As much past research has demonstrated, those social actors who span disparate categories are disadvantaged relative to their more focused peers (Zuckerman 1999; Hsu 2006). Because of the difficulty an audience may have in understanding and making sense of a varied past or because of the operational difficulties one assumes an actor has in combining potential disparate experiences, social actors who are unable to display a narrowly focused identity are either ignored or assumed to be of lower quality (Negro and Leung 2012). Therefore, following past research, we expect as a 1st order prediction:

Hypothesis 3a: Greater spanning (holding constant reputation) decreases appeal.

Yet reputational cues may alter how a generalist is perceived relative to specialists. To the extent that generalist have demonstrated accomplishments in multiple domains, better reputational signals become more difficult to interpret as they are spread across more disparate domains. In these instances, the variety of past experiences may loom larger than ones reputation because they suggest that skills one brings to a particular endeavor may be less specifically relevant but rather more generally relevant. General relevance will not be as valued as specific relevance as with more focused identities, reputation is a clearer measure, as it will be associated with a narrower set of particular past experiences. We therefore expect:

Hypothesis 3b: For those with higher reputations, the positive effect of a good reputation decreases as spanning increases.

Note, this may merely reflect the assumption that people hold of generalists necessarily being less-talented than specialists. Because of this assumption, an evaluator may come to expect generalists to have lower reputation, so they therefore care less about such a measure of quality when encountering it in a market setting – what we call a generalist discount. Second, those that decide to transact with generalists may be looking to fill particular needs which may loom larger than mere reputation signals. For example, people shop at large department store not necessarily because they offer higher quality, but rather due to their convenience. Similarly, restaurants which attempt to cater to a broad range of tastes continue to persist perhaps due to their appeal to groups of heterogeneous patrons who cannot agree on a narrow cuisine. We therefore also expect:

> *Hypothesis 3c: For those with lower reputations, the negative effect of a bad reputation decreases as spanning increases.*

DATA AND METHODS

Empirical Setting

The research questions were tested via the context of Elance, an online marketplace for freelancing services. Self-styled "the future of hiring," Elance provides a venue for would-be employers of freelancers to solicit the skills and services of an online freelancing workforce. This is achieved by posting a "job listing." Potential employers (henceforth *buyers*) provide a job description, detail requirements and set a timeframe for both the job award and deliverables. Freelancers (henceforth *sellers*) proceed to bid on these jobs. After reviewing the past job history and performance of bidding sellers, the buyer then establishes a winner; the job is thus awarded. All deliverables are accomplished online.

Once the contract and deliverables are fulfilled, the buyer has the option to provide feedback on the seller's performance. This is done through a feedback rating system which ranges from 1 to 5. The average feedback ratings and categories of the past jobs are prominently displayed on the profile of the sellers; this establishes a reputation system on which the community builds upon. The reputation system resembles one that makes use of *assessment* signals (Donath 1998) – these signals of quality are not self-reports and have to be earned through both the delivery of high quality work and buyer satisfaction.

In addition, the category of the seller's past work is displayed alongside the feedback ratings. This could be any of the 234 subcategories of work that are externally imposed by Elance. Categories cover a range of services from graphic design to web programming and development. Not all categories are related to web-based services; for instance there are categories that cater to creative writing, financial and legal services. See the Appendix for screen shots of the Elance website.

The computer mediated setting of Elance makes this a compelling setting to test our hypotheses. Firstly, we note the presence of strong, reliable assessment signals side by side categorical differentiation of past work. This lends itself nicely to our independent variables of reputation and categorization. Secondly, the computer mediated setting implies that the *actual quality* of the work is immensely difficult to ascertain *a priori* the award of a job contract. Aside from repeated transactions with a particular seller, buyers have very few data points to infer seller quality apart from what is presented on their profile. The salience of both the online reputation and job categorization system becomes evident as assessment signals of both the seller's ability, quality of work and identity (or lack thereof). Thirdly, the organization of online marketplaces via categorical boundaries in order to reduce search and identification costs mirrors our natural instincts to cluster like objects. Finally, the bid system allows us to construe of each job listing as a quasi-experiment; we observe all information that is available to the buyer, including the bidders that fail to win the job contract. This system thus provides strong counterfactuals to test our hypotheses; such a design will be considerably more difficult to implement in a face-to-face “real world” context.

Data and Variables

To build up an archival dataset of all buyer and seller activity, we gathered bid and transaction data for the time duration of December 2000 to May 2008. These activity data includes both job description and bid information. Job description includes the buyer identification, the job categorization, the listing timestamp and expected delivery deadlines. Bid information includes the bid amount, the bid timestamp and an indicator of contract award. In this data set, there are 643235 unique jobs; of these jobs, only 527513 jobs received any attention from any seller (i.e. was bided upon). In addition, we obtained the text descriptions and categories of all job listings including those that did not receive seller attention. These are on average paragraph length text descriptions that detail job requirements, job specifics and deliverable expectations.

Reputation

The reputation of a seller depends upon the feedback ratings provided by the buyer upon job completion. Building upon our definition of reputation as the expected capability derived from past displays, we operationalize a seller’s reputation score at time t as the cumulative average of feedback scores. Should the buyer choose not to give a feedback score, the observation is dropped for the purposes of the cumulative average; it is not entered into both the numerator and denominators of the reputation score. This measure reflects what is prominently displayed on the seller profile. Thus:

$$reputation_{it} = \frac{1}{N_t}\sum_{\tau<t} feedback_\tau$$

Where $reputation_{it}$ is the reputation score of seller i at time t; $feedback_\tau$ is the feedback score awarded to the seller at time τ; N_t is the number of feedback scores received by the seller up to time t.

Category Space

We employ a novel measure of experience relevance and categorical spanning that utilizes buyer contributed job descriptions. We consider these descriptions as time-dependent snapshots of market expectations, sentiments and requirements of a particular job category. The granularity and richness of text descriptions allow us to abstract a cognitive categorical space upon which the 243 categories are posited. (see the Appendix for a complete list of job categories) As such, the similarities of and differences between categories are reified as distances in this space. This paints a richer picture of categorical clustering. For instance, the job categories of Simple Website and Web Design will be relatively similar compared with the job categories of Simple Website and Academic Writing. In this categorical space, the distance between the two former categories will be smaller than that of the two latter categories. In addition, the time-dependent quality of job descriptions is reflected in the spatial dimension: the focal positions of the categories move in this space as the market updates its conception and definition of the categories. For instance, the categories of Flash Animation and Web Design would be more similar (i.e. closer in categorical space) in 2004 than they are today. Correspondingly their distance will have increased since their association in 2004.

To define the space, we first process the job description texts via the following steps. First, we strip the text of all stop-words, words that do not add any context specific meaning whatsoever (e.g. “is”, “are”, “and”, “the”). Following which, each word is stemmed into stem-tokens. This step implies that related words are collapsed into a single token. For instance, “consulting”, “consultant”, “consultation”, and “consult” will all be stemmed into the token “consult.” Similarly, “managing”, “manager”, management” etc. will all be stemmed into the token “manag.” Following which, duplicate mentions of the words are purged and the word-order disregarded. At this stage, each job description is characterized as a “bag” of unique words. We call the sum of all bags of words the corpus.

To characterize the categorical space, we consider each unique word in the corpus as a dimension in the space. In our corpus, we have a total of 27,080,872 words collected, of which 322,255 are unique. As a 322,255 dimensional space will be a significant computational challenge, we perform a selection of the most prominent features, capitalizing on the fact that word frequencies in language exhibits a power law (Newman 2005). To check this, we fit a power law distribution to our data. The resultant fit is both good and statistically significant ($\alpha = 2.08$, $x_{min} = 8120$, $K\text{-}S.stat = 0.06$, $p = 0.01917$). As such, we selected the 10,102 most frequent words to characterize the categorical space. These 10,102 words comprise 97% of the words in the corpus.

