

Trading across Borders in Online Auctions*

Elena Krasnokutskaya[†]
Johns Hopkins University

Christian Terwiesch[‡]
Wharton School of Business

Lucia Tiererova[§]
Johns Hopkins University

February 15, 2016

Abstract

We invoke the insights from the auction literature to study international trade in services using data from an online market for programming support. We find that the observed clustering of trade between countries can be rationalized through a model featuring endogenous sorting of sellers heterogeneous in quality and costs across projects offered by buyers who differ in willingness to pay for quality and outside options. We extend a single auction entry model to a setting where sellers choose among multiple projects available for bid. This plays an important role in explaining the data and understanding the effects of various trade policies.

Keywords: multi-attribute auctions, quality, participation in auctions, trade in services, trade policy

JEL Classification: C14, C18, D22, D44, D82, L15, L86.

*

[†]Email: ekrasno1@jhu.edu.

[‡]Email: terwiesch@wharton.upenn.edu.

[§]Email: ltierer1@jhu.edu.

1 Introduction

The Internet is transforming service markets by providing platforms that enable transactions between buyers who need a service and a wide range of potential service providers. One key aspect of this transformation is that the provision of many services, particularly professional services, no longer is restricted by the joint location of buyers and providers.¹ Such services become tradable within and across national borders. Hence, the rapid observed rise in international trade in professional services.

The design of online marketplaces for trade in services varies, but the common format involves buyers posting the description of the desired service (project) and then running an auction to select a provider. In turn, potential providers decide what project to bid on. The auction mechanism usually allows the buyer to select a provider based on the value of the submitted bid as well as other considerations, such as the expected quality of the provider.² The emerging organization of international trade in professional services through online auctions suggests that the insights developed in the auction literature could be brought to bear in an effort to understand how these markets function and to develop trade policies tailored to their unique design. The specific features of these markets and the nature of the available data, however, differ from those studied in the traditional empirical auction literature and pose unique methodological challenges. The objective of this paper is to develop a suitable conceptual framework and apply it to the data from a particular service market in an effort to understand the pricing and participation decisions that give rise to observed patterns of international trade and to assess the quantitative impact of several nontraditional trade policies.³

Our analysis is based on data collected over multiple years from a prominent online market for programming services. We observe the universe of around 600,000 buyers submitting projects for bidding, as well as the universe of approximately 50,000 potential sellers.⁴ Both buyers and sellers are widely dispersed across countries. Almost 80 percent of all projects are awarded by buyers to sellers located in foreign countries. As in the case of trade in goods, volume of service trade between countries in our data is proportional to the size of the markets. Even after conditioning on size, however, we observe significant clustering of trade flows, with buyers disproportionately awarding contracts to sellers from specific countries. Importantly, the clustering patterns vary across countries so that the set of seller countries receiving a disproportionate share of contracts from, say, American buyers differs from the set that receives a disproportionate share of awards from, say, British buyers and is different yet from that preferred by Australian buyers.

We find these trade patterns can be rationalized through a model that features endogenous sorting of heterogeneous buyers and sellers into cross-country trading arrangements. The het-

¹Another aspect of this transformation-often referred to as the rise of “on-demand economy,” “the gig economy,” or “1099 economy”- involves individuals forgoing or supplementing regular employment relationships with freelance or contract work through online platforms.

²Some prominent examples include general purpose platforms, such as Upwork, PeoplePerHour, or Guru, that enable online trade in legal, engineering, architecture, programming, marketing, accounting, administrative support, and other services. Specialized platforms include MediBid for medical, Chegg for tutoring, ProZ for translation, or Envato Studio for graphic design services, among numerous others.

³An incomplete list of papers studying the trade in physical products while invoking insights from industrial organization literature include Eaton and Kortum (2002); De Loecker (2011); Goldberg, Khandelwal, Loecker, and Pavcnik (2015); Fieler, Eslava, and Xu (2015). Recent analysis of online auctions is represented by Bajari and Hortacsu (2003); Cabral and Hortacsu (2010); Akerberg, Hirano, and Shahriar (2006); Lewis (2011); Backus and Lewis (2012); Decarolis, Goldmanis, and Penta (2014); Athey and Nekipelov (2012); and, Hendricks and Sorensen (2014). For a survey of the older literature on online auctions, see Bajari and Hortacsu (2004).

⁴A subset of the full market subject to our formal analysis features the same patterns described here.

erogeneity of sellers in the model is captured by their quality and the distribution of private costs. The distribution of sellers' qualities and costs differs across seller countries but does not depend on the prospective buyer's country. Symmetrically, buyers differ in their willingness to pay for quality and in their outside options, which are not observed by sellers. The distribution of buyers' tastes for quality and their outside options vary across buyers' countries but not those of sellers. Despite the fact that the model includes no bilateral preference, cost, or other parameters, it predicts up to one percentage point the market shares of various sellers' countries among the projects auctioned by buyers from a given country. Interestingly, the selection of sellers to participate in auctions (the supply side) turns out to be more important than the award decisions by buyers (the demand side) for generating clustering of trade in the model. To better understand these effects, we now describe the model and the structure of the data that enable more detailed identification.

For each week, our data contain information on all the projects available for bid and the set of all potential sellers participating in the market. For each potential seller, we observe each auction the seller decided to participate in and the bid submitted. Similarly, for each buyer, we observe the set of all sellers participating in the buyer's auction, the buyer's choice among them, and the realized price. This rich data structure enables us to disentangle the impact of demand- and supply-side decisions on the trade in this market.

Transactions in this market are implemented in the form of multi-attribute auctions that allow buyers to deviate from allocation based solely on price (as in standard auctions) in favor of service providers who are attractive on other dimensions—for example, because of higher quality of their services. Indeed, we frequently observe buyers choosing sellers who charge prices above the lowest price submitted in the auction. The design of the market through multi-attribute auctions allows us to infer buyers' preferences for seller attributes.⁵

The online market platform is specifically designed to provide buyers with an opportunity to assess quality of prospective service providers. In particular, buyers have access to the record of seller's past performance, are able to communicate with sellers during the auction and review examples of seller's previous work. This leads us to model sellers' quality as being only partially captured by the available measures of seller past performance and containing an unobserved (to a researcher) component.

To understand the functioning of the market and to be able to conduct informative counterfactual trade policy experiments it is essential to explicitly model the decisions of sellers whether to participate in the market in a given period and, conditional on participation, which auctions to enter.⁶ The key challenge to developing an empirically implementable model of entry in this market is the huge number of potential alternatives that each of the very large number of sellers in the market needs to consider. We find that it is possible to reduce the dimensionality while achieving a good fit to the data by assuming that each seller decides on the type of the project to enter and is indifferent between the project within the type. This allows us to summarize sellers' profits derived from a given auction type as a function of a realized local (type-specific) competition and other model primitives. This approach to modeling entry is novel in the literature but builds on existing work by Seim (2006) who studied location choice in video rental industry. To accurately model participation choices in our market, it is important to also allow for correlation of private entry costs across project types for the same individual. It is natural to expect such

⁵While the multi-attribute auctions format is also prevalent in off-line industry procurement, it is little studied, with the exception of Greenstein (1993, 1995).

⁶Hendricks and Sorensen (2014) and Backus and Lewis (2012)) consider the bidders' choice between several auctions in a different setting.

correlation to be present and modeling it proves essential to fit the data. In this we build on the ideas first developed in Berry (1992). Finally, our approach generalizes ideas developed in the literature on entry into an individual auction to a choice among multiple auctions.

Our estimation strategy exploits the fact that under the multi-attribute auction format buyer’s willingness to pay for quality and buyers’ outside options are not known to sellers. Thus, buyers’ choice set (the set of participating sellers) is exogenous conditional on buyer’s and project’s observable characteristics. This allows us to separate estimation into components that deal with buyer’s choice conditional on the choice set (the demand side) and seller’s optimal participation and pricing strategies (the supply side). To recover demand side parameters, we employ the estimation approach developed in Krasnokutskaya, Song, and Tang (2014b) which overcomes key estimation challenges associated with presence of a very large number of buyers and sellers (so that the number of distinct buyers’ choice sets is comparable to the number of projects); a high turnover of supply side participants; and the lack of full information about sellers’ qualities in the data. This estimation recovers the individual sellers’ unobserved quality as well as the distribution of buyers’ outside options and tastes. With these estimates in hand, we turn to estimation of supply side primitives. In the data we observe all seller types participating in all auction types. This allows us to recover bidding strategies for every type of seller used in every type of auction. Therefore, we are also able to recover the distribution of project cost for every type of the seller. Estimates of the bidding functions and the costs distributions allow us to impute ex-ante profit for every type of seller from participation in every type of project, given the observed competitive conditions in the market in every week in our data. Finally, this allows us to recover the distributions of the entry costs for every seller type to rationalize the observed participation behavior.

The results of estimation reveal that (a) the distributions of the buyers’ willingness to pay for quality and their outside options differ significantly across buyer countries, and (b) the distributions of seller’s qualities and costs differ significantly across seller countries. The combination of these differences rationalizes to high degree the observed trade patterns in the data. Key to understanding this result is the mutual interdependence of demand and supply side responses. The ability of a buyer preferring a particular quality type of service to obtain it is facilitated by the endogenous self-selection of sellers offering that quality into his auctions. The resulting competition between these sellers lowers the price of the quality desired by the buyer, and this dis-incentivizes suppliers of different quality from participating in the auction. The end result is that given buyer’s preference for quality, the endogenous response of supply ensures that he is more likely to procure it. Given the varying distributions of quality across countries, this explains the ability of the model to match the bilateral trade patterns without any pair-specific parameters. In fact, counterfactually shutting down participation channel in the estimated model eliminates 65% of clustering.

We further apply these insights to understand the potential impact from several counterfactual trade policies on the operation of an online market for services such as the one we study. Auction literature usually considers two types of policies that could impact the operation of the market. One type includes policies that directly affect participation (e.g., the bidder has to qualify to participate in the auction, or the bidder has to be invited to participate and the auctioneer decides whom to invite). The other type of policies primarily impact pricing in the market (e.g., the choice of the allocation mechanism such as first price or second price auction, or even less standard mechanisms such as multi-attribute auction or price preference given to a subgroup of bidders). We consider both types of policies and their impact on international trade and the welfare of market participants. The analysis highlights the interaction between pricing

and participation decisions and illustrates the key trade-offs important for trade policy design in online auction markets for services.

To illustrate the effects of trade policies primarily affecting participation, consider first a fairly blunt policy such as imposing a quota on participation of foreign sellers. There are various ways to implement this policy, but a simple mechanical way to think about it is that a foreign seller who decides to submit a bid in a particular auction has to apply for a permission form from the government and only a fraction of foreign seller are allowed to proceed with a bid. If the restriction is imposed irrespectively of the quality of foreign sellers, it has a non-trivial effect since it tends to exclude attractive sellers alongside the unattractive ones. This policy clearly reduces the competitiveness of the market and results in higher prices. Domestic sellers gain substantially from this policy since they are able to both increase their prices and their market share. Despite the policy not depending on quality, it induces a substantial change in the quality composition of sellers participating in the market. Participation of foreign sellers shifts toward higher quality bidders who are more competitive. Interestingly, this shift in composition of foreign participant affects the composition of domestic sellers as well. While domestic sellers of all quality levels participate and win more often, the participation and the market share of higher quality domestic sellers expand significantly more. This change in quality composition mitigates the adverse effect of higher prices on domestic buyers. This is important for a country such as US which is a net importer of services in this market and is thus importantly affected by the impact of the policy on buyers. On the net we actually find the effect of this policy to be slightly positive on US welfare while it is negative if we counterfactually do not allow the quality composition to respond.

A more refined implementation of this policy may condition the quota on quality of foreign sellers. Indeed, an important concern of the government in many service markets is the concern about quality of service providers. A typical policy response involves licensing of service providers. Such a requirement is absent in the market we study but can be easily implemented by, e.g., requiring certain minimum performance on a qualifications test. We now consider the effects of such a policy that imposes a lower bound on quality of foreign sellers permitted to participate in the market. While the specific implications of this policy parallel the effects of the quota described above, overall, we find this policy results in a very small impact on the US market participants because it eliminates the part of the quality distribution that is least attractive to the domestic buyers.

Finally, we consider a trade policy that directly affects pricing in favor of domestic sellers and lets participation adjust accordingly. A version of such policy is commonly used in US defense procurement where the US government imposes a margin by which domestic bids may exceed foreign bids for equivalent products. In the same vein, we consider a policy that levies a fine on a buyer who awards a project to a foreign provider while a domestic provider of similar quality submitted a bid that exceeded the chosen foreign bid by less than a specified margin. In an auction environment pricing always balances markup considerations vs. the probability of winning. Domestic sellers use their price advantage from this policy to increase their markups and to gain market share. In response, foreign sellers lower their prices to mitigate the loss of the market share. This, in turn, prevents domestic sellers from significantly increasing their prices. As a result, the average price in the market changes very little. Thus, the negative price consequences of this policy on domestic buyers are minimal. However, domestic buyers benefit because of the change in composition of foreign sellers towards higher quality (who charge higher markups and are able to lower their prices but still cover their costs with discriminatory policy in place). So the total effect on buyers is positive. The effect on domestic sellers is also clearly

positive. They significantly increase their market share while slightly increasing their markups on average. Overall, price preference policy is able to generate a large domestic welfare gain despite restricting the access to international trade. Note that while government might also raise revenue from collecting the fines, the fines are virtually never paid in equilibrium.

The paper is organized as follows. Section 2 describes the market; the model is developed in section 4. Section 5 discusses the empirical methodology, followed by section 6 summarizing the estimation results; section 7 outlines the results of the clustering decomposition whereas section 8 analyses potential impact of trade policies on this market. Section 9 concludes.

2 Market Description

This paper studies an online market for programming services, in which a platform serves as an intermediary between buyers (the demand side) and potential sellers (the supply side). Buyers procure services such as platform programming, databases, graphics programming and website design by posting job announcements to which interested sellers can respond by submitting a quote for a price at which they would be willing to complete the task. The market serves buyers from a variety of countries around the world by providing them access to sellers who differ in their country of origin and thus possibly in their costs and quality.

