



CIFE CENTER FOR INTEGRATED FACILITY ENGINEERING

**The Impact of Rating Systems on
Subcontracting Decisions
—
A Transaction Cost Analysis**

By

Martin A. Ekström, Hans C. Björnsson,
John C. Kunz, Raymond E. Levitt, and Clifford I. Nass

**CIFE Technical Report #135
February 2003**

STANFORD UNIVERSITY

Copyright © 2003 by

Center for Integrated Facility Engineering

If you would like to contact the authors, please write to:

*c/o CIFE, Civil and Environmental Engineering Dept.,
Stanford University
Terman Engineering Center
Mail Code: 4020
Stanford, CA 94305-4020*

The Impact of Rating Systems on Subcontracting Decisions

—

A Transaction Cost Analysis

Authors: Martin A. Ekström¹, Hans C. Björnsson², John C. Kunz³, Raymond E. Levitt⁴, Clifford I. Nass⁵

1 Abstract

This article investigates how different types of rating mechanisms can affect transaction costs in business-to-business electronic commerce. In consumer electronic marketplaces, where the asset-specificity of the transactions is generally low, simple unweighted rating mechanisms have been shown to increase trust and therefore also decrease the transactions costs associated with opportunistic behavior. However, we argue that if a reputation mechanism is to decrease the transaction costs for highly asset specific transactions, it is necessary that the rating mechanism account for the relationship between the user and the rater. To empirically validate this hypothesis, we created a prototype credibility-weighted rating tool that incorporates a methodology to calculate a weighted rating based on source credibility theory. In an experiment, industry practitioners evaluated bids from service providers using the credibility-weighted rating tool, as well as a standard unweighted tool. For transactions with low asset-specificity, the experiment showed that both types of tools influence the risk premium added to the

¹ Martin A. Ekström, Post-Doctoral Research Fellow, Department of Civil Engineering, Stanford University, CIFE Terman Engineering Center, Stanford University, Stanford, CA 94305-4020, (650)-725 8970, mekstrom@stanford.edu

² Hans C. Björnsson, Professor, Department of Civil Engineering, Stanford University, hansbj@stanford.edu

³ John C. Kunz, Senior Executive Director, Center for Integrated Facilities Engineering, Department of Civil Engineering, Stanford University, kunz@stanford.edu

⁴ Raymond E. Levitt, Professor, Department of Civil Engineering, Stanford University, rel@stanford.edu

⁵ Clifford I. Nass, Professor, Department of Communication, Stanford University, nass@stanford.edu

bids. For highly asset specific transactions, on the other hand, the credibility-weighted ratings significantly influenced the added risk premium, while there was no evidence that the simple average ratings provided by the unweighted tool affected decision-making behavior.

2 Introduction

Observing the rapid adoption of electronic commerce among consumers, industry leaders and researchers believed that the advent of electronic commerce would revolutionize the structure of business processes of all major industries. Internet technologies allow buyers and sellers to find new market partners at a low cost (Bakos, 1997). This should raise competition by lowering the costs of free market business-to-business (B2B) transactions and increase market transparency. By 2004, B2B e-commerce was predicted to hit \$2.7 trillion in the US alone (Kafka et al., 2001). Researchers and industry leaders also predicted that the efficiency gains caused by information technology would lead to the increased use of markets (Malone, 1992) as the mode of transaction at the expense of vertical integration. Rating systems, or reputation mechanisms enhance market transparency and are thus important features of major consumer-to-consumer (C2C) market-places like eBay, as well as Business-to-Consumer (B2C) marketplaces like Amazon.com. Consumers shopping on eBay or Amazon can, with little effort and at no cost, obtain and contribute to the quality judgments made by peer consumers. Today, consumer e-commerce continues to thrive (even though there have been several spectacular failures) and there is even evidence that it has caused the restructuring of industries such as the travel industry. B2B e-commerce, on the other hand, has yet to fulfill its initial expectations.

This paper discusses the impact of different types of rating mechanisms from a transaction cost perspective. Transaction Cost theory provides a useful framework to analyze the slow adoption of electronic commerce in business to business. In his 1937 article "Nature of the Firm" Coase (1937), argues that the structure of a firm is set up to minimize overall transactions costs. Firms should conduct internally only those activities that cannot be procured more cheaply in the market. As a result, a firm will expand precisely to the point where "the costs of organizing an extra transaction within the firm

becomes equal to the costs of carrying out the same transaction by means of an exchange on the open market." Building on Coase's work, Williamson (1975) sums governance (transaction) and production costs to measure the aggregate performance of a governance structure. The two major governance structures are "market" and "hierarchy." "Market" refers to the free market where each production activity is performed by a separate firm while a "Hierarchy" implies a command hierarchy, with vertical integration of production activities within one organization. In addition, Williamson recognizes the importance of hybrid organization forms (like long term contracting, reciprocal trading, and franchising), which are characterized by a mix of markets and hierarchies. He shows how transaction costs are a function of asset-specificity and puts forward transactional hazards as the major cause of transaction costs.

Asset-specificity occurs when investments required by the two parties involved in a transaction cannot be redeployed to alternative uses outside the specific transaction. Furthermore, asset-specificity, in combination with the simultaneous presence of two pairs of factors (bounded rationality-uncertainty/complexity; and opportunism – small numbers), gives rise to transactional hazards that ultimately cause market transactions to fail. Transactional hazards include quality shortfalls, ex-post bargaining over surplus, litigation, hold-up costs, and wasted investments. This framework can help to explain why consumers have been faster than industry buyers to take advantage of the emerging electronic marketplaces. First of all, the asset-specificity is lower when a consumer buys a used tennis racket on eBay than when a general contractor subcontracts the construction of the HVAC system on a new building project. However, it is also important to consider the importance of the two pairs of factors, which differ between B2B and C2C settings.

The pair of factors "Bounded Rationality" – "Uncertainty/Complexity" is present when the bounded rationality of human beings prevents us from making rational decisions in uncertain and/or complex transactions. One explanation to the slower adoption of e-commerce in a business-to-business market is therefore that the goods and services purchased are, in general, more complex than consumer goods. In the future, technology may help to decrease the uncertainty and complexity of business transactions by providing decision makers with structured, accurate, and updated information. Nonetheless, we argue that, in the near term, information technology is more likely to

affect the second pair of factors: “Opportunism – Small Numbers.” Opportunistic behavior is more likely when there are a small number of actors in the market since competition between large numbers of actors generally decreases opportunism. The communities of C2C marketplaces often consist of thousands, even millions, of anonymous users. As a result, buyer and seller transactions are rarely repeated and there is often a large set of alternative transaction partners. In B2B settings customers often conduct a substantial part of their business with a small number of recurring suppliers. However, while it is difficult to affect the number of buyers and sellers in a market, electronic commerce providers have come up with ways to dissuade market participants from opportunistic behavior.

As the 40,000,000 members of eBay’s online community show, the Internet can also support trust by providing a means for market participants to share information about one another using a rating system. In addition, researchers have demonstrated that eBay sellers with high ratings benefit from higher prices (Dewan and Hsu, 2001, Lucking-Reiley et al., 2000, Houser and Wooders, 2000) and increased probability of selling their goods (Resnick and Zeckhauser, 2001). From a transaction economics perspective, Williamson points out that 1) a feedback mechanism can enhance reputation effects (Williamson, 1975) and 2) enhanced reputation effects can “attenuate incentives to behave opportunistically” (Williamson, 1991).

It is important to note that in C2C market places, the asset-specificity of transactions is low. It is therefore not certain that existing rating systems will be adequate to decrease opportunism in B2B market places where the transactions are more asset specific. One apparent pitfall is that the peer practitioners who supply the ratings also act opportunistically by rating dishonestly. We therefore hypothesize that for a rating system to have an impact on highly asset specific transactions, it is important that the decision-maker trust the source of the ratings.

In the large anonymous communities of consumer-to-consumer markets the relationship between the user and the rater may not be a major issue. In business-to-business electronic markets, on the other hand, there are two major reasons that make knowing the identity of the rater more important. First of all, the connectivity between the users is higher than in C2C markets. Secondly, it is more difficult to rate vendors of B2B

goods and services (such as a HVAC subcontractor) than vendors of consumer goods such as the seller of a used tennis racket. Both of these differences reinforce Ratnasingham and Kulmar's (2000) argument that in B2B electronic commerce, researchers and professionals must consider the "role of trust between human actors," a notion that goes back at least to Aristotle (the concept of *ethos*).

