

# QUALITY-ADJUSTED CONSUMER SURPLUS FOR ONLINE LABOR MARKETS WITH ASYMMETRIC INFORMATION

*Completed Research Paper*

**Kevin Yili Hong**

Temple University

1810 N. 13<sup>th</sup> St, Philadelphia, PA 19122

hong@temple.edu

**Paul A. Pavlou**

Temple University

1810 N. 13<sup>th</sup> St, Philadelphia, PA 19122

pavlou@temple.edu

**Pei-yu Chen**

Arizona State University

University Drive and Mill Avenue, Tempe, AZ 85287

peiyu.chen@asu.edu

## Abstract

*Traditional measures of consumer surplus (CS) have implicitly assumed that the quality expected is the same as the quality that is paid for ex ante. However, when product or service quality cannot be perfectly verified ex ante by consumers in markets with asymmetric information, and actual quality received may not necessarily equal quality expected, CS would not be precisely measured. In this paper, we propose a quality-adjusted measure of CS for IT outsourcing e-markets with asymmetric information. We first relax the assumption that consumers always receive the quality they expect ex ante. Second, we leverage expectation-confirmation theory to construct the utility function to derive the proposed quality-adjusted measure of CS. Measurement development is followed by an empirical study of the effects of different measures of CS. We found the quality-adjusted measure better predicts market outcomes, i.e., continuation, subsequent projects and payments.*

**Keywords:** Consumer surplus, online labor markets, asymmetric information, market performance

## Introduction

Consumer surplus (CS), as an important metric of societal welfare, has long been of interest to economists (Breslaw and Smith 1995; Hausman and Newey 1995; Marshall 1920; Song 2007; Vartia 1983). However, how to precisely capture CS still remains a challenge (Bapna et al. 2008; Irvine and Sims 1998). Important as it is, methods to empirically measure CS are surprisingly scarce. The study of CS traces back to the early 20<sup>th</sup> century where it was defined as the difference between the maximum price a consumer is willing to pay (WTP) for a product and the actual price paid ( $p$ ). The literature has mainly focused on demand estimation of commodity products, in which CS is graphically represented by the integral above the equilibrium price and below the demand curve, without substantive consideration for *ex-post* consumer satisfaction. Since WTP is usually conceptualized as a pre-purchase concept, and the equilibrium of supply-demand is based on the WTP at the time of transaction, instead of after consumption, it creates a theoretical problem for CS because it is supposed to capture *ex post* utility derived from a product over the utility the paid price can offer. Indeed, Alfred Marshall's (1920) original definition of CS centered on the concept of satisfaction, as his seminal work (Marshall 1920) stated:

... [the consumer] derives from a purchase a surplus of **satisfaction**. The excess of the price which he would be willing to pay rather than go without the thing, over that which he actually does pay, is the economic measure of this surplus of **satisfaction**. It may be called consumer's surplus. [p. 124]

Later Michael Burns (1973) reconsidered Marshall's terminology of CS, and proposed the expression of "gain in utility" to substitute "surplus of satisfaction" to make it measurable. Generally, demand estimation methods are used to infer CS (Minjae Song 2007). However, demand estimation methods may not be useful for highly heterogeneous niche products or idiosyncratic services.<sup>1</sup> Furthermore, the measurement of CS generally focuses on a "pre-transaction" concept. CS is traditionally measured as  $WTP - p$ , assuming quality to be known *ex-ante* (or actual quality is what is expected and paid for). However, in a scenario with pre-purchase information asymmetry, such as global online markets for outsourcing of IT services, the quality of the product or service the consumer receives may deviate from what was expected, due to at least two reasons. First, the consumer may not have a realistic expectation of the actual product/service quality, leading to a cognitive mismatch of her<sup>2</sup> expectation and the actual quality received. Second, unique products/services are difficult to contract, and incomplete contracts lead to the potential for moral hazard. The source of moral hazard is the asymmetric information between the buyer and the seller because seller actions cannot be perfectly observed and hence contracted upon (Hölmstrom 1979). To make things worse, geographical separation renders a natural remedy to moral hazard – performance monitoring – virtually impossible. For example: imagine a buyer who requests a website development service and is willing to pay \$250 for the website. The request attracts 10 providers with bids ranging from \$100 to \$300. The fact that the chosen provider receives \$200 assumes a \$50 surplus. However, the simple math of deriving \$50 in surplus is not accurate because the *actual* quality of the service may either be better or worse than the *ex ante* expected quality of \$250. More general, a consumer who receives a product that does not meet her expectations suffers a loss of surplus (Lawton 2008), even if she bought the product at a discounted price lower than her *ex ante* WTP. In such scenarios, the traditional measure of CS may be biased, and even at odds with reality; to calculate the true realized value of CS, a quality adjustment for actual *ex post* quality of the product or service received is needed to reflect economic reality.

While these simple examples of measuring CS convey the concerns of the traditional measure of CS, this limitation may be more systematic. Consumer dissatisfaction is common in markets with asymmetric information (Resnick and Zeckhauser 2002). Hence, when actual quality received may not perfectly equal quality expected due to variations in consumption experience, the traditional CS measure may misrepresent (either over-estimate but under-estimate) the true surplus markets offer to consumers, thus creating a need for a quality-adjusted measure of CS. Besides, when there are multiple product or service offers on the markets, consumer willingness to pay would vary based on expected quality of each offer. Indeed, Alfred Marshall's original theorization of CS focused on a market for a single commodity product. An assumption was made that product quality was "common knowledge", meaning that every consumer is homogeneous on

<sup>1</sup> Services are customized products that require consumer involvement and interactions with the service provider, and the ultimate quality of the service depends on the interaction. Hence, for services that are tied to their providers, CS is related to the provider's quality to a larger degree than commodity products.

<sup>2</sup> In this paper, we use "she" to refer to the buyer and "he" the seller.

product quality assessment with full *ex ante* information. Simply put, “what you pay is what you get”. However, for many non-commodity products and services, knowledge about quality may not be common and transferrable; therefore, the fundamental issue the literature has not yet addressed is the simple phrase: “what you pay is not necessarily what you get”. In this paper, we derive a measure of CS that takes into consideration the (a) heterogeneity in product/service quality; and (b) *ex-post* consumer satisfaction.

While extant research on CS has not factored actual quality received and ex-post satisfaction, satisfaction and expectation confirmation theories (ECT) (Anderson and Sullivan 1993; Oliver 1977) offer a theoretical basis to adjust the traditional measure of CS. ECT argues that *ex-ante* expectations, coupled with actual performance, lead to *ex-post* satisfaction; the relationship is mediated by positive or negative confirmation between ex ante expectations and ex post quality. Simply put, when actual ex post quality falls below ex ante expectations, the consumer is likely to be dissatisfied, thus imposing a negative effect on CS. Accordingly, ECT can be used to derive the utility function to derive a quality adjusted measure of CS, thus allowing us to relax the assumption that consumers actually receive what they ex-ante expected. We relax the key assumption made in the CS literature that what a consumer receives equals what he expects. While relaxing this key assumption that *ex-ante* expectation does not equal *ex post* satisfaction is bound to put forth measurement challenges, it is a modest step towards economic reality. Moreover, identifying whether and when the quality-adjusted measure of CS is systematically different from the traditional measure of CS is important in both predicting the performance of markets with asymmetric information and prescribing implications for increasing CS. The following research questions have guided our study:

- How should CS be measured in markets with asymmetric information?
- When does the quality-adjusted measure of CS deviate from the traditional measure of CS?
- Does the quality-adjusted measure of CS predict market performance better than the traditional measure of CS?

To answer these research questions, we first derived a quality-adjusted measure of CS to include the buyer's ex post actual quality received. While it is generally difficult to measure CS (Bapna et al. 2008), using micro-level data from an online market for the outsourcing of IT services allowed us to include measures of both *ex-ante* expected service quality<sup>3</sup> and actual *ex-post* service satisfaction. While not readily available in traditional markets, data on past seller quality and consumer satisfaction are usually documented, aggregated and maintained in online markets by third-parties (e.g., marketplace intermediaries) in the form of feedback ratings. Accordingly, these micro-level data on previous service provider quality and ex post consumer satisfaction with the actual service provided create an opportunity for us to empirically derive a quality-adjusted measure for CS. Furthermore, as online markets for services are vaunted for their benefits to society, this study seeks to empirically quantify the magnitude of their surplus to consumers.