The relative frequency of word use in each category then denotes the category’s position in this space at a particular time *t*. This is achieved through a magnitude normalized, cumulative sum of all jobs that fall into the category before time *t*:

$$\vec{\boldsymbol{r}}_{\boldsymbol{Ct}} = \frac{1}{\left\| \sum_{c \in C, \tau < t} \vec{\boldsymbol{x}}_{\boldsymbol{c\tau}} \right\|} \sum_{c \in C, \tau < t} \vec{\boldsymbol{x}}_{\boldsymbol{c\tau}}$$

Here, $\vec{\boldsymbol{r}}_{\boldsymbol{Ct}}$ denotes the characterization vector of category *C* at time *t*; $\vec{\boldsymbol{x}}_{\boldsymbol{\tau}}$is the characterization vector of each individual job at time τ as given by its bag of words (binomial word count).

To check our characterization vectors for face validity, we calculated the distances between the top nine most transacted categories in Elance. These categories are: (1) Web Content, (2) Application Development, (3) Web Programming, (4) Web Design, (5) Simple Website, (6) Flash Animation, (7) Logos, (8) Graphic Design and (9) Ecommerce Website. The Euclidean distances between the characteristic vectors are calculated. These are then visualized as a heat map in Figure 1.

[Insert Figure 1 about here]

As evidenced in Figure 1, the categorical space construction has good face validity. We observe that while Web Programming, Web Design and Simple Website are relatively close to each other, they are very different from Logos and Flash Animation. App Development and Web Programming evidently are perceived similarly while App Development and Web Design are not. Finally, the graphic design jobs, Graphic Design and Logos, are closer in distance than the programming/web related jobs.

Categorical Relevance and Spanning

Each seller occupies a time-dependent position in the category space depending on his/her past completed job categories. To do this, we create a center of mass measure that seeks to indicate a seller's position as a point in categorical space:

$$\vec{\boldsymbol{R}}_{\boldsymbol{it}} = \frac{1}{M} \sum_{c \epsilon C_{it}} \vec{\boldsymbol{r}}_{\boldsymbol{ct}}$$

Here, $\vec{\boldsymbol{R}}_{\boldsymbol{it}}$indicates the position of seller i at time t. C_{it} is the set of categories that labels seller i's past jobs prior to time τ. M is the number of all jobs with feedback prior to time t. $\vec{\boldsymbol{r}}_{\boldsymbol{ct}}$ is the characterization vector of category c at the bid time t.

Relevance is operationalized as the Euclidean distance between the job category of the current bid and the seller's position in categorical space:

$$relevance_{bid,it} = \left\| \vec{\boldsymbol{R}}_{\boldsymbol{it}} - \vec{\boldsymbol{r}}_{\boldsymbol{c}_{\boldsymbol{bid,t}}} \right\|$$

$\vec{\boldsymbol{r}}_{\boldsymbol{c}_{\boldsymbol{bid,t}}}$ is the characterization vector of the current job of interest that the seller is bidding on at time *t*. A seller that has experience that is completely relevant to the current job he/she is bidding on will have a score of zero. This is a distance measure; the measure of experience relevance is thus reverse coded.

Spanning is operationalized as the average of all Euclidean distances between all past job categories of a particular seller and his point position in categorical space at time of bid.

$$spanning_{it} = \frac{1}{M} \sum_{c \epsilon C_{it}} \left\| \vec{\boldsymbol{r}}_{\boldsymbol{ct}} - \vec{\boldsymbol{R}}_{\boldsymbol{it}} \right\|$$

A seller who only took jobs from a single category will have a spanning score of 0. Note that unlike the relevance measure, the spanning variable is independent of the job category of the bid. Also note that unlike the relevance measure, the spanning measure is *not* reverse coded.

We also note that the categorical spanning and relevance variables are calculated *at the time of bid.* Our measures of categorical spanning and experience relevance therefore capture *time-varying market categorical perception*.

Covariates

Three covariates are included in the analysis. The first is a measure of dyadic tie strength between the buyer and seller. This is coded as the number of prior contracts that were fulfilled between the seller and the focal buyer. The second is a measure of overall seller experience. This is measured as the total number of prior contracts fulfilled by the seller regardless of buyer identity. The final covariate is a monetary measure of the latest bid offered by the seller before the job auction is closed. This is measured in US dollars. Tables 1 and 2 list summary statistics and correlations of the variables used.

[Insert Tables 1 and 2 about here]

MODELS AND HYPOTHESIS TESTING

The binary nature of the dependent variable lends itself to a logistic regression specification which estimates a dependent variable bounded by 0 and 1 (Long 1997). Here, we code 1 as a bid that wins a contract and 0 as one that does not. In addition, the online setting allows us to employ job level fixed effects to account for both buyer and time dependent heterogeneity.

The models run on a subset of the historical data. As we are interested in reputation effects, we remove from the data all bids sellers with no prior reputation whatsoever. In addition, jobs which either fail to award any contracts or awarded contracts to *all* bidders are discarded to facilitate the within job design. This effectively removes jobs that exhibited *no seller competition* to obtain outcome variance within jobs. Note that these jobs are only discarded for the purposes of the regression model; they are included in the categorical definition measures.

In this data subset, there are 137194 unique jobs listed by 59877 unique buyers. These jobs are bided on by 18914 unique sellers; 1498480 bids were made giving an average of 10.92 bids per job. Correspondingly each job on average awards 2.511 contracts amongst the bidders. Summary statistics of this data set are presented in Table 1. A table of correlations is presented in Table 2. As the experience relevance and categorical spanning measure of jobs with no feedback and jobs with feedback are both highly correlated, we only consider the relevance and spanning measures of jobs that were accorded feedback of some kind.

Using a maximum likelihood estimator with job fixed effects, we estimated 4 different models. Model 1 enters each independent variable and covariates in a quasi-linear fashion without interaction effects. Model 2 and 3 enters interaction effects: model 2 examines interaction effects between *reputation* and *relevance*; model 3 examines interaction effects between *reputation* and *spanning*. Finally, Model 4 enters both interaction effects into the regression specification to test for independence. All models are run with job clustered standard errors. The regressions were estimated with a generalized linear mixed model maximum likelihood estimator (Broström & Holmberg 2011).

Results

Estimated model coefficients are reported in Table 3. All models include all the control covariates. To interpret the results, we will first discuss Model 1, which employs no interaction terms. Models 2, 3 and 4 will be discussed more in depth as logistic interaction effects require some interpretive care.

[Insert Table 3 about here]

The coefficients of Model 1 supports Hypotheses 1, 2a and 3a; all independent variables of interest as well as the control variables behave as expected. The effect of reputation on the probability of winning a bid is positive, and that of relevance and spanning negative (recall that relevance is reverse coded). The control variables also exhibit expected directionality: the effect of bid amount is negative, while the effect of seller experience is positive. We do note that the size of these two effects is relatively small in comparison to the other variables in the regression.

Having a strong prior tie to the buyer greatly increases your chances of winning a bid, as is evidenced by the large and significant effect size of the dyadic tie variable.