The intermediary company allocates jobs through multi-attribute auctions, allowing buyers to take into account seller characteristics in addition to the price quote. As a result, the selected seller is not necessarily the one who submits the lowest quote. An important feature of this allocation mechanism is that the award rule is not announced and thus remains unknown to other market participants.

The registry provides limited information on verifiable “outside” credentials as well as information about the on-site performance of the seller. The latter includes a history of performance-related measures such as reputation scores or ratings, buyers’ numerical feedback about working with a given seller, as well as instances of delays and disputes. In the case of a dispute, the company provides professional arbitration services that ensure that a seller is paid if only if the completed job satisfies industry standards.

Sellers often communicate with buyers before posting price quotes, with an average of three messages exchanged between a seller and a buyer prior to submitting a price quote. Sellers can also attach an example of previous work or a sketch of the proposed code. Hence, a buyer has an opportunity to form an opinion about each seller’s quality. These communications are not observable in the data which suggests a possibility of unobserved seller heterogeneity.

Prospective sellers observe neither the exact set of their competitors nor the bids submitted by other sellers.

3 Descriptive Analysis

Our dataset covers the first six years of operation of this on-line market. We focus our analysis on projects associated with graphics- as well as media-related programming. This is one of the most active segments of our programming market which operates regularly and receives projects from multiple buyer countries every week. Our dataset includes 49,334 projects with bids from 10,213 distinct sellers.

3.1 General Summary Statistics

Table 1 provides general summary of the data. It shows that an average project in our dataset is around \$500; on average a buyer expects the project to be completed in slightly less than three weeks although a large fraction of projects have a much shorter deadline (from a couple of days to a week).

Most buyers appear in the dataset only once. In an average week 465 projects are auctioned in this market with 1366 unique sellers expressing interest in participation⁷(in a median week the number of projects is 607; and the number of potential sellers is close to 1500). Majority of projects are allocated to one of the bidding sellers. On average only 3% of buyers choose not to allocate the project in this market after completing an auction.

A large fraction of sellers appears in the market only a few times; a median seller stays in the market for only one week. In our analysis we distinguish between the transitory sellers (those who stayed in the market less than 6 months) and permanent sellers who stayed in the market longer than six months. Majority of permanent sellers have tenure of more than two years and many of them stay in the market in excess of three years. As we document in the table, 25% of bids in an average week are submitted by permanent sellers (in a median week this share is equal to 21%). In our dataset an average permanent seller submits close to 300 bids during his tenure (median permanent seller submits 252 bids). Permanent sellers are quite uniform in their recorded performance: an average score of a permanent seller is 9.8 out of 10 possible points (with median equal to 9.75). Only a small fraction (around 12%) of permanent sellers are ever involved in a dispute, or have a delay (15% of permanent sellers) registered against them.

Multi-attribute feature of an auction is important since only 18% of the projects are allocated to the bidder submitting the lowest bid in an average week (21% in the median week). Specifically, an average buyer chooses to pay 32% (median is 29%) premium above the lowest bid submitted in his auction. Further, the instance of winning appears to be positively correlated with price. A logit regression

$$Y_{i,l} = X_{i,l}\beta_l + \beta_{0,l}B_{i,l} + \mu_{c(i)} + \epsilon_{i,l},$$

where $Y_{i,l}$ is indicator for a winning seller; $X_{i,l}$ are seller i ' performance characteristics and his experience; $B_{i,l}$ is a bid submitted by seller i ; $\mu_{c(i)}$ is a seller country fixed effect, estimates a statistically significant positive coefficient in front of price⁸ which indicates that sellers submitting higher price in a given auction are more likely to win. Such regularity typically is indicative of the presence of omitted variables reflecting (unobserved) seller heterogeneity. We follow methodology developed in Krasnokutskaya, Song, and Tang (2014a) to address this issue in the estimation.

Next, we analyze an international aspect of this market.

3.2 Evidence of International Trade

This is an international market. The demand side is represented by buyers from 170 countries with 55% of the projects submitted by US buyers, 25% of projects originating in UK, Australia and Canada, and a large number of countries responsible for 1% or 2% of the projects. The supply side is served by programmers from 240 countries: India is responsible for 27% of submitted bids; Pakistan, USA and Romania for about 11% each; and a large number of countries are submitting between 1% to 3% each. Further, significant amount of trade in this market occurs

⁷We define a seller as expressing interest if he asks buyer a question about the project or submits a bid.

⁸The estimated coefficient on price to be equal to 1.82 with the standard error equal to 0.56.

Table 1: General Summary Statistics

	Mean	Std. Dev	10%	50%	90%
Projects					
size (\$)	536	237.6	110	500	865
duration (days)	18	13	2	7	32
number of projects per buyer	1.05	0.02	1	1	1
number of projects (per week)	465	117	321	607	866
number of potential sellers (per week)	1366.4	541.2	550	1498	1963
frequency of no award (per week)	0.03	0.015	0.02	0.035	0.06
fraction awarded to lowest bid	0.18	0.03	0.17	0.21	0.23
winning bids (% of lowest bid)	1.32	0.3	0	1.29	1.75
Sellers					
tenure(weeks)	26.7	55.7	0	1	151
share of permanent sellers (per week)	0.26	0.05	0.17	0.21	0.35
Permanent Sellers					
number of bids (total)	324.2	211.9	60	252	736
average score	9.8	0.06	9	9.75	10
number of arbitrations	0.02	0.78	0	0	1
number of delays	0.075	0.496	0	0	1

This table reports summary statistics for the sample of 49,334 projects from graphics, web and media related programming categories.

across geographical borders. For example, only about 9% of US projects are allocated to US sellers, 50% of projects from India are allocated to Indian sellers, less than 1% of projects from UK, Germany or France are allocated to the sellers from these countries respectively and so on.

Table 2: Return to Submitting a Bid

Buyer Country	Seller's Country		
	North America	Eastern Europe	South and East Asia
North America	0.23	0.16	0.13
UK	0.19	0.16	0.12
Western Europe	0.14	0.18	0.12
Southern Europe	0.18	0.18	0.13
Eastern Europe	0.14	0.19	0.11
Australasia	0.17	0.14	0.12
South and East Asia	0.19	0.12	0.17

This table reports the fraction of winning bids among all bids submitted by the sellers from a given sellers country group for the projects from a specific buyer country group.

In the interest of tractability we focus on the trade patterns between the groups of buyers' and sellers' countries. Countries are grouped by geographic proximity and similarity of language

and economic conditions. The demand side of the market is represented by the following seven groups: North America, United Kingdom, Western Europe, Southern Europe, Eastern Europe, Australasia and South and East Asia.⁹ Similarly supply side focuses on seven groups of seller-countries: North America, Latin America, Western Europe, Eastern Europe, Africa and Middle East, South and East Asia, and Australasia.¹⁰ Majority of sellers originate from North America, Eastern Europe or South and East Asia, making these three country groups the focus of our analysis. In our data North American sellers plays less important role in comparison to the Eastern European or Asian groups. Specifically, only about 30% of projects receive bids from North American sellers while close to 90% of projects attract at least one Asian bid and 75% of projects attract at least one Eastern European seller.

We demonstrate next that trade in this market is characterized by pairwise clustering. That is, different buyer countries seem to focus their trade on different seller countries. To see this consider Table 2 which reports a fraction of winning bids among all bids submitted by a given seller country group to the auctions of a specific buyer country group. Clearly, the return to submitting a bid varies across pairs. Specifically, 23% of all bids submitted by North American sellers to North American buyers result in winning in contrast to 14% of North American bids submitted to Western or Eastern European buyers. Similarly, 18-19% of Eastern European bids win if they are submitted to Western or Eastern European buyers but only 12% wins if submitted to Asian buyers. Finally, 17% of bids submitted by Asian sellers result in winning if they are submitted to Asian buyers but only 11% if submitted to Eastern European buyers.

Another way to detect clustering is demonstrated in Table 3. Here we are comparing the conditional distribution of the number of projects allocated to different sellers countries conditional on buyer country to the marginal distribution of the number of allocated projects across sellers countries. Specifically, we are looking at the ratio of conditional distribution to the marginal one. The logic underlying this analysis is as follows. The number of projects from a given buyer country allocated to a specific seller country naturally depends on the economic presence of these countries in the on-line market. Specifically, we expect to see a large number of projects from North America to be allocated to all seller countries just because the overall number of such projects is very large. We eliminate the influence of the economic presence of the buyer country by comparing conditional shares of projects allocated to sellers from different country groups conditional on the buyer country group. Similarly, we expect to see a large share of projects to be allocated to Asian sellers in auctions for all buyers groups just because the number of Asian sellers participating in this market by far exceeds the number of sellers from any other country group. To isolate clustering pattern we study deviation of conditional shares (conditional on a buyer country) from a marginal shares. The results of this analysis are summarized in the left-hand-side panel of Table 3.¹¹

If trade were proportional to the size we would expect the ratios in Table 3 to be close to one. Instead, we find substantial distortions: for example, the share of North American projects

⁹North America combines USA and Canada; Western European group includes Austria, Belgium, Denmark, Finland, Germany, Iceland, Ireland, Netherlands, Norway, Sweden, and Switzerland; Southern Europe consists of France, Greece, Italy, Portugal, and Spain, Eastern Europe includes Belarus, Bosnia and Herzegovina, Bulgaria, Croatia, Czech Republic, Estonia, Hungary, Latvia, Lithuania, Moldova, Poland, Romania, Russia, Serbia, Slovakia, Slovenia, and Ukraine; Australasia consists of Australia and New Zealand; and South and East Asia includes India, Pakistan and Singapore.

¹⁰On the supply side Western Europe is combined with South Europe due to the small number of observations; South and East Asia additionally includes China, Malaysia, Indonesia, and Bangladesh.

¹¹This calculation underlines a standard test used for detecting clustering patterns in statistical literature. The appropriate X^2 test statistics is given in Equation (6) in Section 6.4.

Table 3: Patterns in Participation and Allocation

Buyer Country	Allocation			Participation		
	Seller's Country			Seller's Country		
	North America	Eastern Europe	South and East Asia	North America	Eastern Europe	South and East Asia
North America	1.436	0.975	0.975	1.200	0.937	0.996
UK	1.136	1.065	0.962	1.020	1.058	0.981
Western Europe	0.77	1.188	0.935	1.000	1.066	0.957
Southern Europe	0.927	1.102	0.968	0.940	0.953	0.941
Eastern Europe	0.614	1.231	0.926	0.700	1.062	1.032
Australasia	0.849	1.001	1.099	0.880	0.916	1.053
South and East Asia	0.666	0.803	1.19	0.840	0.809	1.098
Marginal	0.065	0.295	0.473	0.053	0.274	0.564

This table documents patterns in projects' allocation and sellers participation across pairs of buyer-seller country groups. Specifically, the left-hand-side panel reports the ratio between the conditional distribution of the number of bids across seller country groups conditional on buyer country group and the marginal distribution of the number bids across seller country groups. Similarly, the right-hand-side panel reports the ratio between the conditional distribution of the number of projects awarded to different seller country groups conditional on buyer country group and the marginal distribution of the number of awarded projects across seller country groups.

allocated to North American sellers is 44% higher than the share of North American sellers in the marginal distribution whereas the corresponding share associated with projects auctioned by Asian buyers is 33% lower than the share of North American sellers in marginal distribution. Similarly, the share of Eastern European projects allocated to Eastern European sellers exceeds the marginal share of Eastern European sellers by 23% and the marginal share of Asian sellers in Asian projects exceeds their marginal share by 19%.¹²

Additionally, participation in this market is characterized by similar the clustering pattern. This can be seen from the right-hand-side panel of Table 3 which shows the ratio of the conditional distribution of the number of bids submitted to auctions of a given buyer country across different seller countries to the marginal distribution of the number of bids across seller countries. This table reflects clustering pattern similar to the one observed in the allocation of projects. This regularity perhaps is not surprising since in this market a buyer chooses from a set consisting of sellers who selected to participate in the auction conducted by this buyer. If buyers of country c_1 derive higher values from sellers of country c_2 and thus are more likely to choose sellers from this country, then sellers from country c_2 would tend to submit bids for projects from country c_1 with higher probability. Hence, the clustering effect might be re-enforced since the choice sets of buyers from country c_1 would be more likely to include sellers from country c_2 than the choice sets of buyers from other countries. This effect is likely to be further reinforced through pricing since competition between the desirable bidders due to their enhanced participation would be more intense. A model rationalizing clustering of trade in this market has to take strategic participation into account.

¹²Similar patterns have been documented in the literature studying trade in the physical goods where the clustering is often summarized by means of the gravity equation.

The clustering might naturally arise because of the pairwise preferences characterizing the demand side or through pairwise differences in project and entry costs on the supply side. In subsequent analysis we investigate to what extent the clustering in trade could be generated through sorting of heterogeneous sellers across buyers with heterogeneous tastes where neither sellers' attributes or costs nor buyers' preferences depend on their partner's country.

4 Model

The market brings together a number of one-time buyers seeking to procure service for a single project and qualified sellers interested in providing such service. Project l is characterized by buyer's country of origin, $o_b(l)$ with $o_b = \{1, 2, \dots, O_b\}$, and the size, z_l . We denote the set of projects offered in the market by M and summarize this set by a vector, $J_M = (m_1, \dots, m_{O_b})$ reflecting the number of projects across buyers' countries.

Each seller j is characterized by a vector of attributes, $x_j \in \mathcal{X} = \{\bar{x}_1, \cdot, \cdot, \cdot, \bar{x}_P\}$, and a scalar quality index, q_j , which admits values from a discrete set $\mathcal{Q}(x) = \{\bar{q}_1(x), \cdot, \cdot, \cdot, \bar{q}_K(x)\}$. Notice that distribution of quality indices is x -specific in a sense that both the number of quality levels and the quality levels themselves may depend on x .

Sellers are of two types $r = \{p, t\}$ where 'p' denotes permanent and 't' denotes transitory sellers. Permanent seller's quality is known to all market participants; a transitory seller's quality is his private information which is distributed according to the distribution $H_{Q|x} = \{\bar{q}_k(x), \bar{\pi}_k(x)\}$ for a seller characterized by a vector of attributes x . Transitory seller reveals his quality draw to a buyer when submitting a bid. Here we follow a convention of using capital letters to denote random variables (such as quality of transitory seller, Q) and low-case letters to denote realizations of random variables (such as $q_k(x)$).

