Trading across Borders in Online Auctions *

Elena Krasnokutskaya[†]
Johns Hopkins University

Christian Terwiesch[‡]
Wharton School of Business

Lucia Tiererova[§]
Johns Hopkins University

October 3, 2017

Abstract

We invoke the insights from the auction literature to study 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 who are heterogeneous in both quality and costs across projects offered by buyers who differ in outside option and willingness to pay for quality. To accommodate the possibility of such an outcome we extend a single-auction entry model to a setting where sellers choose among multiple projects. This feature plays an important role in explaining the data and understanding the effects of various trade policies.

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

JEL Classification: D22, D44, D82, F14, L15, L86.

*We would like to thank Susan Athey, Dan Ackerberg, Greg Lewis and Marc Rysman for helpful discussions. We are also grateful to seminar participants at the University of Arizona, Carnegie-Mellon University, the London School of Economics, the 2015 Winter Meeting of Econometrics Society, the 2016 International Industrial Organization Conference, the 2016 Society for Economic Dynamics Annual Meeting, and the 2016 Banf Microeconomics Workshop.

[†]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 service providers. One key aspect of this transformation is that the provision of many services, particularly professional ones, is no longer restricted by the proximity of buyers to providers. Such services become tradable within and across national borders. Hence, the observed rapid 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 a desired service (project) and then running an auction to select a provider. The auction mechanism, usually referred to as a multi-attribute auction, 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.[1,2] The emerging organization of trade in professional services through online auctions suggests that the insights developed in the auction literature could be used to study how these markets function and to develop policies tailored to their design.[3] However, the specific features of these settings and the nature of the available data often differ from those studied in the traditional empirical auction literature and thus require new methodological tools. The objective of this paper is to develop a suitable conceptual framework, validate its applicability through an empirical analysis of trade in a particular service market, and apply it to quantify the impact of several policies traditionally used in auction markets on international trade in services.

A key feature of these markets is that multiple projects are offered for bid simultaneously, hence, sellers must decide which auctions to enter. Participation decisions determine the competitive environment in individual auctions and ultimately determine prices. These decisions themselves are based on the differences in bidders' expected profitability across auctions, which is determined, in part, by the differences in expected competition. In equilibrium, the potential bidders' choices of what auctions to participate in, the resulting competitive environment in each auction, submitted bids, and the project award are functions of the set of potential sellers

[1]Some prominent examples range from 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, to specialized ones such as MediBid for medical, Chegg for tutoring, ProZ for translation, or Envato Studio for graphic design services, among many others.

[2]Multi-attribute auctions are similar to scoring auctions, studied by Che (1993), Asker and Cantillon (2008), and Bajari and Lewis (2011), since both mechanisms allow buyers to take into account seller's attributes other than price at the allocation stage. These mechanisms differ in two important aspects. First, in multi-attribute environments, the seller's characteristics are exogenously given as opposed to being part of the seller's strategic choices as in scoring auctions. Second, the allocation rule in scoring auctions is explicitly specified before bidding begins whereas buyer preferences for the sellers' attributes are not known to sellers in a multi-attribute auction.

[3]Recent analysis of online auctions is represented by Bajari and Hortacsu (2003); Cabral and Hortacsu (2010); Ackerberg, 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).

and also of the set of projects available for bid. Specifically, when the set of projects changes, participation, pricing and other outcome variables change as well. This property would arise even in an environment where all projects are identical and all potential sellers are symmetric. In the settings with heterogeneous projects and sellers, such as the one considered in this paper, sellers' participation and pricing decisions would further reflect competitive advantages individual sellers' might have in relation to different projects. Thus, sorting of sellers across projects might arise.

Traditional models of entry focus on one auction at a time. In these settings, potential sellers simply decide whether to participate in a given auction. Such models are not capable of accommodating the two properties described above. At the same time, failing to correctly account for the relationship between the set of available projects and the competitive auction environment may result in biased estimates of the distribution of sellers' costs. Similarly, counterfactual analysis that does not consider possible re-sorting arising as a consequence of the policy may provide poor guidance to policy makers. That is why we introduce a novel modeling framework characterizing participation decisions in the context of an auction environment where multiple projects are simultaneously available for bid and allow this framework to guide our analysis of the process determining trade in the online market for services.[4] While we propose this model in the context of an online setting, it is likely to have wider applicability since the majority of procurement markets involve simultaneous auctioning of multiple (heterogeneous) projects.