To interpret the interaction models, we first consider that the logistic regression model is non-linear. This implies that the cross derivative of the two interaction terms (i.e. the interaction effect) is *not* constant and highly dependent on the covariates (Ai & Norton 2003). Unlike linear interaction models whose cross derivatives are constant and equals the interaction coefficient term, interaction effects cannot be "read off" the model coefficients in the case of the logistic regression. In addition, the significance of the effects is also covariate dependent. Thus, the interaction effect of logistic specifications can be insignificant when the model coefficient is significant, and vice versa.

We calculate the covariate dependent interaction effects of both Model 3 and Model 4 based on Ai and Norton (2003). These are then plotted against the *predicted* probabilities of the data using the model coefficients. These plots are shown in Figure 2 and Figure 3. We observe that both interaction effects resemble a U-shaped curve with the magnitude of the negative interaction effect peaking at around the predicted probability of 0.5. Correspondingly, the logistic interaction effects are insignificant for predicted probabilities that tend towards 0 and 1, implying that reputation and categorization effects are insignificant when the case for a seller to win or lose is overwhelmingly obvious. As such, Hypotheses 2b,c and 3b,c are rejected at these regions.

[Insert Figures 2 and 3 about here]

To ascertain the directionality of the non-linear interaction effects, we plot the functional form of the interaction terms of Model 3 and 4 in Figure 4 and Figure 5 against the predicted probabilities as implied by these terms. Figure 4 shows how reputation affects the probability of winning given three relevance distances. These three distances were selected to represent the first quartile, median and third quartile while the range of reputation scores highlights the middle quartile. Here, we observe that the relationship is monotonic; the gradient of the relationship decreases as relevance distances scores increase. Across each reputation band, we observe that holding high reputation constant (say a score of 4.9), the seller is penalized more when his past experience is relatively irrelevant to the bid category compared with holding a lower quartile reputation constant (e.g. a score of 4.2). This suggests that non-relevant experience penalties are experienced less by lower reputation sellers. This finding supports Hypotheses 2b and 2c.

[Insert Figures 4 and 5 about here]

Similar patterns dictate the interaction effect between reputation and categorical spanning. Figure 5 shows how reputation affects the probability of winning given 3 category spanning scores which represents the first quartile, median and third quartile. Again, the range of reputation scores focuses on the middle quartile. We observe that the relationship is monotonic and similarly, the gradient of the relationship decreases as the categorical spanning increase. Curiously, there is a "flipping point:" lower reputation sellers not only experience less penalty associated with categorical spanning, they are instead perceived to be more favorable than similar sellers who are more specialized. We do note that the rate of gradient decrease is larger than that of the relevance-reputation interaction; interaction effects between categorical spanning and reputation are larger than that between relevance and reputation. This finding supports Hypotheses 3b and 3c.

Finally, we find significance in both interaction coefficients in Model 4, suggesting that categorical spanning and experience relevance interactions with reputation are independent. This is not surprising considering that the spanning and relevance measures are only loosely correlated.

REFERENCES

Ai, C., & Norton, E. C. (2003). Interaction terms in logit and probit models. *Economics letters*, *80*(1), 123-129.

Broström, G., & Holmberg, H. (2011). glmmML: Generalized linear models with clustering. *R package version 0.81-8*.

Carroll, G. and A. Swaminathan. 2000. Why the Microbrewery Movement?: Organizational Dynamics of Resource Partitioning in the U.S. Brewing Industry. *American Journal of Sociology* 106(3) 715-62.

Dobrev, S. and T. Kim and M.T. Hannan. 2001. “Dynamics of Niche Width and Resource Partitioning.” *American Journal of Sociology*, 106: 1299–1337.

Donath, J. S. (1999). Identity and deception in the virtual community. *Communities in cyberspace*, *1996*, 29-59.

Fombrun, C. J. 1996. Reputation: Realizing value from the corporate image. Boston: Harvard Business School Press.

Fombrun, C., & Shanley, M. 1990. What‘s in a name? Reputation building and corporate strategy. Academy of Management Journal, 33: 233-256.

Hannan, Michael T., László Pólos, and Glenn R. Carroll. 2007. *Logics of Organization Theory: Audiences, Codes, and Ecologies*. Princeton: Princeton University Press.

Hsu, G. 2006. “Jacks of all trades and masters of none: Audiences’ reactions to spanning genres in feature film production.” *Administrative Science Quarterly*, 51: 420–50.

Jensen, M., H. Kim, and B.K. Kim. 2012. "Meeting Expectations: A Role-Theoretic Perspective on Reputation." In The Oxford Handbook of Corporate Reputation, edited by Michael L. Barnett and Timothy G. Pollock. Oxford University Press: 140-159.

Kovacs, B. and M.T. Hannan. 2012. “Category Spanning, Distance, and Appeal.” *Working paper*, University of Lugano.

Lange, D., P. M. Lee and Y. Dai. 2011. ‘Organizational reputation: A review’. *Journal of Management,* 37:153-184.

Long, J. Scott. 1997. *Regression Models for Categorical and Limited-Dependent Variables*. Thousand Oaks, CA: Sage.

Negro, G. and M.D. Leung 2012. “’Actual’ and Perceptual Effects of Category Spanning: An Analysis of Critical Ratings of Italian Wines,” *Organization Science* (Articles in Advance), 23(5): pp.1-13.

Newman, M. E. (2005). Power laws, Pareto distributions and Zipf's law. *Contemporary physics*, *46*(5), 323-351.

Podolny, Joel and Damon Phillips. 1996. "The Dynamics of Organizational Status." *Industrial and Corporate Change* 5(2): 453-471.

Rao, H. 1994. The social construction of reputation: Certification contests, legitimation, and the survival of organizations in the American automobile industry. Strategic Management Journal, 15: 29-44.

Rao, H. P. Monin, and R. Durand. 2003. "Institutional change in toque ville: Nouvelle cuisine as an identity movement in French gastronomy." *American Journal of Sociology*, Vol. 108: 795–843.

Rao, H. P. Monin, and R. Durand 2005. "Border crossing: Bricolage and erosion of categorical boundaries in French gastronomy." *American Sociological Review*, 70: 968-991.

Resnick, Paul and Richard Zeckhauser. Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System. The Economics of the Internet and E-Commerce. Michael R. Baye, editor. Volume 11 of Advances in Applied Microeconomics. Amsterdam, Elsevier Science. 2002. pp. 127-157.

Resnick, Paul, Zeckhauser, Richard, Swanson, John, and Kate Lockwood. The Value of Reputation on eBay: A Controlled Experiment. *Experimental Economics*. Volume 9, Issue 2, Jun 2006, Page 79-101.

Roberts, P.W., & Dowling, G.R. 2002. Corporate reputation and sustained superior performance. Strategic Management Journal, 23: 1077-1093.

Shapiro C. 1983. "Premiums for high quality products as returns to reputations," Quarterly Journal of Economics, Vol. 98: 659–679.

Spence, M. (1973). 'Job market signaling'. Quarterly Journal of Economics, 87, pp. 355-375.

Stuart T. E., Hoang, H., Hybels, R.C. "Interorganizational Endorsements and the Performance of Entrepreneurial Ventures" *Administrative Science Quarterly*, Vol. 44, 1999**.**

Zerubavel E. 1997. *Social Mindscapes*. Cambridge, Massachusetts: Harvard University Press.

Zuckerman, E.W., Kim, T., Ukanwa, K., Rittmann, J. 2003. "Robust Identities or Nonentities? Typecasting in the Feature-Film Labor Market." *American Journal of Sociology*, 108 (5): 1018-1074.

Zuckerman, E.W. 1999. "The Categorical Imperative: Securities Analysts and the Legitimacy Discount." *American Journal of Sociology*, 104:1398–1438.