Seller j is further characterized by private marginal costs, (ν_j, c_j) , where $\nu_j = (\nu_{1,j}, \dots, \nu_{O_b,j})$ and $\nu_{o_b,j} z_l$ reflects seller's total participation costs associated with project from a country o_b , and $c_j z_l$ denotes sellers' total cost of performing service. We assume that marginal participation costs and marginal cost of work are independent of each other and across sellers; and are distributed according to the distributions $F_{(x,q)}^\nu$ and $F_{(x,q)}^C$. Note that distributions of costs depend on sellers' attributes and quality but not on the country of project's origin.

Sellers present in the market (potential bidders) are randomly drawn from the general population of sellers. We denote the set of potential bidders by N . Information about this set which is available to all potential bidders is summarized by a vector I_N which reports the numbers of permanent potential bidders for each (x, q) -group, and the numbers of transitory potential bidders for each x -group. In general, we will use notation I_Ω below to reflect public information about the set of sellers Ω .

The time line for the market runs as follows. Each potential seller j observes (J_M, I_N) and a draw of his marginal participation costs ν_j ; he then chooses among $O_b + 1$ alternatives: he either submits a bid for a project of type $o_b \in \{1, \cdot, \cdot, \cdot, O_b\}$ or decides not to participate in the market. If seller decides not to participate he obtains a payoff of zero and leaves the market. If he chooses one of the participation options he observes his marginal cost of work, c_j , and the set of sellers who chose the same type of project as he did. He does not observe specific projects where his competitors submit their bids. The bidder then decides on the value of his marginal bid, \tilde{b}_j ; randomly selects a project within his chosen type; observes project size; and places the bid equal to $z \tilde{b}_j$. In a slight abuse of notation we refer to the set of sellers who decided to submit a bid to a project of type o_b as the set of potential bidders for this type of project and denote

this set by N_{o_b} (information about this set is summarized in I_{o_b}). The seller thus has access to information about I_{o_b} , J_M , and c_j when he decides on his bid. After bids are submitted, buyers decide on project allocation.

Project Allocation. We use A_l to denote the set of sellers who submitted bids for project l and refer to such sellers as *active* bidders. These sellers form buyer's choice set. For each seller $j \in A_l$ the buyer observes attributes, x_j , quality index q_j and his bid $b_{l,j} = \tilde{b}_{l,j}z_l$. Here $\tilde{b}_{l,j}$ is a per unit bid.

Allocative decisions follow the format of multi-attribute auction. Specifically, buyer l associates a value, $\Delta_{l,j}$, with an active seller $j \in A_l$ and awards his project to the active seller with the highest level of $\Delta_{l,j} - z_l\tilde{b}_{l,j}$ if this level exceeds buyer's outside option $U_{l,0}$; otherwise, he leaves the project unassigned.

The buyer's value is a function of seller's attributes with buyer-specific coefficients α_l and $\epsilon_{l,j}$ ($\epsilon_{l,j}$ represents the residual value assigned by buyer to a specific seller), i.e.,

$$\Delta_{l,j} = \alpha_l(q_j + x_j\beta) + \epsilon_{l,j}, \quad (1)$$

We let $\epsilon_l = \{\epsilon_{l,1}, \dots, \epsilon_{l,|A_l|}\}$ and refer to (α_l, ϵ_l) as the vector of buyers' utility coefficients.¹³

In keeping with the definition of a multi-attribute auction, sellers do not observe the utility coefficients or the outside option of a specific buyer, and consider these to be a random draws from the corresponding distributions specific to the population of buyers from country o_b , $F_{\alpha,\epsilon,U_0|o_b}$.

Sellers' Strategies and Market Equilibrium. Sellers' strategies in this game consist of two components: participation component, $d : R_+^{O_b} \times \mathcal{I}_N \times \mathcal{J}_M \rightarrow \{0, 1, \dots, O_b\}$ which maps vectors of marginal entry costs, ν , and information about market competitiveness, (I_N, J_M) , into preferred participation choices; and bidding component, $\sigma(\cdot; o_b) : R_+ \times \mathcal{I}_{N_{o_b}} \times R_+ \rightarrow R_+$ which for each o_b maps total cost of work and information about competitive structure associated with specific type of project, $(I_{N_{o_b}}, m_{o_b})$, into marginal bids.

Let us denote a vector of strategies used by competitors of seller j by (d_{-j}, σ_{-j}) ; then a seller j of type, $\tau_j = (r, x, q)_j$, who chooses to participate in an auction of type o_b ; draws marginal cost of work c_j and submits marginal bid \tilde{b}_j obtains the expected payoff equal to

$$\Pi_{\tau_j}(\tilde{b}_j, c_j | I_{N_{o_b}}, m_{o_b}; o_b, \sigma_{-j}) = \bar{z}(\tilde{b}_j - c_j) \mathcal{P}(j \text{ wins} | \tilde{b}_j, \tau_j, I_{N_{o_b}}, m_{o_b}; o_b, \sigma_{-j}), \quad (2)$$

where $P(j \text{ wins} | \tilde{b}_j, \tau_j, I_{N_{o_b}}, m_{o_b}; o_b, \sigma_{-j})$ is his probability of winning a project of type o_b given the competitive structure for the projects of this type (implied by $(I_{N_{o_b}}, m_{o_b})$ and the fact that competitors are randomly allocated across projects within the type) and provided that his competitors follow bidding strategies σ_{-j} . Here \bar{z} denotes the expected size of the project sold in this market.¹⁴ The payoff expression reflects the fact that when deciding on a bid, a seller is only informed about the set of potential rather than actual competitors for the project where he is submitting a bid.

¹³We have also considered a more general specification where β varies across buyers. However, the specification above fits our data best.

¹⁴We assume that the distribution of sizes is the same across buyer countries. This assumption, indeed, closely approximates the reality.

The expression for $P(j \text{ wins} | \tilde{b}_j, \tau_j, I_{N_{o_b}}, m_{o_b}; o_b, \sigma_{-j})$ is given by

$$\begin{aligned} & \sum_a \Pr(j \text{ wins} | \tilde{b}_j, \tau_j, A_{-j} = a; o_b, \sigma_{-j}) \Pr(A_{-j} = a | I_{N_{o_b}}, m_{o_b}; o_b, \sigma_{-j}) \text{ with} \\ \Pr(j \text{ wins} | \tilde{b}_j, \tau_j, A_{-j} = a; o_b, \sigma_{-j}) &= \Pr(\Delta_j - \tilde{b}_j \geq \max_{p \in a} \{U_0, \Delta_p - \sigma_p(c_p)\} | o_b). \end{aligned}$$

In this expression the expectation is taken with respect to the set of actual competitors (sellers submitting a bid in the same auction as j). The probability of a specific realization a is computed taking into account that each seller in N_{o_b} chooses one out of m_{o_b} auctions at random. Thus the number of sellers from each τ -group entering a specific auction follows a multinomial distribution, i.e. it is distributed as the number of successes in the number of trials equal to the size of this group among potential sellers for auctions of type o_b , $|N_{o_b, \tau}|$, where the probability of success is given by $\frac{1}{m_{o_b}}$.¹⁵ The probability of winning conditional on the set of actual competitors is computed by taking expectation over the distribution $F_{\alpha, \epsilon, U_0 | o_b}$ and over the distributions of the competitors' costs for the set of competitors a (competitors' costs when combined with σ_{-j} account for the distribution of bids submitted by competitors).

Seller's j ex-ante payoff from participating in an auction of type o_b is then given by

$$\bar{\Pi}_{\tau_j}(I_N, J_M; o_b, d_{-j}, \sigma_{-j}) - \nu_{o_b, j} \bar{z}.$$

This payoff obtains from $\Pi_{\tau_j}(\sigma_j(c_j), c_j | I_{N_{o_b}}, m_{o_b}; z, o_b, \sigma_{-j})$ by taking expectation over the possible realizations of c_j as well as possible realizations of ν_{-j} (that together with the participation strategies d_{-j} of j 's competitors from N determines the set of potential competitors for the types of auction o_b , N_{o_b}). Intuitively, seller chooses an option that promises the highest net profit. Profitabilities of various alternatives are determined by the buyers' tastes (price sensitivity and outside option) and expected competition in the corresponding submarket.

In line with the existing empirical auction literature, we assume that the observed outcomes reflect a type-symmetric pure strategy Bayesian Nash equilibrium (psBNE). In such an equilibrium, participants who are *ex ante* identical (i.e. either permanent or transitory and characterized by the same x and q) adopt the same strategies. Formally, an equilibrium of this game is given by a profile of strategies $\{d_\tau^*, \sigma_\tau^*\}$ with $\tau = (r, x, q)$ such that

$$\begin{aligned} d_\tau^*(\nu_j, I_N, J_M) &= \arg \max_{o_b} [\bar{\Pi}_\tau(I_N, I_M; o_b, d_{-j}^*, \sigma_{-j}^*) - \nu_{o_b, j} \bar{z}] \\ \sigma_\tau^*(c_j, I_{N_{o_b}}, m_{o_b}; o_b) &= \arg \max_b \Pi_\tau(b, c_j | I_{N_{o_b}}, m_{o_b}; o_b, \sigma_{-j}^*) \end{aligned}$$

for all possible realizations of (ν_j, c_j) , I_N , J_M , $I_{N_{o_b}}$, and for all $o_b \in \{1, \cdot, \cdot, \cdot, O_b\}$. Existence of the equilibrium of this game can be established by applying results from McAdams (2003) in this environment.

Comment. In this analysis we assume that sellers do not take the size of the project into account when deciding where to submit a bid. Such an assumption is not conceptually necessary since our methodological framework can be easily adjusted to allow sellers choose projects by size and the country of origin. It is, however, convenient from the point of view of implementation since it reduces the number of auxiliary objects (and, ultimately, the number of parameters) that need to be estimated - this point will become clear once we explain our estimation strategy. It

¹⁵So that $\Pr(|A_\tau| = 0) = (m_{o_b} - 1)/m_{o_b}^{|N_{\tau; o_b}|}$, $\Pr(|A_\tau| = 1) = (|N_{\tau; o_b}|/m_{o_b})((m_{o_b} - 1)/m_{o_b})^{|N_{\tau; o_b}| - 1}$, $\Pr(|A_\tau| = 2) = [|N_{\tau; o_b}|/m_{o_b}][(|N_{\tau; o_b}| - 1)/2m_{o_b}]((m_{o_b} - 1)/m_{o_b})^{|N_{\tau; o_b}| - 2}$ and so on.

also does not distort the reality too much since sellers tend to choose projects from the same size bin (small, medium, large or very large) and the numerical size of the project is often revised in which case bids are pro-rated. Finally, this assumption does not preclude us from fitting the data quite well.

5 Estimation Methodology

We have access to weekly observations on the operation of this market. For a given week the data contain information on the set of projects offered in the market, and the set of potential sellers; for each project we know the set of sellers submitting bids, their bids and the buyer's choice. In addition, for each seller who ever appears in the market we observe his vector of x -attributes and whether the seller is of permanent or transitory type. Unlike buyers, we do not observe sellers' quality indices, q . Our environment is thus characterized by unobserved seller heterogeneity.

We need to recover from the data the distributions of buyers' utility coefficients and buyers' outside options conditional on buyer country, $F_{\alpha, \epsilon | o_b}$ and $F_{U_0 | o_b}$; utility coefficient, β , permanent sellers' quality levels, q_j , the distribution of transitory sellers qualities, $H_{Q|x} = \{\bar{q}_k(x), \bar{\pi}_k(x)\}$ on the demand side; and the distributions of sellers' marginal participation and project costs, $F_{(x,q)}^\nu$ and $F_{(x,q)}^C$ on the supply side.

Our estimation strategy exploits the fact that under the multi-attribute auction format buyer's willingness to pay for quality or buyers' outside options are not known to sellers. Thus, buyers' choice set (the set of participating sellers) is exogenous conditional on buyer's and project's observable characteristics. This allows us to separate estimation into components that deal with buyer's choice conditional on the choice set (the demand side) and seller's optimal participation and pricing strategies (the supply side). We begin by discussing the estimation of demand-side primitives.

5.1 Demand Estimation

We make use of the two-step estimation approach developed in Krasnokutskaya, Song, and Tang (2014a). Here we briefly summarize the rationale for using this methodology, identification strategy and the details of estimation procedure.¹⁶

This setting presents several methodological challenges. To see this let us first consider an environment without transitory sellers (or where transitory sellers differ only in their observable attributes, x). In this setting we only need to focus on recovering the quality levels (fixed effects) of permanent sellers. Recall that in a traditional discrete choice setting fixed effects associated with different alternatives are identified from the observed probabilities that a given alternative is chosen conditional on the choice set. In our setting choice sets are buyer-specific since sellers self-select to participate in a given auction. Due to the large numbers of sellers and buyers conditional choice probabilities cannot be precisely estimated. To get a sense of magnitudes consider that the number of permanent sellers present in the market for a given type of work is around 300 to 500 whereas only 2 or 3 permanent sellers appear in any given auction. This

¹⁶Hagedorn, Law, and Manovskii (2012) develop a related methodology for identifying equilibrium sorting on unobservables in the traditional labor markets.

means that the number of possible choice sets is at least $C_{300}^3 = \frac{300!}{3!297!} = \frac{300 \cdot 299 \cdot 298}{3} = 8,910,200$ which exceeds the number of projects we have in our dataset. In fact, in our data no more than five projects are associated with the same set of participating permanent sellers. Because of this, we have to consider probabilities that aggregate over buyers choice sets, a :

$$\Pr(i \text{ wins} | b) = \sum_{a, i \in a} \Pr(i \text{ wins} | b, a) \Pr(A = a).$$

However, such aggregation may not allow us to exploit variation in seller's probability of winning (analogous to product's market share) across choice sets, a traditional approach to identifying the distribution of random coefficients. It might be useful to implement only partial aggregation.

