Efficiency of electronic service allocation with privately known quality

Article *in* European Journal of Operational Research · December 2016

DOI: 10.1016/j.ejor.2016.05.055

2 authors:

Tobias Widmer
University of Hohenheim

Joerg Leukel
University of Hohenheim

Efficiency of electronic service allocation with privately known quality

Tobias Widmer[a,*], Joerg Leukel[a]

[a]*Faculty of Business, Economics and Social Sciences, University of Hohenheim, Schwerzstr. 35, 70599 Stuttgart, Germany*

This is the authors version of the following journal article:

Widmer, T. and Leukel, J. (2016). Efficiency of electronic service allocation with privately known quality. *European Journal of Operational Research, 255(3), pp. 856-868.* http://dx.doi.org/10.1016/j.ejor.2016.05.055

*Corresponding author. Tel.: +49 711 459 22274; Fax: +49 711 459 22961.

Email addresses: tobias.widmer@uni-hohenheim.de (Tobias Widmer), joerg.leukel@uni-hohenheim.de (Joerg Leukel)

Abstract

We characterize how a social planner can design electronic service allocation when the desired service quality of customers and the actual service quality of providers are private information. Because private information is present in our analysis, we derive a second-best allocation mechanism that satisfies incentive compatibility, individual rationality and budget balance. While using the first-best outcome as a benchmark, we study the efficiency properties of the associated optimal allocation rules. In a set of simulation experiments with uniformly and normally distributed private information, we find that the asymptotic efficiency of the second-best mechanism is bounded away from 100% even for a large number of customers and providers. This finding indicates that the agents in our model do not become informationally small as the market size increases.

Keywords: Auctions/bidding, electronic services, incomplete information, double-sided mechanism

1. Introduction

The increasing number of third party vendors offering electronic services has stimulated the growth of information technology (IT) outsourcing in many organizations (Chang and Gurbaxani, 2012). Firms outsource their applications to these vendors that provide access to computing resources at a specified quality of service (QoS) such as availability, throughput, and execution time (Gartner, 2014). Advances in IT facilitate the substitution of traditionally static and long-term relationships by flexible contracts of shorter duration, with Cloud computing being the most recent manifestation of these advances (Armbrust et al., 2010). Recently, marketplaces for Cloud services have emerged, for instance, the Avnet Cloud Marketplace (2016). Such marketplaces might be administrated by government authorities or large corporations, who aim at a socially optimal allocation. The overall objectives are to serve the customers and to best utilize geographically remote data centers. This scenario dates back to grid computing, a predecessor of Cloud computing, for which the social welfare properties have been studied in prior research (e.g., Schnizler et al. (2008); Stösser et al. (2010)).

On markets for electronic services, multiple providers offer their services to multiple customers, who attempt to agree on an exchange of the services for money. However, for electronic services it is not viable to solely account for the price as the single non-functional property because customers usually have different requirements for the quality characteristics (O'Sullivan et al., 2002). On the other hand, service providers use the QoS for differentiation from the competition. Under consideration of the QoS, determining the optimal service allocation is difficult for two reasons. First, each customer's desired QoS is known only to that customer and each provider's actual QoS is known only to that provider. The social planner observes no one's desired or actual QoS and no trader observes the QoS of any other trader. Second, the allocation mechanism must guarantee four specific economic properties, which are common in optimal auction design (Myerson, 1981): (i) The mechanism must provide adequate incentives for the participants because strategic individuals may misreport their true preferences (incentive compatibility), (ii) the mechanism must not force individuals to participate in the market (individual rationality), (iii) the mechanism must omit any independent intermediary but facilitate distributed decision-making among the participants (Egri and Váncza, 2013); this requirement implies that all payments must be distributed among the participants (budget balance), and (iv) the mechanism must maximize the social welfare (ex post optimality). Standard impossibility theorems from mechanism design assert that meeting all four requirements simultaneously is not attainable (Laffont and Maskin, 1979; Myerson and Satterthwaite, 1983). Therefore, the social planner must decide about a viable tradeoff of these requirements. One compromise in the presence of private information is to derive a second-best mechanism and compare its outcome to the associated first-best outcome that would arise if all information were publicly known (e.g., Arya et al. (2015); Babich et al. (2012)). A second-best mechanism is one that maximizes the *expected* social welfare among all incentive compatible, individually rational, and budget-balanced mechanisms (Börgers, 2015).

Current approaches for integrating QoS into electronic service allocation use auction mechanisms to elicit QoS and price attributes for determining the optimal allocation (Bapna et al., 2008; Blau et al., 2010). However, classic auction theory is based on two assumptions: First, private information about reservation values exists on the customer side of the market only, while double-sided information asymmetry is generally not considered (Myerson, 1981). Second, the offered quality is fixed prior to provider selection (Bichler and Kalagnanam, 2005). Online auctions that use quality attributes apart from the price also affect the auction's outcome throughout the allocation process (Bockstedt and Goh, 2011). In settings with double-sided private information, it is not clear that the QoS actually offered by a provider will match the desired QoS of the customer. More specifically, when every provider offers a service of distinct QoS and every customer has distinct needs, facilitating allocations between the right pairs of traders is critical for maximizing the total welfare of a market. Therefore, the privately known desired QoS of customers and the privately known actual QoS of providers must be internalized into the allocation mechanism. However, how this integration affects the allocation outcome is still not known.

We address the problem of optimal service allocation on double-sided markets with private information. We draw on mechanism design to derive a second-best allocation mechanism for electronic services. We analyze a market where gains from trade that can be generated depend on the privately known QoS of the matched customers and providers. While deriving the optimal allocation rules from the perspective of a social planner, we study the efficiency properties of the set of mechanisms that satisfy incentive compatibility, individual rationality and budget-balance. To identify these optimal rules, we focus on direct revelation mechanisms by invoking the revelation principle (Myerson, 1979) in a first step. Then, the issue of real-world relevance is addressed by the implementation through position auctions. Afterwards, we report on a set of simulation experiments.

The objectives of this research are to: (i) derive a second-best mechanism for allocating electronic services with private information about QoS and (ii) evaluate this mechanism in a set of experiments to study its efficiency properties. Our proposal is informed by the work of Johnson (2013), who derived a profit-maximizing matching mechanism with double-sided private information. While the approach examined by Johnson (2013) focuses solely on mechanisms that maximize the expected profit of the auctioneer, we derive the optimal allocation rules from the perspective of a social planner that seeks to maximize the expected social welfare. We are particularly interested in studying the efficiency properties of the second-best allocation mechanism. In our prior research (Widmer et al., 2013), we designed a mechanism for a specific allocation problem in a Cloud computing scenario by integrating energy efficiency as a particular QoS into the preferences of market participants. We advance this mechanism by (i) making customers' desired QoS and providers' actual QoS intrinsic parts of the mechanism and (ii) accounting for the four economic properties in the presence of double-sided private information.

The remainder of this article is organized as follows. In section 2, we discuss the approaches to QoS-aware electronic service allocation. In section 3, we describe our mechanism for service allocation with private information and study its efficiency properties (section 4). In section 5, we report on the experimental evaluation and discuss our findings. We provide our conclusion in section 6.

2. Literature review

We discuss extant literature on mechanism design and examine the integration of QoS as an intrinsic part of the mechanism. The field of mechanism design studies how privately known preferences of multiple individuals, also called agents, can be aggregated toward a social choice (Nisan and Ronen, 2001). For making customers' desired QoS and providers' actual QoS an intrinsic part of the mechanism, research concerned with multidimensional auctions and matching mechanisms is of particular interest.

A basic procurement auction with two dimensions, namely price and quality, was proposed by Che (1993). In this auction a single customer announces a publicly known scoring rule to multiple providers that are competing for winning a project. Price and quality preferences are aggregated into the utility function of each participant. The outcome is optimal for the customer if she can commit to a scoring rule in her best interest. This model was extended by Branco (1997) by integrating correlated cost types of the providers into the utility functions. The objective function of the single customer is the maximization of the social welfare instead of the customer's pay-off as in Che's model. Branco's approach shows that the customer needs to use a two-stage auction to implement the optimal outcome: the customer (i) selects one provider, and (ii) bargains to readjust the level of quality to be provided. However, their model is concerned with a single provider only and does not consider double-sided competition.

Specific auctions for the procurement of electronic services with multiple attributes have also been studied. Bichler and Kalagnanam (2005) proposed solutions to winner determination problems in multidimensional auctions with multiple sourcing and configurable offers. They examined the impact of several business rules such as propositional logic for knowledge representation, that need to be imposed on the winner determination problem in order to obtain an acceptable supply from multiple providers. Although this mechanism maximizes the customer's utility in experimental settings, their approach does not examine the effect of double-sided competition under private information on both sides of the market.

A multidimensional combinatorial auction mechanism for trading electronic grid services among multiple providers and customers was presented by Schnizler et al. (2008). The proposed auction maximizes the social welfare and satisfies incentive compatibility in equilibrium but requires an outside subsidization because it runs a permanent deficit in budget. In our approach, we resort to mechanisms that maximize the expected social welfare to the extent that budget balance can be achieved.

The mechanism designed by Blau et al. (2010) is a multidimensional procurement auction for trading so-called composite electronic services. A composite service is a set of several elementary services and thus represents a domain-specific bundle. Multiple service providers offer composite services to a customer who specifies requirements through a bidding language that considers multiple QoS attributes. The proposed mechanism is incentive compatible in weakly-dominant strategies, individually rational, and optimal ex post, but limited to a single-sided market environment.

Mechanisms of double-sided private information have been subject of inquiry in economic theory. A seminal piece of work stems from Myerson and Satterthwaite (1983), who studied efficient market mechanisms for bilateral trading when reservation values are private information. This model was then extended by McAfee (1991) by integrating a continuous quantity parameter into both traders' private reservation function. Efficiency requires the mechanism to decide when and how much of a certain commodity shall be traded. The approach of Myerson and Satterthwaite (1983) is similar to our work, though limited to bilateral trading, whereas we consider a setting with multiple customers and providers.

The bilateral mechanism of Myerson and Satterthwaite (1983) was generalized by Gresik and Satterthwaite (1989) to allow for multiple customers and providers with double-sided private information.

It was found that market inefficiency disappears as the number of traders becomes large. Yet in this model, the optimal match of customers and providers with privately known QoS cannot be determined because it does not matter which customer trades with which provider. However, in settings with privately known QoS on both market sides matching the right pairs of traders is crucial to ensure that the highest possible match surplus can be elicited by the mechanism.

In general, double-sided matching mechanisms may be used to determine the allocation of electronic services. In these mechanisms, customers and providers only produce mutual surplus if they are matched together. For a setting of double-sided private information, Johnson (2013) invoked the revelation principle to derive the optimal matching and payment rules from the perspective of a profit-maximizing intermediary. One way to implement this mechanism is with position auctions, in which multiple customers and providers simultaneously submit their bids to the mechanism, the bids are ranked against each other and the individuals are allocated accordingly. Johnson (2013) studied two different formats used in position auctions, namely the winners-pay format and the all-pay format. In the winners-pay format only matched individuals pay their bid to the auctioneer. The all-pay format requires all individuals to pay their bid, regardless of a successful match. However, Johnson focused solely on maximizing the auctioneer's profit. In our work, we devise the optimal allocation rules from the perspective of a social planner that seeks to maximize the expected gains from trade. Unlike Johnson, we are particularly interested in studying the efficiency properties of the second-best set of mechanisms for electronic service allocation.