Table 1
Summary Statistics
(N = 1498480)

Variables	Mean	St. Dev.	Min	Max
Reputation	4.21	1.01	0.00	5.00
Feedback Categorical Spanning	0.29	0.14	0.00	0.66
Feedback Relevance	0.37	0.20	0.00	1.22
All Categorical Spanning	0.31	0.13	0.00	0.68
All Experience Relevance	0.36	0.20	0.00	1.21
Overall Experience	131.04	332.55	0.00	5227.00
Buyer Tie Strength	0.04	0.91	0.00	302.00
Bid Amount/[USD]	653.53	8470.43	0.00	10000000.00

Table 2
Correlation Table of Variables

	Reputation	Feedback Spanning	Feedback Relevance	All Spanning	All Relevance	Experience	Buyer Tie Strength
Reputation							
Feedback Spanning	0.41***						
Feedback Relevance	0.31***	0.39***					
All Spanning	0.45***	0.90***	0.48***				
All Relevance	0.32***	0.42***	0.97***	0.49***			
Experience	0.15***	0.14***	0.01***	0.12***	0.02***		
Buyer Tie Strength	0.00***	-0.02***	-0.02***	-0.02***	-0.02***	0.01***	
Bid Amount/[USD]	0.00***	0.00***	-0.01***	0.00***	-0.01***	0.00	0.00

*Note: * p < 0.05, ** p < 0.01, *** p < 0.001*

Table 3
Likelihood of Winning a Bid
Logistic Regression Model Coefficients, Fixed-Effects by Job

	MODEL 1 No interaction effects		MODEL 2 Reputation-Relevance Int.		MODEL 3 Reputation-Spanning Int.		MODEL 4 Test for Independence	
	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.	Coef.	Std. Err.
(Intercept)	-3.173***	(1.690e-02)	-3.961***	(2.643e-02)	-3.950***	(2.588e-02)	-4.285***	(3.112e-02)
Reputation	0.295***	(3.932e-03)	0.483***	(6.024e-03)	0.484***	(5.908e-03)	0.563***	(7.081e-03)
Exp. Relevance Distance	-0.546***	(1.531e-02)	3.470***	(6.886e-02)	-0.608***	(1.530e-02)	2.068***	(8.848e-02)
Spanning Distance	-0.288***	(2.245e-02)	-0.313***	(2.232e-02)	6.649***	(1.143e-01)	4.507***	(1.334e-01)
Overall Experience	0.0003***	(6.055e-06)	0.0003***	(6.041e-06)	0.0003***	(6.078e-06)	0.0003***	(6.079e-06)
Buyer Tie Strength	0.523***	(7.051e-03)	0.530***	(7.045e-03)	0.538***	(7.038e-03)	0.525***	(7.041e-03)
Bid Amount/[USD]	-0.0001***	(3.138e-06)	-0.0001***	(3.125e-06)	-0.0001***	(3.112e-06)	-0.0001***	(3.108e-06)
Reputation: Relevance Distance			-0.934***	(1.585e-02)			-0.611***	(2.010e-02)
Reputation: Spanning Distance					-1.602***	(2.588e-02)	-1.119***	(3.039e-02)
df	1498472		1498471		1498472		1498470	
Aikaike Information Criterion	953200		949700		949200		948300	

*Note: Standard errors in parentheses, * p < 0.05, ** p < 0.01, *** p < 0.001*

Figure 1
Heat-map of categorical distances
Blue implies similarity (small distance); yellow implies difference (large distance)