Our analysis is based on data collected over multiple years from a prominent online market for programming services. We observe some 600,000 buyers submitting projects for bidding, as well as approximately 50,000 sellers.[5] For each week, our data contain information on all projects available for bid and the set of all potential sellers participating in the market. For each potential seller, we observe auctions the seller decided to participate in and the bids he submitted. Similarly, for each buyer, we observe the set of all sellers participating in the buyer's auction, sellers' bids, the buyer's choice and the ex-post measures of performance. This rich data structure enables us to disentangle the impact of demand- and supply-side decisions on trade in this market.

The nature and quality of the data from the online marketplace we study are similar to other procurement auction markets, such as the commonly studied highway procurement market. Yet, it has a number of advantages enabling us to make progress towards understanding the role of sorting in the procurement process. In particular, the market operates at high frequency (weekly) and the data contain a large amount of variation in the set of available projects. This is helpful for assessing the presence and consequences of sorting induced by such variation. The

[4] Hendricks and Sorensen (2014) consider the bidders' choice between several auctions in a different setting. Similarly, Tadelis and Zettlemeyer (2015) study an environment where informing asymmetric bidders about the quality of the auctioned item allows bidders to sort to auctions in a way that is welfare enhancing.

[5] The subset of data we use in our formal analysis features the same patterns described here.

amount of heterogeneity across projects in our data is also more limited and easier to condition on. Specifically, from a potential bidder's point of view, projects differ mainly in the type of programming required, country of the buyer, and the size of the project. This is much more tractable than the heterogeneity across highway procurement projects which typically aggregate a wide variety of sub-projects, that may be complementary to other sub-projects the potential bidder is concurrently working on because of, e.g., geographical proximity. Finally, a major concern in the study of highway procurement auctions is the endogeneity of participation with respect to capacity constraints. Such concerns are much more limited in the market we study because projects are typically of quite short duration relative to the frequency at which the market operates and the vast majority are finished significantly before the deadline.

The key challenge to developing an empirically implementable model of entry in online market is the huge number of potential alternatives that each of the very large number of sellers in the market needs to consider. Krasnokutskaya, Song, and Tang (2017a) proposes a method to reduce the dimensionality on the demand side by aggregating sellers into groups on the basis of observable and unobservable characteristics. Here we build on this insight by considering an equilibrium in group-symmetric strategies whereby a seller from a given group chooses a type rather than a specific project to enter. This approach allows us to summarize sellers' profits as a function of sellers' groups and project types and of realized local (project-type-specific) competition.[6] To accurately model participation choices in our market, it is important to also allow for the correlation of private entry costs across project types for the same individual. It is natural to expect such correlation to be present and modeling it proves essential to fit the data.[7]

Under the multi-attribute auction format extensively used in online markets the buyers' willingness to pay for quality and the buyers' outside options are unknown to the sellers. Thus, the buyers' choice sets (the sets of participating sellers) are exogenous conditional on buyers' and projects' 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 employ the estimation approach developed in Krasnokutskaya, Song, and Tang (2017a) to recover the individual sellers' unobserved quality as well as the distribution of buyers' outside options and tastes. With these estimates in hand, we turn to the 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 seller. Estimates of the bidding functions and the cost 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 for a given week. Finally, this allows us

[6]This approach to modeling entry is novel in the literature but builds on existing work by Seim (2006) who studied location choice in the video rental industry.

[7]In this aspect we build on the ideas first developed in Berry (1992).

to recover the distributions of the entry costs for every seller type to rationalize the observed participation behavior.

In our analysis of the model's empirical performance we particularly emphasize the sorting property of our model. Specifically, in our data both buyers and sellers are widely dispersed across countries. Buyers award almost 80 percent of all projects to sellers located in countries outside their own. As in the case of trade in goods, the 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 a 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 again from that preferred by Australian buyers.