In what follows, we describe our research context of global online markets for the outsourcing of services, and illustrate the limitations of the traditional CS measure in markets with asymmetric information. We next review the literature on CS in markets with asymmetric information and integrate the satisfaction and ECT literatures. We then derive the quality-adjusted measure for CS, and compare the quality-adjusted to the traditional measure of CS. We find that these two measures deviate systematically and we compare their effects on market performance. We conclude by discussing the study's theoretical implications for CS and practical implications for the design of markets with asymmetric information.

## Online Labor Markets with Asymmetric Information

In online markets for the outsourcing of services (herein referred to as “online labor markets”), such as *eLance*,<sup>4</sup> *rent-a-coder*, and *Freelancer* (Malone and Laubacher 1998), people can outsource various IT services, such as website development, graphical design, or creative writing. The potential of these markets was widely touted (Howe 2006); since their inception, they have been expanding at an

<sup>3</sup> Ex-ante quality is usually conceptualized as reputation in the economics literature (George J. Mailath and Larry Samuelson 2001).

<sup>4</sup> As *elance.com* is the first major online market for services that was widely covered by the press, it has become a synonym to “online markets for services”. Please see Malone and Laubacher (1998) for a more detailed description of these markets.

astounding pace despite the economic crisis (Economist 2010). These markets serve as intermediaries for IT services (buyers post request for proposal (RFP) for services and providers offer bids for IT services) that help match buyers with service providers across the globe at low search costs (Yannis Bakos 1999). The Bureau of Labor Statistics estimates that more than 30 million people now work as independent professionals in the United States alone, and the dynamics of these online markets have drawn increased attention in the academic literature (Banker and Hwang 2008; Gefen and Carmel 2008; Snir and Hitt 2003, Hong and Pavlou 2012, Lin and Goes 2012).

In these markets, buyers choose service providers by weighing the expected quality of providers relative to their bids (Banker and Hwang 2008). Due to information asymmetry, buyers may suffer from adverse selection when trying to shield themselves from the uncertainty involved in selecting providers. One important aspect that these markets provide is the quality rating of providers, which is the average of all ratings from buyers from past completed projects. These ratings are usually on an interval scale of 1-10 or 1-5. The availability of these rating data offers the opportunity to factor buyer expectation of provider quality and actual buyer experience into a CS measure. The online marketplace we obtained data has around 3 million active users, over 1.5 million projects contracted and awarded, and over \$100 million in total transaction volume by the end of March 2012.

## **Theoretical Background**

### ***Consumer Surplus***

CS was defined as “the excess of the price which he (the consumer) would be willing to pay rather than go without the thing, over that which he actually does pay” (A. Marshall 1920, p. 124).<sup>5</sup> The classic definition of CS focused on willingness to pay (WTP), and the actual monetary payment, given a known quality. There has been considerable interest in CS in the literature, and there are several theoretical models to extend the original models (Cicchetti and III 1971; Goolsbee and Petrin 2004; Nevo 2003; Randall and Stoll 1980; Turnovsky et al. 1980) along with empirical applications (Bapna et al. 2008; Brynjolfsson 1996; Harris and Blair 2006).

CS has been viewed as a criterion for evaluating the societal impact of technology (Lorin M. Hitt and Erik Brynjolfsson 1996), and how much technology helps enhance total societal welfare (Grover and Ramanlal 1999). Erik Brynjolfsson (1996) examined the contribution of information technology (IT) to CS and found that IT spending generated over \$50 billion in net value and increased economic growth by 0.3% per year in the United States of America. In line with classical economics, their approach to CS was to add up individual surpluses of infra-marginal buyers who paid less than they would be willing to pay (integrating the area under the demand curve between old and new price – before and after an IT investment) (Brynjolfsson et al. 2003; Hitt and Brynjolfsson 1996).

CS has also been measured and estimated with pure-characteristic demand model (Minjae Song 2007) with structural estimation methods developed by Steven Berry, James Levinsohn and Ariel Pakes (also known as BLP, 1995, 2003). Economists also developed other methods. Jerry A. Hausman and Whitney K. Newey (1995) used non-parametric methods and estimated the welfare loss from tax on gasoline (Hausman and Newey 1995). When Hicksian demands are unknown, Ian J. Irvine and William A. Sims proposed a measure based on Slutsky compensated demand (Irvine and Sims 1998). These approaches are appropriate for explaining the societal impact given market equilibrium of a homogeneous product, but less effective for deriving a precise measure of CS.

### ***Customer Satisfaction and Expectation-Confirmation Theories (ECT)***

Customer satisfaction is a *post hoc* evaluation of consumption experience (Oliver 1980). The literature has established that two variables – performance-specific expectation, and expectancy disconfirmation - play a major role in customer satisfaction (Anderson and Sullivan 1993; Oliver 1977). With virtually no exception, the stream of research on consumer satisfaction has reached the conclusion that “satisfaction is a function of an initial standard and some perceived discrepancy from the initial reference point” (Oliver 1980, p.

---

<sup>5</sup> As individual CS aggregates to societal welfare for consumers, research on CS gave rise to welfare economics (Marshall 1920).

460). The literature has also characterized the effects of post-purchase satisfaction, such as complaints and repurchases (Robinson 1979) and behaviors (Fishbein 1967), such as switching (Keaveney 1995). Satisfaction was characterized as a utility-based concept in both the economics (Burns 1973) and marketing (Anderson and Sullivan 1993) literatures. ECT (Oliver 1977; Spreng et al. 1996) is a well-established theory to explain consumer satisfaction. It posits that the “(dis)confirmation” of a consumer’s pre-purchase expectation of a product is the key to his post-purchase satisfaction.

ECT was used in many areas, such as measurement of customer satisfaction (McKinney et al. 2002), the effect of application service provision on satisfaction (Susarla et al. 2003), systems continuance (Bhattacharjee 2001), and business relationships (Kim et al. 2009).

Due to the idiosyncratic nature of services, especially specificity, complexity (Snir and Hitt 2003) and non-contractibility (Mithas et al. 2008), buyers have *ex-ante* uncertainty about the expected service from the provider (Dimoka et al. 2012), and *ex-post* intractability of the contracted provider’s efforts. A natural outcome is customer’s expectations of service quality being disconfirmed, thereby leading to dissatisfaction; or being confirmed, thereby leading to additional satisfaction. We thus argue that *ex-ante* expectations and *ex-post* satisfaction have important implications for measuring CS.

### Markets for IT Labor with Asymmetric Information

Today’s economy has been characterized as a services economy (Vargo and Lusch 2008). Research on services, goes back to 18<sup>th</sup> century when Adam Smith (1961) proposed the term “non-productive economic activities”. Later the revolutionary role of services in the industrial economy was conjectured by scholars (Chandler 1977; Clark 1983). By reviewing the literature on services, Richard B. Chase and Uday M. Apte (2007) concluded that “service performances cannot be guaranteed since they are generally delivered by human beings who are known to be less predictable than machines” (p. 380). Needless to say, services require intense involvement and interactions, and the evaluation of their performance is multi-faceted. Hence, *ex-ante* uncertainty of service quality is significant, since service providers know more about their own characteristics than buyers who can only infer quality from information signals. The literature has also highlighted the characteristics of services. Compared to commodities, services are idiosyncratic, complex (Snir and Hitt 2003), non-contractible (Brynjolfsson and Smith 2000), with highly variable quality (Rust et al. 1999) and hence can be viewed as “highly-customized” products. Unlike the purchasing of commodities in online markets that can be easily contracted on product descriptions, conditions, and warranties, services have multiple complex components of *labor* that cannot be perfectly described (Spence 1973), nor can they be easily contracted (such as the efficiency of a customized website). Still, the proposed quality-adjusted measure of CS applies to both services and products.

### Measure Development

The traditional measure of CS makes the assumption that *ex-ante* quality expectation carries over. We argue that in markets with asymmetric information, what a consumer expects may differ from what the consumer actually receives. Based on Marshall’s (1920) definition of CS, we derive an *ex-post* utility-based measure of CS with two important features: First, we differentiate *ex-ante* quality expectation ( $q$ ) and *ex post* delivered service quality ( $q'$ ). Second, following the satisfaction literature, we incorporate the effect of disconfirmation of expectation and asymmetric disconfirmation to construct the utility function to capture a consumer’s actual experience with the service. We use the following notations:

**Table 1—Notations**

Variable	Description
$\underline{V}$	Buyer’s willingness to pay (lower bound)
$\bar{V}$	Buyer’s willingness to pay (higher bound)
$\bar{q}$	Highest quality a provider can perform the service
$\hat{q}$	Acceptable provider quality
$q$	A buyer’s <i>ex-ante</i> quality expectation of a provider
$q'$	A buyer’s <i>ex-post</i> quality assessment of a provider
$p$	Contract price
$U()$	A buyer’s utility function

## Quality-adjusted Measure of Consumer Surplus

As noted earlier, measuring CS with the difference between *WTP* and price may not be accurate because the assumption in the literature that the service or product a consumer receives is the same level of quality as what she expects and paid for may not hold. This suggests two measures of quality: expected quality versus actual quality and *ex-ante expected utility* versus *ex-post actual utility*. However, the literature mainly focuses on *ex-ante expected utility* of a known quality assuming that *ex-ante* quality is equal to *ex-post* quality. We extend extant research that has focused on price with no substantive considerations for heterogeneities of product or service quality level and buyer ex post satisfaction by deriving a measure of CS that includes (a) heterogeneity in service quality and (b) *ex-post* consumer satisfaction.