In this paper, we present a reputation mechanism grounded in source credibility theory, an area of communication science that explicitly studies and formalizes trust between humans. The next section compares different approaches to construct rating mechanisms and this paper ends by reporting on an experiment that compared the added value of this reputation mechanism relative to a standard rating model in the context of the construction industry. The experiment, in which industry practitioners evaluated bids in a business-to-business electronic marketplace, investigated the effect on bid behavior of different rating mechanisms when decision-makers use an electronic market to procure services that are normally procured using a hybrid governance structure. More specifically, the experiment investigated the interaction effect of bidding behavior of the type of rating system and the asset-specificity of the transaction.

3 Alternative Approaches to Reputation Mechanisms in Electronic Commerce

This section discusses alternative approaches to construct reputation mechanisms that support B2B electronic commerce transactions. We provide an overview of three approaches - 1) statistical analysis of past transactions, 2) network of trust models, and 3) rule-based mechanisms - and present the problems of implementing these solutions in B2B electronic commerce.

Several researchers have presented methods that aggregates rating based on statistical analysis of past transactions. The best known methods in this category are collaborative filtering mechanisms. Resnick et al (1994) and Shardanand and Maes (1995) performed pioneering work in this area by creating Internet-based recommender systems that weighs recommendations by the extent to which users agree. It is important to point out that collaborative filtering has been found to be most applicable when a large set of users rate items that are difficult to quantify and describe (e.g., movie ratings,

books) and may therefore be less applicable in B2B e-commerce settings. Dellarocas (2000) applies statistical analysis and clustering to differentiate between honest vs. dishonest raters. Furthermore, Chen and Pal Singh (Chen and Singh, 2001) propose a reputation hierarchy as a means to explicitly calculate rater reputation. Their model takes into account that a rater's expertise may vary depending on the domain that is being rated and calculates the confidence level that can be associated with a rating. The weights are calculated through a propagation of "endorsement" (a function of discrepancy) across raters and groups organized in a hierarchy.

The major problem with methods that statistically analyze past ratings is that they require a substantial amount of data to obtain useful results. This may not be a problem for C2C e-commerce merchants, such as Amazon.com, but could pose substantial difficulties in B2B e-commerce, especially during a start-up phase. Furthermore, it is not certain that the weights calculated by a collaborative filtering mechanism, for instance, are consistent with user expectations. These models tend to over-emphasize the effect of strangers because they ignore personal trust.

Another approach to the construction of rating filters is the "network of trust." Building upon the assumption that people tend to trust the friend of a friend more than someone unknown, several researchers (e.g. (Zacharia et al., 1999)) have proposed formalizing people's "networks of trust" – the concept of trusting a friend of a friend – into rating applications⁶. The strength of such a solution is that it can build upon existing relationships, which could be important for a B2B community in the process of moving from an online to an offline presence. This solution has problems however. It is difficult to measure the trust that a user attributes to the members of his or her "Network of Trust." In interviews with industry decision-makers, we have found that the attitude towards the idea of a network of trust varies considerably. For some interviewees, the concept of trusting "a friend of a friend" seemed intuitive, while others did not consider it to be relevant whether they and an unknown rater turned out to have a common friend. A second problem is that the most common approach (e.g., (Zacharia et al., 1999)) is to use a single dimension to model trust, combining both a party's trustworthiness as a business

⁶ In practice, Epinions.com has deployed a "Web of Trust" mechanism as a rating system for consumer reviews.

partner (Will he cheat in business?) and credibility as an evaluator (Can I trust what he is saying?).

“Rule-based mechanisms” constitute another important type of rating filters. Abdul-Rahman et al. (2000) propose the deployment of rules to determine and update rater weights. These rules assess whom the user trusts based on outcomes of previous interactions. The problem with this approach is that the rules tend to be ad-hoc. For example, if a user A believes that rater B’s rating of Supplier C is inaccurate, should A’s trust in B decrease by 0.6 or 0.4?

We have presented three important criticisms of the applicability of existing rating mechanisms to rating mechanisms in electronic commerce: 1) they require large datasets of rating/transaction data for calibration (statistical analysis of past ratings); 2) they rely on input parameters that were difficult to measure (network of trust); and 3) they rely on ad hoc operators (rule –based mechanisms). We will next show how a rating system based on source credibility has the potential to mitigate all the of these problems

4 Source Credibility Theory

Trust and credibility are two fundamentally different concepts. Fogg and Tseng (1999) define credibility as “*believability or trust in information.*” In the era of modern communication science, Hovland et al. (1953) identified perceived trustworthiness and expertise as the main dimensions of a source’s credibility. The higher the trustworthiness and expertise a source is judged to have, the higher will be the importance given to information coming from that source. Source credibility has been shown to be applicable in commercial settings (e.g., (Birnbaum and Stegner, 1979)), for the evaluation of organizations (Newhagen and Nass, 1989), as well as for the judgment of web content (Fogg and Tseng, 1999), but, hitherto, little research has investigated its applicability in electronic commerce, which is the subject of this article.

As mentioned, source credibility can serve to overcome the three above listed problems associated with alternative approaches to rating mechanisms. First of all, source credibility theory provides tested frameworks (e.g.,(Birnbaum and Stegner, 1979)) for aggregating ratings from different sources. These frameworks decrease the dependence on ad-hoc operators. It also provides validated scales for measuring a source’s (rater’s)

credibility (McCroskey, 1966); these can serve as the key input parameter in a rating system based on source credibility. Finally, the weights in a rating based on source credibility theory depend on user preferences and not on rater behavior, which decreases the amount of data required to calibrate the rating application. The opportunity to measure the credibility of the rater’s organization as well as the person further decreases the amount of user input needed. Table 1 summarizes the opportunities for solving the three major problems of existing rating mechanisms, using a rating system based on source credibility.

Table 1 Summary of how a rating system based on source credibility theory mitigates major problems of alternative rating mechanisms

Alternative Methodology	Key Problem of deploying Alternative Methodology in B2B e-commerce	Opportunity for solution using source credibility based reputation mechanism
Statistical Analysis of Past Ratings	Need large amounts of clean data for calibration	<ol style="list-style-type: none"> 1. Relying on user preferences rather than rater behavior decreases the amount of data needed for calibration. 2. Measuring credibility of the organization further decreases amount of user input needed
Network of Trust	Difficult to measure input parameters	Scientifically validated scales
Rule based mechanisms	Rely on Ad hoc operators for aggregating ratings	Validated aggregation functions

Given that a rating system based on source credibility can address each of the three problems in the table, we investigate how source credibility theory can support a reputation mechanism in B2B electronic commerce. The remainder of this paper first presents TrustBuilder, a prototype-rating tool that operationalizes source credibility theory. Next we discuss the results of an experiment in which industry practitioners deployed this tool to evaluate bids from service providers. The discussion below shows evidence that a rating system based on source credibility can support trust in highly asset specific transactions associated with business-to-business services.

5 TrustBuilder: A rating tool that leverages source credibility theory

5.1 Overview

Williamson (1975) identifies the possibility to create an internal feedback mechanism as one advantage that an internal organization enjoys over market governance. Moreover, he argues that incentives for dishonesty would make it difficult to put such systems into place across organizational borders. In this research project we propose a rating system that is specifically designed to account for the varying credibility of raters from within and outside the user's organization. TrustBuilder is a reputation mechanism grounded in source credibility theory. It is a prototype-rating tool that calculates the weights of ratings of B2B service providers. In this section, we present the first version of the TrustBuilder tool that supports the specific problem of evaluating subcontractors in the construction industry. It employs a three-step process to help the user transform a set of ratings provided by different raters into information that supports the evaluation of subcontractors performance: 1) Credibility input, 2) Calculation of rater weights, and 3) Display of ratings and rater information.