In summary, the review of the extant literature corroborates that private quality information has not been sufficiently internalized into double-sided electronic service allocation. Our approach is informed by the work of Johnson (2013) and extends it by (i) deriving the optimal allocation rules from a social welfare perspective and (ii) studying the efficiency properties of these optimal rules by simulation.

3. Formal framework

This section introduces a framework providing definitions and formal notations for the design of a double-sided service allocation mechanism. The framework is based on the model of Johnson (2013) but focuses on the social welfare properties of the mechanism for optimal customer-provider matching.

3.1. Agents

The set of customer agents is given by $\mathscr{A}_C$ with $N = |\mathscr{A}_C|$, and the set of provider agents is $\mathscr{A}_P$ with $M = |\mathscr{A}_P|$. Each customer agent $a_i \in \mathscr{A}_C$ requests a single, standardized electronic service (unit demand), while each provider agent $b_j \in \mathscr{A}_P$ offers a single electronic service (unit supply). Standardized services are electronic services of equal functional properties. The maximum number of contracts between a_i and b_j is given by $K = \min\{M, N\}$. Assume $K \geq 2$.

3.2. Private information

In the following, the terms quality and QoS are used interchangeably. Each customer agent privately learns the desired quality θ_i that she has assigned to the requested service. Desired quality is drawn from a probability density function $f(\theta_i)$, being strictly positive on $[0, \bar{\theta}]$ with cumulative distribution function

$F(\theta_i)$. On the supply side, each provider agent privately learns the actual quality σ_j that she can offer for the electronic service. Actual quality is drawn from a probability density function $h(\sigma_j)$, being strictly positive on $[0, \bar{\sigma}]$ with cumulative distribution function $H(\sigma_j)$. Let $\theta = (\theta_1, \dots, \theta_N)$ and $\sigma = (\sigma_1, \dots, \sigma_M)$ be the private information vectors of all customer agents and provider agents, respectively. The vectors $\theta_{-i} = (\theta_1, \dots, \theta_{i-1}, \theta_{i+1}, \dots, \theta_N)$ and $\sigma_{-j} = (\sigma_1, \dots, \sigma_{j-1}, \sigma_{j+1}, \dots, \sigma_M)$ specify the private information except for a_i respectively b_j. Further, let $\mathbb{E}_{\theta_{-i},\sigma}$ denote the expectation over all private information conditional on customer agent a_i's information, and let $\mathbb{E}_{\theta,\sigma_{-j}}$ denote the analogous expectation for provider agent b_j. The unconditional expectation over all private information is denoted by $\mathbb{E}$.

3.3. Valuation and cost

Customer agent a_i's valuation for consuming a service is given by $v(\theta_i, \sigma_j)$. This valuation can be interpreted as a_i's maximum willingness to pay for the service. It depends on a_i's desired quality, as well as on the difference between its own desired quality and provider b_j's actual quality. The definition of $v(\theta_i, \sigma_j)$ captures the fact that provider agent b_j does not know a_i's willingness to pay. We focus on customer valuation functions that are non-monotonic in the quality offered by the provider. For instance, a customer may prefer a service with medium over high computational capacity. A high-capacity service is well able to process many simultaneous requests from the customer's application that uses this service. If, however, this application does not have enough computational power or resources, the application will fail to answer these simultaneous requests in due time. This leads to higher buffering in the application and thus longer response times. Thus, a customer's valuation must take into account the application that uses the service and the tradeoff between being idle or buffering heavily (Hang and Singh, 2010). Therefore, we assume that any mismatch in desired quality and actual quality creates adjustment problems for the customer agent. That is, $v(\theta_i, \sigma_j)$ is maximized when the supplied quality and the desired quality are equal (i.e., when $\theta_i = \sigma_j$). By assumption, the maximal value is increasing in θ_i.

On the supply side, each provider agent b_j's provision cost $c(\theta_i, \sigma_j)$ for its service depends on its actual quality and on the difference between its own actual quality and the customer agent's desired quality. This definition suggests that customer agents do not know the provision costs of provider agents. If a provider produces a quality lower than the quality desired by a customer, this provider incurs higher cost from not fulfilling the requirements. If, in contrast, a provider maintains higher quality than desired, her cost increases due to idle resources (Greenberg et al., 2009). Hence, we assume that $c(\theta_i, \sigma_j)$ is minimized when $\sigma_j = \theta_i$ and that the minimal value is increasing in σ_j. This assumption captures the fact that a mismatch in actual quality and desired quality creates higher provision costs resulting from after-sales customer service cost and missed opportunity cost. Both $v(\theta_i, \sigma_j)$ and $c(\theta_i, \sigma_j)$ are assumed to be thrice differentiable.

3.4. Expected utilities

Similar to common auction settings, we assume risk-neutral agents with quasi-linear utility functions. For describing an allocation of a service we use a decision variable. Let $x_{ij}(\theta, \sigma) \in [0, 1]$ denote the probability that a service is allocated from provider agent b_j to customer agent a_i. For example, if $x_{11}(\theta, \sigma) = 1$, the service is provided by provider agent b_1 and consumed by customer agent a_1 in the final allocation. For ease of exposition, we use vectorized function arguments without further reference and, for example, write $x_{ij}(\theta_i, \theta_{-i}, \sigma) \equiv x_{ij}(\theta_1, \ldots, \theta_N, \sigma_1, \ldots, \sigma_M)$ in any appropriate context.

Using this notation, each customer agent's expected utility function is given by

$$U_C(\theta_i) = \mathbb{E}_{\theta_{-i},\sigma}\Big[\sum_{j=1}^{M} v(\theta_i, \sigma_j) x_{ij}(\theta_i, \theta_{-i}, \sigma) - t_C(\theta_i, \theta_{-i}, \sigma)\Big], \tag{1}$$

where $t_C(\theta_i, \theta_{-i}, \sigma)$ is the monetary transfer made by a_i conditional on all other agents' private information. Similarly, the expected utility of provider agent b_j is given by

$$U_P(\sigma_j) = \mathbb{E}_{\theta,\sigma_{-j}}\Big[t_P(\theta, \sigma_j, \sigma_{-j}) - \sum_{i=1}^{N} c(\theta_i, \sigma_j) x_{ij}(\theta, \sigma_j, \sigma_{-j})\Big], \tag{2}$$

where $t_P(\theta, \sigma_j, \sigma_{-j})$ corresponds to the monetary compensation provider agent b_j receives for providing a service to customer agent a_i.

3.5. *Incentive compatibility and individual rationality*

We characterize the proposed mechanism in terms of a direct revelation mechanism (Myerson, 1979). In a direct revelation mechanism all agents are induced to reveal their private information (i.e., desired and actual quality respectively) to the mechanism, which then dictates the allocation of services and the respective monetary transfers. In equilibrium, all agents will report truthfully. Hence, a direct mechanism is defined by the decision variables, as well as the payments for service allocation. Formally, a direct mechanism is a set of functions $\{x_{ij}(\theta, \sigma), t_C(\theta, \sigma), t_P(\theta, \sigma)\}$ for all agents a_i and b_j.

In order to determine the optimal equilibrium outcome from among all possible mechanisms, we adopt the standard approach in mechanism design theory by invoking the revelation principle (Gibbard, 1973; Harris and Raviv, 1981; Myerson, 1981). The application of the revelation principle allows us to restrict our attention to incentive compatible direct mechanisms without loss of generality. Formally, the mechanism is (Bayesian) incentive compatible if and only if for each customer agent's reported desired quality $\hat{\theta} \neq \theta_i$

$$U_C(\theta_i) \geq \mathbb{E}_{\theta_{-i},\sigma}\Big[\sum_{j=1}^{M} v(\theta_i, \sigma_j) x_{ij}(\hat{\theta}, \theta_{-i}, \sigma) - t_C(\hat{\theta}, \theta_{-i}, \sigma)\Big], \tag{3}$$

and for each provider agent's reported actual quality $\hat{\sigma} \neq \sigma_j$

$$U_P(\sigma_j) \geq \mathbb{E}_{\theta,\sigma_{-j}}\Big[t_P(\theta, \hat{\sigma}, \sigma_{-j}) - \sum_{i=1}^{N} c(\theta_i, \sigma_j) x_{ij}(\theta, \hat{\sigma}, \sigma_{-j})\Big]. \tag{4}$$

In addition to incentive compatibility, the mechanism must satisfy individual rationality, so that each agent willingly participates in the mechanism. This constraint ensures that conditional on its private information every agent imputes an expected utility from participating in the mechanism that is greater than or equal to the utility of its outside option. Formally, the mechanism is individually rational if and only if $U_C(\theta_i) \geq 0$ for customer agent a_i and $U_P(\sigma_j) \geq 0$ for provider agent b_j.

3.6. Optimization problem

This research takes the perspective of a social planner, who is interested in distributing all payments among the agents. Hence, the ultimate objective of the mechanism is the maximization of the expected social welfare subject to budget balance. The expected social welfare is defined as the sum of all agents' expected utilities. The mechanism balances the budget if all expected transfers made among the agents add up to zero. That is, (ex ante) budget balance is defined as

$$\sum_{i=1}^{N} \int_0^{\bar{\theta}} \mathbb{E}_{\theta_{-i},\sigma}[t_C(\theta_i, \theta_{-i}, \sigma)] f(\theta_i) d\theta_i - \sum_{j=1}^{M} \int_0^{\bar{\sigma}} \mathbb{E}_{\theta,\sigma_{-j}}[t_P(\theta, \sigma_j, \sigma_{-j})] h(\sigma_j) d\sigma_j = 0. \tag{5}$$

This constraint is necessary because the mechanism must not depend on any external source of funds. By accumulating the expected utilities in (1) and (2) over all agents and substituting for the constraint of budget balance (5), we get the expected social welfare of the service allocation as follows:

$$\begin{aligned}
&\sum_{i=1}^{N} \int_0^{\bar{\theta}} U_C(\theta_i) f(\theta_i)\, d\theta_i + \sum_{j=1}^{M} \int_0^{\bar{\sigma}} U_P(\sigma_j) h(\sigma_j)\, d\sigma_j \\
&\quad = \sum_{i=1}^{N} \int_0^{\bar{\theta}} \mathbb{E}_{\theta_{-i},\sigma}\Big[\sum_{j=1}^{M} v(\theta_i, \sigma_j) x_{ij}(\theta_i, \theta_{-i}, \sigma) - t_C(\theta_i, \theta_{-i}, \sigma) \Big] f(\theta_i)\, d\theta_i \\
&\qquad + \sum_{j=1}^{M} \int_0^{\bar{\sigma}} \mathbb{E}_{\theta,\sigma_{-j}} \Big[t_P(\theta, \sigma_j, \sigma_{-j}) - \sum_{i=1}^{N} c(\theta_i, \sigma_j) x_{ij}(\theta, \sigma_j, \sigma_{-j}) \Big] h(\sigma_j)\, d\sigma_j \\
&\quad = \mathbb{E}\Big[\sum_{i=1}^{N} \sum_{j=1}^{M} (v(\theta_i, \sigma_j) - c(\theta_i, \sigma_j)) x_{ij}(\theta, \sigma) \Big].
\end{aligned} \tag{6}$$

The mechanism's objective is the maximization of the expression obtained in (6) over all agents. Therefore, the mechanism faces the following optimization problem:

$$\max_{x_{ij}} \mathbb{E}\Big[\sum_{i=1}^{N} \sum_{j=1}^{M} (v(\theta_i, \sigma_j) - c(\theta_i, \sigma_j)) x_{ij}(\theta, \sigma) \Big], \tag{7}$$

which must be solved subject to incentive compatibility, individual rationality and the feasibility constraints $0 \leq \sum_j x_{ij}(\theta, \sigma) \leq 1$ for all a_i and $0 \leq \sum_i x_{ij}(\theta, \sigma) \leq 1$ for all b_j.