Consider a setting with service buyer (consumer)  $i$ , who posted a project  $i$ , with a pre-defined *budget*, received  $N$  bids and finally chose service provider  $j$ . Buyer  $i$  contracted with provider  $j$  on bid price  $p_j$ . Buyer  $i$  expects provider  $j$  to have quality  $q_j$ . Provider  $j$  delivered the service with quality  $q_j'$ . The intuitive way to look at the process is to see how much the buyer is willing to pay for a service at each stage.

Before contracting, the buyer may have an *ex-ante* budget estimate (or *WTP*) for a service with a certain quality, denoted as *ex-ante WTP*. Upon contracting, her *WTP* becomes specific to the expected quality of the chosen service provider. When the service is provided, she has an *ex-post WTP* reflecting actual quality provided and satisfaction of the service. Given our notation, the traditional measure of CS, according to the CS literature, is:

$$(1) \quad CS_j = V_j - p_j$$

$V_j$  is the expected value of the service performed by provider  $j$ , i.e., the buyer's *WTP* for provider  $j$  for completing the project.

$p_j$  is the price that provider  $j$  bids to offer the requested service.

Here both  $V_j$  and  $p_j$  are monetary measures in a unit of a currency, such as US dollars. Since before the transaction, the buyer does not know the real value of the service to be offered by provider  $j$ , it can only be an expected value, it is possible to approach  $V_j$  as the buyer's *WTP* for the service performed by provider  $j$ .

Because *WTP* is an *ex ante* concept, Equation [1] can only precisely measure CS if a consumer would be willing to pay the same amount  $V_j$  when she received the service (or product). The assumption is prone to be violated when the actual quality  $q_j'$  is susceptible to deviate from expected quality  $q_j$ .

it is reasonable to argue that a buyer's *WTP* is tied to provider quality ( $q_j$ ), The literature has either made the linearity assumption (Banker and Hwang 2008), or a concave functional assumption used in utility theory (Fishburn 1970) and the satisfaction literature (Oliver 1980). Therefore, without any assumptions, the general form of traditional measure of CS becomes:

$$(2) \quad CS_j = V(q_j) - p(q_j)$$

And the *ex-post* quality-adjusted measure of CS is given by:

$$(3) \quad CS_j' = V(q_j') - p(q_j)$$

Since we seek to reflect economic reality, we need to consider an appropriate form for the utility function with regard to expected or experienced quality. Among many studies, Banker and Hwang (2008) used a linear formulation for the utility function. The limitation of a linear function is that it assumes that each unit of increase in quality will result in the same amount of utility increase without decrease in marginal utility. ECT provides the rationale for deriving a monotonically increasing and concave function to represent risk-averse consumers.

According to the analytical framework of Anderson and Sullivan (1993), it is *ex-post* perceived quality ( $q$ ) and the relative size between expectation of quality  $q$  and  $q'$  that determines overall satisfaction. ECT predicts that the consumer's positive or negative confirmation, coupled with expected performance, determines her satisfaction level. We incorporate the ECT logic to adjust CS to reflect reality. As per ECT, we propose integrating ex post confirmation/disconfirmation for the CS model to reflect the consumer's

actual utility. Another important aspect of ECT discussed in the literature is asymmetric disconfirmation. As prospect theory (Kahneman and Tversky 1979) would argue, people place more emphasis on losses than gains. Similarly in the literature on satisfaction (Anderson and Sullivan 1993), quality that falls below expectations has a greater impact on satisfaction than quality that exceeds expectations. Thus, we construct a measure that would reflect smaller effect of positive confirmation (when  $q' > q$ ) relative to negative confirmation (when  $q' < q$ ). Thus, a monotonically increasing and concave function is deemed to be suitable to represent the WTP function.

With a non-linear concave utility formulation and asymmetric disconfirmation, we follow Rust et al. (1999) and Anderson and Sullivan (1993), assuming that utility is a function of expected quality, is continuous, twice differentiable, monotonically increasing and concave. The assumption implies that consumers are risk-averse, and that consumers suffer more from not having their expectations met than exceeding their expectations. This assumption is borne out by abundant empirical evidence in the satisfaction and service quality literature (Anderson and Sullivan 1993; DeSarbo et al. 1994; Inman et al. 1997; Rust et al. 1995). A commonly used functional form that meets the assumption is a log utility function<sup>6</sup>:  $V(q_j) = v_0 + v_1 \cdot \ln(q_j)$ , where  $v_0$  is a constant that indicates the reservation utility, and  $v_1$  is a constant that indicates the sensitivity to service quality. Therefore:

$$(4) \quad CS_j = v_0 + v_1 \cdot \ln(q_j) - p_j$$

Equation [4] is what the literature usually assumes, which we herein denote as the traditional measure of CS. However, since *ex-ante* quality may not carry over because of asymmetric information (consumer choosing the wrong service provider) and moral hazard (provider slacking off after contracting), and the actual service received maybe either worse or better than the *ex-ante* service quality expectation, we construct the following adjusted measure for CS:

$$(5) \quad CS_j' = v_0 + v_1 \cdot \ln(q_j') - p_j$$

Equation [5] is the general form of quality-adjusted measure for CS, denotes as the quality-adjusted measure of CS. In practice, when a consumer posts an RFP, it is plausible to assume that she has a *a priori* valuation (or WTP) (e.g. the consumer is willing to pay \$250 if the provider delivers a website of a reasonable quality level). This valuation is usually in the form of a budget (Snir and Hitt 2003), which captures a range of values a consumer is willing to pay for a given service. This *ex-ante* WTP could have an upper and lower bound due to quality variations. Therefore, we make a second assumption that a buyer's posted budget indicates her willingness to pay for a service of a certain level. The upper bound of budget assumes a fully satisfying expected quality, while the lower bound of budget assumes an acceptable level of expected quality; the consumer is willing to pay zero if the provider has the lowest possible quality on the marketplace.

We denote  $\bar{V}$  as the WTP for the service with fully satisfying quality (i.e., the quality that achieves the highest WTP for the consumer). It is rational to assume that the consumer's upper bound of WTP  $\bar{V}$  can either be tied to the highest quality a provider can have, or the average provider quality on the marketplace. We denote this corresponding quality as  $\bar{q}$  and use it to illustrate the measurement development (in reality any meaningful value for satisfying quality can be used). By the same token, we argue that the lower bound of WTP  $\underline{V}$  would correspond to a lower but acceptable quality on the marketplace,  $\hat{q}$  (the same also applies here that any meaningful value for acceptable quality can be plugged in). We plotted the distribution of average rating of providers who won projects, and over 90% have an average rating of 8 or above. Therefore we assume that on average, the lowest acceptable quality rating is 8, and further, buyers are homogenous in terms of sensitivity to quality. Therefore, we have the following constraints:  $\bar{V}_j = v_0 + v_1 \cdot \ln(\bar{q})$  and  $\underline{V}_j = v_0 + v_1 \cdot \ln(\hat{q})$ . For each service, we derive

<sup>6</sup> Other commonly used quasi-linear functions include quadratic utility functions.

$\bar{V}_j - V_j = v_1 \cdot \ln\left(\frac{\bar{q}}{\hat{q}}\right)$  and to consolidate (derivation omitted for brevity):

$$(6) \quad CS_j = \bar{V}_j + \frac{(\bar{V}_j - V_j)}{\ln\left(\frac{\bar{q}}{\hat{q}}\right)} \cdot \ln\left(\frac{q}{\bar{q}}\right) - p_j$$

Where  $q$  is the buyer's expectation of provider  $j$ 's quality;  $CS_j$  is expected CS. If we adjust for quality, similarly we can derive a quality-adjusted CS measure.

$$(7) \quad CS_j' = \bar{V}_j + \frac{(\bar{V}_j - V_j)}{\ln\left(\frac{\bar{q}}{\hat{q}}\right)} \cdot \ln\left(\frac{q'}{\bar{q}}\right) - p_j.$$

Compared with existing measures, the proposed quality-adjusted measure of CS takes into consideration the *actual* service provided and the *ex-post* satisfaction, thus integrating the actual value the service renders to a buyer. Besides, the quality-adjusted measure of CS has face validity by capturing economic reality (potential for ex post dissatisfaction or extra satisfaction). Besides services, this measure is expected to generalize to products and commodities since *ex-ante* expectation of product quality and *ex-post* satisfaction with the product applies to products.