5.2 Step 1: Credibility Input

TrustBuilder applies the validated McCroskey (McCroskey, 1966) twelve-item semantic differential seven-point Likert scale to measure rater credibility. The items on this scale measure two key dimensions of a source's credibility: "authoritativeness" (which corresponds to Hovland's (Hovland et al., 1953) "expertise"), and "character" (Hovland's "trustworthiness"). Based on the results of an earlier experiment (Ekstrom and Bjornsson, 2002), we also included two additional factors in our model of rater credibility. First, TrustBuilder controls for whether the rater is known to the user. Second, TrustBuilder notes whether the rater is in the same organization as the user, even if the two do not know each other. A procurement manager may regard a competitor as not being very reputable, but may still trust those of the competitor's employees whom he knows on a personal basis. Conversely, a manager who does not know a rater may trust a rater more

if she shares the same organizational affiliation. In sum, TrustBuilder uses four different factors to model user *i*'s estimate of rater *j*'s credibility (C_{ij}):

- **Know Rater (KR_{ij}):** Does user *i* know rater *j*? This is a binary measure entered by the user.
- **Same Organization (SO_{ij}):** Do user *i* and rater *j* work for the same organization? The model calculates this binary measure based on the organizational affiliation of the user and the rater.
- **Rater Expertise (X_{ij}):** What is the expertise of rater *j* in the opinion of the user *i*? The calculation of X_{ij} is shown in Table 2 below.
- **Rater Trustworthiness (TW_{ij}):** What is the trustworthiness of rater *j* in the opinion of user *i*? The calculation of TW_{ij} is shown in Table 2 below.

While TrustBuilder models “Know Rater” and “Same Organization” using binary variables, it applies the interval McCroskey scale (McCroskey, 1966) to measure “expertise” and “trustworthiness.” Table 2 shows the scale and its operationalization in TrustBuilder.

Table 2: The McCroskey scale and its operationalization in the TrustBuilder rating tool to model rater expertise and trustworthiness.

Factor	Scale items	Operationalization:
Expertise	Reliable-Unreliable Uninformed – Informed Unqualified – Qualified Intelligent – Unintelligent Valuable – Worthless Expert – Inexpert	$X_{ij} = \sum_{k=1}^{k=6} x_{ijk}$
Trustworthiness	Honest – Dishonest Unfriendly - Friendly Pleasant - Unpleasant Selfish - Unselfish Awful - Nice Virtuous -Sinful	$TW_{ij} = \sum_{k=1}^{k=6} tw_{ijk}$

As Table 2 shows, evaluating rater expertise and trustworthiness is straightforward when the user knows the rater. However, TrustBuilder also calculates

rater expertise and trustworthiness in the event that the user does not know the rater. In this case, the system asks the user to rate two types of “typical” but unknown raters:

1. “**Typical project manager working for contractor X**”: The user rates the expertise and trustworthiness of typical project managers working for each of the contractors whom 1) the user knows, and 2) has supplied ratings to the system. Based on these prototypical ratings, the system can assign a value to raters who are unknown to the user but who works for a contractor with the user is familiar.
2. “**Typical project manager working for a typical California contractor**”: The system uses this baseline rating to assign expertise and trustworthiness values to raters when both the organization and the individual are unknown to the user.

The system calculates the overall scores for all raters on the four factors KR, SO, X, and TW, before normalizing them as z-scores.⁷ The normalization ensures that for each user, all factors will have a mean of zero and a standard deviation of one.

The next step involves converting the credibility measures into an overall value that reflects the user’s assessment of the credibility (or weight) of the different raters. As Equation 1 shows, TrustBuilder employs an exponential function to model rater j’s credibility from user’s i’s perspective (C_{ij}):

$$C_{ij} = \exp(-1 + \beta_{KR_i} KR_{ij} + \beta_{SO_i} SO_{ij} + \beta_{X_i} X_{ij} + \beta_{TW_i} TW_{ij}) \quad (1)$$

where: KR, SO, X and TW are user i’s z-scores for rater i on each of the four factors; and β_{KR_i} , β_{SO_i} , β_{X_i} and β_{TW_i} are coefficients associated with each factor.

5.2.1 Step 2: Calculation of rater weights

The next step is to estimate the coefficients of Equation 1 to calculate rater weights. TrustBuilder uses a methodology of pair-wise comparisons. Figure 1 shows a user interface where the performance of a painting subcontractor (“PaintA”) has been rated by two of the seven raters (see Figure 1).

⁷ z-scores measure a scale reading’s distance from the mean in terms of standard deviations. In this case the mean and standard deviation were calculated for each user and scale item.

UserForm3

Calibrate Weight of Ratings

Instructions
Please assess painting subcontractor Paint A's performance based on two ratings from peer general contractors. Based **only** on the following two ratings of Paint A, please indicate Paint A's expected performance.

First Rating

Rater
Jim Murray
Chief Estimator
Boulder & Whitney

Rating of Paint A
Very Poor

Second Rating

Rater
Andrew Rice
PM
NGC Construction

Rating of Paint A
Very Good

Paint A's expected performance

Very Poor ○ ○ ● ○ ○ ○ ○ ○ ○ ○ **Very Good**

Press OK when you have selected the appropriate level OK

Figure 1: User interface to calibrate weight of ratings through pair-wise comparisons of divergent ratings from different raters.

Rater 1 rated PaintA’s performance as “Good” and Rater 2 rated it as “Poor”. Participants submit their evaluations by clicking a 10-point Likert scale between the values “Very Poor” and “Very Good.” The value $w_{i1,2}$ corresponds to the weight that the user i attributes to Rater 1’s ratings vis-à-vis Rater 2’s. By modeling the credibility of each rater as an exponential function, we obtain the following model for $\hat{w}_{i1,2}$:

$$\hat{w}_{i1,2} = \frac{C_{i1}}{C_{i1} + C_{i2}} \quad (2)$$

where:

$$C_{ij} = \exp(-1 + \beta_{KR_i} KR_{ij} + \beta_{SO_i} SO_{ij} + \beta_{X_i} X_{ij} + \beta_{TW_i} TW_{ij})$$

TrustBuilder can then estimate β_{KR} , β_{SO} , β_X and β_{TW} by minimizing the sum of squares of the errors associated with all pairs (k,l) of raters included in the pair-wise comparisons.

The overall rating (R_{im}) of a subcontractor m from the user i 's perspective will equal the ratings provided by each rater (j) multiplied by i 's estimate of j 's credibility. The result is the following straightforward formula:

$$R_{im} = \sum_j R_{jm} * C_{ij} / \sum_j C_{ij} \tag{3}$$

$$R_{jm} \neq 0$$

5.3 Step 3: Display Ratings and Rater Information

TrustBuilder also displays ratings and rater information. Figure 2 shows an example of the TrustBuilder user interface, which provides the overall ratings for one subcontractor (service provider).

Input I: Ratings and Bids

Trade: *Paving* CSI-Code: 2500
 Bidder: Sigma Marble & Granite, Inc. Bid (\$): 151,400

Overall Ratings (weighted by rater credibility)- Scale 1-10

Criteria	Rating
Schedule	8
Quality	8
Collaboration	7
Change Orders	7
Administration	7
Experience	8
Hire Again	8

Bids

Competitor	Bid (\$)
Sigma Marble & Granite, Inc.	\$151,400
Competitor 1	\$164,944
Competitor 2	\$186,035
Competitor 3	\$166,471
Competitor 4	\$185,709

Input II: Rater Information

The CredRate ratings above are calculated based on ratings from the following raters:

Name	Title	Company	Rater Weight in Overall Ratings
Jim Murray	Chief Estimator	Boulder & Whitney	56%
Paul Owen	Project Manager	Boulder & Whitney	7%
Philip Holmes	Project Manager	NGC Construction	4%
Charlene Lindgren	Estimator	NGC Construction	4%

Overall Rater Agreement: **High**

Task I: Evaluation

How qualified is Sigma Marble & Granite, Inc. to do this job?
 Very Unqualified Very Qualified

How confident are you in your judgement?
 Very Unconfident Very Confident

How comfortable are you hiring Sigma Marble & Granite, Inc. to do this job?
 Very Uncomfortable Very Comfortable

Task II: Contingency Adjustment

Bid (\$): 151,400 Please Enter Contingency (%): 5 Final Estimate (\$): 158970

Figure 2: User interface showing the ratings of a subcontractor on seven criteria. The user evaluates the subcontractor (Task I) and can add contingency (risk premium) to the bid (Task II).