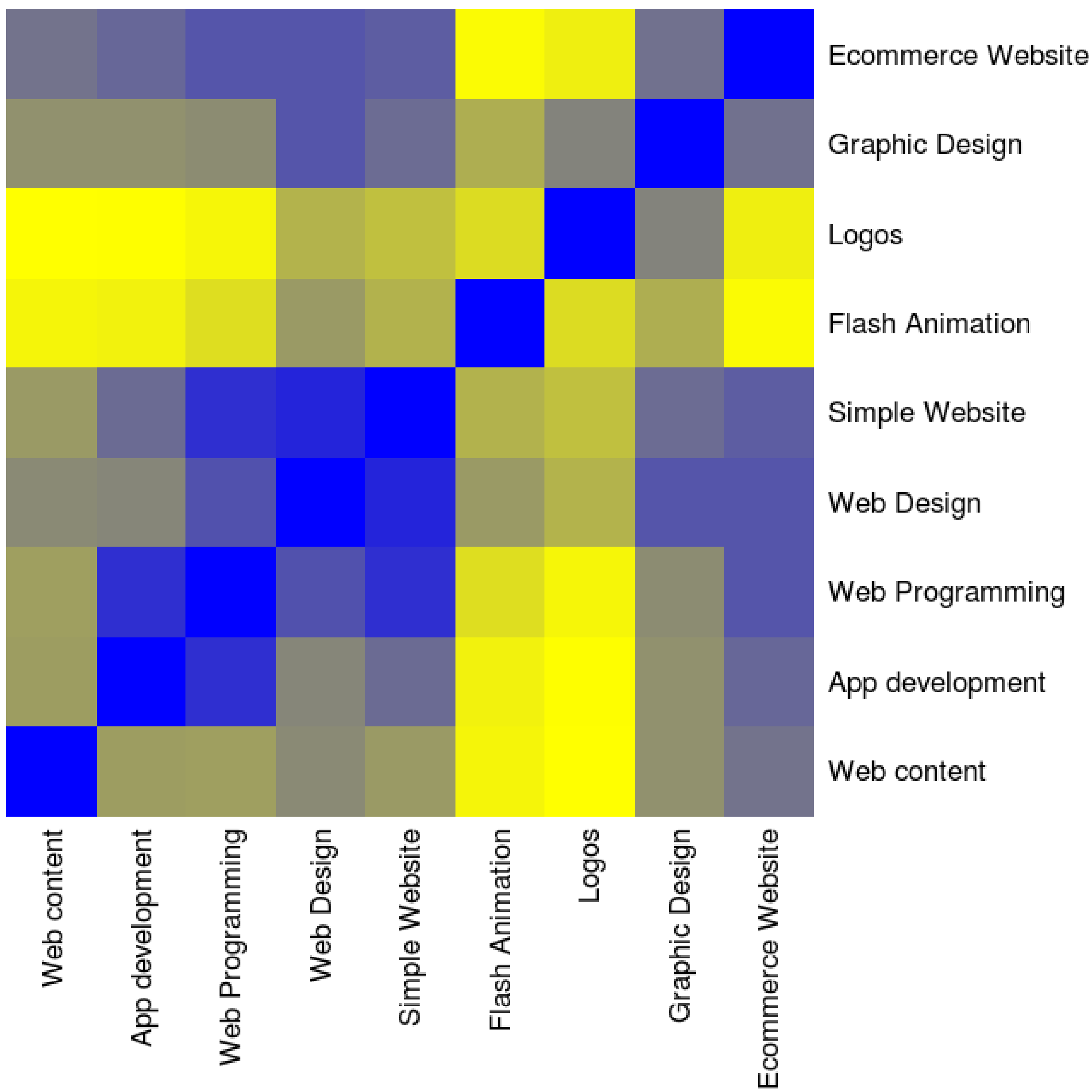


Figure 2

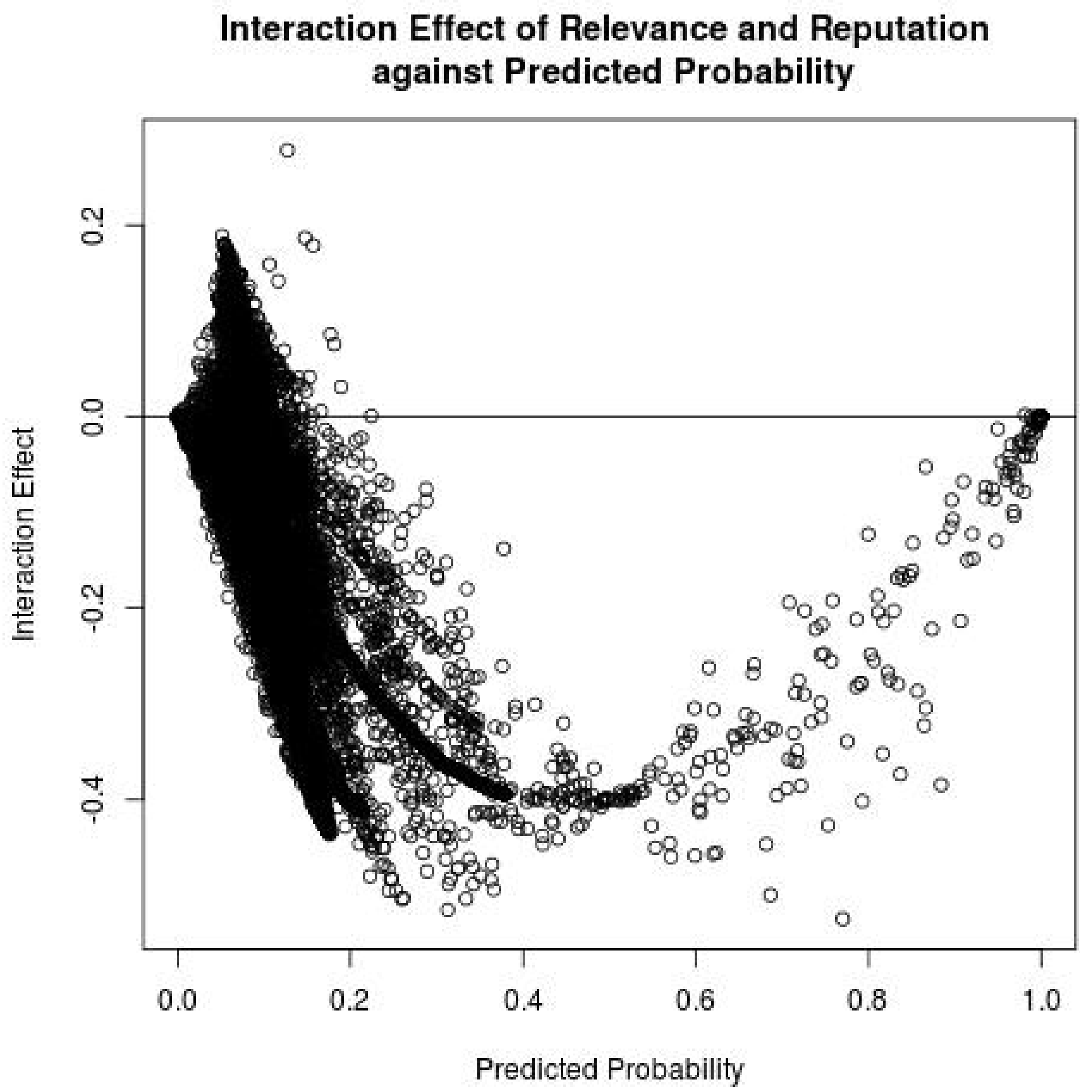

Interaction Effect of Relevance and Reputation
against Predicted Probability
Interaction Effect
0.2
0.0
-0.2
-0.4
0.0
0.2
0.4
0.6
0.8
1.0
Predicted Probability

Figure 3

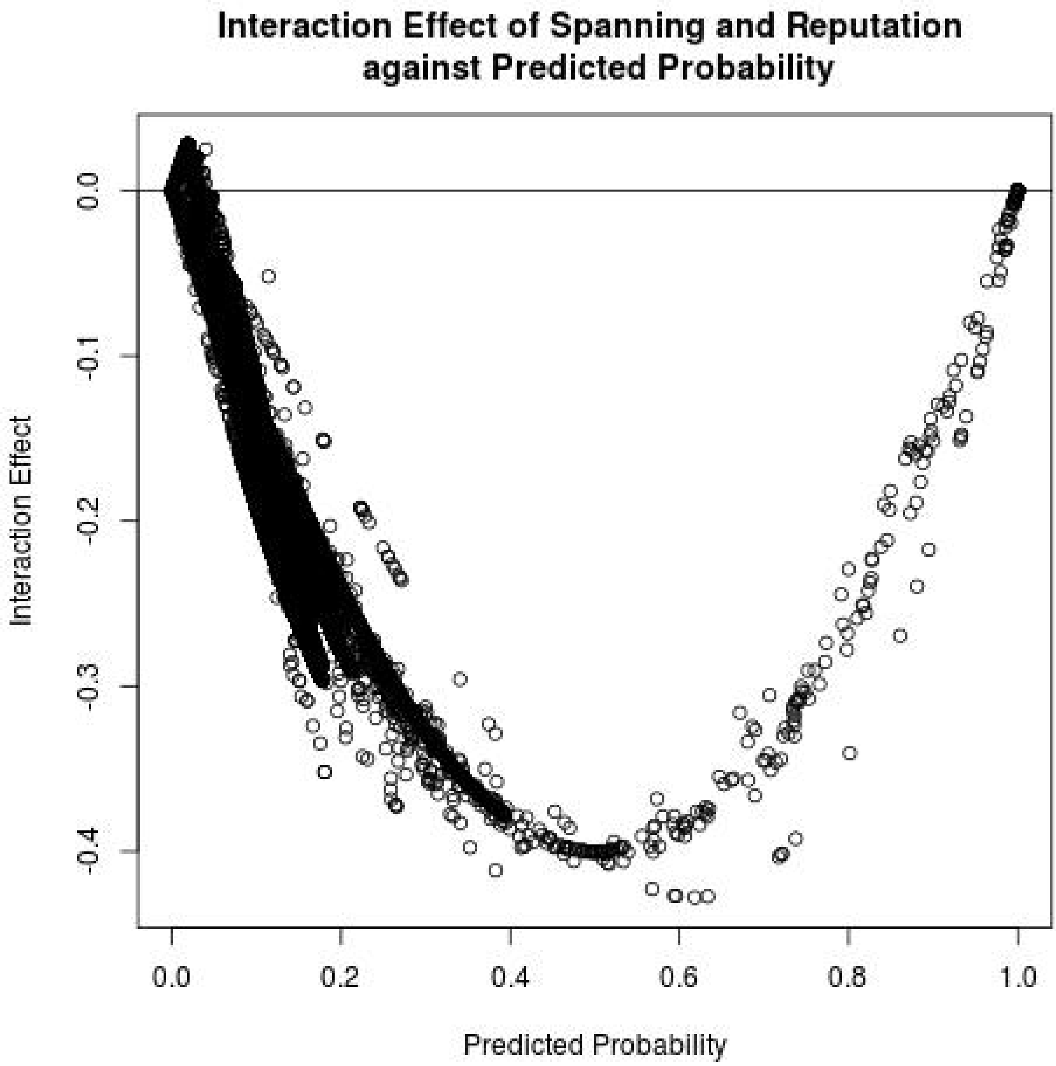

Interaction Effect of Spanning and Reputation
against Predicted Probability
Interaction Effect
0.0
-0.1
-0.2
-0.3
-0.4
0.0
0.2
0.4
0.6
0.8
1.0
Predicted Probability

Figure 4
Interaction effect of relevance distance and reputation at upper and lower quartiles of relevant distance
Range of reputation score selected to represent middle quartiles