### Comparing the Performance Effects of the Two CS Measures

Since CS captures the extra utility a buyer derives from a product or service, to test the predictive power of the quality-adjusted measure of CS relative to the traditional CS measure on *market performance*. We identify three indices of market performance: *continuation* (whether the consumers continues posting RFPs on the marketplace after the first project), *subsequent projects* (total number of RFPs posted after the first project), and *total transaction volume* (total amount paid in the marketplace after the first project).

Consumer retention is a key driver of consumer lifetime value and profitability (Gupta and Zeithaml 2006). Understanding the effect of the two competing CS measures on market performance can shed light on the relative predictive power of the quality-adjusted CS measure. Also, these performance effects can be seen as the criterion variables that can compare the quality-adjusted versus the traditional measure of CS. Since the quality-adjusted measure of CS includes post-purchase satisfaction and is likely to capture whether the buyer is satisfied with the service after experiencing the service. Since a satisfied buyer is more likely to continue posting RFPs in the marketplace, we propose:

*H1a: The quality-adjusted measure of CS better predicts market performance: (a) continuation, (b) subsequent contracts, (c) total transaction volume, than the traditional measure of CS.*

The deviation in ex-post quality from ex ante expected quality leads to a difference between the quality-adjusted CS from the traditional measure of CS ( $\Delta CS$ ).<sup>7</sup> This value measures the extent of disconfirmation of expectations. A positive value of  $\Delta CS$  implies a pleasant surprise a buyer receives, and a negative value of  $\Delta CS$  implies that buyer expectation is disconfirmed.

$$(8) \quad \Delta CS = CS' - CS = V(q') - V(q)$$

Drawing upon the satisfaction and ECT literatures,  $\Delta CS$  is more likely to determine consumer behavior than either CS or  $CS'$ . We expect  $\Delta CS$  to further predict market performance beyond CS or  $CS'$ . As derived earlier,  $\Delta CS$  lies in the relative level of *ex-ante* quality expectation versus *ex-post* quality assessment.

*H1b: The difference between the quality-adjusted and the traditional CS measure ( $\Delta CS$ ) will be*

<sup>7</sup> In what follows, we use  $\Delta CS$  to denote "difference between the quality-adjusted measure and traditional measure of CS".



positively associated with market performance: (a) continuation, (b) subsequent projects, (c) total transaction volume.

## Empirical Methodology

### Sample and Integrated Datasets

Our samples come from two archival sources. First, transaction data consisting of a sample of 159,748 RFPs from 38,315 buyers, that took place between February 4, 2004 and September 24, 2010, was obtained from the online market with MySQL query. The outsourcing services were across ten project categories. Second, we obtained the PPP adjusted GDP per capita indices and the official language data from the CIA World Factbook<sup>8</sup>, which is widely used in the literature (Gefen and Carmel 2008). We also obtained PPP data from the International Monetary Fund<sup>9</sup> and the World Bank<sup>10</sup> to amend the CIA World Factbook data because PPP data were not available in the CIA World Factbook. We compared indices from the three data sources to ensure consistency (Table 3). We combined the PPP-adjusted GDP data and English-speaking country data with the first dataset. Buyers and providers are from 213 countries or regions, and the English language is the official language for 83 countries or regions.

Table 3—Summary Statistics of Relevant Variables (N=159, 730)

Variable	Mean	Std. Dev.
Budget (higher bound)	316.000	364.09
Budget (lower bound)	69.358	138.70
Ex-ante average rating	9.841	0.37
Ex-post rating	9.838	0.91
rating difference	0.3	0.9
Contracted Price	136.660	245.51

### Empirical Task 1: Estimating the Level of CS

#### Operationalization of CS Measures

As explained in Section III, Equations [6] and [7] are the general forms of the proposed quality-adjusted measure of CS. For empirical tractability, we employed several proxies for our variables. Similar to Anderson and Sullivan (1993), we used feedback ratings to capture service quality. We use the average rating of the provider received in the market as a proxy for *ex-ante* expectation of quality  $q$ , and use the rating the buyer left for him for a service indicating his performance in that specific service,  $q'$ . Similar to Anderson and Sullivan, the ratings are on a discrete 1-10 interval scale (1 indicates lowest quality and 10 highest quality). For WTP, the literature has used “maximum bid price” as a proxy (Bapna et al. 2008). The main difference between maximum budget and maximum bid is that the latter is the actual amount the buyer pays for the service, while the former reveals *ex ante* WTP. Practically, the service provider can bid over the maximum budget, thus leading to negative CS. The budget is chosen as a proxy for WTP since it not only reveals the amount the buyer is willing to pay, but it also affects the providers bid to offer for the service. Although either deflating or inflating the budget could potentially be the buyer’s strategic decision to maximize CS, we argue that either choice is not rational. As per our debriefing with several respondents after a survey, all outsourcers are aware that posting an inflated upper bound of budget would attract higher bids, which works against their intention to reap a higher CS; posting a low budget would result in

<sup>8</sup> CIA World Factbook: <http://www.cia.gov/cia/publications/factbook/geos/ae.html/>.

<sup>9</sup> International Monetary Fund, World Economic Outlook Database. 2010 data.

<sup>10</sup> World Development Indicators database, World Bank - 29 September 2010

fewer or no bids. Therefore, rational outsourcers (consumers) are motivated to post their true WTP.

To make sure that our results are not dependent on the values we choose as acceptable quality level, we experimented with different values, including lowest rating (most liberal), mid-point of the scale, and a high scale (most conservative) as acceptable quality level (any value between 1-10 may be acceptable quality), and the highest possible rating (10) as highest quality<sup>11</sup>. Since most buyers chose a provider with an average rating of 8 or higher, 8 is chosen as the acceptable ex-ante quality.

As we calculated, when assuming an acceptable level quality of 8 on a scale of 1 to 10, average CS is 159.3 per project (st.d.=229.5), and average quality adjusted CS is 139.3 (st.d.=551.6). We also find that technical projects which are presumably more complicated than other projects, have a higher difference between the traditional measure and quality-adjusted measure of CS, indicating that for highly technical projects, actual CS is more likely to deviate from ex ante expected CS, since *ex-post* quality assessment is more likely to deviate from the *ex-ante* quality expectation.

Figure 1 has three implications for CS in markets with asymmetric information. First, over time, CS has been increasing, indicating that the market is functioning well as the market matures (when market size increases), validating Joel Waldfogel's intuition that through the proliferation of service providers (sellers), consumer satisfaction will be enhanced (2007, p 102). Second, over time,  $\Delta CS$  decreases, indicating that consumers, on average, are getting more satisfied with services. Third, the market is not fully correcting the bias since we observe a large variance of  $\Delta CS$  over time. Therefore, it is important to examine under what scenarios the quality-adjusted CS deviates from the traditional CS measure.

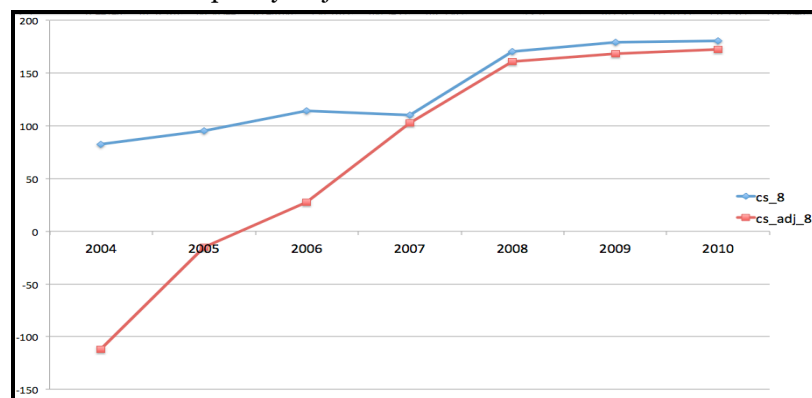


FIGURE 1—TREND OF CS BY YEAR

We employed paired *t*-tests to examine the difference between the mean of CS with the two measures.<sup>12</sup> As Table 4 shows, on average, across all services, the traditional CS measure consistently yields a significantly higher surplus than the quality-adjusted CS measure. The results indicate that, although these two measures yield a qualitatively comparable level of CS, the traditional measure is still systematically biased toward inflating the level of CS with all paired *t*-tests ( $p < 0.0001$  level). Also it should be noted that some projects are highly satisfactory, thus allowing buyers to reap extra CS while others are less than satisfactory thus buyers suffer from a loss of CS. However, at the aggregate level, positive and negative terms cancel out. Finally, given that the average bid price for a project is \$160 and the average winning bid is \$137, the level of CS is surprisingly high, indicating notable economic benefits these marketplaces offer.