TrustBuilder displays ratings for seven different criteria which all involve qualitative judgment: schedule, quality, collaboration, change orders, administration, experience, and “hire again”. Peer industry practitioners provide the ratings by indicating on ten-point Likert scales the extent to which they agree with statements such as: “I would be willing to hire SubA to work for me again.” The TrustBuilder tool also shows the identity of each rater along with his/her relative weight in the overall ratings, as well as measures of rater agreement and total rater credibility.

6 An experiment investigating the impact of rating systems and asset-specificity

6.1 Introduction

Below we present the results of an experiment that investigated the effect of different rating mechanisms when decision-makers use an electronic market to procure services normally are normally procured in a hybrid governance structure. The purpose of the experiment was to study the effect on procurement decisions of a rating system based on source credibility for transactions with varying degrees of asset-specificity. The experiment compared the performance of two different rating models:

- *Credibility-weighted tool*: The TrustBuilder model
- *Unweighted Rating tool*

In this experiment we have chosen the construction industry as the field of study. We therefore begin by discussing a construction project from a transaction cost perspective. The next section presents an experiment in which construction industry professionals based procurement decisions on actual ratings of subcontractors.

6.2 Transaction Costs and Governance Structure in the Construction Industry

One industry in which entrepreneurs, as well as investors, envisioned reaping substantial benefits from electronic commerce is the large and fragmented Architecture

Engineering Construction (AEC) industry. In April 2000, approximately \$1 billion had been invested in approximately 200 AEC e-commerce start-ups (Bass, 2000). These companies planned to facilitate transactions between the approximately 700,000 U.S. contractors and subcontractors. Similar to other industries, the number of transactions conducted in these new AEC e-marketplaces turned out to be very low, and only a handful of them were still in business as of April 2002 (Fuscaldo, 2002). It is therefore interesting to investigate the impact of rating applications on the transaction costs in the construction industry.

Gunnarson & Levitt (Gunnarson and Levitt, 1982) argue that an optimum governance system in the construction industry can be viewed as a function of asset-specificity. A first step towards analyzing the subcontracting of the trades that constitute a construction project from a transaction cost perspective is to categorize them in terms of asset-specificity. Williamson differentiates between six types of asset-specificity several of which can apply for each of many physical intellectual and human assets needed to complete a construction project. The six types include 1) site or location specificity, 2) physical asset-specificity (e.g., a specialized crane), 3) human asset-specificity or “learning by doing”, 4) brand name capital, 5) dedicated assets in the form of discrete investments specific to the relation with a particular transaction partner, and 6) temporal specificity. In the construction industry, brand name capital (4) and discrete investments (5) are less applicable, while the remaining four types of asset-specificity are commonplace. When asset-specificity is high, one or both parties is “locked into” the transaction. One noteworthy example of physical asset-specificity in the AEC-industry is the customized exterior wall paneling of the Guggenheim museum in Bilbao, Spain. The custom-made wall panels would have little value on the open market. An example of location asset-specificity is a job site where there is only one company that can deliver ready-mixed concrete within an acceptable delivery radius (Gunnarson and Levitt, 1982). The general contractor would then have the choice of making the concrete on-site or face a possible “hold-up” situation where the concrete manufacturer can take advantage of being a monopoly. Temporal specificity is similar to Thompson’s (1967) definition of sequential interdependence. In the construction industry a hybrid governance structure is the typical mode in which contractors subcontract services. Eccles (Eccles, 1981) has

shown that, even though subcontractors and contractors are legally independent business entities, the participants tend to form close long-term relations (Costantino and Pietroforte, 2002), a relationship that almost constitutes a “quasi-firm.”

However, a typical construction project comprises a large number of transactions that concern assets of varying asset specificities. Table 3 provides a classification in terms of asset-specificity of the procurement activities required in a construction project. The procurement of commodity products, such as lumber or kitchen appliances, involves little or no asset-specificity. The relationship between the buyer and the seller is at an arm’s length basis where the buyer, for each transaction, chooses the supplier with the best trade off between product, price and availability. Certain services, such as painting, can also be seen as commodities. However, there is still an element of asset-specificity as the buyer is constrained by the availability of service providers in the local market. The governance structure is therefore a mix of hybrid and market. Other services are somewhat more specialized and the transactions give rise to human asset-specificity. For instance, key personnel at a general contractor and a controls subcontractor can leverage the lessons learned from working together on one project to the next project. A general contractor often employs the same subcontractor repeatedly which gives rise to the “quasi-firm” (Eccles, 1981) hybrid structure. Subcontracts that are on the critical path, such as structural steel, also involve temporal asset-specificity of sequential interdependence (or reciprocal dependence if the dependence is bi-directional). Schedule interdependence has empirically been shown to be a very important source of asset-specificity in the shipbuilding industry (Masden et al., 1991), which is very similar to AEC. In view of the potential impact of these trades on overall project profitability, some general contractors have opted to internalize them.

Another example of a hierarchical organization of highly asset specific transactions is a marine subcontractor, who purchases their own high capacity floating cranes to avoid potential hold-up (Gunnarson and Levitt, 1982). For highly specialized trades, we therefore categorize the typical governance structure as in between hybrid and market. Finally, in the case of integrated product design and construction the asset-specificity becomes even higher. For a specialized design-build general contractor, the design is a key determinant of project performance in terms of time and cost, as well as

quality. Furthermore, it is difficult to specify and measure the quality of the architect’s design until construction is completed. As a result, it is not surprising that many specialized design-build contractors choose to internalize the design.

Table 3: Categorization of transactions in the construction industry in terms of asset-specificity and governance structure. Hybrid is the typical mode of governance but there are variations.

Type of AEC procurement	Commodity Products	Commoditized services	Somewhat specialized services	Specialized services	Integrated product design and construction
Examples in AEC transaction	Procurement of lumber	Sub-contracting of Painting	Sub-contracting of Controls	Sub-contracting of Structural Steel	Specialized Design/Build Contracts
Asset-specificity	Low	Medium/Low	Medium	Medium/High	High
Sources of significant asset-specificity	None	Locational	Locational Human	Locational Human, Temporal Physical	Locational Human, Temporal
Typical Governance Structure	Free Market	Hybrid/Market	Hybrid	Hybrid/Integrated	Hierarchy (Internal Organization)

In this paper we investigate whether the use of different rating tools can decrease the costs of free market transactions of varying asset specificity for subcontracting in construction. Such a decrease could in turn lead to an increased use of free market governance. Having presented the setting of the experiment we will now present the specific research hypotheses, which investigate the impact of different rating mechanisms on procurement decisions involving varying asset-specificity.

6.3 Research Hypotheses

The experiment investigated three fundamental research hypotheses. The first hypothesis relates to the validity of the credibility-weighted rating model while the last two hypotheses investigate user behavior given the type of rating tool used and the asset-specificity of the transaction.

Hypothesis 1: The factors used in the credibility weighted model influence rater weight.

Insignificant coefficients for a factor in the credibility-weighted model (that is, those that are close to zero in Equation 1) indicate that it is not relevant to the model or is heavily correlated with other factors.

Secondly, the experiment investigates how the output of the two rating applications affects the perceived transaction cost of hiring an unknown subcontractor. In the construction industry, standard practice is for the estimator to add a contingency or risk premium to a subcontractor's bid. In this experiment, we used the contingency estimates as a proxy for transaction costs since they reflect the participants' estimates of the added cost associated with hiring an unknown subcontractor. As a result, it was possible to study the extent to which the output of the different rating mechanisms affected the estimated transaction costs. Or, in other words, would the fact that the subcontractor had received a high rating lead the user to add a lower contingency? Of particular interest is whether there are any differences depending on the asset-specificity of the transaction involved. For transactions with low asset-specificity, we expect both types of rating tools to mitigate the transaction costs:

Hypothesis 2: For transactions with low asset-specificity, evaluations based on ratings from both a credibility-weighted tool and an unweighted tool will be negatively correlated to the contingency added to the bids.