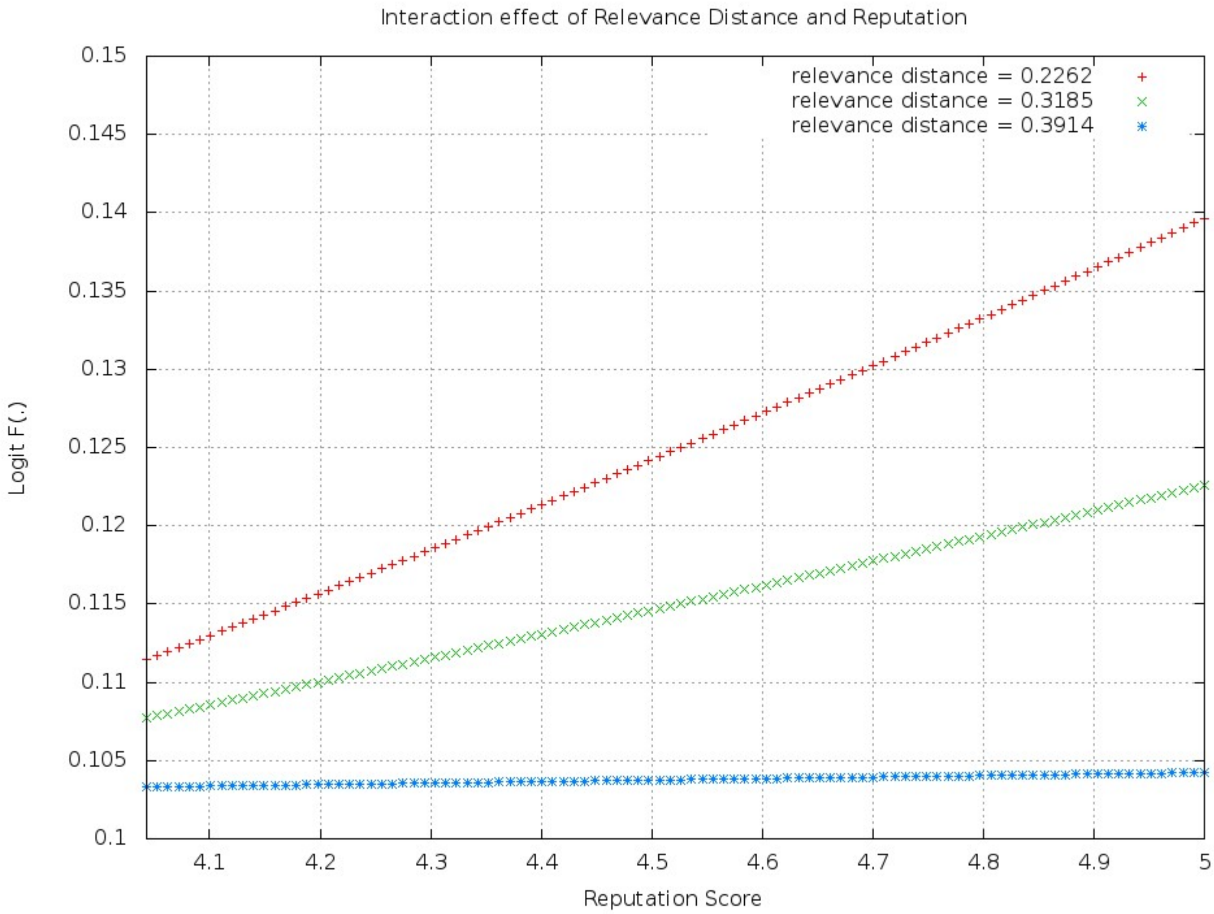
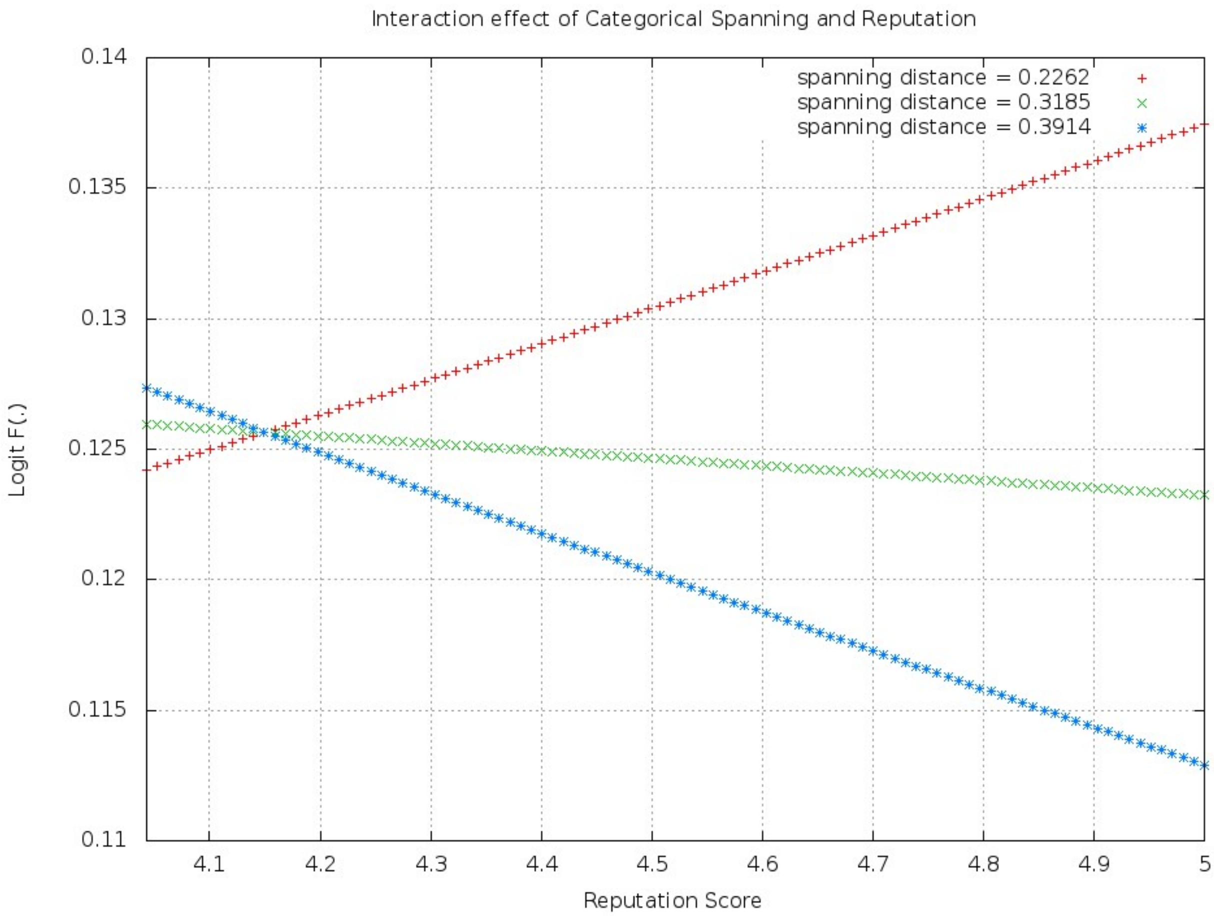


Figure 5
Interaction effect of categorical spanning and reputation
at upper and lower quartiles of spanning distance measures
Range of reputation score selected to represent middle quartiles