## Empirical Task 2: Effects and Predictors of Consumer Surplus

### Variables Definition and Measures

**Dependent Variables.** In the effects model, we examined the level of CS on the three dimensions of market performance:

<sup>11</sup> Regression results are robust to different acceptable quality levels.

<sup>12</sup> To test whether the CS level changes significantly when potential outliers are dropped from the analysis, we analyzed the holdout samples with top and lowest 5%. The result is shown to be robust to potential outliers.

**Continuity:** The buyer's continuity in the market was measured with a binary variable indicating whether the buyer posted any new projects after the first project. A satisfied buyer is likely to post projects on the marketplace again.

**Subsequent Projects:** It was measured with a discrete variable capturing the total number of subsequent projects the buyer sought on the market. We employed a natural logarithmic transformation for transaction volume to achieve a less skewed distribution. We denote the transformed variable  $\ln(\text{sub\_projects})$

**Transaction Volume:** It was measured with a continuous variable capturing the total amount of money paid on the marketplace by the buyer besides the first project. We employed a natural logarithmic transformation for the variable to achieve a less skewed distribution. We denote the transformed variable as  $\ln(\text{sub\_payments})$ .

In the predictor model, the dependent variable is  $\Delta\text{CS}$ .  $\Delta\text{CS}$  was measured by subtracting the quality-adjusted measure by the traditional measure we derived in Equation (8). We employed a natural logarithmic transformation for this variable to achieve a less skewed distribution.

Table 4—Description and Measurement of Control Variables<sup>13</sup>

<i>Project Level Control Variables</i>
<b>Project Size:</b> It is harder for providers to estimate the cost for finishing a larger project since project size is associated with uncertainty; therefore, bid prices would be more dispersed than for smaller projects, offering buyers the chance to reap a higher surplus. Project size was measured with natural logarithmic transformation average bid price for a project (Snir and Hitt 2003). Although different proxies may be used, such as the length of a project and project client value (budget).
<b>Project Category:</b> Different categories of projects may have different price levels. Project Category was coded into three dummy variables - Writing, Design and Data Entry. We set IT projects as the base group, and the effects of all three variables are therefore relative to IT projects.

## Econometric Specification and Estimation Methods

To examine how the two measures of CS and  $\Delta\text{CS}$  affect market performance (continuity, subsequent projects, and total transaction volume), given that continuation (whether a buyer continues posting RFPs after her first project) is a binary variable, logistic specification with maximum likelihood estimation was chosen (Wooldridge 2002).

$$(9) \Pr(\text{cont}_{it_2} = 1 | \ln(\text{cs}_{it_1}), \text{seeker\_type}_{it_1}, \text{project\_type}_{it_1}, \text{entrance\_time}_{it_1}, \text{controls}_{it_1}) = \frac{1}{1 + e^{-\left(\beta_0 + \beta_1 \ln(\text{cs\_diff}_{it_1}) + \beta_2 * \text{seeker\_type}_{it_1} + \beta_3 * \text{project\_type}_{it_1} + \beta_4 * \text{entrance\_time}_{it_1} + \beta_{5-11} * \text{controls}_{it_1} + u_{it_1}\right)}}$$

Number of subsequent projects is a non-negative, discrete variable. Over-dispersion was observed for the distribution of the number of subsequent projects. Thus, a properly specified count model was needed, and a negative binomial model was chosen to estimate the relative effect of CS, CS', and  $\Delta\text{CS}$  on the number of subsequent projects. This is considered a more flexible form for overly dispersed count data.

$$(10) \text{sub\_projects}_{it_2} = \beta_0 + \beta_1 * \ln(\text{cs}_{it_1}) + \beta_2 * \text{seeker\_type}_{it_1} + \beta_3 * \text{project\_type}_{it_1} + \beta_4 * \text{entrance\_time}_{it_1} + \beta_{4-11} * \text{controls}_{it_1} + u_{it_1}$$

For transaction volume, we used an additional test since subsequent payments was highly correlated with subsequent projects ( $r=0.78$ ). Total transactions is a non-negative ratio variable; therefore, we employed ordinary least squares estimation with robust standard error for coefficient estimation.

<sup>13</sup> Some control variables are omitted because of page limit. Measurement of all control variables are available upon request from authors.

$$(11) \ln(sub\_payments_{it_2}) = \beta_0 + \beta_1 * \ln(cs_{it_1}) + \beta_2 * seeker\_type_{it_1} + \beta_3 * project\_type_{it_1} + \beta_4 * entrance\_time_{it_1} + \beta_{5-11} * controls_{it_1} + u_{it_1}$$

$\Delta CS$  was modeled as a function of buyer experience, repeat transactions, labor arbitrage, and language barrier (plus a set of control variables) as covariates. Consistent with the analytical derivation, we used  $i$  to index project and  $t$  to index time (sequence of projects). Since our data contains information of the same buyer's different projects at different times, there might be correlations among the projects of the same provider because of time-invariant factors. To account for unobserved time invariance of buyers, such as buyer education with panel data, several estimation methods were used. First, a pooled OLS estimation with clustered standard error corrected the standard errors. Second, a fixed-effects or a random-effects approached was used to account for the unobserved time invariant in the error terms. These estimation models were used to control for unobserved heterogeneity in buyer characteristics when constant over time and correlated with other independent variables. Thus, we specified our antecedent model as:

$$(12) \ln(\Delta CS_{it}) = \beta_0 + \beta_1 * ppp_i + \beta_2 * common\_language_i + \beta_3 * ppp_i * common\_language + \beta_{4-n} * controls_i + \alpha_i + u_{it}$$

We performed resampling for the needs of the analysis of the effects models. We retrieved all the first-time contracts, calculated the continuation, subsequent projects and subsequent payment following those first-time contracts, and matched the data. Overall 38,307 first/one-time projects are obtained for the analysis. Maximum likelihood criterion (Logistic analysis) and ordinary least squares with robust standard errors are employed to perform the estimation.

### Estimation Results for the Effects Model

**Table 5. Logistic Regression Results (DV=*continuation*)**

Variables	Model 1	Model 2	Model 3
$\ln(cs)$	0.007(0.005)	-	-
$\ln(cs\_adj)$	-	0.019*** (0.004)	-
$\ln(cs\_diff)$	-	-	0.030*** (0.005)
<i>seeker_type</i>	1.265*** (0.135)	1.261*** (0.135)	1.260*** (0.135)
<i>project_type</i>	-0.862*** (0.081)	-0.864*** (0.081)	-0.846*** (0.081)
Constant	-0.151(0.154)	-0.180(0.154)	-0.107(0.153)
Wald Chi <sup>2</sup> (12)	888.61	903.30	920.85
Pseudo R <sup>2</sup>	0.0187	0.019	0.0193

Notes: The total number of observations is 38, 307. We report the robust standard errors in parentheses.

\*\*\* Significant at 0.1 percent level. \*\* Significant at 1 percent level. \* Significant at 5 percent level.

**Table 6. Negative Binomial Regression Results (DV=*sub\_projects*)**

VARIABLES	Model 1	Model 2	Model 3
$\ln(cs)$	0.010(0.008)	-	-
$\ln(cs\_adj)$	-	0.028*** (0.007)	-
$\ln(cs\_diff)$	-	-	0.043*** (0.007)
<i>seeker_type</i>	1.745*** (0.167)	1.740*** (0.166)	1.746*** (0.164)
<i>project_type</i>	-0.945*** (0.095)	-0.943*** (0.096)	-0.922*** (0.095)
Constant	1.577*** (0.213)	1.533*** (0.211)	1.646*** (0.208)
Lnalpha	1.310*** (0.015)	1.308*** (0.015)	1.307*** (0.015)
Wald Chi <sup>2</sup> (12)	1040.29	1053.85	1098.72

Notes: total number of observations is 38, 307. We report the robust standard errors in parentheses.