Let us now return to the realistic setting where transitory sellers differ in their qualities and these qualities are observable to buyers. In such setting, transitory sellers participation and bids depend on their (unobserved) qualities. Due to this regularity the distribution of qualities of transitory sellers in a given auction depends on transitory sellers' bids and attributes (which makes these later variables endogenous from econometrics point of view). In other words, a mixture problem obtains. Notice that standard methodologies developed in other mixture settings (Kasahara and Shimotsu (2009), Berry, Levinsohn, and Pakes (1995), and Berry, Levinsohn, and Pakes (2004) and others) do not work here since mixture probabilities depend on the variables which appear in the payoff expression and the number of observations per transitory seller is really small.

One might attempt to deal with this problem by solving for $\Pr(Q_h = \bar{q}_k | b_h, x_h)$ from the model within the estimation routine for a given vector of a parameter values. However, solving one such bidding and participation game would take a long time and solutions can be very fragile if parameter values are far from the truth. Further, a large number of sellers and projects result in a huge number of possible choice sets for which the problem would have to be solved. These issues combined make this approach computationally infeasible.

Identification Strategy. Because of the reasons outlined above we adopt identification strategy and a two-step estimation approach where in the first step we classify permanent sellers characterized by a common vector of x -attributes into groups of equal quality. Such grouping, once constructed, facilitates recovery of model's primitives in several ways. First, the task of recovering permanent sellers' qualities is reduced since now we only need to recover the quality levels associated with different groups rather than quality level for every permanent seller. Further, buyers' choice sets may now be represented in terms of participating sellers' (x, q) -group memberships for permanent sellers and x -group memberships for transitory sellers rather than in terms of their identities. Thus, representing sellers in terms of group memberships offers a natural way for partial aggregation of buyers' choice sets thus facilitates recovery of the distributions of the utility coefficients and buyers' outside options. Finally, such representation permits imposing the restriction that the distributions of permanent and transitory sellers qualities have the same support. The later feature allows us to separate identification of the payoffs associated with various bundles of (x, q) attributes from the identification of the mixing probabilities. Formal identification argument can be found in Krasnokutskaya, Song, and Tang (2014a).

Briefly, the distributions of utility coefficients are identified through the variation in the buyers' choice sets (specifically, through the variation in the composition of the set of permanent active bidders) and the variation in prices of permanent bidders. For this we need to consider choice sets that include at least two permanent sellers. This identification mechanism is very

similar to the one traditionally used in discrete choice estimation. Further, notice that for a given distribution of utility coefficients, the distribution of the payoffs associated with transitory sellers is known up to a mixing probability (which depends on transitory seller’s bid). To see how the mixing probability is identified consider auctions where a set of participants consists of a permanent and a transitory sellers. The variation in the winning probability of the permanent seller in response to the variation in his bid and holding the bid and the x - group of the transitory seller fixed identifies conditional mixing probabilities and the distribution of outside option. To separate mixing probabilities from the distribution of outside options we need to separately consider sets of auctions where the set of participants includes transitory sellers from different x -groups.

The estimation procedure we use consists of two components: classification algorithm; and GMM estimation of the distributions of utility coefficients, outside options, quality differences, and the conditional distribution of transitory sellers’ quality, $\Pr(Q_h = \bar{q} | b_h, x_h)$. We now describe each of these components below.

Recovering Quality Group Structure. Let us denote the set of permanent sellers with a vector of observable attributes x by $S(x)$. Intuitively, consider sellers i and j from $S(x)$ who participate in two separate but ex-ante identical auctions (i.e., the project characteristics and the realized set of competitors are the same) and submit equal bids. Under such circumstances the seller with the higher value of q has the higher chance of winning. Note that the winner is not deterministic in the presence of uncertainty about buyers’ utility coefficients and outside option. The ranking of winning probabilities is preserved when aggregating over the projects with the same distribution of buyers utility coefficients and outside options (i.e., projects with the same o_b) and / or over possible sets of competitors as long as the probability of encountering a given set of competitors is the same for both sellers. This condition holds if, for example, the pool from which competitors are drawn does not include either i or j .

Formally, we rely on the following pair-specific index:

$$r_{i,j}(b) = \Pr(i \text{ wins} \mid B_i = b, i \in A, j \notin A, i, j \in N_{o_b}; o_b). \quad (3)$$

This index reflects the probability that seller i wins an auction of type o_b when submitting a bid b and when the set of his direct competitors does not include j . Proposition 1 in Krasnokutskaya, Song, and Tang (2014a) establishes pairwise ranking of bidders i and j on the basis of indices $r_{i,j}(b)$ and $r_{j,i}(b)$. Further, Krasnokutskaya, Song, and Tang (2014b) demonstrate that if pairwise ranking of every pair of sellers within $S(x)$ can be established (the set of pairwise connections is complete) then the group structure of this set of sellers is identified and proposes a classification procedure to recover quality group structure of a given set $S(x)$.

Briefly, the algorithm works as follows. For each seller i in $S(x)$, we first divide the remaining sellers in $S(x)$ into two groups, one with sellers likely to have higher quality than i and the other with sellers likely to have lower quality than i . We obtain this division by comparing the p -values from two pairwise bootstrap tests of the inequality restrictions $r_{i,j} \geq r_{j,i}$ and $r_{i,j} \leq r_{j,i}$. Next, we place seller i in one of the two groups depending on whether seller i is likely to have the same quality as the other sellers in the group. Thus we obtain one group structure for each seller i , and choose one of these structures that has strongest empirical support (in terms of average p -values). This gives the first division of $S(x)$ into two subgroups.

We then sequentially select a subgroup with sellers most likely to have heterogeneous qualities, and divide the group similarly as before. To prevent overfitting (i.e., ending up with too many

subgroups), we stop the division process when a goodness-of-fit measure defined in terms of average p -values is dominated by a penalty term.

This classification algorithm estimates the whole group structure at once instead of reconstructing it from pairwise comparisons sequentially. The advantage of such approach is that it imposes in estimation transitivity of sellers' ordering which otherwise may be violated in finite samples.

GMM Estimation. Next, we proceed with GMM estimation of other demand-side primitives treating the recovered permanent sellers' group memberships as given.¹⁷

The moment conditions are primarily built around the permanent seller's winning probability given the seller attributes and quality group affiliation and on some features of the set of active competitors. We construct moment conditions along the identification strategy we outlined above. Let $e_\tau(b, I, J; o_b)$ denote the conditional winning probability of seller from group τ in the auction of type o_b with competitive conditions summarize by vectors I (composition of N and N_{o_b}) and J (composition of the set of projects); $W_{\tau,l}$ is an indicator variable which is equal to one if auction winner is of type τ . Then we construct a moment condition as follows:

$$\mathbf{E}[h(b, I, J)(\bar{W}_\tau - \bar{e}_\tau(b, I, J; o_b))] = 0,$$

where $h(\cdot, I)$ is a function of bid vector in the auction with competitive conditions summarized by I and J .¹⁸

We use two sets of moments. The first set includes (1a) moments that are based on the permanent seller's probability of winning in an auction where two or more active permanent bidders belong to the same group; (1b) moments that are based on the permanent seller's from a group (p, x, q) probability of winning in an auction where he competes with one or more permanent bidders belonging to a different group. As for the choice of $h(\cdot)$, we consider the following functions: constant (equal to 1); the difference between the permanent seller j 's (purported winner) bid and a bid of another permanent seller (in this example, seller i), $B_j - B_i$; the bid of seller j , B_j ; the squared difference between the permanent seller j 's bid and a bid of another permanent seller from the same group, $(B_j - B_i)^2$; and the squared bid of sellers j , B_j^2 ; product of the difference and the bid of seller j , $(B_j - B_i)B_j$; the product of seller j 's bid and the bid of a transitory seller, $B_j B_h$; the product of the square of seller j 's bid and the transitory sellers bid, $B_j^2 B_h$; as well as the product of seller j 's bid and transitory seller's attributes, $B_j x_h$; the product of the square of seller j 's bid and the transitory seller's attributes, $B_j^2 x_h$. We consider such set of moments for each buyer country group.

Accommodating Unobservability of Transitory Sellers' Qualities. To use these moment conditions, we need to evaluate $\bar{e}_\tau(b, I, J; o_b)$. However, note that a buyer observes (x, q) -group affiliations of all sellers in his choice set, whereas the econometrician does not observe

¹⁷The estimation error due to using the estimated quality groups does not affect the asymptotic distribution of the GMM estimator because it has arbitrarily fast convergence rate due to the finite number of quality groups.

¹⁸Notice that the moments are formulated in terms of the events related to the probability that a certain type of seller rather than a specific individual seller wins. Specifically, for a given auction l :

$$W_{\tau,l} = \frac{1}{|A_{\tau,l}|} \sum_{j \in A_{\tau,l}} W_{j,l}$$

$$\bar{e}_{\tau,l}(b, I, J; o_b) = \frac{1}{|A_{\tau,l}|} \sum_{j \in A_{\tau,l}} \bar{e}_{j,l}(b, I, J; o_b).$$

Here $A_{\tau,l}$ denotes the set of active sellers of type τ in auction l ; whereas $W_{j,l}$ and $\bar{e}_{j,l}(b, I, J; o_b)$ denote the indicator and the theoretical probability that a specific seller j wins respectively.

transitory sellers' qualities. This means that $\bar{e}_\tau(b, I, J; o_b)$ is the winning probability after integrating out the vector of participating transitory sellers' qualities. Let Q_t be the participating transitory sellers' quality vector in the auction, and take $\bar{e}_\tau(b, I, J; Q_t = \bar{q}, o_b)$ to be the same as $\bar{e}_\tau(b, I, J; o_b)$ except that it indicates the winning probability for group τ when the auction has participating transitory sellers with quality vector \bar{q} . The latter winning probability $\bar{e}_\tau(b, I, J; Q_t = \bar{q}, o_b)$ reflects buyers' decisions, and is determined by the distribution of buyers' utility coefficients and outside option. We write

$$\bar{e}_\tau(b, I, J; o_b) = \sum_{\bar{q}} \bar{e}_\tau(b, I, J; Q_t = \bar{q}, o_b) \Pr(Q_t = \bar{q} | b_t, I, J; o_b), \quad (4)$$

where b_t is the vector of transitory seller bids submitted in the auction. Note that the conditional probability of $Q_t = \bar{q}$ given b and I does not depend on the group characteristic (x, q) of permanent bidders.

To obtain the expression for $e_\tau(b, I, J; Q_t = \bar{q}, o_b)$ we parametrize the distributions of $\epsilon_{i,l}$ and (α, β, U_0) in a standard way. However, it is not immediately obvious how to parametrize $\Pr(Q_h = \bar{q} | b, I, J; o_b)$, because it involves the transitory sellers behavior. Following Krasnokutskaya, Song, and Tang (2014a) we exploit the following representation of these objects:

$$\Pr(Q_t = \bar{q} | b_t, I, J; o_b) = \frac{\prod_{j \in A^t} \Pr(Q_j = \bar{q}_j) \Pr(j \in A_t | \tau, I_N, J) f_\tau^B(b_j | I_{o_b}, J; o_b)}{\sum_{\bar{q}'} \prod_{j \in A^t} \Pr(Q_j = \bar{q}'_j) \Pr(j \in A_t | \tau', \bar{I}_N, J) f_{\tau'}^B(b_j | I_{o_b}, J; o_b)},$$

and parameterize $f_\tau^B(b | I_{o_b}, J; o_b)$ and $\Pr(j \in A | \tau, I_N, J; o_b)$ which correspond to the density of bids and the probability of participation associated with a transitory seller characterized by $\tau = (t, x, q)$.

Identification strategy presented in previous section permits nonparametric recovery of $\Pr(Q_t = \bar{q} | b_t, I, J; o_b)$. We make use of additional restrictions in estimation that would allow us to identify $\Pr(h \text{ is active} | x_h, q_h, I_N, J; o_b)$, $f_b(b_h | x_h, q_h, I_{o_b}, J; o_b)$, and $\Pr(Q_h = q_h | x_h)$ separately. Specifically, we match empirical and theoretical means and variances of the permanent and transitory sellers' bid distributions, as well as the frequencies with which potential permanent and transitory sellers submit a bid to auctions from different buyer countries aggregated to the level observed in the data. In some specifications, we additionally impose optimality of the sellers participation decisions.¹⁹

5.2 Supply Side Estimation

Recovering the Distributions of Projects Costs. We extend the standard methodology for recovering the distribution of private costs to the multi-attribute auction environment. Our approach builds on that of Guerre, Perrigne, and Vuong (2000) developed for standard first price auctions. Specifically, we recover the distributions of the sellers' costs conditional on sellers' attributes by combining the bid distributions of permanent sellers with the corresponding inverse bid functions:

$$F_{(x,q)}^C(c) = F_{(p,x,q); o_b}^B(\xi_{(p,x,q); o_b}^{-1}(b | I_{o_b}, m_{o_b}) | I_{o_b}, m_{o_b}).$$

¹⁹In these specifications the distribution of utility coefficients, buyers' outside options, and quality levels are estimated jointly with the distributions of sellers costs.

The inverse bid function, $\xi_{(\tau; o_b)}(b | I_{o_b}, m_{o_b})$ with $\tau = (p, x, q)$, is derived from the first order condition of the corresponding permanent seller's optimization problem:

$$\xi_{\tau; o_b}(b | I_{o_b}, m_{o_b}) = b - \frac{\Pr(i \text{ wins} | b, \tau, I_{o_b}, m_{o_b}; o_b)}{\frac{\partial}{\partial b} \Pr(i \text{ wins} | b, \tau, I_{o_b}, m_{o_b}; o_b)}.$$

Both $\Pr(i \text{ wins} | b, \tau, I_{o_b}, m_{o_b}; o_b)$ and $\frac{\partial}{\partial b} \Pr(i \text{ wins} | b, \tau, I_{o_b}, m_{o_b}; o_b)$ can be expressed from the model once the primitives on the demand side are recovered. Notice, that both the distribution of bids and the inverse bid function characterize the bids submitted by permanent sellers of type (x, q) in an auction of type o_b . Further, both of these objects depend on the competitive conditions as summarized by the number of sellers bidding for the projects of a given type, I_{o_b} , as well as the number of projects of this type available for bid in a given week, m_{o_b} . However, the distribution of project costs does not depend on either of these objects. Thus, the estimate of a distribution can be obtained on the basis of any of the (o_b, I_{o_b}) values or by averaging over all the values observed in the data.