APPENDIX

FIGURE 1

JOB INFORMATION

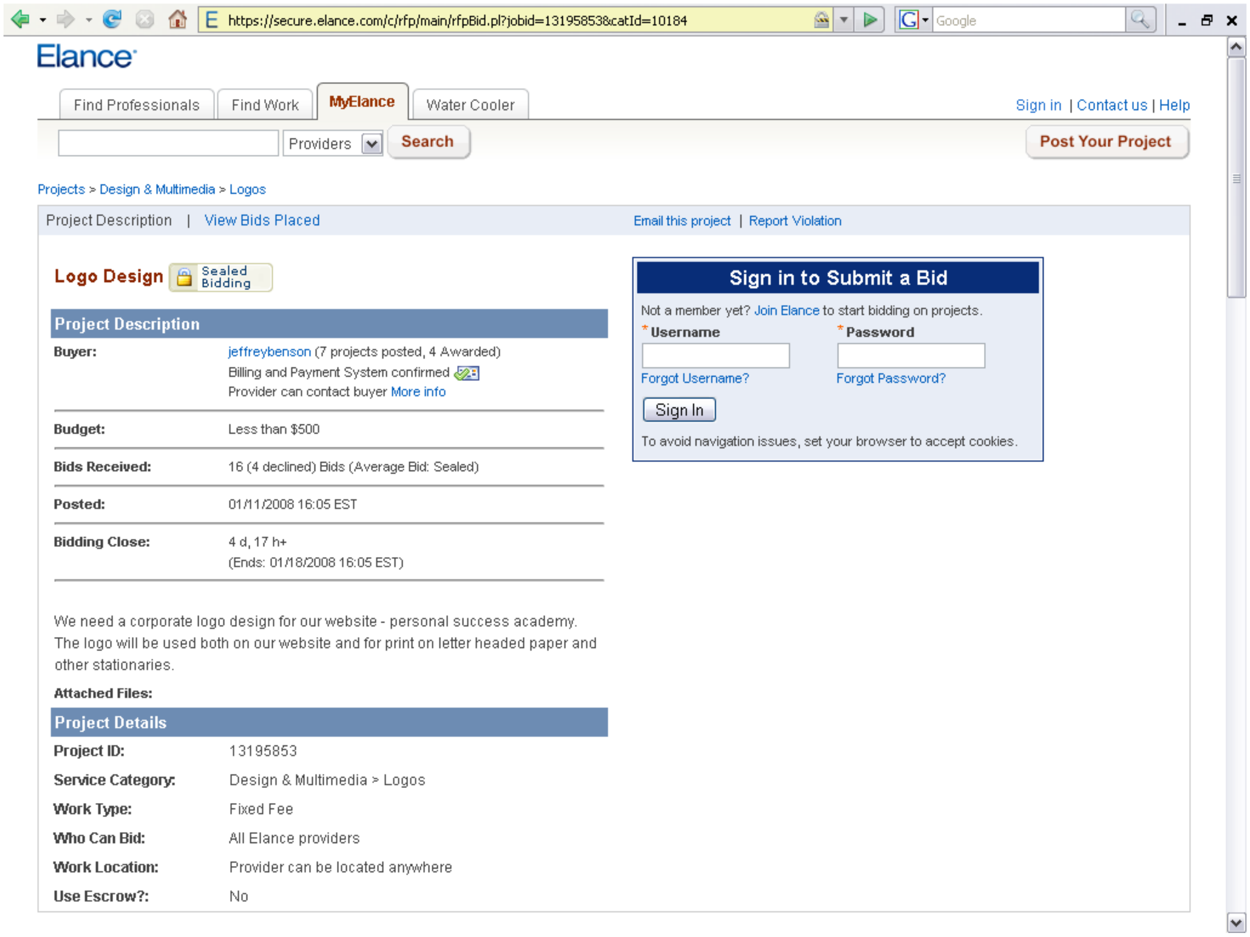
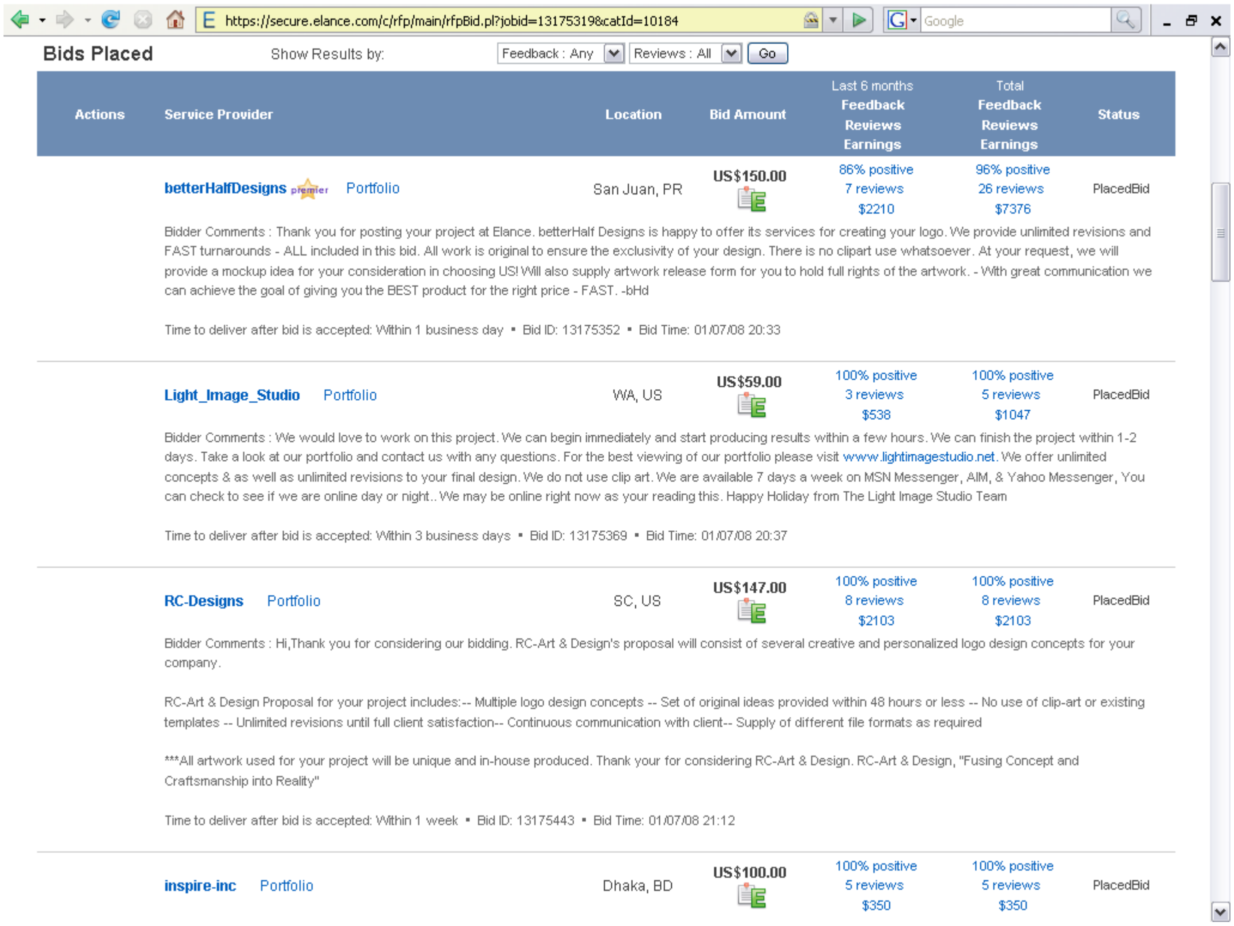