4. Pricing mechanisms

In this section, we first characterize incentive compatibility and individual rationality, followed by defining the mechanism's allocation rule and expected monetary transfers. Then, we specify conditions sufficient for this mechanism to maximize the expected social welfare. Finally, we provide an illustrative example for a market with uniformly distributed private information.

4.1. Characterizing incentive compatibility and individual rationality

According to the Myerson-Satterthwaite-Theorem (Myerson and Satterthwaite, 1983), it is impossible to reach ex post efficiency when incentive compatibility, individual rationality and budget balance are required. Although other collective objective functions could be considered, this article focuses on the problem of maximizing the expected social welfare, which weights utility gains from each type of customer the same as utility gains from each type of provider. We adopt the approach of Myerson and Satterthwaite (1983), as well as Gresik and Satterthwaite (1989), and define the following two functions. Let the virtual valuation of customer agent a_i be given by

$$\psi_C(\theta_i,\sigma_j) = v(\theta_i,\sigma_j) - \frac{1-F(\theta_i)}{f(\theta_i)}\frac{\partial v(\theta_i,\sigma_j)}{\partial \theta_i}. \tag{8}$$

The virtual valuation function expresses the strategic behavior of customer agents in an indirect mechanism implementation. Since $v(\theta_i,\sigma_j)$ is maximized for $\theta_i = \sigma_j$ and increasing in the maximal value, $\psi_C(\theta_i,\sigma_j)$ is strictly smaller than the actual valuation. Hence, customer agents have an incentive to understate their true valuation. They behave strategically in this fashion in order to influence their transaction prices.

For provider agents, let the virtual provision cost be defined by

$$\psi_P(\theta_i,\sigma_j) = c(\theta_i,\sigma_j) + \frac{H(\sigma_j)}{h(\sigma_j)}\frac{\partial c(\theta_i,\sigma_j)}{\partial \sigma_j}. \tag{9}$$

The virtual provision cost reflects the strategic behavior of provider agents in an indirect mechanism implementation. Since $c(\theta_i,\sigma_j)$ is minimized when $\sigma_j = \theta_i$ and increasing in the minimal value, the virtual cost is strictly greater than the actual provision cost. Thus, a provider agent has an incentive to overstate her true cost. In such a case, provider agents try to raise their transaction prices for the offered service by reporting a higher provision cost.

These virtual reservation functions play a crucial role in deriving the second-best mechanism for electronic service allocation. The following Lemma characterizes the set of all incentive compatible and individually rational mechanisms for our allocation problem.

Lemma 1. *Let $x_{ij}(\theta,\sigma)$ be the probability that provider agent b_j is allocated to customer agent a_i. Then transfer functions $t_C(\theta,\sigma)$ and $t_P(\theta,\sigma)$ exist such that $\{x_{ij}(\theta,\sigma), t_C(\theta,\sigma), t_P(\theta,\sigma)\}$ is incentive compatible and individually rational if and only if $\mathbb{E}_{\theta_{-i},\sigma}[x_{ij}(\theta_i,\cdot)]$ is non-decreasing, $\mathbb{E}_{\theta,\sigma_{-j}}[x_{ij}(\cdot,\sigma_j)]$ is non-increasing and*

$$\mathbb{E}\Big[\sum_{i=1}^{N}\sum_{j=1}^{M}(\psi_C(\theta_i,\sigma_j) - \psi_P(\theta_i,\sigma_j))x_{ij}(\theta,\sigma)\Big] \geq 0. \tag{10}$$

Proofs of this and all other Lemmas and Theorems are provided in the Appendix A.

Lemma 1 is crucial for constructing ex ante optimal mechanisms because it states that if the decision variables $x_{ij}(\theta,\sigma)$ satisfy the monotonicity properties, as well as inequality (10), then transfers $t_C(\theta,\sigma)$ and $t_P(\theta,\sigma)$ exist, such that $\{x_{ij}(\theta,\sigma), t_C(\theta,\sigma), t_P(\theta,\sigma)\}$ is an incentive compatible and individually rational mechanism. Notice that this framework explicitly allows for randomized transfers. However, in the presence of risk-neutral agents, we can use deterministic transfers equal to the corresponding expectations without loss of generality.

Apart from the condition for incentive compatibility and individual rationality, inequality (10) in Lemma 1 is the ex ante budget balance condition. The next Lemma establishes that the mechanism distributes all payments among the agents such that no external source of funds is necessary.

Lemma 2. *Any incentive compatible, individually rational mechanism satisfies ex ante budget balance.*

4.2. Allocation rule

The definition of an appropriate allocation rule with transfers for our mechanism design problem requires some additional notation. The difference in virtual valuation and virtual cost is defined as $\psi(\theta_i,\sigma_j) = \psi_C(\theta_i,\sigma_j) - \psi_P(\theta_i,\sigma_j)$. The reserve functions of customer agent a_i and provider agent b_j are respectively given by

$$R_C(\theta_i) = \begin{cases} \sup\{\sigma_j \in [0,\bar{\sigma}] : \psi(\theta_i,\sigma_j) \geq 0\} & \text{if} \quad \{\sigma_j : \psi(\theta_i,\sigma_j) \geq 0\} \neq \emptyset \\ 0 & \text{otherwise,} \end{cases} \tag{11}$$

$$R_P(\sigma_j) = \begin{cases} \inf\{\theta_i \in [0,\bar{\theta}] : \psi(\theta_i,\sigma_j) \geq 0\} & \text{if} \quad \{\theta_i : \psi(\theta_i,\sigma_j) \geq 0\} \neq \emptyset \\ \bar{\theta} & \text{otherwise.} \end{cases} \tag{12}$$

Reserve function $R_C(\theta_i)$ is used by a customer agent to identify the worst possible provider agent (i.e., the one with the highest actual quality) under the constraint $\psi(\theta_i,\sigma_j) \geq 0$, given her own desired quality θ_i. Reserve function $R_P(\sigma_j)$ is used by a provider agent to identify the worst possible customer agent (i.e., the one with the lowest desired quality) under the same constraint, given her own actual quality σ_j. In addition, the lowest possible desired quality among all customer agents and, respectively, the highest possible actual quality among all provider agents are defined as

$$\underline{\theta}_i = \inf\{\theta_i \in [0,\bar{\theta}] : \{\sigma_j : \psi(\theta_i,\sigma_j) \geq 0\} \neq \emptyset\}\} \quad \text{and} \tag{13}$$

$$\bar{\sigma}_j = \sup\{\sigma_j \in [0,\bar{\sigma}] : \{\theta_i : \psi(\theta_i,\sigma_j) \geq 0\} \neq \emptyset\}\}. \tag{14}$$

The customer agent with desired quality $\underline{\theta}_i$ is the worst-off customer agent in the market, while the provider agent with actual quality $\bar{\sigma}_j$ is the worst-off provider agent. Let $\rho_\theta(\theta_i) = |\{\theta_k \in \theta : \theta_k \geq \theta_i\}|$ be the rank of desired quality θ_i within the vector of all customer agents' desired qualities $\theta = \{\theta_1, \ldots, \theta_N\}$. Define $\rho_\sigma(\sigma_i)$ similarly for provider agents. The quantity

$$w_k^C(\theta_i) = \frac{(N-1)!}{(N-k)!(k-1)!} F(\theta_i)^{N-k}(1-F(\theta_i))^{k-1} \tag{15}$$

gives the probability that customer agent's desired quality θ_i has rank k within vector θ. Define $w_k^P(\sigma_j)$ similarly for provider agents. Further, let $f_{(k)}(\cdot)$ and $h_{(k)}(\cdot)$ be the probability density functions of the k-th order statistics of customer agents and provider agents respectively. The associated cumulative distribution functions are denoted by $F_{(k)}(\cdot)$ and $H_{(k)}(\cdot)$.

Consider the following direct revelation mechanism. The *Truncated Positive Assortative Allocation (TPAA) mechanism* is defined by the allocation rule

$$x_{ij}(\theta, \sigma) = \begin{cases} 1 & \text{if} \quad \rho_\theta(\theta_i) = \rho_\sigma(\sigma_i) = k \quad \text{and} \quad \psi(\theta_i, \sigma_j) \geq 0, \\ 0 & \text{otherwise} \end{cases} \tag{16}$$

with expected payments of customer agents

$$T_C(\theta_i) = \sum_{k=1}^{K} w_k^C(\theta_i) \int_0^{R_C(\theta_i)} v(\theta_i, r) h_{(k)}(r) dr - \int_{\underline{\theta}_i}^{\theta_i} \sum_{k=1}^{K} w_k^C(s) \int_0^{R_C(s)} \frac{\partial v(s,r)}{\partial s} h_{(k)}(r) dr ds \tag{17}$$

and expected compensation for provider agents

$$T_P(\sigma_j) = \sum_{k=1}^{K} w_k^P(\sigma_j) \int_{R_P(\sigma_j)}^{\bar{\theta}} c(r, \sigma_j) f_{(k)}(r) dr + \int_{\sigma_j}^{\bar{\sigma}_j} \sum_{k=1}^{K} w_k^P(s) \int_{R_P(s)}^{\bar{\theta}} \frac{\partial c(r,s)}{\partial s} f_{(k)}(r) dr ds. \tag{18}$$

The allocation rule of the TPAA mechanism captures two distinct features. First, a service is allocated from a provider agent to a customer agent if and only if the rank of a_i's desired quality is equal to the rank of b_j's actual quality. This rule implies that the service from the provider agent with the highest actual quality is allocated to the customer agent with the highest desired quality, the second-highest, and so on. Such mechanisms are positively assortative (Shimer and Smith, 2000). Second, service allocation from a provider agent to a customer agent takes place if and only if the difference in a_i's virtual valuation and b_j's virtual provision cost is positive.

Notice that the associated transfers are the same as the transfers (A.8) and (A.9) derived in the proof of Lemma 1, but evaluated at the allocation rule (16). These transfers are similar to the payments in other mechanism design settings. Customer agent's expected payment is her expected surplus less her cost that incurs due to information revelation (informational rent). Similarly, provider agent's expected compensation equals her expected surplus plus a term she receives for revealing her information. Notice that transfers are made, even if no service is allocated.