Recovering the Distribution of Participation Costs. We make use of the model's predictions concerning sellers' participation choices in order to recover the distribution of participation costs. To accommodate the specifics of our environment we allow that seller i 's costs for entering various types of auctions in a given time period may be correlated. Specifically, the cost of participating in an auction of type o_b in period t is given by $\nu_{i,t,o_b} = \nu_{i,t,0} + \tilde{\nu}_{i,t,l}$; components $\nu_{i,t,0}$, and $\{\tilde{\nu}_{i,t,1}, \dots, \tilde{\nu}_{i,t,O_b}\}$ are independent from each other and across sellers and periods and distributed according to the exponential distributions with parameters λ_0 and $\{\tilde{\lambda}_1, \dots, \tilde{\lambda}_{O_b}\}$ respectively.²⁰

Heuristically, the joint distribution of $\{\nu_{i,t,0} + \tilde{\nu}_{i,t,1}, \nu_{i,t,0} + \tilde{\nu}_{i,t,2}, \dots, \nu_{i,t,0} + \tilde{\nu}_{i,t,O_b}\}$ is nonparametrically identified through variation in the realized market structure, $\bar{I}_{N,t}$ and $\bar{J}_{M,t}$, across time periods. Indeed, the probability of seller with $\tau = (p, x, q)$ entering an auction of type o_b is given by

$$\Pr(i \in N_{o_b} | \theta, I_N = \bar{I}_{N,t}, J_M = \bar{J}_{M,t}; o_b) = \Pr(\nu_{i,t,0} + \tilde{\nu}_{i,t,o_b} \leq \bar{\pi}_\tau(\bar{I}_{N,t}, \bar{J}_{M,t}; o_b), \\ \tilde{\nu}_{i,t,o'_b} - \tilde{\nu}_{i,t,o_b} \geq \bar{\pi}_\tau(\bar{I}_{N,t}, \bar{J}_{M,t}; o'_b) - \bar{\pi}_\tau(\bar{I}_{N,t}, \bar{J}_{M,t}; o_b)).$$

Thus, the variation in expected profits induced by variation in $\bar{I}_{N,t}$ and $\bar{J}_{M,t}$ traces out the joint distribution of entry costs. This in turn implies that marginal distributions of independent components $\{\nu_{i,t,0}, \tilde{\nu}_{i,t,1}, \tilde{\nu}_{i,t,2}, \dots, \tilde{\nu}_{i,t,O_b}\}$ are also identified.²¹

Let us use $\bar{\pi}_\tau(o_b)$ to denote $\bar{\pi}_\tau(\bar{I}_{N,t}, \bar{J}_{M,t}; o_b)$, and $\Delta_{l,o_b} \bar{\pi}_\tau$ to denote $\bar{\pi}_\tau(\bar{I}_{N,t}, \bar{J}_{M,t}; l) - \bar{\pi}_\tau(\bar{I}_{N,t}, \bar{J}_{M,t}; o_b)$ for brevity. Then, under the assumption that components $\nu_{i,t,0}$ and $\tilde{\nu}_{i,t,o_b}$ are exponentially dis-

²⁰Such entry cost structure is similar to the one used in Berry (1992) and Jia (2008).

²¹This is implied by well-known results from the literature on competing risks such as Kotlarski theorem (see, for example, Rao (1992)).

tributed the equation above can be re-written as below.

$$\Pr(i \in N_{o_b} | \tau, \bar{I}_{N,t}, \bar{J}_{M,t}; o_b) = \frac{-\lambda_0 \tilde{\lambda}_{o_b}}{\sum_{l=1}^{O_b} \tilde{\lambda}_l} \exp\left(-\sum_{l=1}^{O_b} \tilde{\lambda}_l \Delta_{l,o_b} \bar{\pi}_\tau\right) \left(\exp\left(-\sum_{l=1}^{O_b} \tilde{\lambda}_l \bar{\pi}_\tau(o_b)\right) \times \right. \quad (5)$$

$$\left. \frac{\exp\left(\left(\sum_{l=1}^{O_b} \tilde{\lambda}_l - \lambda_0\right) \bar{\pi}_\tau(o_b)\right) - 1}{\sum_{l=1}^{O_b} \tilde{\lambda}_l - \lambda_0} - \exp\left(-\sum_{l=1}^{O_b} \tilde{\lambda}_l \max(0, \max_l(-\Delta_{l,o_b} \bar{\pi}_\tau))\right) \frac{\exp(-\lambda_0 \bar{\pi}_\tau(o_b)) - 1}{\lambda_0} \right).$$

The probability on the left-hand-side of this equation can be computed from the data whereas the expression on the right-hand side could be obtained using the model, the distribution of project costs, the inverse bid functions, and the demand-side primitives recovered as explained above. We thus estimate parameters $\lambda_0, \tilde{\lambda}_1, \tilde{\lambda}_2, \dots, \tilde{\lambda}_{O_b}$ by minimizing the distance between the left-hand side and right-hand side expressions in (5) for different seller groups and for several values of (I_N, J_M) .

6 Estimation Results

This section presents the results of estimation. We begin by summarizing the quality group structure recovered in the classification step. Next we discuss the estimated quality levels associated with different quality groups, the estimated distributions of buyers' outside options and willingness to pay for quality, as well as the estimated distributions of sellers' costs. We conclude with the discussion of the fit to the data.

6.1 Estimation of Quality Group Structure

The quality group structure is recovered for the population of permanent sellers characterized by a vector of observable attributes x , $S(x)$, for all $x \in \bar{x}_1, \dots, \bar{x}_P$.

Implementation Details. We focus on the set of projects associated with graphics-related programming (media, computer games, computer-generated animation) of medium (\$400-\$600) and medium-large (\$600-\$800) size with duration of 10-20 days.

We define a seller to be permanent if he stayed with the platform for more than six months. In contrast, a seller who left the market in less than six months is considered to be transitory. To minimize ambiguity of this definition we discard the first three years of the market operation. Majority of permanent sellers in our data are past their six months cut-off in this dataset.²²

In our preferred specification we recover permanent sellers' quality groups structure conditional on the seller's country group and the long-run average of his reputation scores. Sellers are divided into three groups: average score less than 9.7 (low score), average score above 9.7 and below 9.9 (medium score), and average score above 9.9 (high score). This definition allocates approximately 30%, 30%, and 40% across the three seller groups. We have also considered specifications where the vector of x attributes also included the number of arbitrations (documented conflicts between the sellers and buyer) as well as the number of low scores (below seven out

²²Those permanent sellers who reach the cut-off during the remaining three years constitute less than 5% of the total number of permanent sellers. We treat them as transitory during the first six months of their tenure and as permanent after that.

of ten). The estimation results for such specifications are very similar to those reported in the paper even if less precise.

The classification index for a pair of sellers is constructed on the basis of projects for which both sellers belong to the set of potential bidders. We consider a seller i to be a potential bidder for project l if he is active during the week of auction for project l and has submitted bids for projects related to graphics programming in the past.

We implement classification procedure using only the sample of the US-based buyers. This is because the classification procedure requires that a large number of observations for each pair of permanent sellers should be available for projects from a given buyer country and with a specific vector of project characteristics. This condition is not satisfied for other buyer country groups considered in this paper. Thus we estimate classification from the sample of US buyers and then impose that buyers from other countries agree with this classification. We demonstrate in the results section that this approach allows us to achieve good fit to the data.

The Results of Classification Analysis. Classification of permanent sellers from the sample of US-based projects have been implemented in Krasnokutskaya, Song, and Tang (2014a). We summarize the main finding here. The reader is referred to Krasnokutskaya, Song, and Tang (2014a) for details and the summary of the robustness checks performed by the authors.

The results of classification step for the sellers from North America, Eastern Europe and South and East Asia country groups are presented in Table 8. The table reports for every seller-country and reputation score bin the total number of permanent sellers with these characteristics who participated in the projects with graphics-related programming as a main task, the estimated number of the distinct quality groups as well as the number of permanent sellers by quality group and the number of sellers in the corresponding confidence set.

The number of the distinct quality levels conditional on observable seller characteristics tend to be larger than one for all country groups and all levels of the reputation scores. The later holds even for the high and medium reputation scores which correspond to very narrow intervals of possible scores. Further, the confidence sets for the estimated quality groups are rather small and thus group structure estimates are fairly precise.

The classification procedure does not assign a quality level to the estimated groupings of sellers. This limits the comparison between the recovered quality groups associated with different values of observable seller characteristics. Such comparison is deferred to the next section. At this point one can only note that the quality structures appear to vary both across seller country groups and across performance indicators within the country group. This is manifested, for example, in the variation of the estimated number of groups and in the allocation of mass across the recovered quality groups for different values of sellers' observed characteristics.

6.2 Results of Estimation: Demand Side Parameters

We begin by commenting on our choice of parameterization of model's components.

Parameterization. For the purpose of estimation we divide buyer's net valuation for a specific seller by the quality coefficient α . This obtains a specification which similar to the one often used in the estimation of differentiated product models:

$$\begin{aligned} u_{l,j} &= \bar{q}_{\kappa(j)}(x_j) + x_j\beta - \tilde{\alpha}_l b_{l,j} + \tilde{\epsilon}_{l,j} \text{ or} \\ u_{l,j} &= \tilde{\bar{q}}_{\kappa(j)}(x_j) - \tilde{\alpha}_l b_{l,j} + \tilde{\epsilon}_{l,j}. \end{aligned}$$

Table 4: Estimated Quality Groups by Supplier Covariates

Country Group	Average Score	Total Number of Suppliers	$Q = L$	$Q = M$	$Q = H$
North America	low	12	4 (6)	8 (10)	
North America	medium	13	4 (6)	9 (11)	
North America	high	17	12 (13)	5 (6)	
Eastern Europe	low	18	6 (8)	12 (14)	
Eastern Europe	medium	52	33 (37)	12 (14)	7 (9)
Eastern Europe	high	83	6 (7)	65 (69)	12 (15)
East Asia	low	91	62 (68)	18 (22)	11 (13)
East Asia	medium	66	6 (8)	53 (57)	7 (9)
East Asia	high	58	50 (53)	8 (11)	

This table is reproduced from Krasnokutskaya, Song, and Tang (2014a). It shows the estimated group structure and a consistently selected number of groups for each cell determined by covariate values. Column 3 indicates the total number of the suppliers in the cell. Columns 4-6 report the size of the estimated quality group. The size of the corresponding confidence set with 90% coverage is reported in parenthesis. Note that the confidence set with the level $(1 - \alpha)$ for a given quality group is defined to be a random set whose probability of containing this quality group is ensured to be asymptotically bounded from below by $(1-\alpha)$.

Here $\kappa(j)$ denotes quality group of bidder j ; $\bar{q}_{\kappa(j)}$ denotes quality level associated with this group. Additionally, $\tilde{q}_{\kappa(j)} = \bar{q}_{\kappa(j)} + x_j\beta$, $\tilde{\alpha} = \frac{1}{\alpha}$, and $\tilde{\epsilon} = \frac{\epsilon}{\alpha}$.²³

We assume that buyer-seller specific component follows the extreme value distribution with the standard deviation σ_1 . Further, the buyer-specific outside option is represented as $U_l = \gamma_0 + \epsilon_{0,l}$ where γ_0 is (in some specifications) buyer-country specific constant and $\epsilon_{0,l}$ is distributed according to the extreme value distribution with the standard deviation σ_2 . We further assume that the buyer-specific price sensitivity is distributed according to a normal distribution with (in some specifications) buyer-country specific mean and the standard deviation σ_α .²⁴ Since we estimate the standard deviation of the error term we normalize the price sensitivity of the North American buyers to be equal to one. We also estimate the mean of the outside option and that

²³We could be worried about such re-parameterization in the case when zero belongs to the support of α . However, this would only mean that infinity belongs to the supports of $\tilde{\alpha}$ and $\tilde{\epsilon} = \frac{\epsilon}{\alpha}$, the case that can be easily accommodated.

²⁴Strictly speaking, the distribution of α should have been chosen to have a non-negative support. However, we estimate the standard error of this distribution to be quite small so that this assumption does not make any practical difference. The same comment applies to our assumption on the distribution of bids below.

is why normalize one of the quality levels (quality level 1 of the low average score group, the South and East Asian country group) to be equal to zero.

Since the majority of transitory sellers complete only one or two projects their long-run average reputation scores are not observed in the data. We assume that buyers use public information to form beliefs about the probability that a beginning seller with a given number and sum of scores belongs to a particular long-run average score group. We recover these beliefs non-parametrically using beginning of career and long-run data on permanent bidders.

We assume that transitory and permanent sellers' bid distributions are well approximated by normal distributions $N(\mu_{B^t}, \sigma_{B^t}^2)$ and $N(\mu_{B^p}, \sigma_{B^p}^2)$,²⁵ respectively. The means of the bid distribution depend on the buyer's country group, the seller's quality, country group, and long-run average reputation score group (or the number and average of reputation scores for transitory sellers), and on the number of potential permanent competitors by group and the numbers of projects by buyer country. Similarly, we approximate permanent and transitory bidders' respective probabilities of participation by normal cumulative distribution functions that depend on linear indices of the buyer's country group, seller's quality, country group, long-run average score (or the number and average of reputation scores for transitory sellers), the numbers of potential competitors by group and the numbers of projects by buyer country.

Results of Estimation. Tables 5 and 6 report the estimated parameters for three specifications.²⁶ The first specification assumes that the distribution of buyers' price sensitivities and the distribution of their outside options is the same across all buyer countries. The second specification allows the mean of the distribution of the price sensitivities to vary across buyer countries while maintaining that the mean of the outside option is constant. The third specification allows both the means of the distributions of the price sensitivities and the means of the distributions of outside options to vary across buyer countries. Notice, that all three specification assume that buyers from a specific country group may be willing to pay more or less for quality but their willingness to pay does not depend on seller's country group.