For highly asset specific transactions, on the other hand, we expect the credibility-weighted model to perform better than the simple unweighted model:

Hypothesis 3: For transactions with low asset-specificity, evaluations based on ratings from a credibility weighted will affect bid contingency more than do evaluations based on ratings from an unweighted tool.

Having posited the three research hypotheses, we will now discuss the method used in our experiment to investigate them.

6.4 Experimental Method

Fifteen construction industry professionals who worked for three California general contractors participated in the experiment. All of the participants were actively involved in evaluating bidding subcontractors, and each participant knew at least two of the other participants.

The users evaluated a set of real bids from the subcontractors that had been hired to construct a San Francisco office building in 2001. In total, the experiment involved twenty-six subcontractors bidding to perform the sixteen different trades that were subcontracted on the \$3M office building.

To provide a set of ratings of the subcontractors' performance, the participants had, prior to the experiment, rated the twenty-six subcontractors on seven different criteria using ten-item Likert scales. The experiment was a within-participant design that was carried out on an individual basis using a personal computer. The participants calibrated the TrustBuilder tool by first rating their peer raters on the McCroskey credibility scale (see Table 2), and then making the pair-wise comparisons to allow the tool to calculate rater weights.

In the next step, the participant used the two tools to evaluate the subcontractors. Half of the participants used TrustBuilder (see Figure 2) to evaluate thirteen of the subcontractors before using the unweighted tool to evaluate the remaining half; the other participants used the tools in reverse order. The unweighted rating tool was a simplified version of the weighted tool. It showed only the average ratings and rater agreement along with the number of raters. In both tools, the subcontractors' low bids that were presented to the participants were roughly equal to those of the original project. The two tools also displayed the bids of four competing subcontractors. It is important to note that the name of the low bidding subcontractor had been changed to prevent the participants from recognizing the subcontractors and thus evaluating them based on previous experience. In both tools, the participants evaluated overall subcontractor quality, stated how confident they were in their judgment as well as how comfortable they were hiring the subcontractor, and adjusted the subcontractor's low bid by adding a risk buffer.

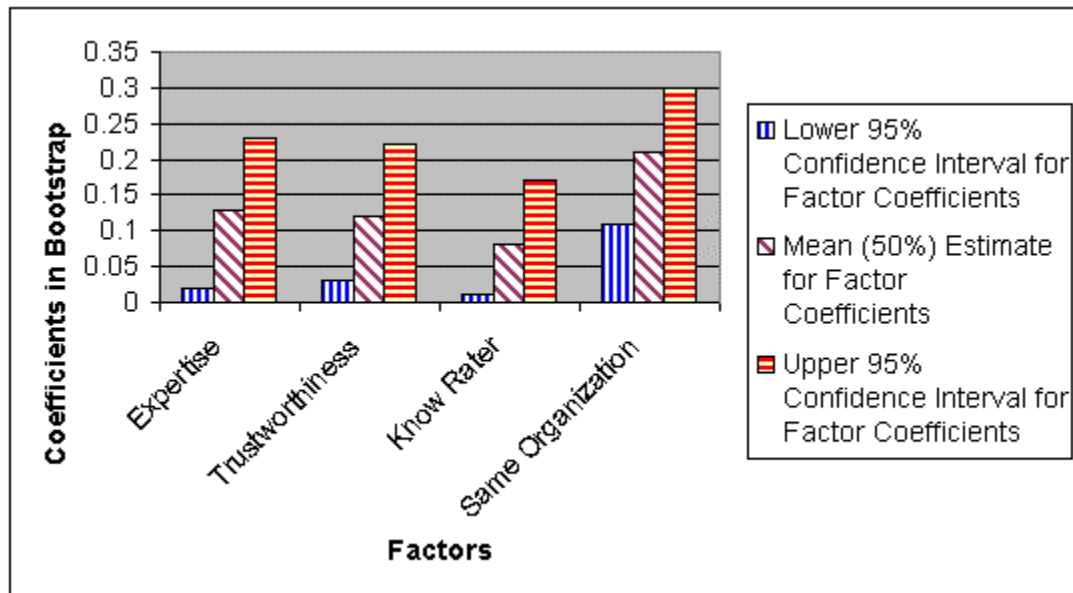
6.5 Results

6.5.1 Significance of factors in credibility model

Consistent with Hypothesis 1, the results showed, that the four factors (“Know Rater”, “Same Organization”, “Trustworthiness”, and “Expertise”) proposed in the TrustBuilder model were all statistically significant. This conclusion was the outcome of a bootstrap analysis in which a set of fifteen users was randomly sampled (with replacement). The program running the analysis then performed the exponential regression (to estimate the coefficients in Equation 1) based on the 315 comparisons provided by the fifteen users in the sample. The program performed this procedure 2000 times to test the statistical significance of the estimates. Figure 3 shows that all four factors were positive within a 95% confidence interval in the bootstrap analysis.

The current results provide evidence that the two classical factors in source credibility theory, perceived expertise and trustworthiness, contribute to the prediction of rater weights in an AEC rating application. The new factors included in this study, whether the rater knows the user and whether the user and rater are in the same organization, also influenced user assessments of credibility. These results are particularly striking given that the different factors are by nature correlated. For example, the fact that the user knows a rater increases the likelihood that the two will work for the same organization, and makes it more probable that the user will find the rater trustworthy and competent.

Figure 3 Results from a bootstrap analysis, showing the factor coefficients in the exponential regression of rater weights. As shown, all coefficients are positive in the 95% confidence interval. The results show that all factors in the model (including perceived expertise and trustworthiness from source credibility theory) are significant predictors of rater weight. The result indicates that the proposed operationalization of source credibility is valid.



6.5.2 Asset-specificity and rating tools

To analyze the impact of asset-specificity, we first classified the different trades by their asset specificities by rating each trade on a 3-point scale (H/M/L) in terms of the four types of asset-specificity (locational, human, physical, and temporal) that come into play when a construction project is subcontracted. To obtain an overall measure of asset-specificity, we simply totaled the scores for the three measures, having converted the H/M/L to the numerical scores 0/1/2. We then classified the eight bids with the highest overall scores as transactions with high asset-specificity. These highly asset specific transactions (belonging to the categories “specialized services” and “somewhat specialized services” in Table 3) were associated with the trades Plumbing and HVAC,

Controls, HVAC units, and Electrical. Similarly, the eight bids⁸ with the lowest asset-specificity were classified as transactions with low asset-specificity. The bids with low asset-specificity belonged to the trades Door frames, Glazing, Carpet, Painting, and Fire Extinguishers (categorized as “somewhat specialized services” and “specialized services” in Table 3). The appendix shows the ranking of the different trades in terms of asset-specificity. We will now discuss how the output of the credibility-weighted and unweighted rating tools affected the contingency (or risk buffer) that the participants added to the bids from the unknown subcontractors.

6.5.2.1 Trades with low asset-specificity

For transactions with low asset-specificity, we can make two interesting observations. Firstly, in accordance with Hypothesis 2, the output of both types of rating tools influences decision-makers purchasing decisions. Secondly, the effects appear to be equivalent for both types of rating systems. Figure 4 shows the relationship between contingency and overall qualification, or the output of the two types of rating systems, as predicted by a linear regression model for transactions with low asset-specificity.

In order to obtain the same number of transactions in the two categories, we randomly selected 2 of the 5 bids which were tied for 7th place in terms of lowest asset-specificity and characterized them in the categorized them as transactions with low asset-specificity.