4.3. Maximizing the expected social welfare

In this section, we show that the allocation rule in equation (16) maximizes the expected social welfare among all incentive compatible and individually rational mechanisms characterized in Lemma 1. First, we demonstrate that the monotonicity properties required in Lemma 1 are satisfied. Then we specify conditions for the TPAA mechanism to maximize the expected social welfare.

Taking the partial derivative of the expected allocation probability and evaluating this variable at the allocation rule (16) for customer agents gives

$$\begin{aligned}\frac{\partial}{\partial\theta_i}\mathbb{E}_{\theta_{-i},\sigma}[x_{ij}(\theta_i,\cdot)] &= \frac{\partial}{\partial\theta_i}\sum_{k=1}^{K} w_k^C(\theta_i)\int_0^{R_C(\theta_i)} h_{(k)}(r)dr \\ &= \frac{\partial}{\partial\theta_i}\sum_{k=1}^{K} w_k^C(\theta_i)H_{(k)}(R_C(\theta_i)) \\ &= \sum_{k=1}^{K}\frac{\partial w_k^C(\theta_i)}{\partial\theta_i}H_{(k)}(R_C(\theta_i)) + \sum_{k=1}^{K} w_k^C(\theta_i)h_{(k)}(R_C(\theta_i))R_C'(\theta_i). \qquad (19)\end{aligned}$$

Observe that the second summand in (19) is always positive because $R_C(\theta_i)$ is increasing. For the first summand we use the same argument as in Johnson (2011): Fix index k^*. Then increase all terms above $k > k^*$ where $\partial w_k^C(\theta_i)/\partial\theta_i$ is negative, and decrease all terms below $k < k^*$ with $\partial w_k^C(\theta_i)/\partial\theta_i$ being positive. By the binomial theorem, we have $\sum_{k=1}^{K} w_k^C(\theta_i) = 1$, and thus $\sum_{k=1}^{K}\partial w_k^C(\theta_i)/\partial\theta_i = 0$ (Abramowitz and Stegun, 1972, p. 10). Hence, equation (19) is estimated as follows:

$$\begin{aligned}\frac{\partial}{\partial\theta_i}\mathbb{E}_{\theta_{-i},\sigma}[x_{ij}(\theta_i,\cdot)] &> \sum_{k=1}^{K}\frac{\partial w_k^C(\theta_i)}{\partial\theta_i}H_{(k)}(R_C(\theta_i)) \\ &> H_{(k^*)}(R_C(\theta_i))\sum_{k=1}^{K}\frac{\partial w_k^C(\theta_i)}{\partial\theta_i} = 0, \qquad (20)\end{aligned}$$

where the last inequality follows from the binomial theorem. Therefore, $\mathbb{E}_{\theta_{-i},\sigma}[x_{ij}(\theta_i,\cdot)]$ is increasing in θ_i. For provider agents allocation rule (16) yields

$$\begin{aligned}\frac{\partial}{\partial\sigma_j}\mathbb{E}_{\theta,\sigma_{-j}}[x_{ij}(\cdot,\sigma_j)] &= \frac{\partial}{\partial\sigma_j}\sum_{k=1}^{K} w_k^P(\sigma_j)\int_{R_P(\sigma_j)}^{\bar\theta} f_{(k)}(r)dr \\ &= \frac{\partial}{\partial\sigma_j}\sum_{k=1}^{K} w_k^P(\sigma_j)(1-F_{(k)}(R_P(\sigma_j))) \\ &= -\Big(\sum_{k=1}^{K}\frac{\partial w_k^P(\sigma_j)}{\partial\sigma_j}F_{(k)}(R_P(\sigma_j)) + \sum_{k=1}^{K} w_k^P(\sigma_j)f_{(k)}(R_P(\sigma_j))R_P'(\sigma_j)\Big). \qquad (21)\end{aligned}$$

Notice that $R_P(\sigma_j)$ is increasing. Labeling the indices k as k^* and using a similar argument as above shows that the expression in brackets in (21) is always positive. Therefore, the whole term is negative, and thus $\mathbb{E}_{\theta,\sigma_{-j}}[x_{ij}(\cdot,\sigma_j)]$ is decreasing in σ_j. This proves that the monotonicity properties of Lemma 1 are satisfied. In addition, the allocation rule also fulfills constraint (10) because the rule is positive assortative as long as the difference between virtual valuation and virtual cost is positive.

The following Theorem summarizes the main results.

Theorem 1. *Let customer agent a_i's valuation $v(\theta_i, \sigma_j)$ be maximized when $\theta_i = \sigma_j$ with the maximal value increasing in θ_i. Further, let $v(\theta_i, \sigma_j)$ be supermodular in both arguments and concave in θ_i. Let the distribution function $F(\cdot)$ be log-concave and let*

$$\begin{aligned} &\frac{\partial^3 v(\theta_i, \sigma_j)}{\partial \theta_i^2 \partial \sigma_j} \geq 0, \\ &\frac{f(\theta_i)}{1 - F(\theta_i)} \geq \frac{\partial}{\partial \theta_i} \log\Big(\frac{\partial v(\theta_i, \sigma_j)}{\partial \sigma_j}\Big). \end{aligned} \tag{22}$$

On the supply side, let provider agent b_j's provision cost $c(\theta_i, \sigma_j)$ be minimized when $\sigma_j = \theta_i$ with the minimal value increasing in σ_j. Further, let $c(\theta_i, \sigma_j)$ be submodular in both arguments and convex in σ_j. Let the distribution function $H(\cdot)$ be log-concave and let

$$\begin{aligned} &\frac{\partial^3 c(\theta_i, \sigma_j)}{\partial \theta_i \partial \sigma_j^2} \leq 0, \\ &\frac{h(\sigma_j)}{H(\sigma_j)} \geq \frac{\partial}{\partial \sigma_j} \log\Big(\frac{\partial c(\theta_i, \sigma_j)}{\partial \theta_i}\Big). \end{aligned} \tag{23}$$

Then, the TPAA mechanism maximizes the expected social welfare among all incentive compatible and individually rational mechanisms. Further, budget balance is satisfied.

Theorem 1 specifies conditions sufficient for the TPAA mechanism to maximize the expected social welfare, subject to the incentive compatibility and individual rationality constraints. The TPAA mechanism is a direct revelation mechanism in which:

1. The mechanism induces each customer agent to submit her desired service quality θ_i and each provider agent to submit her actual quality σ_j to the mechanism.

2. The mechanism sorts all quality values on both sides in descending order (i.e., positively assortative).

3. The mechanism allocates the customer agent with the highest desired quality to the provider agent with the highest actual quality, the second highest, and so on, but excludes all allocations for which the difference in virtual valuation and virtual cost is negative.

4. All agents pay/receive their associated expected transfers, regardless of whether they are allocated.

4.4. Illustrative example

We provide an example to illustrate the design of the TPAA mechanism. Suppose each customer agent's desired quality θ_i and each provider agent's actual quality σ_j are identically and uniformly distributed over the unit interval. All customer agents have valuations of $v(\theta_i, \sigma_j) = 1 + \sqrt{\theta_i} - (\theta_i - \sigma_j)^2$ and all provider agents have provision costs of $c(\theta_i, \sigma_j) = \sigma_j^2 + (\theta_i - \sigma_j)^2$. These functions satisfy the requirements in Theorem 1. Since all private information is independently drawn from the uniform distribution, the densities are $f(\theta_i) = h(\sigma_j) = 1$, with cumulative distributions $F(\theta_i) = \theta_i$ and $H(\sigma_j) = \sigma_j$ over the interval $[0, 1]$. Given these functions, the virtual valuation of customer agents is $\psi^C(\theta_i, \sigma_j) = v(\theta_i, \sigma_j) - (1 - \theta_i)(2\sigma_j - 2\theta_i + \frac{1}{2\sqrt{\theta_i}})$, while the virtual cost of provider agents is $\psi^P(\theta_i, \sigma_j) = c(\theta_i, \sigma_j) + \sigma_j(4\sigma_j - 2\theta_i)$.

Figure 1 illustrates the relationship between actual and virtual reservation values of the agents. The fact that a customer's virtual valuation is less than or equal to her actual valuation indicates that she will in general have an incentive to understate her desired quality in order to favorably influence the transaction price. The customer with the strongest incentive to understate her desired quality is the one with the highest desired quality. However, since this customer will generate more social welfare (i.e., gains from trade) than a customer with a low desired quality, an efficient mechanism will want the highest quality customer to be matched with a provider of a lower actual cost whenever doing so is efficient. Hence, the virtual valuation of the highest quality customer is equal to her actual valuation. To ensure that this customer reports her desired quality truthfully, the mechanism must make lower reports unprofitable. It does so by lowering the probability that a customer with a lower desired quality will be matched with a lower probability. This reduced probability is inefficient but the loss in social welfare is smaller from reducing the probability of trade for customers with low desired qualities than with high desired qualities. The reverse explanation applies to the pattern observed for providers. Hence, this example illustrates what the revelation principle does: It embeds the strategic behavior of the agents in the mechanism so that in equilibrium the agents no longer engage in strategic misrepresentation.

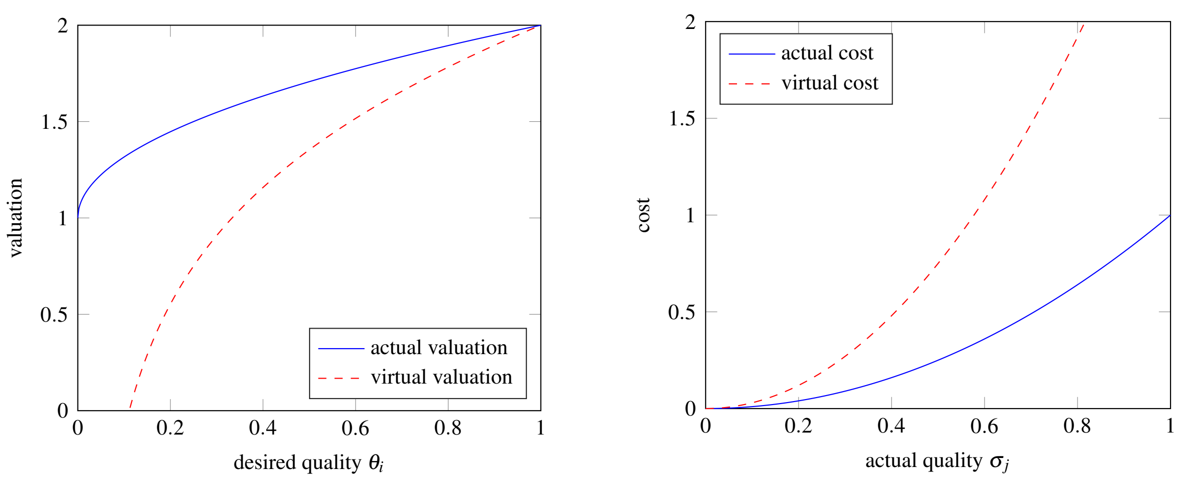

Figure 1: Actual and virtual reservation functions of customer agents (left) and provider agents (right).