**Table 7. Regression Results<sup>14</sup> (DV =  $\ln(\text{sub\_payments})$ )**

Variables	Model 1	Model 2	Model 3
$\ln(\text{cs})$	-0.001 (0.008)	-	-
$\ln(\text{cs\_adj})$	-	0.023*** (0.007)	-
$\ln(\text{cs\_diff})$	-	-	0.046*** (0.007)
<i>seeker_type</i>	1.404*** (0.094)	1.400*** (0.094)	1.397*** (0.093)
<i>project_type</i>	-1.029*** (0.077)	-1.032*** (0.077)	-1.003*** (0.077)
Constant	2.751*** (0.144)	2.686*** (0.143)	2.783*** (0.142)
F statistic	145.66	146.68	149.64
R-squared	0.033	0.033	0.034

Notes: total number of observations is 38, 307. We report the robust standard errors in parentheses.

Overall, we find full support for H1. As shown in Table 6, only quality-adjusted CS and  $\Delta\text{CS}$  positively predicted buyer continuity in the market ( $\beta > 0$ ,  $p < 0.001$ ), subsequent projects ( $\beta > 0$ ,  $p < 0.001$ ), and total transaction volume ( $\beta > 0$ ,  $p < 0.001$ ). Surprisingly, the traditional measure of CS does not predict any of the three market performance outcomes. The results testify to the claim that a quality-adjusted measure of CS is needed and support the notion that the difference of the two measures ( $\Delta\text{CS}$ ) is consequential for the marketplace in terms of its performance.

The quality-adjusted measure of CS not only helps predict market performance, but it also has implications for the design of markets with asymmetric information. As  $\Delta\text{CS}$  is the amount of utility (satisfaction) gain or loss due to ex-ante expectation of service/product quality being positively or negatively confirmed ex post, identifying what leads to  $\Delta\text{CS}$  can shed light on how different parameters may have an effect on additional ex post gain or loss of utility. ECT predicts that the difference between expectation would affect consumer satisfaction, which in turn has an effect on consumer loyalty. As we do observe that the two CS measures are quantitatively different, one question that naturally arises is what is the driving force of the difference. We identify three possible factors that may have an effect on CS, CS' and  $\Delta\text{CS}$ : *project size*, *global labor arbitrage*, and *common language*.

**Project Size:** We categorize projects based on their size to small, medium and large projects.

**Global labor arbitrage** was measured with a continuous variable capturing the PPP-adjusted GDP per capita of the service provider. We denote this variable *ppp*.

**Common Language** was measured with a binary variable to indicate whether the provider and the buyer are from countries that use the same official language (e.g., India and U.S. use the English language).

**Table 8. Regression Results for the Antecedents Model<sup>15</sup> (DV =  $\ln(\text{cs\_diff})$ )**

VARIABLES	Coefficient Estimates		
	DCS	CS	CS'
Provider PPP	-7.74e-06*** (1.15e-06)	2.21e-06* (1.30e-06)	-2.56e-07 (1.07e-06)
Common Language	0.361*** (0.0189)	-0.158*** (0.0213)	-0.108*** (0.0175)
Provider PPP* Common Language	<b>-3.20e-06**</b> <b>(1.27e-06)</b>	-1.24e-06 (1.43e-06)	1.94e-07 (1.18e-06)
Constant	5.672*** (0.0511)	5.190*** (0.0422)	-0.00712 (0.0454)
Observations	159,675	159,675	159,675
R-squared (within)	0.022	0.014	0.019
Number of panels	38,296	38,296	38,296

Notes: We report cluster-robust standard errors in parentheses. Standard errors are adjusted for 38, 296 clusters in seeker identifier.

<sup>14</sup> Control variables were omitted from the main results for brevity. The full regression results are available upon request.

<sup>15</sup> Control variables are omitted from the main results for brevity.

Overall the three estimations tried to correct the time invariant unobserved heterogeneity of buyers over time. Even though the estimations yield comparable level of coefficients, for robustness, we tested whether a fixed effects model should be preferred over pooled OLS and a random effects model. We employed the Hausman test (Hausman 1978) to identify whether a fixed effects model is needed to obtain consistent estimates. Since we saw a difference between the cluster-robust standard errors and the default standard errors of the random effects estimators, the crucial assumption that  $\alpha_i$  and  $u_{it}$  are i.i.ds may be invalid. Therefore, we corrected the limitation of standard Hausman test by employing the approach suggested by Wooldridge (2002) to construct a robust Hausman test. We thus used a user-written command (Schaffer and Stillman 2006) following Cameron and Trivedi (2009).

Since both the Hausman test and the robust Hausman test for fixed effects strongly reject the null hypothesis that random effects estimation provides consistent estimates, we are obliged to interpret the results based on fixed effects estimations. First, provider's PPP-adjusted GDP per capita ( $\beta = -0.063$ ,  $p < 0.001$ ) negatively affected  $\Delta CS$ . This supports our claim that with the same amount of money, on average, providers from poorer countries will work harder and deliver better results than those from richer countries. Global labor arbitrage is an economic phenomenon where, as a result of the removal of barriers to international trade, jobs move to nations where the cost of labor is relatively cheap. This phenomenon was observed by the New York Times (Roach 2004). Global labor arbitrage affects buyers' decisions since buyers are more likely to choose a service provider from poorer country to take advantage of labor arbitrage, thus increasing the buyer's CS. From a utility perspective, the valuation of \$1 dollar varies across countries. The utility a service provider can derive from a certain monetary amount depends on the purchasing power of that amount. Gefen and Carmel (2008) found that buyers tend to choose providers from poorer countries; their argument is that for the same amount of monetary incentive, buyers would expect service providers from poorer countries to work harder to reap higher producer surplus. We argue that the purchasing power of service provider increases  $\Delta CS$  because providers from lower PPP countries are likely to work harder towards satisfaction of the buyers.

Second, common language ( $\beta = +0.198$ ,  $p < 0.001$ ) positively affected  $\Delta CS$ . Language barrier is a major issue on global online markets since effective communication relies on good command of a common language. People speak same language usually communicate ideas better. Accurate usage of words can eliminate ambiguity, reduce communication costs, and avoid redundant work. Since English is the primary language in global online markets, whether both parties fluently speak English would determine whether the service requirements can be effectively communicated. Therefore, we would expect language barrier to be negatively associated with buyers' *ex-post* utility resulting in lower  $\Delta CS$ .

The interaction effects in the analysis show that the positive effect of English on  $\Delta CS$  is attenuated by the increase of PPP (e.g., India providers will provide a higher DCS than US/UK providers). Therefore, countries like India will be ideal for reaping *ex-post* extra WTP (satisfaction), while providers from low PPP non-English speaking (China) or English speaking high PPP countries (US) may be similar, and providers from non-English speaking high PPP countries are least favorable.

## **Robustness Tests**

### **Additional Robustness Checks**

We performed several robustness checks to insure the validity of our results. First, scatter plots of observations and dependent variables did not show any pattern for the consequences model, indicating independence of observations (i.i.d.s). The role of multicollinearity was checked with variation inflation factors (VIFs) (Hair Jr et al. 1995), and VIFs were below the suggested threshold ( $VIF < 3$ ). We also checked the correlations between variables to amend the low power of VIF test in large datasets, and no threat to validity was detected. These robustness checks lend further credence to our findings.

### **Economic Effect Sizes<sup>16</sup>**

We examined the economic effects (effect sizes) of independent variables in these models because the

---

<sup>16</sup> The economic effects should be interpreted cautiously since there might be non-linear effects.

statistical significance may be an artifact of the sample size. For the consequences model, on average, a \$10 increase in  $\Delta CS$  in the first project will increase the odds of a buyer continuing to post RFPs on the market by 1.25%. The quality-adjusted measure of CS is more than three times more predictive than the traditional measure of CS. A \$10 increase in  $\Delta CS$  in the first project will increase 2% of total future payments.

For the predictive model, on average, 1 more project completed in the market increase  $\Delta CS$  by 0.11%. And for a given project, if the provider's PPP-adjusted GDP per capita increases by \$10000,  $\Delta CS$  will be reduced by 10%. If a project is a repeat (versus a first-time project),  $\Delta CS$  will increase by 26.8%. And if the provider is from a country where English is the official language,  $\Delta CS$  will be increased by almost 32.7%.

In sum, the analysis of the economic effects sizes shows that our predictors are economically significant.

## Discussion

### Key Findings and Contributions

This study proposed a quality-adjusted measure for CS, which we believe, is particularly salient in markets with asymmetric information. Due to asymmetric information (since providers cannot be physically monitored after both parties are contracted), actual *ex post* quality may deviate from *ex ante* expected quality, which would impact consumer utility and thereby CS. The study drew upon the satisfaction and ECT literatures to revisit the concept of CS and extend existing research on measuring CS. Although the distinction between the traditional measure and quality-adjusted measure of CS ( $\Delta CS$ ) may sound intuitive at first brush, it indeed needs to be formalized and empirically tested.