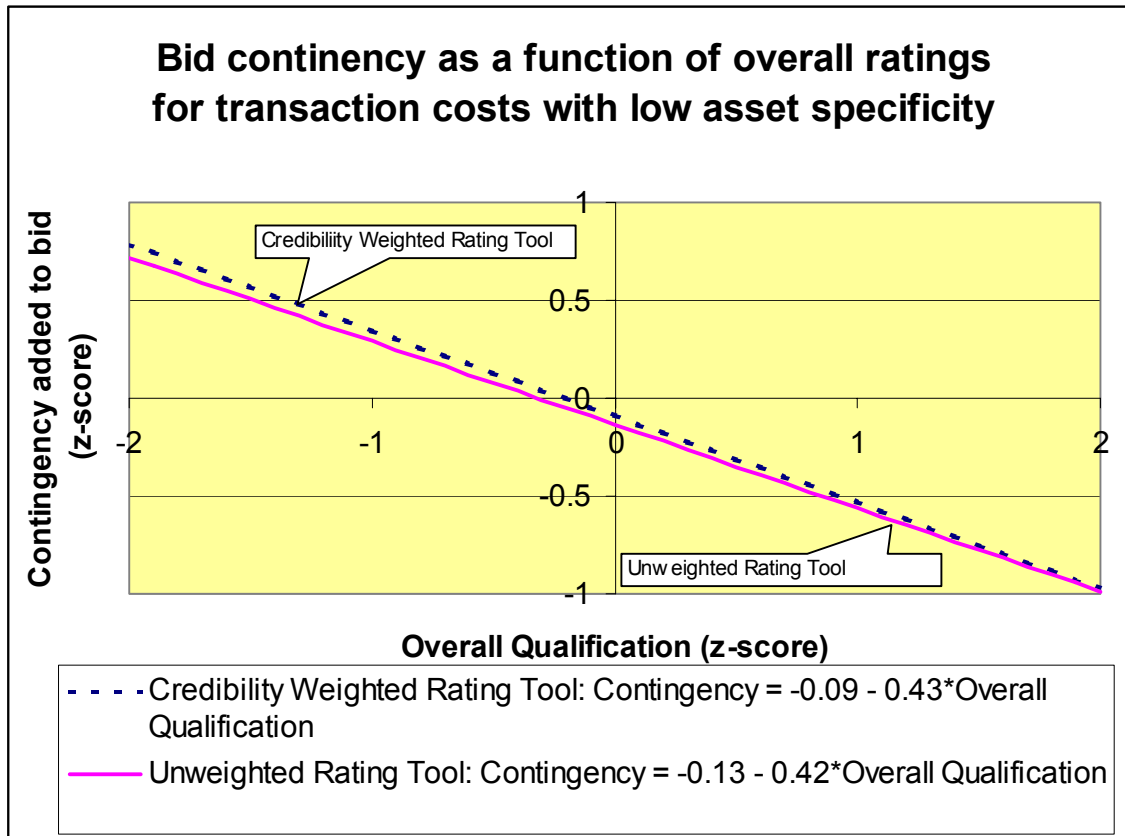


Figure 4 Relationship between contingency (risk buffer) and overall qualification as predicted by linear regression for services with low asset-specificity. Both types of rating tools significantly impact the contingency added to the bids. For transaction with low asset-specificity, both types of rating tools can help to decrease the transaction costs associated with opportunism.

The regression showed a statistically significant relationship between contingency and overall qualification for both the credibility-weighted tool (Contingency = $-0.09 - 0.43 \cdot \text{Overall Qualification}$, $t(46) = -4.66$, $p < .0005$) and the unweighted tool (Contingency = $-0.13 - 0.42 \cdot \text{Overall Qualification}$, $t(62) = -4.23$, $p < .0001$.) We can therefore conclude that Hypothesis 2 does not seem to be valid. For transactions with low asset-specificity, the output of a rating system seems to influence purchasing decisions independently of the type of rating system that is used.

6.5.2.2 Highly Asset-Specific Transactions

For transactions with high asset-specificity, the results showed decision-maker behavior to depend on the type of rating tool in use. The ratings produced by the credibility-weighted tool had a significant impact on bid contingency. The coefficient of bid contingency as a function of overall qualification was strictly negative (t-test: $\beta = -.32$, $t(62) = -2.56$, $p < .05$). For the unweighted tool, on the other hand, the analysis showed no significant impact on user decisions ($\beta = -.005$). We performed a bootstrap analysis, to further investigate the difference in the impact of the different rating mechanisms for highly asset specific transactions. In each iteration, the bootstrap algorithm produced two random samples (one for each type of tool) on which it calculated a linear regression model of bid contingency as a function of overall qualification. Next, it calculated the difference between the coefficients for the credibility weighted and the unweighted model. The bootstrap analysis showed that the difference between the coefficients in the models of bid contingency as a function of rating output was strictly positive (Bootstrap (1000 iterations), $M = 0.32$, $p < 0.05$). Consistent with Hypothesis 3, the output of the credibility weighted rating system had a significantly greater effect on the contingency added to the bids than did the output of the unweighted system.

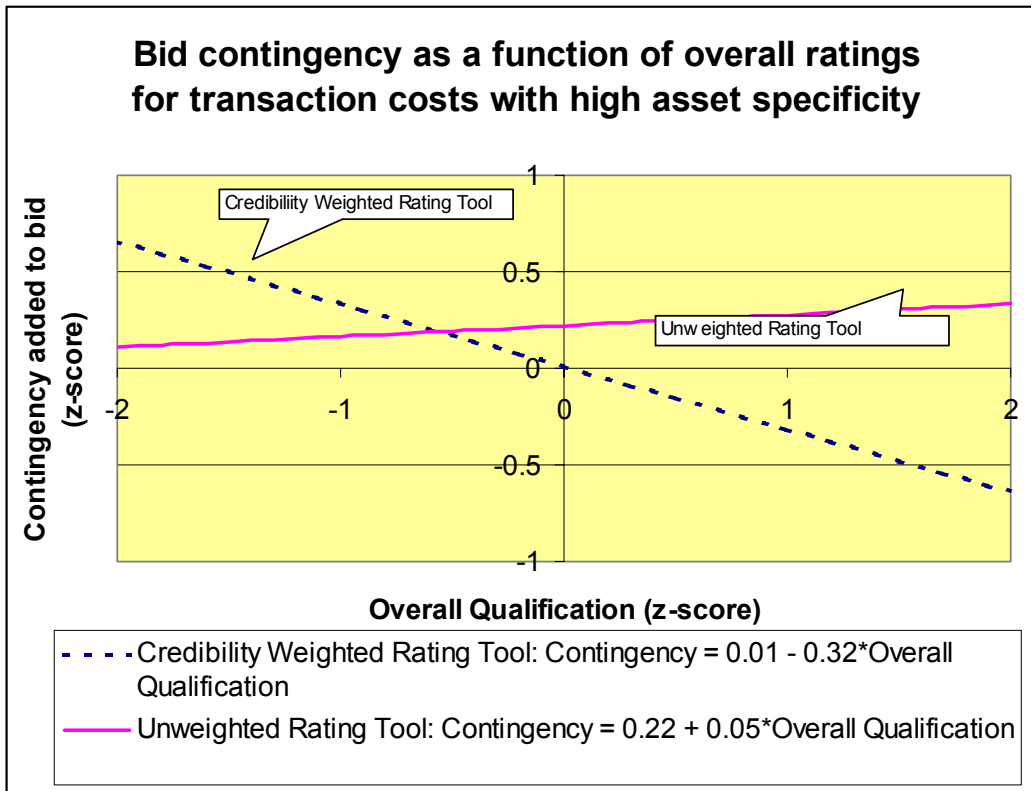


Figure 5 Relationship between contingency (risk buffer) and overall qualification as predicted by linear regression for services with high asset-specificity. Only the credibility-weighted rating tool impact the contingency added to the bids. We conclude that for transaction with high asset-specificity, only the credibility weighted rating tool can help to decrease the transaction costs.

6.5.2.3 Predicted impact of different rating applications on transaction costs in the construction industry

Given that we have used contingency as a proxy for transaction costs, the above presented results form the basis for an analysis of the impact of the two different types of rating tools in a construction industry setting. We will provide a numerical as well as graphical illustration of the impact of these results.