Because private information is present in our analysis, the TPAA mechanism is a second-best mechanism (Börgers, 2015). Its match condition is based on the difference of the agents' virtual reservation functions. The associated first-best mechanism assumes public information among all agents. In the first-best case, matching always occurs as long as a customer agent's actual valuation exceeds a provider agent's actual provision cost. The boundaries for successful matching in the first-best and the second-best mechanism are shown in figure 2 for uniform random variables (left) and normal random variables (right). The shaded area in solid green contains the eligible quality pairs within the domain $[0,1] \times [0,1]$ that warrant service allocation in the first-best mechanism. The dotted region within the solid area contains all quality pairs that are allocated in the second-best mechanism. The significantly smaller area in the second-best case illustrates the distortion due to the presence of private information concerning service quality.

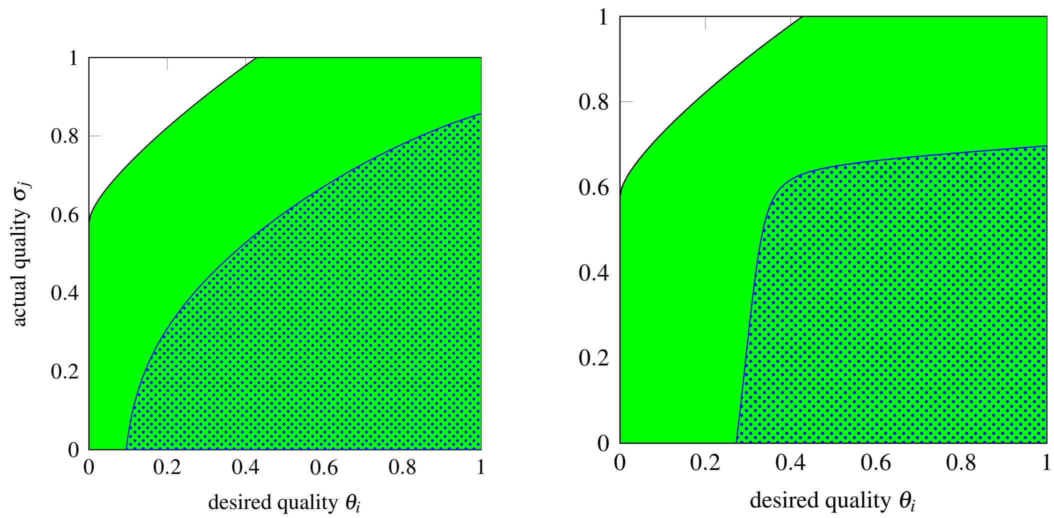

Figure 2: The mechanism's first-best and second-best boundaries for uniform (left) and normal (right) distributions.

Since the TPAA mechanism is positively assortative, it converges to deterministic quantile matching (Johnson, 2011). Therefore, in the limit as $K \to \infty$, the empirical quantile function is asymptotically equivalent to the k-th order statistic (van der Vaart, 2000). This means that $\sigma \to H^{-1}(F(\theta))$ in the allocation rule, and therefore $\sigma \to \theta$ for uniformly distributed θ and σ on $[0,1]$. Solving $\psi(\theta,\theta) \geq 0$ gives the lower bound $a = 0.1157$ and the upper bound $b = 0.7604$. Hence, in the limit all quality realizations less than a and greater than b are truncated by the second-best mechanism. Its efficiency as $K \to \infty$ is given by

$$\frac{\int_a^b (v(t,t) - c(t,t))\, dt}{\int_0^1 (v(t,t) - c(t,t))\, dt} \approx 0.6859, \tag{24}$$

since $f(t) = 1$ for the uniform distribution. Therefore, the TPAA mechanism's asymptotic efficiency for uniformly distributed quality realizations is 68.59%. The numerical simulation in section 5.2 confirms this upper bound (cf. figure 3). In section 5.3, we discuss the insights that can be obtained from this result.

5. Experimental evaluation

This section reports an experimental evaluation of the TPAA mechanism developed in this proposal. We describe the setup, report the results, and discuss the findings.

5.1. Experimental setup

The experimental setup is based on the setting used in the example presented in section 4.4. All customer agents used valuation functions $v(\theta_i, \sigma_j) = 1 + \sqrt{\theta_i} - (\theta_i - \sigma_j)^2$, and all provider agents had provision cost $c(\theta_i, \sigma_j) = \sigma_j^2 + (\theta_i - \sigma_j)^2$. The agents' private information was described by random variables in two variants: In the first setting, both customer agents' desired qualities and provider agents' actual qualities were drawn from the uniform distribution over the unit interval; that is, $\theta_i \sim \sigma_j \sim U(0,1)$ for all agents a_i and b_j. In the second variant, desired and actual qualities were drawn from the normal distribution truncated to the unit interval with mean $\mu = 0.5$ and standard deviation $\sigma = 0.1$; that is, for all agents a_i and b_j in the second setting, $\theta_i \sim \sigma_j \sim N(0.5, 0.01)$. For each variant, the experiment was conducted 10^5 times, and average values were calculated. The steps were:

1. For each agent desired and actual quality realizations were drawn independently from the respective random distribution (uniform and normal).
2. The mechanism arranged these quality realizations positive assortatively and allocated the services if the difference between virtual valuation and virtual cost was non-negative.
3. Once an allocation was arranged, the allocation surplus was calculated as $v(\theta_i, \sigma_j) - c(\theta_i, \sigma_j)$.
4. The mechanism calculated the sum of all allocation surpluses as defined in the maximand of (7).

In this article, we took the social welfare achieved by the first-best mechanism as a benchmark and compared it to the outcome produced by the second-best mechanism in the presence of private information. Hence, the efficiency of the TPAA mechanism was defined as the ratio between the second-best and the first-best outcome.

5.2. Results

The left hand side of figure 3 shows the average behavior of the mechanism's efficiency as a function of the number of customer agents for uniformly distributed desired and actual quality values. The number of provider agents in the market was fixed to 10, 50, and 100. For ten provider agents the efficiency was at its minimum of 0.1310, followed by an efficiency increase as the number of customer agents grew. At about 15 customer agents the efficiency flattened and remained constant at 0.8397. When 50 provider agents offered their services to less than 20 customer agents, the efficiency vanished completely. For 20 to 70 customer agents the efficiency increased. It then flattened and remained constant at 0.8397. With 100 provider agents in the market the efficiency was zero for less than 50 customer agents. It then increased until the same limit of 0.8397 was reached (not shown in the graph). Analytically, this limit was calculated by

$$\frac{\int_a^b (v(1,t) - c(1,t))\,dt}{\int_0^1 (v(1,t) - c(1,t))\,dt} \approx 0.8397 \tag{25}$$

where $a = 0$ and $b = 0.8571$ were the bounds of integration that arose from the exclusion condition $\psi(1,t) \geq 0$.

The right hand side of figure 3 shows the mechanism's efficiency as K was increased. For $N = M = 30$ agents (first data point), the efficiency had its lowest value at 0.6078. The curve then increased monotonically and approached the asymptote at 0.6859 from below.

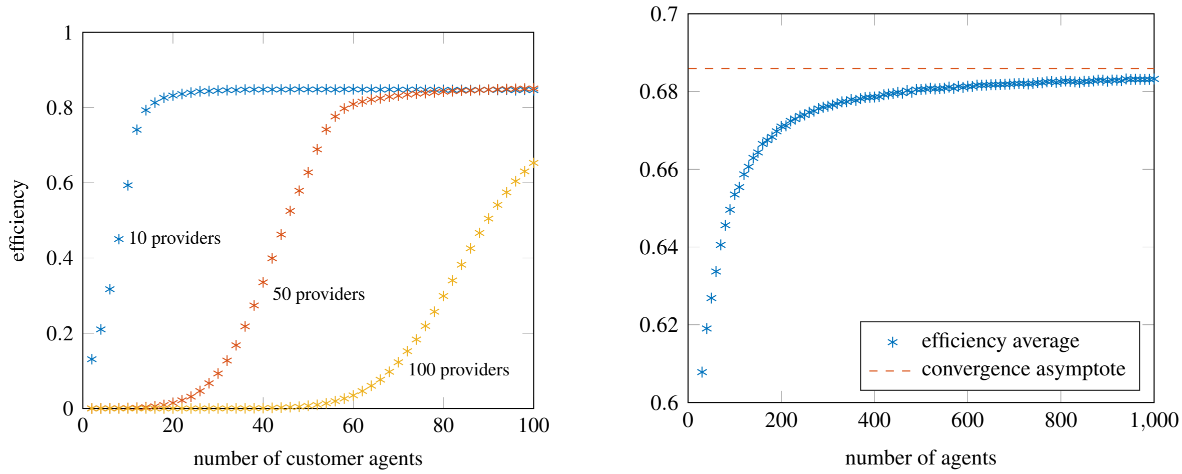

Figure 3: Mechanism efficiency for uniformly distributed quality realizations. Left: Number of provider agents fixed to 10, 50, and 100. Right: Number of customer and provider agents increased equally ($N = M$).

Figure 4 shows the results for normally distributed quality values truncated to the unit interval. The left hand side of the figure illustrates the mechanism's efficiency in a market with 10, 50, and 100 provider agents. With ten provider agents in the market, the lowest efficiency was 0.6336 when two customer agents requested for services. As the number of customer agents increased, the efficiency quickly increased and remained constant at 0.9735 for ten and more customer agents. When 50 provider agents were in the market, the lowest efficiency was 0.0769 for two customer agents. The efficiency then increased and also remained constant at 0.9735 for 50 or more customer agents. When 100 provider agents offered their services to two customer agents, the efficiency was at its minimum of 0.0053. It then increased and also remained constant at 0.9735 for 100 or more customer agents (not shown in the graph). Again, this asymptote was calculated analytically as

$$\frac{\int_a^b (v(1,t) - c(1,t)) f(t)\, dt}{\int_0^1 (v(1,t) - c(1,t)) f(t)\, dt} \approx 0.9735, \tag{26}$$

where $a = 0$ and $b = 0.6967$ were the bounds of integration due to the exclusion condition and $f(t)$ was the density of the normal distribution with $\mu = 0.5$ and $\sigma = 0.1$.

The right hand side of figure 4 shows the mechanism's efficiency as K was increased. For $N = M = 10$ agents (first data point), the efficiency had its lowest value at 0.9257. The curve then increased monotonically and approached the asymptote at 0.9315.