FIGURE 2
BID INFORMATION



https://secure.elance.com/c/rfp/main/rfpBid.pl?jobid=13175319&catId=10184

Bids Placed Show Results by: Feedback : Any Reviews : All Go

Actions	Service Provider	Location	Bid Amount	Last 6 months Feedback Reviews Earnings	Total Feedback Reviews Earnings	Status
	betterHalfDesigns premier Portfolio	San Juan, PR	US$150.00	86% positive 7 reviews $2210	96% positive 26 reviews $7376	PlacedBid
	Light_Image_Studio Portfolio	WA, US	US$59.00	100% positive 3 reviews $538	100% positive 5 reviews $1047	PlacedBid
	RC-Designs Portfolio	SC, US	US$147.00	100% positive 8 reviews $2103	100% positive 8 reviews $2103	PlacedBid
	inspire-inc Portfolio	Dhaka, BD	US$100.00	100% positive 5 reviews $350	100% positive 5 reviews $350	PlacedBid

betterHalfDesigns

Bidder Comments : Thank you for posting your project at Elance. betterHalf Designs is happy to offer its services for creating your logo. We provide unlimited revisions and FAST turnarounds - ALL included in this bid. All work is original to ensure the exclusivity of your design. There is no clipart use whatsoever. At your request, we will provide a mockup idea for your consideration in choosing US! Will also supply artwork release form for you to hold full rights of the artwork. - With great communication we can achieve the goal of giving you the BEST product for the right price - FAST. -bHd

Time to deliver after bid is accepted: Within 1 business day • Bid ID: 13175352 • Bid Time: 01/07/08 20:33

Light_Image_Studio

Bidder Comments : We would love to work on this project. We can begin immediately and start producing results within a few hours. We can finish the project within 1-2 days. Take a look at our portfolio and contact us with any questions. For the best viewing of our portfolio please visit www.lightimagestudio.net. We offer unlimited concepts & as well as unlimited revisions to your final design. We do not use clip art. We are available 7 days a week on MSN Messenger, AIM, & Yahoo Messenger, You can check to see if we are online day or night.. We may be online right now as your reading this. Happy Holiday from The Light Image Studio Team

Time to deliver after bid is accepted: Within 3 business days • Bid ID: 13175369 • Bid Time: 01/07/08 20:37

RC-Designs

Bidder Comments : Hi,Thank you for considering our bidding. RC-Art & Design's proposal will consist of several creative and personalized logo design concepts for your company.

RC-Art & Design Proposal for your project includes:-- Multiple logo design concepts -- Set of original ideas provided within 48 hours or less -- No use of clip-art or existing templates -- Unlimited revisions until full client satisfaction-- Continuous communication with client-- Supply of different file formats as required

***All artwork used for your project will be unique and in-house produced. Thank your for considering RC-Art & Design. RC-Art & Design, "Fusing Concept and Craftsmanship into Reality"

Time to deliver after bid is accepted: Within 1 week • Bid ID: 13175443 • Bid Time: 01/07/08 21:12

FIGURE 3
SELLER INFORMATION

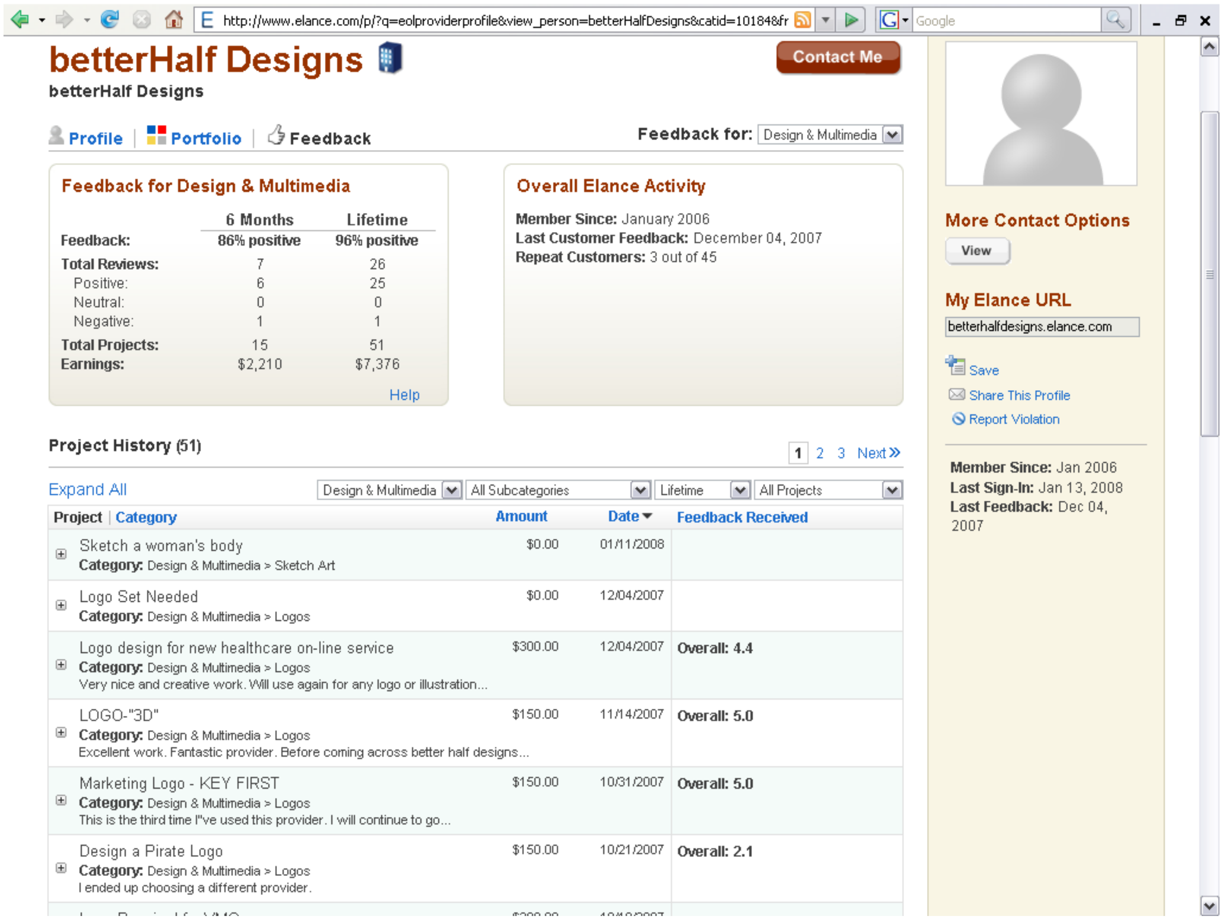


TABLE 1
E*LANCE HIGH LEVEL CATEGORIES

High Level Categories
Admin Support
Design & Multimedia
Engineering & Manufacturing
Finance & Management
Legal
Sales & Marketing
Test Writing
Web & Programming
Writing & Translation

TABLE 2
SUBCATEGORIES OF DESIGN & MULTIMEDIA

Design & Multimedia Sub-Categories
3D Graphics
Animation
Banner Ads
Brochures
Card Design
Cartoons & Comics
Catalogs
CD & DVD Covers
Commercials
Corporate Identity Kit
Digital Image Editing
Direct Mail
Displays & Signage
Emails & Newsletters
Embedded Video/Audio
Graphic Design
Illustration
Label & Package Design
Logo Design
Menu Design
Music
Other - Design
Other - Multimedia Services
Page & Book Design
Photography & Editing
Podcasts
Presentation Design
Print Ads
Radio Ads & Jingles
Report Design
Sketch Art
Stationery Design
Videography & Editing
Viral Videos
Voice Talent