Our results confirm importance of the unobserved quality component highlighted in Krasnokutskaya, Song, and Tang (2014a). The estimates for quality levels indicate that two or three distinct quality levels are associated with each value of sellers' observable characteristics and the distributions of qualities vary with seller's observable characteristics. Buyers are willing to pay a substantial premium for the improvement in seller's quality: an average buyer would be prepared to pay a premium equal to 60% of the project size to move from the lowest to the highest quality level of the medium score Eastern European seller. Same buyer would pay about 45% of the project size to move from the lowest to the medium quality level of a medium score Eastern European seller. The estimated standard errors of stochastic components appear to be of a moderate size which indicates that buyers' choices are driven by the variation in observable covariates and quality to the important degree.

The estimates have reasonable magnitudes and are broadly consistent across specifications. The specifications that allow for observed heterogeneity in buyers' utility coefficients and outside options indicate that non-trivial differences exist across buyer countries. Thus, buyers from UK, Eastern Europe, Australia and South and East Asia appear to be more price sensitive than buyers from North America, or Western Europe. Similarly, Southern European buyers and the buyers from the United Kingdom appear to have the lowest outside option whereas South and

²⁵See the comment for the distribution of α above.

²⁶The estimates for the auxiliary objects such the distribution of bids and participation frequencies for various buyer-seller country groups pairs are available from authors upon request.

Table 5: Sellers' Quality Levels

Country	Score	Quality	Specifications		
			(I)	(II)	(III)
North America	Low	1	0.173 (0.301)	-0.18 (0.31)	0.061 (0.302)
North America	Low	2	0.711*** (0.030)	0.418*** (0.031)	0.448*** (0.033)
North America	Medium	1	0.311 (0.517)	0.116 (0.221)	0.123 (0.221)
North America	Medium	2	0.729*** (0.152)	0.515*** (0.101)	0.554*** (0.111)
North America	High	1	0.124*** (0.051)	0.143*** (0.043)	0.127*** (0.051)
North America	High	2	0.906*** (0.211)	0.519*** (0.204)	0.636*** (0.211)
Eastern Europe	Low	1	-0.121 (0.051**)	-0.097*** (0.041)	-0.011** (0.052)
Eastern Europe	Low	2	0.899*** (0.133)	0.405*** (0.131)	0.505*** (0.123)
Eastern Europe	Medium	1	0.321*** (0.101)	0.162 (0.102)	0.117 (0.101)
Eastern Europe	Medium	2	0.716*** (0.022)	0.708*** (0.021)	0.523*** (0.021)
Eastern Europe	Medium	3	0.804*** (0.061)	0.757*** (0.063)	0.707*** (0.066)
Eastern Europe	High	1	-0.240*** (0.111)	-0.063 (0.090)	-0.012 (0.071*)
Eastern Europe	High	2	0.733*** (0.011)	0.601*** (0.010)	0.627*** (0.021)
Eastern Europe	High	3	0.953*** (0.121)	0.882*** (0.113)	0.845*** (0.131)
South and East Asia	Low	1	normalized to 0		
South and East Asia	Low	2	0.566*** (0.211)	0.235** (0.132)	0.353*** (0.142)
South and East Asia	Low	3	0.807*** (0.215)	0.572*** (0.101)	0.674*** (0.102)
South and East Asia	Medium	1	-0.310 (0.221)	0.041 (0.211)	-0.054 (0.321)
South and East Asia	Medium	2	0.569*** (0.021)	0.235*** (0.032)	0.259*** (0.024)
South and East Asia	Medium	3	0.765*** (0.032)	0.675*** (0.043)	0.723*** (0.033)
South and East Asia	High	2	0.893*** (0.171)	0.482*** (0.144)	0.307*** (0.123)
South and East Asia	High	3	0.972*** (0.031)	0.681*** (0.033)	0.618*** (0.031)

This table reports the estimates of quality levels for various quality groups that correspond to the three specifications we consider. The estimates are based on a sample of 11,170 projects.

East Asian buyers appear to have the highest outside option.

The findings on outside options are consistent with statistics on the rates charged by soft-

Table 6: Parameters of the Distribution of Buyers' Tastes

Parameters	Specifications		
	(I)	(II)	(III)
Standard Deviations of Unobservables:			
$\log(\sigma_\epsilon)$	-0.458*** (0.021)	-0.293*** (0.023)	-0.496*** (0.021)
$\log(\sigma_{v_0})$	-0.221*** (0.101)	-0.245*** (0.100)	-0.336*** (0.111)
$\log(\sigma_\alpha)$	-0.490*** (0.010)	-0.368*** (0.011)	-1.310*** (0.010)
Mean of Price Sensitivity:			
Constant	1	1	1
United Kingdom		0.188*** (0.076)	0.177*** (0.071)
Western Europe		-0.017 (0.021)	-0.229*** (0.043)
Southern Europe		-0.220* (0.121)	-0.176 (0.132)
Eastern Europe		-0.090*** (0.112)	0.212** (0.111)
Australia		0.120*** (0.030)	0.161*** (0.044)
South and East Asia		0.307*** (0.021)	0.299*** (0.025)
Mean of an Outside Option:			
Constant	-1.221*** (0.322)	-1.264*** (0.432)	-1.66*** (0.443)
United Kingdom			-0.992*** (0.341)
Western Europe			-0.519*** (0.213)
Southern Europe			-1.745 (1.261)
Eastern Europe			0.576*** (0.221)
Australia			0.321*** (0.101)
South and East Asia			0.771*** (0.042)

This table reports the estimated parameters of the distributions of buyers' utility coefficients and outside options. The estimates are based on the sample of 11,170 projects.

ware developers around the world reported in various media sources²⁷ which indicate that the programmers in US or United Kingdom are paid almost ten times more than the programmers in South and East Asia or in Eastern Europe; with Western Europe and Australia in-between with programmers salaries five times higher than those in Asia or Eastern Europe. Thus, it appears that outside option of Eastern European or Asian buyers should be more similar in value to that

²⁷See, for example, <http://www.bloomberg.com/visual-data/best-and-worst//highest-paid-software-engineers-countries>.

delivered through an on-line market than the outside option of US, UK or Western European buyers.

It is a bit harder to evaluate findings concerning price sensitivity. The data on average household incomes across countries are widely available and they indicate that the individual's purchasing power in US, Australia, UK and Western Europe is much higher than purchasing power of individuals from Southern Europe, South and East Asia or Eastern Europe. It is unclear, however, how well these numbers reflect paying power of the technologically-savvy small businesses in these countries.

Table 7: Distribution of Project Costs

S: Q:	Low 1	Low 2	Low 3	Medium 1	Medium 2	Medium 3	High 1	High 2	High 3
North American Sellers									
μ_C	0.82 (0.056)		0.73 (0.023)	0.81 (0.056)		0.79 (0.26)		0.88 (0.034)	0.88 (0.023)
σ_C	0.12 (0.032)		0.16 (0.043)	0.2 (0.023)		0.2 (0.023)		0.16 (0.034)	0.17 (0.023)
Eastern European Sellers									
μ_C	0.72 (0.044)	0.61 (0.032)		0.87 (0.031)	0.59 (0.025)	0.78 (0.036)	0.94 (0.043)	0.55 (0.033)	0.73 (0.023)
σ_C	0.13 (0.011)	0.17 (0.032)		0.19 (0.027)	0.21 (0.017)	0.15 (0.034)	0.16 (0.033)	0.19 (0.043)	0.17 (0.024)
South and East Asia Sellers									
μ_C	0.71 (0.034)	0.57 (0.023)	0.76 (0.034)	0.81 (0.033)	0.55 (0.023)	0.68 (0.041)		0.57 (0.023)	0.75 (0.032)
σ_C	0.18 (0.021)	0.18 (0.043)	0.14 (0.032)	0.16 (0.023)	0.17 (0.032)	0.18 (0.041)		0.18 (0.021)	0.25 (0.023)

This table reports the estimated means and standard deviations of the distribution of project costs for different groups of sellers. Standard errors are constructed through bootstrap procedure.

6.3 Results of Estimation: Supply Side Parameters

Table 7 reports the estimated means and the standard deviations of the project cost distributions. Our modeling approach maintains that that the costs differ only across seller countries rather than across country pairs. However, this restriction is not imposed in the estimation. In fact, we recover the costs distributions for various buyer-seller country group pairs and then test the equality of the recovered cost distributions for a given seller country group and across various buyer country groups. Testing results indicate that the equality of costs distributions cannot be rejected in most cases. The Asian seller group is an exception where the equality of the cost distribution associated with Asian buyer group and those associated with other buyer groups is borderline rejected (with p-value = 0.13).

The results indicate that North American sellers tend to have higher costs relative to Eastern European and Asian sellers whereas the later two groups have comparable costs. Interestingly, the costs are *U*-shaped in quality, i.e., across different seller country and score groups the lowest

quality sellers appear to have high cost whereas the costs of medium quality is generally lower than the costs of high quality sellers. This illustrates the heterogeneity of sellers who participate in on-line market.

Table 8: Participation Costs

Common Entry Cost Components		Buyer-Country-Specific Deviations	
$\lambda_{0,NA}$	0.195*** (0.031)	λ_{US}	0.501*** (0.031)
$\lambda_{0,EE}$	0.082*** (0.028)	λ_{UK}	0.125*** (0.012)
$\lambda_{0,SEA}$	0.099*** (0.022)	$\lambda_{Western\ Europe}$	0.076*** (0.023)
		$\lambda_{Eastern\ Europe}$	0.033*** (0.014)
		$\lambda_{Australia}$	0.105*** (0.022)
		$\lambda_{SE\ Asia}$	0.024*** (0.010)

This table reports the estimated parameters for the distributions of participation costs associated with different buyer-country groups.

Table 8 reports the estimates for the distribution of entry costs. These estimates correspond to the specification where the distribution of the $\epsilon_{j,0}$ is seller-country-specific whereas the distribution of vector $\tilde{\epsilon}_j = \{\tilde{\epsilon}_{j,1}, \dots, \tilde{\epsilon}_{j,O_b}\}$ remains the same across seller countries. To interpret the results it is useful to keep in mind that the expectation of the exponential distribution is given by the reciprocal of the parameter λ . Therefore, North American sellers generally have lower costs (relative to the sellers from other countries). Further, the entry costs for Western and Eastern European as well as Asian auctions appear to be higher than the cost of entering North American, UK or Australian auctions.

Table 9 reports the average entry costs conditional on participation for various country pairs. The numbers reported in this table reflect equilibrium outcomes and, thus, are not directly informative about any specific primitive. We include them here to provide an idea about the magnitude of entry costs incurred in this market. As table indicates the entry cost constitute about 7% of the project costs on average.

Last, we would like to comment on the limitations of the analysis presented in this section. In this analysis, we take seller's reputation score as given and ignore the possible dynamic considerations associated with reputation building. To mitigate this concern, we base our estimation of the distribution of sellers' costs on the optimization problem of a permanent seller. While permanent sellers may still take reputation-related concerns into account, the incentives associated with these concerns are likely to be quite weak. A single score does not make a large impact on average reputation score once a seller has completed ten or more projects. Indeed, in the data a bad score does not make a statistically significant impact on the probability of winning or on the bid of an established seller.

Table 9: Expected Entry Costs Conditional on Participation

Buyer Country	Sellers		
	North America	Eastern Europe	SE Asia
North America	0.132	0.120	0.099
UK	0.111	0.064	0.039
Western Europe	0.105	0.071	0.031
Eastern Europe	0.130	0.063	0.020
Australia	0.129	0.073	0.046
SE Asia	0.096	0.013	0.019

This table reports the implied entry costs (expected conditional on participation) for different buyer-seller country group pairs.

6.4 Model Fit

Tables 10 and 11 summarize the fit of the model in terms of sellers' participation decisions and project allocations respectively. Both tables are based on the estimates that allow for the distribution of buyers' price sensitivities and outside options to differ across buyer country groups.

Table 10: Participation Fit

Buyer Country	Data			Model		
	Northern America	Eastern Europe	SE Asia	Northern America	Eastern Europe	SE Asia
Shares Conditional on Seller Country						
North America	0.526	0.481	0.48	0.527	0.484	0.476
UK	0.12	0.121	0.117	0.108	0.118	0.129
Western Europe	0.079	0.096	0.089	0.055	0.087	0.109
Eastern Europe	0.026	0.043	0.047	0.035	0.042	0.039
Australia	0.074	0.104	0.107	0.092	0.088	0.104
SE Asia	0.026	0.039	0.078	0.029	0.051	0.066

This table reports the average (across weeks) frequencies with which sellers from different country groups enter auctions conducted by buyers from various buyer country groups. This table differs from table 3 which reports the overall share of bids from a given seller country group submitted to the auction from a given buyer country group.

We summarize the fit of participation component by comparing empirical and implied frequencies with which sellers from different country groups enter auctions conducted by buyers from various buyer country groups in any given week. The results of this analysis is recorded in Table 10. This table demonstrates that predictions of the model reflect behavior in the data quite well. The main discrepancy is associated with North American sellers – the model slightly underpredicts their participation in UK and Western European auctions while overpredicting

their participation in Australian auctions. Additionally, the model overpredicts participation of Eastern European and underpredicts participation of Asian sellers in South and East Asian auctions. It is worth mentioning that the fit of the participation component is quite difficult to achieve since it rationalizes sellers' participation and pricing choices which in turn incorporate sellers' perception about buyers' equilibrium behavior. Nevertheless, the outcomes predicted from the model appear to be quite close to those observed in the data.

The top panel of Table 11 describes allocation patterns generated by the model. The conditional shares of seller-country groups for a given buyer-country group generated by the model are within one or two percentage points of those implied by the data. The largest discrepancy arises in the case of American sellers: our model slightly overpredicts allocation of projects towards these sellers for all country groups except North America.