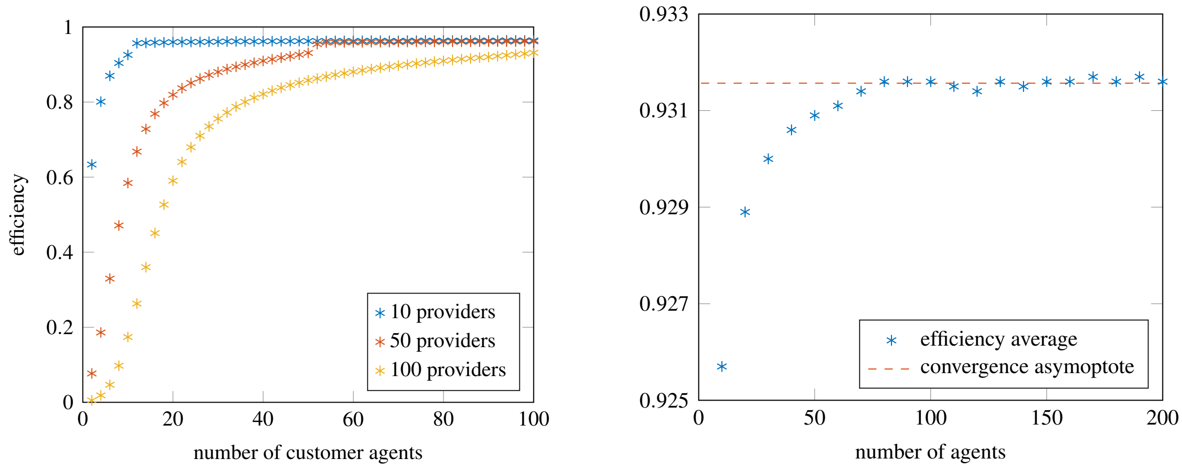

Figure 4: Mechanism efficiency for normally distributed quality realizations. Left: Number of provider agents fixed to 10, 50, and 100. Right: Number of customer and provider agents increased equally ($N = M$).

5.3. Discussion

Our experiments demonstrate that the mechanism's inefficiency arising from privately known QoS does not disappear even when the number of agents becomes large. Nonetheless, the results suggest that the efficiency can be quite high. This finding provides evidence for the efficacy of the proposed allocation mechanism. In the following paragraphs, the insights that can be obtained from our research are discussed.

First, in settings with many agents on both market sides, we find that although the mechanism's asymptotic efficiency can be quite high, it is bounded away from 100%. This result indicates that even the presence of many agents does not eliminate the mechanism's inefficiency caused by the asymmetry of information among agents. The reason for this phenomenon is related to the agents' informational smallness, which has been studied by Gul and Postlewaite (1992) in the context of replicated economies as well as by McLean and Postlewaite (2002) in more general frameworks. Loosely speaking, an agent is informationally small if the incremental impact of that agent's private information (i.e., desired/actual QoS) on the demand of every electronic service is "small", given the information of other agents (a precise formalization of this notion can be found in McLean and Postlewaite (2002)).

In principle, the asymptotic inefficiency observed for our mechanism could disappear if certain conditions were satisfied. Gul and Postlewaite (1992) identify such conditions sufficient for eliminating inefficiencies in replicated economies. However, in certain problems with asymmetric information, these conditions fail to hold. Therefore, the associated mechanisms cannot achieve full efficiency even when the number of agents is large. One such example is the replication of the classic bilateral trading mechanism studied by Myerson and Satterthwaite (1983), in which only a specific match produces pairwise private surpluses. The two-trader case illustrates the impossibility of designing incentive compatible mechanisms that are ex post efficient. In fact, the associated second-best mechanism reaches an efficiency of 84.36% for uniformly distributed types (Gresik and Satterthwaite, 1983). Now, if this bilateral case were replicated with independent valuations across pairs such that only specific matches generate mutual value, the inefficiency due to asymmetric information may not vanish (Gul and Postlewaite, 1992). The presence of additional pairs in the market would not have an impact on the problem faced by any particular pair for allocating the object between the two traders. This example demonstrates that even in the presence of many agents (and consequently many objects), the mechanism's asymptotic efficiency may still be bounded away from 100%.

The fact that the asymptotic inefficiency does not disappear in the TPAA mechanism is related to the preceding discussion. The proposed model of the TPAA mechanism can be regarded as a specific replication of the Myerson-Satterthwaite bilateral trading model. In our model, surplus is produced only when a specific pair of two agents is matched together. It is only this particular agent pair that cares about the electronic service to be allocated. Adding more agent pairs to this economy leaves unchanged the incremental impact of each agent's private information on the demand of the electronic service. In the proposed model, each provider agent sells a different service and each customer agent desires a different QoS. It is this feature that prevents the agents from becoming informationally small as the market becomes large. The absence of informational smallness is the reason why the asymptotic inefficiency due to the asymmetry of information does not disappear.

The illustrative example discussed earlier shows that the efficiency of the TPAA mechanism for large economies is 68.59%, given uniformly distributed QoS on the unit interval. This efficiency, however, is significantly lower than that for the bilateral trading mechanism (84.36% as reported by Gresik and Satterthwaite (1983)). If the latter economy were to be replicated, the number of commodities would go to infinity while the agents' valuations would still remain independent. We, however, focus on a framework in which an agent's utility depends upon the types of other agents. In particular, a customer agent's utility is maximized when its desired QoS matches the actual QoS delivered by a provider agent. Similarly, provider agents maximize their utility by delivering exactly the QoS desired by customer agents. It is this interdependence of the agents' utilities that causes the additional efficiency loss of over 15.77% as compared to the outcome in Myerson and Satterthwaite (1983). This observation suggests that agents with interdependent utilities have greater incentives to misreport their QoS in order to win a better allocation than the agents in the (replicated) bilateral trading mechanism with independent valuations.

Second, although the structure of our formal framework differs from that of McLean and Postlewaite (2002), we argue that their notion of an agent's informational size can, to some extent, still be applied to our model. In fact, the informational rent of an agent quantifies the amount of compensation that must be paid to that agent to induce honest reporting of private information. The agents' informational rents are linked to their informational size (McLean and Postlewaite, 2004). In our proposed model, the agents' informational rents are given by the second summand in the transfer functions (17) and (18). Applying a similar argument as Johnson (2011) to our illustrative example for $K \to \infty$, the informational rent $S_C^\infty(\theta)$ of customer agents and the informational rent $S_P^\infty(\sigma)$ of provider agents in the limit market are respectively given by

$$S_C^\infty(\theta) = \int_0^\theta \frac{\partial v(t,t)}{\partial t} dt = \sqrt{\theta} \qquad \text{and} \qquad S_P^\infty(\sigma) = \int_\sigma^1 \frac{\partial c(t,t)}{\partial t} dt = 1 - \sigma^2. \tag{27}$$

Hence, both quantities are bounded away from zero when the number of agents becomes large. This example with uniformly distributed QoS on $[0,1]$ shows that the amount of compensation for honest reporting does not disappear even when the market size goes to infinity. In this sense, the informational size of the agents does not go to zero as the market becomes large.

Third, it is worth noting that the difference in virtual valuations required by the allocation rule (16) does not vanish even when the number of traders is large. The necessary condition for any two agents to match in the limit market is $\psi(\theta,\theta) \geq 0$ or, equivalently,

$$v(\theta,\theta) - c(\theta,\theta) \geq 2\theta^2 + \frac{1-\theta}{2\sqrt{\theta}} > 0.7915 \tag{28}$$

for uniformly distributed QoS. This difference in valuation and provision cost must be satisfied to warrant service allocation in large markets. In particular, the right side of (28) is bounded away from zero by 0.7915 for all $\theta \in (0,1]$ even when $K \to \infty$. In the asymptotically efficient mechanism proposed by Gresik and Satterthwaite (1983), the difference in reservation values vanishes at the same rate as $1/2K$ approaches zero. Again, the presence of this non-vanishing threshold (28) confirms that the TPAA mechanism cannot attain full efficiency even when the number of traders is large.

Fourth, the results obtained in the setting with normally distributed QoS (mean of 0.5 and standard deviation of 0.1) are different. As in the uniform case, the asymptotic efficiency is also bounded away from 100%. However, it is significantly higher than that for the uniform distribution. Normally distributed QoS entails an asymptotic efficiency of 93.15% as $K \to \infty$. When the agents' types are drawn from the normal distribution, the associated realizations are scattered around its mean of 0.5 with exclusion lines beyond the standard deviation of ± 0.1 (cf. right hand side of figure 2). Thus, there is more mass in the middle, which means that the agents' incentive to distort their QoS for receiving a better allocation is smaller. The left side of figure 2 depicts this distortion due to normally distributed private information.

A last comment is on the implementability of the proposed mechanism in dominant strategies. In general, imposing incentive compatibility constraints for dominant-strategy mechanisms instead of Bayesian mechanisms restricts the set of implementable mechanisms considerably. However, Mookherjee and Reichelstein (1992) identify mechanism design problems for which Bayesian incentive compatibility can be replaced by the more stringent requirement of dominant strategy incentive compatibility without any losses. Since the allocation rule defined in Theorem 1 satisfies the monotonicity conditions of Mookherjee and Reichelstein (1992), there is no loss from replacing Bayesian equilibrium constraints by dominant strategy requirements. Hence, the TPAA mechanism proposed in this article is implementable in dominant strategies.

6. Conclusion

The contribution of this research is a second-best mechanism for allocating electronic services with privately known QoS. Using a mechanism design framework, we derived the mechanism and studied its efficiency properties. The proposed mechanism internalizes the QoS desired by multiple customer agents and the QoS actually offered by multiple provider agents. By providing a theoretically-sound extension of the model of Johnson (2013), we were able to design a mechanism that maintains the properties of the underlying model. In our extension, we studied the social welfare properties of the second-best mechanism for matching customers and providers. Specifically, the allocation mechanism (i) is incentive compatible, (ii) is individually rational, (iii) balances the budget, and (iv) maximizes the expected social welfare. Our first experiment considered a market with uniformly distributed QoS, while our second experiment assumed normally distributed QoS. In all experiments, we found that the asymptotic efficiency of the second-best mechanism is bounded away from 100% even when the number of agents goes to infinity. This finding is significant because it indicates that the agents in our model are informationally large. We explained why the agents' informational size does not vanish even in large markets. Our research provides important insights into the efficiency properties of the proposed mechanism. Although full efficiency cannot be attained, our results suggest that in settings with normal random QoS, the mechanism's inefficiency due to the asymmetry of information is reduced to about 7% as the market grows larger.

From a managerial perspective, our study can support a social planner in designing competitive service allocation when each participant is privately informed about their QoS. On one hand, the social planner can ensure that every decision is made by the participants only. Hence, the emerging distributed mechanism implementation eludes the need for an external, independent decision maker. On the other hand, the social planner can take into account that full efficiency may not be attained as long as the participants are informationally large. If the incremental impact of each participant's private information on the demand of the electronic service remains unchanged, inefficiencies may not vanish even in the presence of many participants. However, the results of the simulation study in this proposal can help the social planner to obtain a more accurate estimation of efficiency losses due to asymmetric information.

Future research might be pursued in two directions. First, in the current model, QoS is a one-dimensional parameter, which could be extended to multiple quality attributes. This change would require extending the model and revisiting its properties. Second, while the properties of the model have been assessed through analytical evaluation and two experiments, a more comprehensive quantitative evaluation could be required to study the efficacy on a greater scale. For this purpose, the mechanism could be studied in simulation experiments with actual or synthetic data that more realistically reflect markets for electronic services including QoS and interdependent utility functions.

Acknowledgements

This work has been partially supported by the German Federal Ministry of Education and Research under grant 16KIS0083.

References

Abramowitz, M., Stegun, I. A. (Eds.), 1972. Handbook of Mathematical Functions with Formulas, Graphs, and Mathematical Tables. Dover Publications.