We validate the proposed model of participation by showing that it is capable of generating a clustering of trade observed in the data. The estimation results 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 sellers' qualities and costs differ significantly across seller countries. The combination of these differences rationalizes the observed trade patterns in the data to a high degree. It is important to emphasize that sorting is not mechanically hardwired into the model. Specifically, the heterogeneity of sellers in the model is captured by their quality and the distribution of private costs. The distributions of sellers' qualities and costs differ across seller countries but do not depend on the prospective buyer's country. Analogously, buyers differ in their willingness to pay for quality and in their values of outside option, which are not observed by sellers. The distribution of these objects vary across buyers' countries but are not tied to the countries of the 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.

The key to understanding this result is the mutual interdependence of demand and supply side responses. The ability of a buyer to obtain service of a preferred quality level is facilitated by the endogenous self-selection of sellers offering that quality in his auction. 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 have greater access to 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. Indeed, our results indicate that sellers' participation decisions significantly contribute to the observed clustering of trade. In fact, counterfactually shutting down the participation channel in the estimated model eliminates 70% of the clustering.

Further, relative to the large literature on entry in non-auction markets, an attractive feature of our data is that we are able to tie sellers' participation decisions directly to their ex-post profitability since we observe prices submitted by sellers as well as subsequent allocations (buyers' choices). Hence, our framework imposes tight discipline on the participation component of the model.[8] We find that the model not only fits the data well but it does so under a plausible set of estimated parameters despite the tight restrictions imposed on the model by the data. This is indicative of the suitability of our modeling framework. Finally, even though the sorting in our model is centered around the buyers' and sellers' countries, the framework can be easily extended to allow for other dimensions of heterogeneity, e.g. size of the project.

It is worth noting that these empirical findings also contribute to the literature analyzing international trade. The focus of trade literature, represented by Eaton and Kortum (2002) or Melitz (2003) models, tends to be on explaining aggregate (typically manufacturing) trade flows which makes it necessary to abstract away from many features of individual markets. Our paper is less ambitious in scope because we focus on one particular service market but the payoff is that we are able to model the key economic mechanisms of that market's operation in much greater detail. Specifically, we are able to highlight sellers' heterogeneity conditional on costs and to establish the quality dimension as an important determinant of trade in online markets. The novel implications of these features in combination with endogeneity of participation resulting in sorting of coders across auctions conducted by buyers from different country groups are the findings that we hope will be of interest to scholars using the international trade framework.[9] These features motivate our choice of counterfactual experiments.

Specifically, we consider the potential impact of several regulatory policies affecting seller participation. While the regulation of online markets is perhaps not yet at the forefront of the policy debate, it may soon be there given the fast growing importance of these markets. We also believe that an understanding of the mechanisms underlying market responses to various interventions may inform further market design improvements. More generally, however, this analysis highlights the crucial role of resorting as a potential consequence of the policy. This insight applies not only to the online markets that are the focus of this paper but also other procurement markets which typically run simultaneous auctions. Further, this analysis also contributes to the international trade literature. Indeed, the traditional trade analysis (such as that summarized in Eaton and Kortum (2002)) does not permit resorting of sellers in response to policy since

[8]The common modeling approach in non-auction literature is to postulate an ad hoc profit function and to choose its parameters to rationalize observed entry patterns. Such an approach has many degrees of freedom as the profit function is chosen to rationalize entry but is not compared to actual ex post profits from entry (which are often not observed in data). In contrast, our hands are tied because we use a full model of pricing and allocation subsequent to participation and we require both of these decisions predicted by the model to be consistent with direct observations in the data.

[9]The focus of the trade literature has also been recently directed at understanding the role of producers' quality in determining flows of trade between countries. Among others Fieler, Eslava, and Xu (2015) and Balat, Brambilla, and Sasaki (2016) emphasize the role of quality in explaining observed patterns of trade in products markets.

producers are essentially assumed to be homogeneous conditional on costs.

Auction literature usually analyzes two types of policies: (a) 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); (b) policies that primarily impact pricing in the market (e.g., the choice of the allocation mechanism such as a first price or a second price auction, or even less standard mechanisms such as a multi-attribute auction or a 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.