A numerical example, in which we convert the results from z-scores to percentages, serves to quantify the difference in impact of the two rating models for transactions of high and low asset-specificity. On average the standard deviations for the

participants estimates equaled 1.68 (units) for overall subcontractor and 3.51 (%) for the bid contingency added to the bids. Starting with a transaction with high asset-specificity, let us assume that the typical user is evaluating an unknown subcontractor that is bidding on a \$200,000 HVAC subcontract. We also assume that the output of the credibility-weighted rating tool makes the user assess the subcontractor's overall qualification to be 7 rather than 5 that she typically assigns to subcontractors, for which she has no ratings (i.e., an increase of $\frac{7-5}{1.68}$ standard deviations.) We would then expect the user to alter the contingency added to bid by $\frac{7-5}{1.68} * -0.31 * 3.51\% = -1.30\%$ as result of the output of the rating tool. For the \$200,000 HVAC subcontractor this difference would equal a decrease in the estimated transaction costs by \$2,400. If the user had instead used the unweighted tool, it is unlikely that there would have been any impact on the contingency for the highly asset specific HVAC transaction (as Figure 5 illustrates). If, on the other hand, painting (a trade of low asset-specificity) had constituted the subcontract we would expect identical impacts for both types of rating tools. In the scenario, the regression coefficient of -.42 for the unweighted tool would then imply an 1.76% decrease in the added contingency, which is basically identical to the result for the credibility-weighted tool (coefficient = -.43, change in contingency = -1.80%).

Figure 6 illustrates graphically, how the two rating systems can affect the transactional hazards by enhancing the reputation effects and thus decrease the contingency added to the bids. It shows the transaction costs associated with market governance structures as a function of the asset-specificity (k) and the type of rating tool. For the market structure (TC Market (k)) there are three cost curves: 1) without rating mechanism, 2) with an unweighted rating mechanism, and 3) with a credibility-weighted rating mechanism. These three cost curves illustrate the major findings of the experiment:

- 1) For trades with low asset-specificity (e.g., painting) both an unweighted and a credibility-weighted tool decrease transaction costs (Hypothesis 2).
- 2) For trades with high asset-specificity (e.g, HVAC) only the credibility-weighted rating system decreases transaction costs (Hypothesis 3).

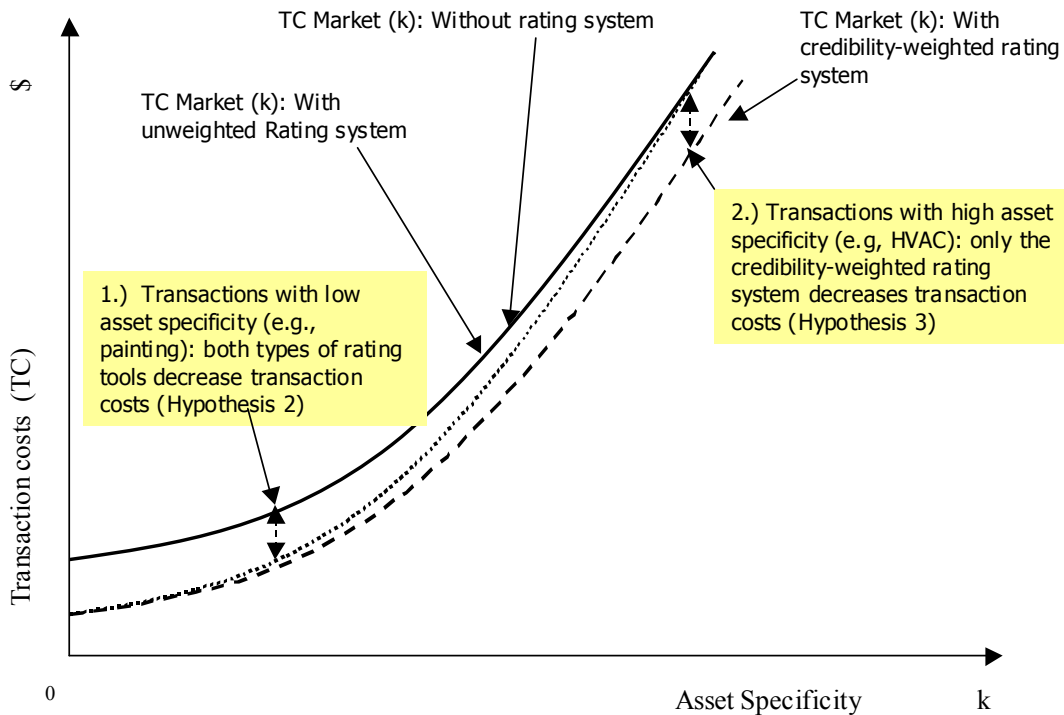


Figure 6 Predicted impact of the adoption of rating applications on the transaction costs in a market governance structure. 1) For trades with low asset-specificity (e.g., painting) both an unweighted and a credibility-weighted tool decrease transaction costs. 2) For trades with high asset-specificity (e.g., controls) only the credibility-weighted rating system decreases transaction costs.

One important conclusion is that B2B electronic commerce providers of highly asset specific product and services should consider supporting their customers' decisions through more refined rating mechanism than the simple unweighted solutions, which are found in consumer electronic commerce.

7 Discussion

This study casts light on the relationship between the asset-specificity of a transaction and the impact of different rating applications on decision-maker behavior. It also suggests an explanation for the slower adoption rate for rating applications in business-to-business electronic commerce compared to consumer electronic commerce. The rating models that have been successfully adopted in consumer electronic commerce

do not address the role of trust between humans. These solutions may be applicable for business transactions involving low asset-specificity but are less likely to be applicable in highly asset specific business-to-business transactions, given that they do not sufficiently account for the importance of human relationships. This study points to the possibility of basing a rating system on source credibility as a means to support decision maker behavior for transactions with high asset-specificity. We argue that the strength of a rating tool operationalizing source credibility theory is that it 1) incorporates tested frameworks for aggregating information, 2) applies validated scales for measuring the input parameters, and 3) does not require large amounts of data for calibration.

It is also interesting to discuss the effects that the introduction of different rating mechanisms can have on the governance structure of a market. Figure 7 shows how the introduction of rating systems could affect the governance structure of the construction industry. The curve TC Hybrid I (k) represents the transaction costs of the currently prevailing hybrid governance structure in the construction industry. As the introduction of a rating system brings down the cost of free market transactions, Figure 7 shows a move towards an increased free market governance at the expense of hybrid governance.

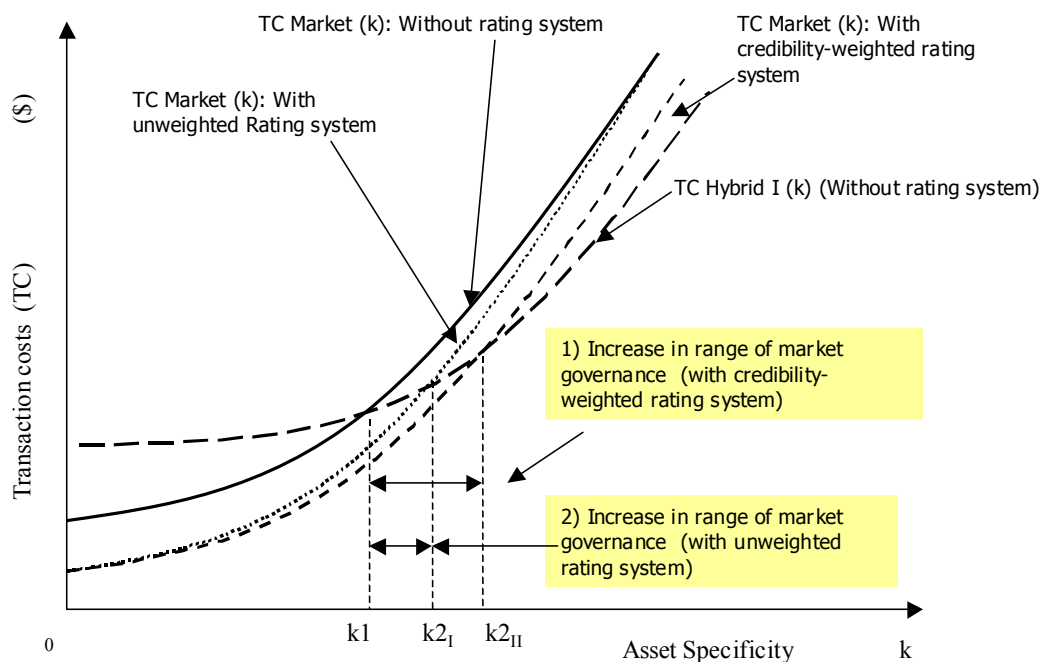


Figure 7 Predicted impact of different types of rating mechanisms on market governance structure in the construction industry. The deployment of a credibility-weighted rating system (1) will lead to a

relatively increase in the range of market governance than does the deployment of an unweighted rating system (2). .(Adapted from Williamson (1991))

However this effect will not be equal for both types of rating systems. A likely consequence is that the deployment of a credibility-weighted rating system will lead to a relatively larger increase in the range of market governance than does the deployment of an unweighted rating system. We therefore conclude that rating mechanisms, which account for trust between humans, is a likely prerequisite for electronic commerce to bring about a substantial shift towards increased market governance.