The results testify to the economic role of markets with asymmetric information and the importance of the quality-adjusted measure of CS: First, our results provide empirical evidence for the level of CS in online markets. On average, the buyer reaps an average CS of \$94.4-\$175.5 per project, which is about 69%-128% of the winning bid price (\$137), the amount that the buyer actually paid, which is considerable, implying that buyers are paying much less than what they are willing to pay for the services they outsource in online markets.

Second, we theorized CS in the context of online markets for the outsourcing of services. Given the nature of services, we stress the difference between *ex ante* quality expectation and *ex-post* quality confirmation; nevertheless, we expect our quality-adjusted measure of CS to generalize to other markets with asymmetric information (such as eBay.com) since every purchase has a major part of *ex-ante* expectation and *ex-post* evaluation (satisfaction with the product/service). As nearly all products and services have a post-purchase evaluation aspect, it is important to take consumer *ex post* satisfaction into account when estimating CS.

Third, our results also stand to confirm the significant difference between the quality-adjusted measure and traditional measure of CS ( $\Delta CS$ ) in terms of their level, effects, and predictors. Though markets with asymmetric information offer great benefits to buyers, given the economic reality of consumer post-purchase satisfaction, we provide evidence that online markets may actually be producing lower surplus when using the quality-adjusted than the traditional CS measure. Our results support the predictive superiority of the quality-adjusted CS measure. Most important, only the quality-adjusted measure of CS and  $\Delta CS$  are shown to predict market performance. Finally, buyer's experience, repeat transactions, and a common language are shown to positively affect the *ex-post* extra utility ( $\Delta CS$ ) received by the buyer.

### Implications for the Predictive Power of the Quality Adjusted CS Measure

The natural question is why *ex post* quality was *not* factored into CS models? Anecdotal evidence shows many online transactions to end up in dissatisfaction, product returns, and disputes, due to products not matching buyer needs, product defects, and low quality service (such as fulfillment delays) (WSJ 2008); thus, total product returns are very high (Blanchard 2005; Blanchard 2007). Therefore, when buyer *ex-post* satisfaction is taken into account, previous research employing the traditional CS measure may have overestimated the true level of CS. By leveraging ECT to integrate the notions of *ex-post* quality and satisfaction to CS theories, we conceptually show that the traditional measure of CS may either overestimate or underestimate the level of CS because it is only an *ex-ante* concept. Prior studies may have either inflated or deflated CS because the implicit assumption was that buyers are fully satisfied with the products they receive, albeit the reality is that buyer's satisfaction is dependent on the confirmation / disconfirmation of their expectations of product/service quality. Empirically, the traditional measure of

CS is shown to inflate the true CS and not predict market performance. Despite our focus on online labor markets for outsourcing of services, our quality-adjusted CS measure could apply to virtually all products. Thus, it is important to include consumption experience and ex post satisfaction when calculating CS to avoid a biased estimation of societal welfare.

Since the function and sustainability of markets with asymmetric information depend to a large extent on “job supply”, it is important for these markets to increase CS with superior designs. First, since experienced buyers enjoy higher CS, the intermediary should educate new buyers how to navigate the marketplace so they do not suffer from low surplus that may cause them to exit the market. Second, since global labor arbitrage seems to further increase CS for buyers, markets can facilitate service projects between richer and poorer countries. Third, since the language barrier is shown to reduce CS, it is important for signals about the provider’s language proficiency (English or other languages) to be visible.

### ***Limitations and Future Research***

This paper makes several assumptions that may limit its generalizability. First, in the empirical part, as far as the budget as a proxy for *ex-ante* WTP is concerned, it is possible that the buyer would strategically enter an amount to attract the optimal number of providers to bid for services. Future research could examine whether and how different types (such as experienced or inexperienced outsourcers) would strategically design their RFPs to maximize their CS.

Second, we made an assumption that provider average feedback rating is a buyer’s *ex-ante* expectation of quality. We believe it makes sense, as according to the prior literature and our analyses, the aggregate measure of reputation is the most important aspect of quality. Future research could re-collect the data and cluster the data by buyer, and measure each buyer’s willingness to pay individually. We can also use buyer preference estimation (conditional logit model) to estimate buyer willingness to pay, and then use buyer preference function to construct a willingness to pay measure.

Third, we also observed that variance of CS with quality-adjusted measure is larger than the traditional CS measure. This could imply that buyer risk (uncertainty) has not been fully compensated. Therefore, future research can leverage uncertainty theories to examine that buyers in online markets for IT services may extract a lower surplus than they expected on the market level.

### ***Concluding Remark***

This study makes an initiative in better understanding how and by how much CS is accrued in markets with asymmetric information where *ex ante* expectations are susceptible to deviation from the actual *ex-post* quality received. Our work contributes to the literature by quantifying the economic benefits of markets with asymmetric information. The growing value for consumers is the main reasons that these markets can survive and thrive despite information asymmetry. Our study calls for a more precise measure of CS that reflects economic reality. The quality-adjusted measure helps make more accurate predictions in terms of the level, effects, and predictors of CS in markets with asymmetric information.



## REFERENCES

- Anderson, E.W., and Sullivan, M.W. 1993. "The Antecedents and Consequences of Customer Satisfaction for Firms," *Marketing Science* (12:2), pp 125-143.
- Banker, R.D., and Hwang, I. 2008. "Importance of Measures of Past Performance: Empirical Evidence on Quality of E-Service Providers," *Contemporary Accounting Research* (25:2), pp 307-337.
- Bapna, R., Jank, W., and Shmueli, G. 2008. "Consumer Surplus in Online Auctions," *Information Systems Research* (19:4), pp 400-416.
- Bhattacharjee, A. 2001. "Understanding Information Systems Continuance: An Expectation-Confirmation Model," *MIS Quarterly*, pp 351-370.
- Blanchard, D. 2005. "Moving Forward in Reverse Logistics Today," *Industry Week* ([http://logisticstoday.com/logistics\\_services/outlog\\_story\\_7301/](http://logisticstoday.com/logistics_services/outlog_story_7301/)).
- Blanchard, D. 2007. "Supply Chains Also Work in Reverse," *Industry Week* (<http://www.industryweek.com/ReadArticle.aspx?ArticleID=13947>).
- Breslaw, J.A., and Smith, J.B. 1995. "A Simple and Efficient Method for Estimating the Magnitude and Precision of Welfare Changes," *Journal of Applied Econometrics* (10:3), pp 313-327.
- Brynjolfsson, E. 1996. "The Contribution of Information Technology to Consumer Welfare," *Information Systems Research* (7:3), pp 281-300.
- Brynjolfsson, E., Hu, Y.J., and Smith, M.D. 2003. "Consumer Surplus in the Digital Economy: Estimating the Value of Increased Product Variety at Online Booksellers," *Management Science* (49:11), pp 1580-1596.
- Brynjolfsson, E., and Smith, M.D. 2000. "Frictionless Commerce? A Comparison of Internet and Conventional Retailers," *Management Science* (46:4), pp 563-585.
- Burns, M.E. 1973. "A Note on the Concept and Measure of Consumer's Surplus," *The American Economic Review* (63:3), pp 335-344.
- Cameron, A.C., and Trivedi, P.K. 2009. *Microeconometrics Using Stata*. Stata Press College Station, TX.
- Chandler, A.D. 1977. *The Visible Hand: The Managerial Revolution in American Business*. Belknap Press.
- Chase, R.B., and Apte, U.M. 2007. "A History of Research in Service Operations: What's the Big Idea?," *Journal of Operations Management* (25:2), pp 375-386.
- Cicchetti, C.J., and III, A.M.F. 1971. "Option Demand and Consumer Surplus: Further Comment," *The Quarterly Journal of Economics* (85:3), pp 528-539.
- Clark, C. 1983. *The Conditions of Economic Progress*. Garland Pub. (New York).
- DeSarbo, W.S., Huff, L., Rolandelli, M.M., and Choi, J. 1994. "On the Measurement of Perceived Service Quality: A Conjoint Analysis Approach." Thousand Oaks, CA: Sage Publications, p. 199-220.
- Dimoka, A., Hong, Y., and Pavlou, P.A. 2012. "On Product Uncertainty in Online Markets: Theory and Evidence," *MIS Quarterly* (32:3), pp. 395-426..
- Economist, T. 2010. "A Clouded Future: Online Services That Match Freelancers with Piecework Are Growing in Hard Times.," in: *The Economist*. Available online at: <http://www.economist.com/node/16116919>.
- Fishbein, M. 1967. "Attitude and the Prediction of Behavior," *Readings in attitude theory and measurement*, pp 477-492.
- Fishburn, P.C. 1970. "Utility Theory for Decision Making," DTIC Document.
- Gefen, D., and Carmel, E. 2008. "Is the World Really Flat? A Look at Offshoring at an Online Programming Marketplace," *MIS Quarterly* (32:2), pp 367-384.
- Goolsbee, A., and Petrin, A. 2004. "The Consumer Gains from Direct Broadcast Satellites and the Competition with Cable Tv," *Econometrica* (72:2), pp 351-381.