To illustrate the impact of trade policies directly affecting participation, consider first a fairly blunt policy such as imposing a quota on the 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 from the government and only a fraction of foreign sellers are allowed to proceed with a bid. If the restriction is imposed irrespective of the quality of the foreign sellers, it has a non-trivial effect since it tends to exclude attractive sellers as well as the unattractive ones. This policy clearly reduces the competitiveness of the market and results in higher prices. 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 participants 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 expands to a greater degree. This change in quality composition mitigates the adverse effect of higher prices on domestic buyers. This is important for a country such as the US which is a net importer of services in this market and is thus significantly affected by the impact of policies on buyers. On the whole, we 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. Notice that the traditional trade framework would predict that consumer welfare should unambiguously decline, as a result of such policy, since consumers do not have access to the lowest possible price some of the time. This conclusion would arise because the traditional trade approach does not allow for systematic differences among producers conditional on costs.

A more refined implementation of this policy may condition the quota on the quality of foreign sellers. Indeed, an important concern of the government in many service markets is the quality of providers. A typical policy response involves the licensing of service providers. Such a requirement is absent in the market we study but can be easily implemented, e.g., by requiring a certain minimum performance on a qualification test. We consider the effects of such a policy that imposes a lower bound on the 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 to result in a very small impact on US market participants because it eliminates the portion of the quality distribution that is least attractive to the domestic buyers. The profitability of attractive (high quality) sellers is only slightly affected since only a small set of weak competitors is eliminated while their costs and the probability of being considered remain unchanged. Thus, the high quality sellers shift their participation towards the US market only slightly and the resulting price changes are also quite small.

Finally, we consider a trade policy that directly affects pricing in favor of domestic sellers and lets participation adjust accordingly. A version of this 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 when 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 the 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 market share. This, in turn, prevents domestic sellers from significantly increasing their prices. As a result, the average price in the market changes little. Thus, the negative price consequences of this policy on domestic buyers are minimal. On the contrary, domestic buyers benefit because of the change in the composition of foreign sellers towards higher quality (because they have low costs and still find this market attractive despite the discriminatory policy in place). So the total effect on buyers is positive. Overall, the price preference policy is able to generate a large domestic welfare gain despite restricting the access to international trade.

The paper is organized as follows. Section 2 describes the market; Section 3 provides summary statistics and trends observed in the data. 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 analyzes potential impact of trade policies on this market. Section 9 offers conclusions.

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. Interested sellers can respond by submitting a quote for the price (bid) at which they would be willing to complete the task. The market serves buyers from countries around the world by providing 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. As a result, the selected seller is not necessarily the one who submits the lowest bid. An important feature of this allocation mechanism is that the award rule is not announced and thus remains unknown to other market participants.

The platform maintains a registry of participating sellers which 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, which reflect buyers' numerical feedback about working with a given seller,[10] 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 quotes, with an average of three messages exchanged between a seller and a buyer prior to submitting a bid. Sellers can also attach an example of previous work or a sketch of the proposed code. Hence, a buyer has the opportunity to form an opinion about each bidder's quality. The content of 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 competing bids.

3 Descriptive Analysis

We have access to the first six years of the operation of this online 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 originating from many different countries every week. It is also a highly specialized area so that it does not attract providers actively participating in auctions for other types of programming. We restrict our analysis to years four through six of the market operation to avoid concerns about market stability and participants' lack of experience. Finally, we focus on medium to medium-large projects offered in this market because the small projects segment (below $200) and the very large project segment (above $1,000) are characterized by qualitatively different patterns of participation and bidding.[11] Our final dataset includes 49,334 projects with bids from 10,213 distinct sellers.

3.1 General Summary Statistics

Table 1 provides a general summary of the data. It shows that an average project in our dataset is worth 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).

[10]Reputation score is a number between 1 and 10.

[11]The very small projects tend to attract a very large number of providers (20-50 per auction) who participate in the market just a few times. The very large projects appear infrequently. They constitute less than 0.001% of the projects offered in this market. They tend to be longer and may potentially impose large capacity requirements on providers.

Most buyers appear in the dataset only once. In an average week 465 projects are auctioned in this market with 1,366 unique sellers expressing interest in participation[12](in a median week the number of projects is 607; and the number of potential sellers is close to 1,500). The majority of projects are allocated to one of the bidders. On average only 3% of buyers in this market choose not to allocate the project after completing an auction.