This study opens several avenues for further research. First, this experiment investigated the influence of two types of rating mechanisms in a setting where decision-makers were asked to purchase free market services that they would normally procure using some type of hybrid governance such as partnerships or repeat contracting. However, Williamson (1991) argues that reputation effects will decrease the cost of governance for transactions involving a medium to high degree of asset-specificity. As a result, *ceteris paribus*, the impact of the improved reputation effect will lead to increased hybrid contracting relative to hierarchies. In AEC, a general contractor may choose to form a partnership with a structural steel subcontractor instead of performing the structural steel in-house.

It is also important to distinguish between asset-specificity and the size of the purchasing contract. In this case, the data did not allow us to control for the size of the contract as they were substantially correlated. An alternative explanation is therefore, that decision makers become more risk averse as the size of the contract increases, and therefore only let the more advanced rating tool affect their decisions. Further studies, should therefore control for contract size by providing data in which asset-specificity is independent of contract size.

This research project shows that is possible to empirically investigate the added value of rating mechanisms from a user perspective. While theoretical rating models that support electronic commerce abound, researchers have given little attention to how these can support the end user's decisions. Experiments with industry practioners, would further contribute to knowledge in this area.

Another interesting avenue for further research would be to combine a rating filter based on source credibility with data analytical methods. A rating mechanism could evaluate raters that the user knows based on source credibility, while applying collaborative filtering or statistical methods to differentiate between unknown raters. Researchers and practitioners should be aware of the opportunity for incorporating frameworks developed in social science in technologies that support human interaction across the Internet. There is no reason that the substantial body of research studying areas, such as trust, credibility, and reputation, should not be valid also in online situations.

Finally, the proposed mechanism can also potentially prevent rater dishonesty, since any deceitful ratings will primarily hurt people who know and trust the rater. However, the purpose of a credibility weighted rating mechanism is not to enforce truthful behavior in an Internet rating system. A prerequisite for a functioning rating system is that a substantial fraction of the participants in the system supply honest evaluations. The behavior of the participating industry practitioners during our experiment shows that this assumption is not unrealistic. Therefore, by rating the rater, the participants of a business-to-business electronic market can come one step closer to trusting new business partners, which is a prerequisite for leveraging Internet technologies to increase competition and market transparency.

8 Appendix

Table 4 Operationalization of asset specificity for the subcontracted trades included in the experiment.

Asset-specificity	Architectural Woodwork	Metal Fabrications (Tube and Ornamental)	Metal Fabrications (PolyCarb Doors)	DoorFrames/ -Hardware	Glazing	Drywall and PolyCarb walls	Carpet	Painting, including Floors	Fire Extinguishers	Window Treatments	Fire Sprinklers	Plumbing and HVAC	Controls	Install D/F/H	HVAC units	Electrical	Lighting
Locational	M	M	M	L	M	M	M	M	L	M	L	H	H	M	M	M	M
Human	M	M	M	L	M	L	L	M	L	M	L	H	H	L	M	H	M
Physical	M	M	M	M	L	M	L	L	L	L	M	M	M	M	H	M	L
Temporal	L	L	L	L	L	L	L	L	L	L	M	M	M	L	M	M	M
Total Asset-specificity	3	3	3	1	2	2	1	2	0	2	2	6	6	2	5	5	3
Rank	5	5	5	15	9	9	15	9	17	9	9	1	1	9	3	3	5

9 References

- Abdul-Rahman, A. and Hailes, S. (2000) In *Hawaii's International Conference on Systems Sciences* IEEE, Maui, Hawaii.
- Bakos, Y. J. (1997) Reducing Buyer Search Costs: Implications for Electronic Commerce, *Management Science*, **43**, 1676-1992.
- Bass, C. (2000) .
- Birnbaum, M. H. and Stegner, S. E. (1979) Source Credibility in Social Judgment: Bias, Expertise and the Judge's point of view, *Journal of Personality and Social Psychology*, **37**, 48-74.
- Chen, M. and Singh, J. P. (2001) In *ACM Conference on Electronic Commerce (EC'01)*(Ed, Shoham, Y.) Tampa.
- Coase (1937) The Nature of the Firm,.
- Costantino, N. and Pietroforte, R. (2002) Subcontracting practices in USA homebuilding - an empirical verification of Eccle's findings 20 years later, *European Journal of Purchasing and Supply Management*, **8**.
- Dellarocas, C. (2000) In *Proceedings of the 2nd ACM Conference on Electronic Commerce* Minneapolis, Mn.
- Dewan, S. and Hsu, V. (2001) .
- Eccles, R. G. (1981) The Quasifirm in the construction industry, *Journal of Economic Behavior and Organization*, **2**, 335-337.
- Ekstrom, M. and Bjornsson, H. (2002) In *CIB W65 Symposium September 2002*Cincinnati, Ohio, USA.
- Fogg, B. and Tseng, H. (1999) In *CHI 99*ACM, Pittsburgh.
- Fuscaldo, D. (2002) Building a Market, *The Wall Street Journal*, April 15, 2002.
- Gunnarson, S. and Levitt, R. E. (1982) In *7th Internet Congress* Copenhagen.
- Houser, D. and Wooders, J. (2000) Department of Economics, University of Arizona, .
- Hovland, C. I., Janis, I. L. and Kelley, H. H. (1953) *Communication and Persuasion*, Yale University Press, New Haven.

- Kafka, S. J., Temkin, B. D., Sanders, M. R., Sharrard, J. and Brown, T. O. (2001) Forrester Research, Cambridge.
- Lucking-Reiley, D., Bryan, D., Prasad, N. and Reeves, D. (2000) .
- Malone, T., Rockart, J., (1992) Information Technology and the New Organization, *IEEE*, , 636-643.
- Masden, S. E., Jr, J. W. M. and Snyder, E. A. (1991) The Costs of Organization, *Journal of Law ,Economics and Organization*, **4**, 181-198.
- McCroskey, J. C. (1966) Scales for the measurement of ethos, *Speech Monographs*, **33**, 65-72.
- Newhagen, J. and Nass, C. (1989) Differential Criteria for Evaluating Credibility of Newspapers and TV News, *Journalism Quarterly*, **66**, 277-284.
- Ratnasingham, P. and Kulmar, K. (2000) In *International Conference on Information Systems (ICIS-2000)* Brisbane, Australia.
- Resnick, P., Iacovou, N., Suchak, M., Bergstrom, P. and Riedl, J. (1994) In *Proceedings of ACM 1994 Conference on Computer Supported Cooperative Work*.
- Resnick, P. and Zeckhauser, R. (2001) Trust Among Strangers in Internet Transactions: Empirical Analysis of eBay's Reputation System, *Journal of Industrial Economics*, **Forthcoming**.
- Shardanand, U. and Maes, P. (1995) In *ACM Special Interest Group on Computer-Human Interaction*, Vol. 37 ACM Press/Addison-Wesley Publishing Co.
- Thompson, J. D. (1967) *Organizations in action*, McGraw-Hill Book Company, New York.
- Williamson, O. E. (1975) *Markets and Hierarchies: Analysis and Antitrust Implications*, The Free Press, New York.
- Williamson, O. E. (1991) Comparative Economic Organization: The Analysis of Discrete Structural Alternatives, *Administrative Science Quarterly*, **36**, 269-296.
- Williamson, O. E. (1991``) Comparative Economic Organization: The Analysis of Discrete Structural Alternatives", *Administrative Science Quarterly*, **36**, 269-296.
- Zacharia, G., Moukas, A. and Maes, P. (1999) In *Thirty-second Annual Hawaii International Conference on System Sciences (HICSS-32)* Wailea, Hawaii.