- Grover, V., and Ramanlal, P. 1999. "Six Myths of Information and Markets: Information Technology Networks, Electronic Commerce, and the Battle for Consumer Surplus," *MIS quarterly* (23:4), pp 465-495.
- Gupta, S., and Zeithaml, V. 2006. "Customer Metrics and Their Impact on Financial Performance," *Marketing science*), pp 718-739.
- Harris, J., and Blair, E.A. 2006. "Consumer Preference for Product Bundles: The Role of Reduced Search Costs," *Journal of the Academy of Marketing Science* (34:4), pp 506-513.
- Hausman, J.A. 1978. "Specification Tests in Econometrics," *Econometrica: Journal of the Econometric Society*), pp 1251-1271.
- Hausman, J.A., and Newey, W.K. 1995. "Nonparametric Estimation of Exact Consumers Surplus and Deadweight Loss," *Econometrica: Journal of the Econometric Society*), pp 1445-1476.
- Hitt, L.M., and Brynjolfsson, E. 1996. "Productivity, Business Profitability, and Consumer Surplus: Three Different Measures Of," *MIS Quarterly*, pp 121-142.
- Hölmstrom, B. 1979. "Moral Hazard and Observability," *The Bell Journal of Economics* (10:1), pp 74-91.
- Hong, Y., and Pavlou, P. (2012). "An Empirical Investigation on Provider Pricing in Online Crowdsourcing Markets for IT Services." *International Conference on Information Systems (ICIS)*, Orlando, FL.
- Howe, J. 2006. "The Rise of Crowdsourcing" in: *Wired*.  
<http://www.wired.com/wired/archive/14.06/crowds.html>.
- Inman, J.J., Dyer, J.S., and Jia, J. 1997. "A Generalized Utility Model of Disappointment and Regret Effects on Post-Choice Valuation," *Marketing Science*, pp 97-111.
- Irvine, I.J., and Sims, W.A. 1998. "Measuring Consumer Surplus with Unknown Hicksian Demands," *The American Economic Review* (88:1), pp 314-322.
- Kahneman, D., and Tversky, A. 1979. "Prospect Theory: An Analysis of Decision under Risk," *Econometrica: Journal of the Econometric Society* (47:2), pp 263-291.
- Keaveney, S.M. 1995. "Customer Switching Behavior in Service Industries: An Exploratory Study," *The Journal of Marketing* (59:2), pp 71-82.
- Kim, D.J., Ferrin, D.L., and Rao, H.R. 2009. "Trust and Satisfaction, Two Stepping Stones for Successful E-Commerce Relationships: A Longitudinal Exploration," *Information Systems Research* (20:2), pp 237-257.
- Lawton, C. 2008. "The War on Returns. Wall Street Journal. Available Online At: <http://Online.Wsj.Com/Article/Sb121020824820975641.Html> Published May 8. Accessed on April 19, 2010,").
- Lucking Reiley, D., Bryan, D., Prasad, N., and Reeves, D. 2007. "Pennies from Ebay: The Determinants of Price in Online Auctions," *The Journal of Industrial Economics* (55:2), pp 223-233.
- Malone, T.W., and Laubacher, R.J. 1998. "The Dawn of the E-Lance Economy," *Harvard Business Review* (76).
- Marshall, A. 1920. *Principles of Economics: Eighth Edition*. 8th ed. MacMillan and Company Ltd., London.
- McKinney, V., Yoon, K., and Zahedi, F.M. 2002. "The Measurement of Web-Customer Satisfaction: An Expectation and Disconfirmation Approach," *Information Systems Research* (13:3), pp 296-315.
- Melnik, M.I., and Alm, J. 2005. "Seller Reputation, Information Signals, and Prices for Heterogeneous Coins on Ebay," *Southern Economic Journal* (72:2), pp 305-328.
- Mithas, S., Jones, J.L., and Mitchell, W. 2008. "Buyer Intention to Use Internet-Enabled Reverse Auctions? The Role of Asset Specificity, Product Specialization, and Non-Contractibility," *MIS Quarterly* (32:4), pp 705-724.

- Nevo, A. 2003. "New Products, Quality Changes, and Welfare Measures Computed from Estimated Demand Systems," *Review of Economics and Statistics* (85:2), pp 266-275.
- Oliver, R.L. 1977. "Effect of Expectation and Disconfirmation on Postexposure Product Evaluations: An Alternative Interpretation," *Journal of Applied Psychology* (62:4), pp 480-486.
- Oliver, R.L. 1980. "A Cognitive Model of the Antecedents and Consequences of Satisfaction Decisions," *Journal of Marketing Research* (17:4), pp 460-469.
- Randall, A., and Stoll, J.R. 1980. "Consumer's Surplus in Commodity Space," *The American Economic Review* (70:3), pp 449-455.
- Resnick, P., and Zeckhauser, R. 2002. "Trust among Strangers in Internet Transactions: Empirical Analysis of Ebay's Reputation System," *Advances in Applied Microeconomics: A Research Annual* (11), pp 127-157.
- Roach, S.S. 2004. "More Jobs, Worse Work," *the New York Times*), pp <http://www.nytimes.com/2004/2007/2022/opinion/more-jobs-worse-work.html>.
- Robinson, L.M. 1979. "Consumer Complaint Behavior: A Review with Implications for Further Research," *New Dimensions of Consumer Satisfaction and Complaining Behavior* (3), pp 41-50.
- Rust, R.T., Inman, J.J., Jia, J., and Zahorik, A. 1999. "What You Don't Know About Customer-Perceived Quality: The Role of Customer Expectation Distributions," *Marketing science*), pp 77-92.
- Rust, R.T., Zahorik, A.J., and Keiningham, T.L. 1995. "Return on Quality (Roq): Making Service Quality Financially Accountable," *The Journal of Marketing*), pp 58-70.
- Schaffer, M.E., and Stillman, S. 2006. "Xtoverid: Stata Module to Calculate Tests of Overidentifying Restrictions after Xtreg, Xtivreg, Xtivreg2, Xthtaylor," *Statistical Software Components*).
- Smith, A. 1961. "An Inquiry into the Nature and Causes of the Wealth of Nations,").
- Snir, E.M., and Hitt, L.M. 2003. "Costly Bidding in Online Markets for It Services," *Management Science* (49:11), pp 1504-1520.
- Song, M. 2007. "Measuring Consumerwelfareinthe Cpu Market: Anapplication of the Pure,Â&characteristics Demand Model," *The RAND Journal of Economics* (38:2), pp 429-446.
- Spence, M. 1973. "Job Market Signaling," *The quarterly journal of Economics* (87:3), pp 355-374.
- Spreng, R.A., MacKenzie, S.B., and Olshavsky, R.W. 1996. "A Reexamination of the Determinants of Consumer Satisfaction," *The Journal of Marketing* (60:3), pp 15-32.
- Susarla, A., Barua, A., and Whinston, A.B. 2003. "Understanding the Service Component of Application Service Provision: An Empirical Analysis of Satisfaction with Asp Services," *MIS quarterly*), pp 91-123.
- Turnovsky, S., Shalit, H., and Schmitz, A. 1980. "Consumer's Surplus, Price Instability, and Consumer Welfare," *Econometrica: Journal of the Econometric Society* (48:1), pp 135-152.
- Vargo, S.L., and Lusch, R.F. 2008. "Service-Dominant Logic: Continuing the Evolution," *Journal of the Academy of Marketing Science* (36:1), pp 1-10.
- Vartia, Y.O. 1983. "Efficient Methods of Measuring Welfare Change and Compensated Income in Terms of Ordinary Demand Functions," *Econometrica: Journal of the Econometric Society*), pp 79-98.
- Wooldridge, J.M. 2002. *Econometric Analysis of Cross Section and Panel Data*. The MIT press.
- WSJ. 2008. "The War on Returns," in: *Wall Street Journal*. <http://online.wsj.com/article/SB121020824820975641.html> Published May 8. Reporter: Christopher Lawton.