A large fraction of sellers appear in the market only a few times; the median time a seller stays in the market is one week. In our analysis, we distinguish between the transitory sellers (those who stayed in the market less then 6 months) and permanent sellers (who stayed in the market longer than six months). The majority of permanent sellers have tenure of more than two years and many stay in the market in excess of three years. As we document in the table, 26% of bids in an average week are submitted by permanent sellers (in a median week this share is equal to 21%); permanent sellers win 31% of projects in an average week (in the median week this share is 23%). In our dataset an average permanent seller submits over 300 bids during his tenure (the 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.

The multi-attribute feature of the 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 a 32% premium (the median is 29%) 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 equal to one if seller i wins auction l and zero otherwise; $X_{i,l}$ are seller i's performance characteristics and his experience; $B_{i,l}$ is the bid submitted by seller i; $\mu_{c(i)}$ is a seller country's fixed effect, estimates a statistically significant positive coefficient in front of price[13] which indicates that sellers submitting higher prices 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 a methodology developed in Krasnokutskaya, Song, and Tang (2017a) to address this issue in the estimation.

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 the UK, Australia and Canada, and a large number of countries each reponsible for 1% or 2% of the

[12]We define a seller as expressing interest if he inquires about the details of the project.

[13]The estimated coefficient on price is equal to 1.82 with a 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 the lowest bidder	0.18	0.03	0.17	0.21	0.23
winning bids (% of the 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
share of projects allocated to permanent sellers (per week)	0.31	0.07	0.15	0.23	0.39
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.

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 submit between 1% to 3% each. Further, a significant amount of trade in this market occurs 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 the UK, Germany or France are allocated to the sellers from their home countries, and so on.

In the interest of tractability as well as precision of subsequent analysis we focus on the patterns of trade between groups of countries. Countries are grouped by geographic proximity, similarity of language, and economic conditions. The demand side of the market is represented by the following seven groups: North America, the United Kingdom, Western Europe, Southern Europe, Eastern Europe, Australasia, and South and East Asia.[14] Similarly, the supply side

[14]North America combines USA and Canada; Western European group includes Austria, Belgium, Denmark, Finland, Germany, Iceland, Ireland, the 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.

Table 2: Return on 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 seller country group for the projects from a specific buyer country group.

focuses on seven groups of seller countries: North America, Latin America, Western Europe, Eastern Europe, Africa and the Middle East, South and East Asia, and Australasia.[15] In the remainder of the paper we interchangeably use 'buyer country' and 'buyer country group' to refer to these groupings. The 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 play a 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.

Next, we demonstrate 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 the 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, while in contrast only 14% of North American bids result in winning when submitted to projects auctioned by 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% win 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% win 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 seller countries conditional

[15] On 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.

Table 3: Patterns in Participation and Allocation

Buyer Country	Allocation: Seller's Country: North America	Allocation: Seller's Country: Eastern Europe	Allocation: Seller's Country: South and East Asia	Participation: Seller's Country: North America	Participation: Seller's Country: Eastern Europe	Participation: Seller's Country: 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 project allocation and seller 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 projects awarded across seller country groups.

on buyer country to the marginal distribution of the number of allocated projects across seller countries. Specifically, we are looking at the ratio of the 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 online 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 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 from all buyer groups just because the number of Asian sellers far exceeds those from any other country group. To isolate the clustering pattern, we study deviation of conditional shares (conditional on a buyer country) from the marginal shares. The results of this analysis are summarized in the left panel of Table 3.[16]

If trade were proportional to 'size' we would expect the ratios in Table 3 to be close to 1. Instead, we find substantial distortions: for example, the share of North American projects 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

[16]This calculation underlines a standard test used for detecting clustering patterns in statistical literature. The appropriate X^2 test statistics is given in Equation (9) in Section 6.4.

Asian buyers is 33% lower than the share of North American sellers in the 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%.[17]

Next, we investigate participation patterns in this market. We say that seller i participates in auction l if he submits a bid in this auction. Table 3 indicates that participation in this market is also characterized by a clustering pattern. This can be deduced from the right panel 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. The clustering pattern captured in this panel is similar to the one observed in the allocation of projects.