Armbrust, B., Griffith, R., Joseph, A. D., Katz, R., Konwinski, A., Lee, G., Patterson, D., Rabkin, A., 2010. A view of cloud computing. Communications of the ACM 53 (4), 50–58.

Arya, A., Löffler, C., Mittendorf, B., Pfeiffer, T., 2015. The middleman as a panacea for supply chain coordination problems. European Journal of Operational Research 240 (2), 393–400.

Avnet Cloud Marketplace, 2016. The Avnet Cloud Marketplace. [Online; accessed 12-February-2016]. URL http://www.avnetpromotions.com/cloudmarketplace/

Babich, V., Li, H., Ritchken, P., Wang, Y., 2012. Contracting with asymmetric demand information in supply chains. European Journal of Operational Research 217 (2), 333–341.

Bapna, R., Das, S., Garfinkel, R., Stallaert, J., 2008. A market design for grid computing. INFORMS Journal on Computing 20 (1), 100–111.

Bichler, M., Kalagnanam, J., 2005. Configurable offers and winner determination in multi-attribute auctions. European Journal of Operational Research 160 (2), 380–394.

Blau, B., Conte, T., van Dinther, C., 2010. A multidimensional procurement auction for trading composite services. Electronic Commerce Research and Applications 9 (5), 460–472.

Bockstedt, J., Goh, K., 2011. Seller strategies for differentiation in highly competitive online auction markets. Journal of Management Information Systems 28 (3), 235–267.

Börgers, T., 2015. An Introduction to the Theory of Mechanism Design. Oxford University Press.

Branco, F., 1997. The design of multidimensional auctions. RAND Journal of Economics 28 (1), 63–81.

Chang, Y. B., Gurbaxani, V., 2012. Information technology outsourcing, knowledge transfer, and firm productivity: An empirical analysis. MIS Quarterly 36 (4), 1043–1063.

Che, Y.-K., 1993. Design competition through multidimensional auctions. RAND Journal of Economics 24 (4), 668–680.

Egri, P., Váncza, J., 2013. A distributed coordination mechanism for supply networks with asymmetric information. European Journal of Operational Research 226 (3), 452–460.

Gartner, 2014. Forecast Analysis: IT Outsourcing, Worldwide, 1Q14 Update. URL https://www.gartner.com/doc/2707918/forecast-analysis-it-outsourcing-worldwide

Gibbard, A. F., 1973. Manipulation of voting schemes: A general result. Econometrica 41 (4), 587–602.

Greenberg, A., Hamilton, J., Maltz, D. A., Patel, P., 2009. The cost of a cloud: research problems in data center networks. ACM SIGCOMM Computer Communication Review 39 (1), 68–73.

Gresik, T. A., Satterthwaite, M. A., 1983. The number of traders required to make a market competitive: the beginnings of a theory. Discussion Papers 551, Center for Mathematical Studies in Economics and Management Science, Northwestern University.

Gresik, T. A., Satterthwaite, M. A., 1989. The rate at which a simple market converges to efficiency as the number of traders increases: An asymptotic result for optimal trading mechanisms. Journal of Economic Theory 48 (1), 304–332.

Gul, F., Postlewaite, A., 1992. Asymptotic efficiency in large economies with asymmetric information. Econometrica 60 (6), 1273–1292.

Hang, C.-W., Singh, M. P., 2010. From quality to utility: Adaptive service selection framework. In: Proceedings of the 8th International Conference on Service-Oriented Computing (ICSOC 2010). San Francisco, California, USA, pp. 456–470.

Harris, M., Raviv, A., 1981. A theory of monopoly pricing schemes with demand uncertainty. American Economic Review 71 (3), 347–365.

Johnson, T., 2011. Matching through position auctions. Working Papers 001, Department of Economics, University of Notre Dame.

Johnson, T., 2013. Matching through position auctions. Journal of Economic Theory 148 (4), 1700–1713.

Laffont, J. J., Maskin, E., 1979. A differentiable approach to expected utility maximizing mechanisms. In: Laffont, J. J. (Ed.), Aggregation and Revelation of Preferences. North-Holland.

McAfee, R. P., 1991. Efficient allocation with continuous quantities. Journal of Economic Theory 53 (1), 51–73.

McLean, R., Postlewaite, A., 2002. Informational size and incentive compatibility. Econometrica 70 (6), 2421–2453.

McLean, R., Postlewaite, A., 2004. Informational size and efficient auctions. Review of Economic Studies 71 (3), 809–827.

Milgrom, P., Segal, I., 2002. Envelope theorems for arbitrary choice sets. Econometrica 70 (2), 583–601.

Mookherjee, D., Reichelstein, S., 1992. Dominant strategy implementation of bayesian incentive compatible allocation rules. Journal of Economic Theory 56 (2), 378–399.

Myerson, R. B., 1979. Incentive compatibility and the bargaining problem. Econometrica 47 (1), 61–73.

Myerson, R. B., 1981. Optimal auction design. Mathematics of Operation Research 6 (1), 58–73.

Myerson, R. B., Satterthwaite, M. A., 1983. Efficient mechanisms for bilateral trading. Journal of Economic Theory 29 (2), 256–281.

Nisan, N., Ronen, A., 2001. Algorithmic mechanism design. Games and Economic Behavior 35, 166–196.

O'Sullivan, J., Edmond, D., ter Hofstede, A., 2002. What's in a service? Distributed and Parallel Databases 12 (2), 117–133.

Schnizler, B., Neumann, D., Veit, D., Weinhardt, C., 2008. Trading grid services - a multi-attribute combinatorial approach. European Journal of Operational Research 187 (3), 943–961.

Shimer, R., Smith, L., 2000. Assortative matching and search. Econometrica 68 (2), 343–369.

Stösser, J., Neumann, D., Weinhardt, C., 2010. Market-based pricing in grids: On strategic manipulation and computational cost. European Journal of Operational Research 203 (1), 464–457.

van der Vaart, A. W., 2000. Asymptotic Statistics. Cambridge University Press.

Widmer, T., Premm, M., Karaenke, P., 2013. Sourcing strategies for energy-efficient virtual organisations in cloud computing. In: Proceedings of the 15th IEEE Conference on Business Informatics (CBI 2013). Vienna, Austria, pp. 1147–1161.

Appendix A. Proofs

Lemma 1. *Let $x_{ij}(\theta,\sigma)$ be the probability that provider agent b_j is allocated to customer agent a_i. Then transfer functions $t_C(\theta,\sigma)$ and $t_P(\theta,\sigma)$ exist such that $\{x_{ij}(\theta,\sigma),t_C(\theta,\sigma),t_P(\theta,\sigma)\}$ is incentive compatible and individually rational if and only if $\mathbb{E}_{\theta_{-i},\sigma}[x_{ij}(\theta_i,\cdot)]$ is non-decreasing, $\mathbb{E}_{\theta,\sigma_{-j}}[x_{ij}(\cdot,\sigma_j)]$ is non-increasing and*

$$\mathbb{E}\Big[\sum_{i=1}^{N}\sum_{j=1}^{M}(\psi_C(\theta_i,\sigma_j)-\psi_P(\theta_i,\sigma_j))x_{ij}(\theta,\sigma)\Big]\geq 0. \tag{10}$$

PROOF. Suppose $\{x_{ij}(\theta,\sigma),t_C(\theta,\sigma),t_P(\theta,\sigma)\}$ is incentive compatible. We start to derive our argument for provider agents. For any quality pair $\hat{\sigma}\neq\sigma_j$ we must have

$$U_P(\sigma_j)\geq\mathbb{E}_{\theta,\sigma_{-j}}\Big[t_P(\theta,\hat{\sigma},\sigma_{-j})-\sum_{i=1}^{N}c(\theta_i,\sigma_j)x_{ij}(\theta,\hat{\sigma},\sigma_{-j})\Big]\quad\text{and} \tag{A.1}$$

$$U_P(\hat{\sigma})\geq\mathbb{E}_{\theta,\sigma_{-j}}\Big[t_P(\theta,\sigma_j,\sigma_{-j})-\sum_{i=1}^{N}c(\theta_i,\hat{\sigma})x_{ij}(\theta,\sigma_j,\sigma_{-j})\Big]. \tag{A.2}$$

These two inequalities imply that

$$\int_{\sigma_j}^{\hat{\sigma}}\mathbb{E}_{\theta,\sigma_{-j}}\Big[\sum_{i=1}^{N}\Big(x_{ij}(\theta,t,\sigma_{-j})-x_{ij}(\theta,\hat{\sigma},\sigma_{-j})\Big)\frac{\partial}{\partial t}c(\theta_i,t)\Big]dt\geq 0. \tag{A.3}$$

Therefore, if $\hat{\sigma}>\sigma_j$, we must have $\mathbb{E}_{\theta,\sigma_{-j}}[x_{ij}(\theta,\sigma_j,\sigma_{-j})]\geq\mathbb{E}_{\theta,\sigma_{-j}}[x_{ij}(\theta,\hat{\sigma},\sigma_{-j})]$ and thus $\mathbb{E}_{\theta,\sigma_{-j}}[x_{ij}(\cdot,\sigma_j)]$ is non-increasing. A similar argument for customer agents shows that $\mathbb{E}_{\theta_{-i},\sigma}[x_{ij}(\theta_i,\cdot)]$ must be non-decreasing. Corollary 1 in Milgrom and Segal (2002) provides expressions for the indirect utility of each agent in any incentive compatible mechanism:

$$U_C(\theta_i)=U_C(0)+\int_0^{\theta_i}\mathbb{E}_{\theta_{-i},\sigma}\Big[\sum_{j=1}^{M}\frac{\partial v(r,\sigma_j)}{\partial r}x_{ij}(r,\theta_{-i},\sigma)\Big]dr\quad\text{and} \tag{A.4}$$

$$U_P(\sigma_j)=U_P(\bar{\sigma})+\int_{\sigma_j}^{\bar{\sigma}}\mathbb{E}_{\theta,\sigma_{-j}}\Big[\sum_{i=1}^{N}\frac{\partial c(\theta_i,r)}{\partial r}x_{ij}(\theta,r,\sigma_{-j})\Big]dr, \tag{A.5}$$

where $U_C(0)$ and $U_P(\bar{\sigma})$ are the expected utilities evaluated at the lower and upper quality bounds, respectively. By substituting the indirect utilities (A.4) and (A.5) into the sum of all agents' expected utilities given in (1) and (2), we obtain an alternative expression for the expected social welfare defined within the maximization problem in (7):

$$\sum_{i=1}^{N}\int_{0}^{\bar\theta} U_C(\theta_i)f(\theta_i)\,d\theta_i + \sum_{j=1}^{M}\int_{0}^{\bar\sigma} U_P(\sigma_j)h(\sigma_j)\,d\sigma_j$$
$$= \sum_{i=1}^{N}\int_{0}^{\bar\theta}\Big(U_C(0) + \int_{0}^{\theta_i}\mathbb{E}_{\theta_{-i},\sigma}\Big[\sum_{j=1}^{M}\frac{\partial v(r,\sigma_j)}{\partial r}x_{ij}(r,\theta_{-i},\sigma)\Big]dr\Big)f(\theta_i)\,d\theta_i$$
$$+ \sum_{j=1}^{M}\int_{0}^{\bar\sigma}\Big(U_P(\bar\sigma) + \int_{\sigma_j}^{\bar\sigma}\mathbb{E}_{\theta,\sigma_{-j}}\Big[\sum_{i=1}^{N}\frac{\partial c(\theta_i,r)}{\partial r}x_{ij}(\theta,r,\sigma_{-j})\Big]dr\Big)h(\sigma_j)\,d\sigma_j. \quad \text{(A.6)}$$

Expression (A.6) is the expected social welfare expressed by the agents' indirect utilities. Therefore, it must equal the expected social welfare obtained by the agents' direct utilities in the maximand of (7). Equating these two expressions, followed by some basic algebraic manipulations, integration by parts, as well as rearranging and collecting similar terms yields

$$\sum_{i=1}^{N}U_C(0) + \sum_{j=1}^{M}U_P(\bar\sigma) = \mathbb{E}\Big[\sum_{i=1}^{N}\sum_{j=1}^{M}(\psi_C(\theta_i,\sigma_j) - \psi_P(\theta_i,\sigma_j))x_{ij}(\theta,\sigma)\Big]. \quad \text{(A.7)}$$

Since individual rationality holds, we must have $U_C(0) \geq 0$ and $U_P(\bar\sigma) \geq 0$, which gives us expression (10) in Lemma 1.