Table 11: Model Fit

Buyer Country	Data			Model		
	Northern America	Eastern Europe	SE Asia	Northern America	Eastern Europe	SE Asia
Conditional Shares:						
North America	0.093	0.288	0.461	0.092	0.299	0.458
UK	0.074	0.314	0.455	0.078	0.307	0.459
Western Europe	0.05	0.351	0.442	0.061	0.346	0.445
Southern Europe	0.06	0.325	0.457	0.07	0.323	0.456
Eastern Europe	0.04	0.363	0.438	0.052	0.354	0.457
Oceania	0.055	0.295	0.52	0.062	0.305	0.513
South and East Asia	0.043	0.237	0.563	0.048	0.235	0.556
Marginal	0.065	0.295	0.473	0.068	0.298	0.473
Relative to Marginal Distribution:						
North America	1.436	0.975	0.975	1.348	1.003	0.969
UK	1.136	1.065	0.962	1.145	1.031	0.971
Western Europe	0.77	1.188	0.935	0.9	1.16	0.94
Southern Europe	0.927	1.102	0.968	1.036	1.082	0.964
Eastern Europe	0.614	1.231	0.926	0.765	1.189	0.965
Oceania	0.849	1.001	1.099	0.905	1.023	1.084
South and East Asia	0.666	0.803	1.19	0.702	0.789	1.175

This summarizes fit of the allocation patterns implied by the model to the data on the basis of the estimates associated with specification (III) which allows for the distribution of buyers' price sensitivities and outside options to differ across buyer country groups.

The bottom panel of this table investigate the clustering patterns generated by model. It shows the distortion of conditional seller-country shares among the projects allocated by a given buyer country relative to the marginal distribution of projects across seller countries. This exercise confirms that model reproduces clustering pattern observed in the data to a large degree. Specifically, it predicts that North American sellers tend to submit bids to North American auctions at a rate that substantially exceeds the rate implied by the marginal distribution whereas they tend to participate at a rate which is lower than that implied by marginal distribution in the

auctions from other countries. Similarly, North American buyers tend to choose North American sellers at a rate that substantially exceeds the rate implied by the marginal distribution whereas all the other buyer countries tend to substantially underhire them. The model also generates clustering for European buyers and Eastern European sellers as well as for Australian and Asian buyers and South and East Asian sellers.

At the same time the tables indicates that our model do not perfectly capture this clustering. Specifically, it closely follows the clustering pattern for Asian sellers but tends to underpredict distortion in the case of Eastern European sellers and underpredicts distortions in both directions in the case of North American sellers. We quantify the gap between the empirical distortion and the distortion implied by the model using the following index:

$$R = \sqrt{\frac{1}{M_B} \sum_{c_2} \frac{1}{M_S} \sum_{c_1} \left(\frac{s_{c_1|c_2}}{s_{c_1}} - 1 \right)^2} \quad (6)$$

where M_B , and M_S denote the numbers of buyer- and seller-country groups; c_1 index seller groups; c_2 index buyer groups; $s_{c_1|c_2}$ and s_{c_1} represent conditional and marginal frequencies respectively of different seller-country groups. We find that the value of the index which reflects clustering generated by the model captures 80% of the value of the index computed from the data. Thus, our model is capable of generating clustering patterns close to those observed in the data despite the absence of bilateral preferences or costs.

7 Dissecting the Clustering Pattern

Having established that the model is capable of generating the clustering pattern we next decompose clustering into the “demand” and “supply” components. We isolate the impact of sellers participation decisions on the clustering in trade by considering a setting where buyers are presented with a random choice sets rather than a set of sellers who self-select into participation with a given buyer.

Specifically, we re-compute equilibrium outcomes for the auctions in our dataset under random (or non-strategic) participation. We solve model under counterfactual scenario outlined above using an extension of the computational method proposed by Marshall, Meurer, Richard, and Stromquist (1994). Our computational algorithm is summarized in the Appendix. In this exercise we hold the global market conditions – set of potential sellers and the set of projects offered for sale – constant for each time period. To generate a dataset with a random allocation of sellers across projects we proceed in the following way: every potential seller who appears in the market in a given period is allocated to one of the projects offered during this period at random. We compute bidding strategies sellers would use in the environment with such non-strategic participation and use them to simulate bids and auction outcomes under the new participation regime.

Table 12 reports the conditional shares and the ratios of conditional shares to marginal shares generated by the model with non-strategic random participation and compares them to the conditional shares and the ratios generated by the model under the participation strategies used by sellers in the data. The results indicate that the previously observed clustering in trade between North American and UK buyers and North American sellers; European buyers and Eastern European sellers; as well as between Australasian buyers and Asian sellers is diminished.

Table 12: Conditional Market Shares: Clustering Under Non-Strategic Participation

Buyer country	Benchmark Model			Random Participation		
	Northern America	Eastern Europe	SE Asia	Northern America	Eastern Europe	SE Asia
North America	1.348	1.003	0.969	1.127	1.038	0.959
UK	1.145	1.031	0.971	1.139	0.981	0.964
Western Europe	0.9	1.16	0.94	1.078	1.055	0.913
Southern Europe	1.036	1.082	0.964	0.753	0.945	1.078
Eastern Europe	0.765	1.189	0.965	0.595	1.061	1.097
Australasia	0.905	1.023	1.084	0.944	1.094	1.023
South and East Asia	0.702	0.789	1.175	0.591	1	1.061
Marginal	0.068	0.298	0.473	0.079	0.358	0.402

The right-hand-side panel of this table reports the results of simulation analysis which studies clustering under non-strategic participation. The bidding strategies are re-computed to account for non-strategic participation.

In contrast, the conditional allocation in excess of marginal now arises in Western Europe for North American sellers, in Australasian buyers for Eastern European sellers, and in Eastern European/South European buyers for Asian sellers. These patterns indicate that participation decisions are driven by demand in some cases while they work against the demand in other circumstances. In either case participation decisions play an important role in determining the trade patterns in this market. In fact, the difference in clustering of trade between strategic and random participation accounts for 70% of all clustering generated by the model.

8 Analysis of Trade Policies

In this section we investigate the effects of restricting accessibility of international trade and the channels through which such restrictions affect the market.

Specifically, we consider three types of counterfactual scenarios: (1) exclusion of low quality foreign providers, for example, through licensing; (2) general restriction on participation of foreign providers (quota on foreign participation); and (3) a preferential pricing of domestic providers which enables domestic sellers to charge higher prices in equilibrium. Under the last restriction domestic buyers are penalized for purchasing services from foreign sellers unless domestic price exceeds the foreign one by a specified price margin. The policy indirectly promotes reduction in the availability of foreign sellers as we explain later in this section.

In order to implement this analysis we need to solve for the equilibrium bidding and participation strategies in the multi-attribute auction environment under a variety of counterfactual restrictions. The details of the numerical algorithm that we use are summarized in the Appendix. The algorithm is computationally quite costly. Specifically, it takes from one to several hours (depending on the number of groups) to obtain a set of bidding strategies that correspond to a given set of participation strategies. Since we need to achieve convergence both in the participation and bidding components, the computation can take a very long time. Since our objective is not to develop specific policy prescriptions but to investigate the mechanisms through which these policies affect the market, we can simplify the setting to reduce dimensionality of the

problem.

Table 13: Benchmark Solution

Buyer Country	Conditional Shares Seller Country			Relative to Marginal Seller Country		
	North American	Eastern Europe	SE Asia	North American	Eastern Europe	SE Asia
Participation						
North America	0.071	0.316	0.491	1.168	0.996	0.960
Europe	0.061	0.352	0.487	1.000	1.110	0.950
Pacific Region	0.054	0.265	0.538	0.885	0.836	1.090
Marginal	0.061	0.317	0.495			
Allocation						
North America	0.079	0.309	0.475	1.298	0.994	0.965
Europe	0.055	0.372	0.456	0.876	1.196	0.885
Pacific Region	0.048	0.259	0.566	0.754	0.833	1.171
Marginal	0.061	0.311	0.482			

This table summarizes results of simulation for the benchmark model. Here we demonstrate that the properties of interest survive under the aggregation of buyer and seller types which is imposed to ensure computational feasibility of the counterfactual exercise.

Specifically, we first discard auctions held by South European buyers due to their small number and because preferences of these buyers differ significantly from the preferences of buyers from other countries and are thus not amenable to aggregation. Then, we aggregate buyers into three larger country groups: North America (unchanged), Europe (combine UK and Western and Eastern Europe), Pacific region (combine Australasia with South and East Asia).²⁸ Further, we reduce seller heterogeneity by eliminating differences in reputation scores.²⁹ Thus, in the end we have 9 different seller groups – 3 country groups with 3 quality groups per country (with the exception of North America where we have only two quality groups) plus the residual group comprised of all other countries – and three buyer country groups. We additionally ignore the distinction between the permanent and transitory sellers in this analysis. Instead we impose that all sellers are permanent. We have experimented with a specification which allows for a universal (non country-specific) transitory type and have obtained very similar results. In general, allowing for transitory sellers is costly since it requires maintaining additional seller type(s) when solving for bidding strategies.³⁰

We begin by solving a benchmark game which corresponds to the case without trade restrictions. In this analysis we replicate the market structure observed in the data. That is, we impose

²⁸We associate an aggregated group of buyers with the distribution of the tastes for quality and the distribution of the outside options that are equal to the weighted averages of the corresponding distributions for included original groups.

²⁹We replace the quality level of each seller with the average of quality levels of sellers from the same country group who have the same quality rank (1, 2 or 3) but possibly belong to different reputation score group.

³⁰In order to explore a possibility of multiple equilibria in this analysis we solved each game using 100 different starting points. In each case we converged to the same solution. Since we use contraction-style mapping to search for equilibria we conclude that other equilibria if they exist must be of non-stable type.

Table 14: Probabilities of Entry

		Benchmark			Quota		
Seller Country	Q	Buyer Country			Buyer Country		
		North America	Europe	Pacific Region	North America	Europe	Pacific Region
North America	1	0.47	0.13	0.07	0.48	0.10	0.07
	3	0.56	0.22	0.11	0.80	0.19	0.09
Eastern Europe	1	0.48	0.14	0.08	0.32	0.19	0.10
	2	0.69	0.28	0.12	0.62	0.30	0.12
	3	0.64	0.24	0.11	0.59	0.27	0.12
South and East Asia	1	0.53	0.17	0.10	0.42	0.19	0.13
	2	0.74	0.22	0.15	0.65	0.23	0.17
	3	0.73	0.19	0.13	0.68	0.21	0.16

In this table we compare participation probabilities for different seller groups in the benchmark and counterfactual settings.

that the numbers of projects by buyer country and the numbers of potential entrants by seller group for each time period should be as observed in the data. Table 14 reports the distribution of the number of bids (participation behavior) and the allocation of projects across seller countries conditional on buyer country. The table replicates clustering highlighted earlier. That is, the fraction of bids submitted by North American sellers to the projects of North American buyers, and by European sellers to the projects of European buyers, and by Asian sellers to the projects of Pacific region buyers are higher than the corresponding shares implied by the marginal distribution. Similar patterns are present in the distributions of allocated projects conditional on buyer country. Further, table 13 which reports probabilities of participation (that is probability to submit a bid to a given buyer country group) for different seller groups demonstrates that majority of bids for each seller group is submitted to North American buyers. Note that we do not scale the participation by the number of available projects in any way. This pattern arises from the model as a result of the expected competitiveness of various types of auctions generated by the demand and expected supply. Thus, the solution generates all the features of interest to our analysis.

Next, we solve the model under the scenarios that impose restrictions on trade. Specifically, we focus on the case when the restrictions are imposed unilaterally by US (North America) and perform robustness analysis that allows such restrictions to be imposed by all buyer country groups wherever appropriate. The results are summarized in the tables 15 and 16 which report conditional market shares (participation and allocation), and the welfare measures respectively.

Quality Restriction on Participation. First, we consider a setting where restriction on trade excludes foreign sellers with the quality levels below the lowest quality level of North American sellers. Under this scenarios, the lowest quality Eastern European sellers are prevented from participation in North American auctions. Such restriction has low impact on market allocations since the ‘banned’ group is small and holds a modest market share even in the unrestricted setting. Low quality Eastern Europeans reduce their overall participation and shift their bids towards European and Asian auctions, whereas higher quality Eastern European and North American sellers move to North American auctions compensating for the absence of

Table 15: Participation and Allocation

Buyer Country	Participation Seller Country			Allocation Seller Country		
	North America	Eastern Europe	South and East Asia	North America	Eastern Europe	South and East Asia
Benchmark						
North America	0.071	0.316	0.491	0.079	0.309	0.475
Europe	0.061	0.352	0.487	0.055	0.372	0.456
Pacific Region	0.054	0.265	0.538	0.048	0.259	0.566
Quality Regulation						
North America	0.080	0.291	0.525	0.085	0.275	0.494
Europe	0.059	0.372	0.499	0.049	0.391	0.452
Pacific Region	0.05	0.277	0.533	0.45	0.263	0.569
Quota on Participation						
North America	0.101	0.285	0.493	0.135	0.263	0.431
Europe	0.047	0.372	0.514	0.035	0.408	0.464
Pacific Region	0.041	0.278	0.561	0.023	0.261	0.583
Price Preference						
North America	0.104	0.282	0.491	0.121	0.261	0.436
Europe	0.048	0.382	0.509	0.038	0.398	0.466
Pacific Region	0.039	0.262	0.557	0.029	0.265	0.589

This table summarizes participation and allocation across different groups of sellers under several counterfactual scenarios. The numbers reported in the table are simulated averages across sample time periods.

excluded group. The changes in welfare for all market participants except low quality European sellers are rather small.

Quota on Foreign Participation. Next, we consider the case when availability of foreign sellers is restricted without regard to the quality. In this experiment foreign sellers have to obtain a permission in order to submit a bid in North American auction. The number of issued permissions is set to be 30% below the number of foreign bids submitted to North American auctions in unrestricted setting.^{31,32} Under quota restriction the number of bids submitted for North American auctions decreases. Indeed, the participation of foreign sellers would have to decrease unless prices increase. At the same time domestic sellers would adjust their participation so as to maintain higher prices since the increase in their probability of winning is guaranteed by the reduced foreign participation. Thus, domestic participation would not increase sufficiently to offset the reduction in the number of foreign bids. As a result prices in North American auctions will increase. Domestic sellers benefit from this policy since they are able win more projects at higher prices. North American buyers are hurt by the increase in prices. However,

³¹To streamline the analysis we assume that the seller does not get a chance to switch to a different market if his request is denied.