This regularity is perhaps not surprising since, in this market, a buyer is constrained to allocate the project to one of the sellers who endogenously chose to participate in his auction inducing a correlation between seller participation and buyer allocation. The correlation between seller participation and buyer allocation is further reinforced by enhanced price competition since sellers of similar quality participate in the same auction and thus cannot leverage quality differentials to mitigate price competition. A model rationalizing the clustering of trade in this market hence has to take strategic participation into account.

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 the 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 the buyer's country of origin, $o_b(l)$ with $o_b = \{1, 2, ..., O_b\}$, and the size, z_l. We denote the set of projects offered in the market by M and summarize this set by a vector, $J_M = (m_1, ..., m_{O_b})$ reflecting the number of projects across buyers' countries.

Each seller j in the population 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 $\mathcal{Q}(x) = \{\bar{q}_1(x), \cdot, \cdot, \cdot, \bar{q}_{K(x)}(x)\}$. Notice that distribution of quality indexes 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 sellers and 't' denotes transitory ones. A permanent seller's quality is known to all market participants; a transi-

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

tory 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. A transitory seller reveals his quality draw to a buyer when submitting a bid. Here we follow the convention of using capital letters to denote random variables (such as quality of a transitory seller, Q) and lower-case letters to denote the realizations of random variables (such as $q_k(x)$).

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 is available to all potential bidders and is summarized by a vector I_N, which reports the number 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.

Seller j who is present in the market is further characterized by private unit costs, $(\nu_j,\, c_j)$, where $\nu_j = (\nu_{1,j}, ..., \nu_{O_b,j})$; $c_j z$ denotes the seller's total cost of performing the service and $\nu_{o_b,j} z$ reflects the seller's total participation costs if the project is from country o_b and is of size z. We assume that the unit participation costs and the unit cost of work are independent of each other and across sellers; they are distributed according to the distributions $F^{\nu}_{(x,q)}$ and $F^{C}_{(x,q)}$. Note that the distributions of costs depend on a seller's x attributes and quality but not on the country of a project's origin.

The timeline for the market runs as follows. Each potential seller j observes $(J_M,\, I_N)$ and a draw of his unit participation costs ν_j; he then chooses among $O_b + 1$ alternatives: he either decides to submit a bid for a project of type $o_b \in \{1, \cdots, O_b\}$ or decides not to participate in the market. If a 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 unit 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 for which his competitors submit their bids. The bidder then decides on the value of his unit 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 the buyer's choice set. For each seller $j \in A_l$ the buyer observes attributes, x_j, quality index q_j and 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 a 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 $\Delta_{l,j} - z\tilde{b}_{l,j}$ 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 a seller's attributes with buyer-specific coefficients α_l and $\epsilon_{l,j}$ ($\epsilon_{l,j}$ represents the residual value assigned by this buyer to a specific seller), i.e.,

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

Notice that $(q_j + x_j\beta)$ plays the role of an *effective* quality index. The component $x_j\beta$ accounts for the impact of observables on quality whereas q_j captures the *residual* quality component. We let $\epsilon_l = \{\epsilon_{l,1}, ...\epsilon_{l,|A|}\}$ and refer to (α_l, ϵ_l) as the vector of the buyers' utility coefficients.[18]

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 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: a participation component, $d : R_+^{O_b} \times \mathcal{I}_N \times \mathcal{J}_M \to \{0, 1, \cdot, \cdot\cdot, O_b\}$ which maps vectors of unit entry costs, ν, and information about market competitiveness, $(I_N,\, J_M)$, into preferred participation choices; and bidding components, $\sigma(\cdot;\, o_b) : R_+ \times \mathcal{I}_{N_{o_b}} \times R_+ \to R_+$ which for each o_b map total cost of service and information about competitive structure associated with specific type of project, $(I_{N_{o_b}},\, m_{o_b})$, into unit 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 unit cost of service c_j and submits unit bid $\tilde{b}_j$, obtains the expected payoff equal to

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

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.[19] The payoff expression reflects the fact that when deciding what to bid, a seller is only informed about the set of potential competitors for the project where he is submitting a bid.