Suppose now that $x_{ij}(\theta,\sigma)$ satisfies (10) and that $\mathbb{E}_{\theta_{-i},\sigma}[x_{ij}(\theta_i,\cdot)]$ is non-decreasing and $\mathbb{E}_{\theta,\sigma_{-j}}[x_{ij}(\cdot,\sigma_j)]$ is non-increasing. Consider the following expected payments made by customer agent a_i

$$t_C(\theta_i,\theta_{-i},\sigma) = \sum_{j=1}^{M}v(\theta_i,\sigma_j)x_{ij}(\theta_i,\theta_{-i},\sigma) - \int_{0}^{\theta_i}\sum_{j=1}^{M}\frac{\partial v(r,\sigma_j)}{\partial r}x_{ij}(r,\theta_{-i},\sigma)dr \quad \text{(A.8)}$$

and the expected compensation received by provider agent b_j

$$t_P(\theta,\sigma_j,\sigma_{-j}) = \sum_{i=1}^{N}c(\theta_i,\sigma_j)x_{ij}(\theta,\sigma_j,\sigma_{-j}) + \int_{\sigma_j}^{\bar\sigma}\sum_{i=1}^{N}\frac{\partial c(\theta_i,r)}{\partial r}x_{ij}(\theta,r,\sigma_{-j})dr. \quad \text{(A.9)}$$

These expected transfers are obtained by equating the direct and indirect utilities, as well as setting the worst-off quality payoffs to zero; that is, $U_C(0) = U_P(\bar\sigma) = 0$.

To check incentive compatibility of (A.8), observe that

$$
\begin{aligned}
U_C(\theta_i) - U_C(\hat{\theta}) &= \mathbb{E}_{\theta_{-i},\sigma}\Big[\sum_{j=1}^{M} v(\theta_i,\sigma_j)(x_{ij}(\theta_i,\theta_{-i},\sigma) - x_{ij}(\hat{\theta},\theta_{-i},\sigma)) - (t_C(\theta_i,\theta_{-i},\sigma) - t_C(\hat{\theta},\theta_{-i},\sigma))\Big] \\
&= \mathbb{E}_{\theta_{-i},\sigma}\Big[\sum_{j=1}^{M} v(\theta_i,\sigma_j)\int_{\hat{\theta}}^{\theta_i} \frac{\partial}{\partial r} x_{ij}(r,\theta_{-i},\sigma)dr \\
&\quad - \Big(\sum_{j=1}^{M} v(\theta_i,\sigma_j)x_{ij}(\theta_i,\theta_{-i},\sigma) - v(\hat{\theta},\sigma_j)x_{ij}(\hat{\theta},\theta_{-i},\sigma) \\
&\quad + \int_{\theta_i}^{\hat{\theta}} \sum_{j=1}^{M} \frac{\partial v(r,\sigma_j)}{\partial r} x_{ij}(r,\theta_{-i},\sigma)dr\Big)\Big] \\
&= \mathbb{E}_{\theta_{-i},\sigma}\Big[\sum_{j=1}^{M} v(\theta_i,\sigma_j)\int_{\hat{\theta}}^{\theta_i} \frac{\partial}{\partial r} x_{ij}(r,\theta_{-i},\sigma)dr - \sum_{j=1}^{M}\int_{\hat{\theta}}^{\theta_i} v(r,\sigma_j)\frac{\partial}{\partial r} x_{ij}(r,\theta_{-i},\sigma)dr\Big] \\
&= \mathbb{E}_{\theta_{-i},\sigma}\Big[\int_{\hat{\theta}}^{\theta_i} \sum_{j=1}^{M} (v(\theta_i,\sigma_j) - v(r,\sigma_j))\frac{\partial}{\partial r} x_{ij}(r,\theta_{-i},\sigma)dr\Big] \geq 0. \qquad \text{(A.10)}
\end{aligned}
$$

The last inequality is a consequence of $\mathbb{E}_{\theta_{-i},\sigma}[x_{ij}(\theta_i,\cdot)]$ being non-decreasing. Recall that customer agent's valuation is increasing in its maximal value. Therefore, customer agent a_i would do better reporting θ_i instead of $\hat{\theta}$. The proof of incentive compatibility for provider agents is analogous.

Because we have assumed (10), the sum over all expected utilities evaluated at the lowest and highest qualities respectively must be non-negative. Further, equations (A.4) and (A.5) imply that $U_C(\theta_i)$ is increasing in θ_i and $U_P(\sigma_j)$ is decreasing in σ_j. Due to these monotonicity properties and because of (10), it suffices to verify individual rationality for customer agent's lowest desired quality $\theta_i = 0$ and provider agent's highest actual quality $\sigma_j = \bar{\sigma}$. This yields $U_C(\theta_i) \geq 0$ and $U_P(\sigma_j) \geq 0$.

Lemma 2. *Any incentive compatible, individually rational mechanism satisfies ex ante budget balance.*

PROOF. To check ex ante budget balance we must show that the net amount of payments made by customer agents and compensations received by provider agents never runs a deficit. Subtracting the sum of expected compensations from the sum of expected payments yields

$$
\begin{aligned}
&\sum_{i=1}^{N}\int_0^{\bar{\theta}} \mathbb{E}_{\theta_{-i},\sigma}[t_C(\theta_i,\theta_{-i},\sigma)]f(\theta_i)d\theta_i - \sum_{j=1}^{M}\int_0^{\bar{\sigma}} \mathbb{E}_{\theta,\sigma_{-j}}[t_P(\theta,\sigma_j,\sigma_{-j})]h(\sigma_j)d\sigma_j \\
&= \mathbb{E}\Big[\sum_{i=1}^{N}\sum_{j=1}^{M} v(\theta_i,\sigma_j)x_{ij}(\theta_i,\theta_{-i},\sigma) - \sum_{i=1}^{N}\sum_{j=1}^{M}\frac{1-F(\theta_i)}{f(\theta_i)}\frac{\partial v(r,\sigma_j)}{\partial r}x_{ij}(r,\theta_{-i},\sigma)\Big] \\
&\quad - \mathbb{E}\Big[\sum_{j=1}^{M}\sum_{i=1}^{N} c(\theta_i,\sigma_j)x_{ij}(\theta,\sigma_j,\sigma_{-j}) + \sum_{j=1}^{M}\sum_{i=1}^{N}\frac{H(\sigma_j)}{h(\sigma_j)}\frac{\partial c(\theta_i,r)}{\partial r}x_{ij}(\theta,r,\sigma_{-j})\Big] \\
&= \mathbb{E}\Big[\sum_{i=1}^{N}\sum_{j=1}^{M}(\psi_C(\theta_i,\sigma_j) - \psi_P(\theta_i,\sigma_j))x_{ij}(\theta,\sigma)\Big] \geq 0. \qquad \text{(A.11)}
\end{aligned}
$$

The inequality in line (A.11) is a consequence of (10) in any incentive compatible and individually rational mechanism. Hence, the net balance of what customer agents pay and what provider agents receive in the mechanism is greater or equal zero in expectation. Exact equality can always be achieved by subtracting a constant from customer agents' payments.

Theorem 1. *Let customer agent a_i's valuation $v(\theta_i, \sigma_j)$ be maximized when $\theta_i = \sigma_j$ with the maximal value increasing in θ_i. Further, let $v(\theta_i, \sigma_j)$ be supermodular in both arguments and concave in θ_i. Let the distribution function $F(\cdot)$ be log-concave and let*

$$\begin{aligned} &\frac{\partial^3 v(\theta_i, \sigma_j)}{\partial \theta_i^2 \partial \sigma_j} \geq 0, \\ &\frac{f(\theta_i)}{1 - F(\theta_i)} \geq \frac{\partial}{\partial \theta_i} \log\Big(\frac{\partial v(\theta_i, \sigma_j)}{\partial \sigma_j}\Big). \end{aligned} \tag{22}$$

On the supply side, let provider agent b_j's provision cost $c(\theta_i, \sigma_j)$ be minimized when $\sigma_j = \theta_i$ with the minimal value increasing in σ_j. Further, let $c(\theta_i, \sigma_j)$ be submodular in both arguments and convex in σ_j. Let the distribution function $H(\cdot)$ be log-concave and let

$$\begin{aligned} &\frac{\partial^3 c(\theta_i, \sigma_j)}{\partial \theta_i \partial \sigma_j^2} \leq 0, \\ &\frac{h(\sigma_j)}{H(\sigma_j)} \geq \frac{\partial}{\partial \sigma_j} \log\Big(\frac{\partial c(\theta_i, \sigma_j)}{\partial \theta_i}\Big). \end{aligned} \tag{23}$$

Then, the TPAA mechanism maximizes the expected social welfare among all incentive compatible and individually rational mechanisms. Further, budget balance is satisfied.

PROOF. The inner term of the maximand in (7) is the difference between a_i's valuation and b_j's cost. Because $v(\theta_i, \sigma_j)$ is supermodular and $c(\theta_i, \sigma_j)$ is submodular (i.e., $-c(\theta_i, \sigma_j)$ is supermodular), their difference as a linear combination is again supermodular. If the production function is supermodular, the optimal match function is positively assortative (Shimer and Smith, 2000). By Lemma 1, the TPAA mechanism is incentive compatible and individually rational if and only if the monotonicity properties are satisfied. These are satisfied if the reserve functions $R_C(\theta_i)$ and $R_P(\sigma_j)$ are increasing. The hazard rate (22) for customer agents guarantees that a higher desired quality θ_i unambiguously results in a better lottery over the actual qualities of provider agents. Similarly, the hazard rate (23) on the provider agent side ensures that a higher actual quality σ_j unambiguously entails a lower likelihood to find a customer agent. Therefore, the reserve functions are increasing. Lemma 2 establishes budget balance.