³²The numbers reported in the tables are obtained under the assumption that entry costs are paid when applying for the permission to submit a bid. We have also experimented with settings where only part of the entry cost is paid when applying for permission and the rest is paid if permission is granted. The results are qualitatively similar across such settings.

Table 16: Welfare Effects

Buyer Country	Sellers' Profits (\$)	Buyers' Surplus (\$)	Total Welfare (\$)
Benchmark			
North America	22,292	215,361	237,653
Europe	69,008	53,553	122,561
Pacific Region	89,180	48,456	137,636
Quality Regulation			
North America	23,991	220,525	244,516
$\Delta W/W_{bench}$	(0.08)	(0.02)	(0.03)
Europe	67,231	52,832.5	120,064
$\Delta W/W_{bench}$	(-0.03)	(-0.01)	(-0.02)
Pacific Region	70,640	93,226	163,866
$\Delta W/W_{bench}$	(-0.01)	(-0.01)	(-0.01)
Quota on Participation			
North America	33,159	209,529	246,688
$\Delta W/W_{bench}$	(0.49)	(-0.03)	(0.04)
Europe	65,688	48,887	114,576
$\Delta W/W_{bench}$	(-0.05)	(-0.09)	(-0.07)
Pacific Region	87,952	48,641	136,593
$\Delta W/W_{bench}$	(-0.01)	(0.00)	(-0.01)
Fixed Participation			
North America	36,801	188,827	225,628
$\Delta W/W_{bench}$	(0.65)	(-0.12)	(-0.05)
Price Preference			
North America	31,764	226,429	258,192
$\Delta W/W_{bench}$	(0.42)	(0.051)	(0.09)
Europe	58,659	48,525	107,184
$\Delta W/W_{bench}$	(-0.15)	(-0.09)	(-0.13)
Pacific Region	76,053	49,497	125,550
$\Delta W/W_{bench}$	(-0.15)	(0.02)	(-0.09)
Fixed Participation			
North America	32,381	207,800	240,181
$\Delta W/W_{bench}$	(0.45)	(-0.04)	(0.01)

This table summarizes welfare effects of several counterfactual restrictions on trade in the graphic programming segment of on-line market. The numbers reported in the table are simulated averages per time period.

this effect is mitigated by the change in the mix of sellers participating in North American market. Specifically, participation of foreign sellers shifts towards higher quality levels that are inherently more competitive (they are more attractive to buyers and have lowers). This regularity also holds in the case of North American sellers with higher quality group gaining more in the market share relative to the low quality group. These effects can be observed in table 13 which reports probabilities of participation across seller groups under participation quota and in Table 16 which reports welfare changes resulting from quota restriction both under fixed participation

and when participation is allowed to adjust. As the results indicate under fixed participation the welfare loss to buyers is substantial (-12%) and overall welfare is reduced (-5%). The adjustment in participation lowers negative impact on buyers (to -3% only) and thus leads to overall modest increase in North American welfare (+4%).

We draw several conclusions from these results. First, participation decisions play an important role in this market since they impact the mix of participants in addition to their number. Second, changes in quality mix are able to substantially offset price increases associated with the reduced competitiveness of the market. This is because higher quality sellers are characterized by the costs distributions with relatively low mean and variance. Thus, in addition, to delivering higher utility to buyers their increased participation also brings down the prices in the equilibrium relative to the case with fixed participation. Third, since the domestic supply of providers is somewhat limited relative to the very high volume of services requested on the demand side the welfare impact associated with any policy is dominated by the effect on the demand side of the market.

Domestic Price Preference Policy. An important consequence of quota restriction is a substantial increase in prices brought about by the reduction in the competition at auction level. We consider now restriction on trade (domestic price preference policy) in the form resembling bid preference which has been shown to limit the impact on price from the reduction in competition in the context of auction markets.³³ Under this policy a buyer is penalized (pays a fine equal to $x\%$ of the project size) if he purchases service from a foreign rather than domestic seller unless the domestic provider of equivalent quality ranking is not available or the difference in price between foreign and domestic providers of equivalent quality ranking exceeds $x\%$ of the project size. We choose the size of the price preference so that the foreign participation under this policy is similar to the foreign participation under the quota policy. Specifically, we impose that the number of permission requests in the later case should coincide with the number of bids submitted in the former case. To achieve this we set $x\%$ equal to 9% of project value. The impact of this policy which is well understood in auction literature is somewhat different from the impact of the quota.

Indeed, the price preference permits domestic sellers to increase their price above the levels charged in the environment without preferential treatment. However, this tendency is limited by competitive pressure from own-country and foreign providers. Foreign providers are forced to reduce their prices in order to maintain a reasonable chance of winning. In fact, foreign sellers of lowest quality levels and with high costs realizations are effectively driven out of the market since they are not able to win sufficiently often at any reasonable price. The same regularity reduces effective and actual participation of foreign sellers at all quality levels. As in the case of quota, the quality mix of foreign participants shifts towards higher quality levels. Domestic sellers improve their welfare due to the expansion of their market share. However, their welfare does not reach the level achieved under the quota restriction since they are prevented from increasing their prices to the same level. Domestic buyers pay lower prices relative to the case of quota and they are presented with the improved quality mix relative to the benchmark case. As a result the buyers' surplus and the overall welfare increase. The later by 7%. In this case the difference between the fixed and adjusted participation is less important since price preference induces substantial effective adjustment in participation.

It is worth noting that the burden of price preference policy is born by foreign sellers whose

³³Detailed analysis of this policy can be found in McAfee and McMillan (1987) and Krasnokutskaya and Seim (2011).

profitability is substantially reduced by this policy. Additionally, even if other buyer countries impose similar policy of price preference on the transactions of their buyers they are not able to impact North American welfare to a large degree since foreign participation of North American sellers is very limited.³⁴ This experiment, of course, holds the conditions in other markets fixed. In reality foreign country may retaliate through the policies impacting other markets.

9 Conclusion

This paper makes a two-fold contribution to the literature. First, we develop a tractable model of an on-line market for services which rationalizes sorting of heterogeneous sellers across heterogeneous projects and realized volume of transactions. We specifically focus on the ability of our model to rationalize observed international trade since this market is representative of an increased importance that electronic space plays in facilitating international trade in services. Second, we investigate the impact of policies restricting trade on these markets. Insights of this analysis are novel since it concerns markets that have not been considered previously in the context of international trade policies (specifically, those characterized by private variation in costs with transactions implemented through auction mechanism).

The main conceptual difficulty which arises in the study of on-line markets is that these markets involve a very large number of sellers that are potentially competing with each other for a large number of projects; and yet the competition for a specific project is localized to a relatively small number of sellers. Thus, the model of such market has to be able to explain how project-level competition arises from the market-level conditions. This paper is quite successful in achieving this goal. Specifically, the model we propose explains quite well both sellers participation decisions (where to bid), prices they set as well as eventual buyers' choices. Key insight that allowed us to reduce the complexity of the model while obtaining a good fit to the data is that it is sufficient to model seller's choice of a type of project where to bid rather than a choice between specific individual projects. The analysis takes into account variation in sellers choices associated with seller heterogeneity (including unobserved heterogeneity in quality). This feature is enabled by the methodology developed in Krasnokutskaya, Song, and Tang (2014a).

In our model sellers differ in quality and the distribution of private costs (and therefore prices) while buyers are heterogeneous in their valuations for quality and in their outside options. The fit of the model to the data is supported by non-trivial differences in the distributions of preferences across buyer countries as well as differences in the distribution of qualities across seller countries implied by our estimates. We rely on patterns of winning conditional on relevant observable attributes to infer sellers' residual quality from the data. Quality groupings are recovered from auction data involving US buyers since estimation has to be performed conditional on project characteristics including buyers' country and the numbers of projects available for other buyer countries were not sufficiently large. However, the quality group structure recovered in such a way appears to work quite well in explaining choices made by buyers from other countries. This indicates that the fit of the model to the data is quite robust since the estimation methodology essentially employs an estimates obtained from one part to rationalize another part of the data.

From the policy prospective we find that impact on prices has to be carefully considered in US market where the welfare impact is largely determined by the demand side. To this end the policy of domestic price preference appears to be an effective tool in restricting foreign participation

³⁴The results of this experiment are available from the authors upon request.

while improving domestic welfare. On other hand endogeneity of seller participation should not be ignored as it may play an important role in determining policy outcomes. Specifically, the model which does not allow for participation to adjust would predict that the welfare of North American buyers would decrease as a result of the policy restricting foreign participation since such policy results in higher prices paid by North American buyers and thus leads to the reduction in buyer surplus. However, the model which takes into account participation adjustment would account for the fact that the mix of the participants would shift towards higher quality which mitigates the negative effect of price increase. Such model would predict that overall welfare would in fact modestly improve.

Overall, we believe that this paper delivers some important insights both on the international trade in electronic space and on the trade in service markets that will pave the way for the future research in these areas.

References

- ACKERBERG, D., K. HIRANO, AND Q. SHAHRIAR (2006): “The Buy-it-Now Option, Risk Aversion and Impatience in Empirical Model of eBay Bidding,” Working Paper, University of Michigan.
- ATHEY, S., AND D. NEKIPELOV (2012): “A Structural Model of Sponsored Search Advertizing Auctions,” Working Paper, Stanford University.
- BACKUS, M., AND G. LEWIS (2012): “A Demand System for a Dynamic Auction Market with Directed Search,” Working Paper, Harvard University.
- BAJARI, P., AND A. HORTACSU (2003): “The Winner’s Curse, Reserve Prices, and Endogenous Entry: Empirical Insights from eBay Auctions,” *RAND Journal of Economics*, 34(2), 329–355.
- (2004): “Economic Insights form Internet Auctions: A Survey,” *Journal of Economic Literature*, 42(2), 457–486.
- BERRY, S., J. LEVINSOHN, AND A. PAKES (1995): “Automobile prices in market equilibrium,” *Econometrica*, 63, 841–890.
- (2004): “Differentiated Products Demand Systems from a Combination of Micro and Macro Data: The New Vehicle Market,” *Journal of Political Economy*, 112(1), 68–104.
- BERRY, S. T. (1992): “Estimation of a Model of Entry in the Airline Industry,” *Econometrica*, 60(4), 889–917.
- CABRAL, L., AND A. HORTACSU (2010): “Dynamics of Seller Reputation: Theory and Evidence from eBay,” *Journal of Industrial Economics*, 58(1), 54–78.
- DE LOECKER, J. (2011): “Product Differentiation, Multi-Product Firms and the Impact of Trade Liberalization on Productivity,” *Econometrica*, 79 (5), 1407–1451.
- DECAROLIS, F., M. GOLDMANIS, AND A. PENTA (2014): “Common Agency and Coordinated Bids in Sponsored Search Auctions,” Working Paper, Boston University.

- EATON, J., AND S. KORTUM (2002): “Technology, Geography and Trade,” *Econometrica*, 70 (5), 1741–1779.
- FIELER, A. C., M. ESLAVA, AND D. XU (2015): “Trade, Technology and Input Linkages: Theory and Evidence from Colombia,” Working Paper, Duke University.
- GOLDBERG, P., N. KHANDELWAL, J. D. LOECKER, AND N. PAVCNIK (2015): “Prices, Markups and Trade Reform,” *Econometrica*, Forthcoming.
- GREENSTEIN, S. (1993): “Did Installed Base Give an Incumbent any (Measurable) Advantage in Federal Computer Procurement,” *Rand Journal of Economics*, 24(1), 19–39.
- (1995): “Sole-Sourcing versus Competitive Bidding: U.S. Government Agencies’ Procedural Choices for Mainframe Computer Procurement,” *Journal of Industrial Economics*, XLIII(2), 125–140.
- GUERRE, E., I. PERRIGNE, AND Q. VUONG (2000): “Optimal Nonparametric Estimation of First-Price Auctions,” *Econometrica*, 68(3), 525–574.
- HAGEDORN, M., T. LAW, AND I. MANOVSKII (2012): “Identifying Equilibrium Models of Labor Market Sorting,” Working Paper 18661, National Bureau of Economic Research.
- HENDRICKS, K., AND A. SORENSEN (2014): “Evaluating the Role of Intermediary in On-line Auctions,” Working Paper, University of Wisconsin-Madison.
- JIA, P. (2008): “What Happens when Wal-Mart Comes to Town: An Empirical Analysis of Discount Retail Industry,” *Econometrica*, 76(6), 1263–1316.
- KASAHARA, H., AND K. SHIMOTSU (2009): “Nonparametric Identification of Mixture Models of Dynamic Discrete Choices,” *Econometrica*, 77, 135–175.
- KRASNOKUTSKAYA, E., AND K. SEIM (2011): “Bid Preference Programs and Participation in Highway Procurement,” *American Economic Review*, 101.
- KRASNOKUTSKAYA, E., K. SONG, AND X. TANG (2014a): “The Role of Quality in On-line Service Markets,” Working Paper, University of Pennsylvania.
- (2014b): “Uncovering Unobserved Agent Heterogeneity through Pairwise Comparisons,” Working Paper, University of Pennsylvania.
- LEWIS, G. (2011): “Asymmetric Information, Adverse Selection and Online Disclosure: The Case of eBay Motors,” *American Economic Review*, 101(4), 1535–1546.
- MARSHALL, R. C., M. J. MEURER, J.-F. RICHARD, AND W. STROMQUIST (1994): “Numerical Analysis of Asymmetric First Price Auctions,” *Games and Economic Behavior*, 7(2), 193–220.
- MCADAMS, D. (2003): “Isotone Equilibrium in Games of Incomplete Information,” *Econometrica*, 71(4), 1191–1214.
- MCAFEE, R., AND J. MCMILLAN (1987): “Competition for Agency Contracts,” *The RAND Journal of Economics*, 18, 296–307.

RAO, B. (1992): *Identifiability in Stochastic Models: Characterization of Probability Distributions*. Academic Press, New York.

SEIM, K. (2006): “An Empirical Model of Firm Entry with Endogenous Product-type Choices,” *RAND Journal of Economics*, 37(3), 619–640.