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

[19] We assume that the distribution of sizes is the same across buyer countries. This assumption, indeed, closely approximates 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

$$\begin{aligned}\textstyle\sum_{a\subset N_{o_b}} \Pr(j\,\text{wins}|\,\tilde{b}_j, \tau_j,\, A_{-j}=a;\,o_b,\,\sigma_{-j})\Pr(A_{-j}=a|I_{N_{o_b}}, m_{o_b}; o_b,\,\sigma_{-j})\ \text{ with} \\ \Pr(j\,\text{wins}|\tilde{b}_j, \tau_j, A_{-j}=a;\,o_b,\,\sigma_{-j}) = \Pr(\Delta_j - \tilde{b}_j \geq \max_{p\in a}\{U_0, \Delta_p - \sigma_p(c_p)\}|o_b).\end{aligned} \tag{3}$$

In this expression the expectation is taken with respect to the set of actual competitors (sellers submitting bids in the same auction as j). The probability of winning conditional on the set of actual competitors is computed by taking the 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). Finally, 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 $\frac{1}{m_{o_b}}$.[20]

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

$$\bar{\Pi}_{\tau_j}(I_N, J_M;\,o_b, d_{-j}, \sigma_{-j}) - \nu_{o_b,j}\bar{z}. \tag{4}$$

This payoff is obtained from $\Pi_{\tau_j}(\sigma_j(c_j),\,c_j|\,I_{N_{o_b}}, m_{o_b};\,z, o_b, \sigma_{-j})$ by taking the 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 type of auction o_b, N_{o_b}). Intuitively, a seller chooses the 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

$$\begin{aligned} d^*_\tau(\nu_j, I_N, J_M) &= \arg\max_{o_b}[\bar{\Pi}_\tau(I_N, I_M;\,o_b, d^*_{-j}, \sigma^*_{-j}) - \nu_{o_b,j}\bar{z}] \\ \sigma^*_\tau(c_j, I_{N_{o_b}}, m_{o_b}; o_b) &= \arg\max_b \Pi_\tau(b, c_j|\,I_{N_{o_b}}, m_{o_b};\,o_b, \sigma^*_{-j})\end{aligned} \tag{5}$$

for all possible realizations of $(\nu_j,\,c_j)$, I_N, J_M, $I_{N_{o_b}}$, and for all $o_b \in \{1, \cdot, \cdot, \cdot, O_b\}$.

The existence of equilibrium of this game can be established by the following two-step ap-

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

proach. First, the results from McAdams (2003) can be applied to establish existence of the equilibrium in the bidding strategies for each sub-game summarized by $I_{N_{o_b}}, m_{o_b}$. In the second step, the standard contracting mapping argument could be used to establish the existence of equilibrium in participation strategies which is consistent with the bidding strategies characterized in the first step. While the equilibrium in the pricing strategies conditional on $I_{N_{o_b}}, m_{o_b}$ is unique, this property does not apply to the equilibrium in participation strategies. The potential non-uniqueness of the equilibrium in participation strategies does not impact our estimation of the demand-side parameters or of the distributions of the projects' costs. However, in order to recover the distribution of participation costs, we need to have that the same equilibrium is played in all markets characterized by the same value of $(I_N,\ J_M)$. This assumption is quite standard in the literature.

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 to 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 also does not distort 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 prevent 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 a permanent or transitory type. Unlike buyers, we do not observe sellers' quality indexes, 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, $F^{\nu}_{(x,q)}$ and $F^{C}_{(x,q)}$ on the supply side.

Our estimation strategy exploits the fact that, under the multi-attribute auction format, neither buyers' willingness to pay for quality nor buyers' outside options are known to sellers. Thus, buyers' choice sets (the set of participating sellers) are exogenous conditional on buyers' and projects' observable characteristics. This allows us to separate the 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 the demand-side primitives.

5.1 Demand Estimation

We make use of the two-step estimation approach developed in Krasnokutskaya, Song, and Tang (2017a). The methodological challenges that are addressed by this methodology and the details of the identification strategy are summarized in the Online Appendix.[21] Under this methodology, in the first step permanent sellers characterized by a common vector of x-attributes are classified into groups of equal quality. In the second step, these groupings are imposed in GMM estimation which recovers the distributions of utility coefficients, outside options, quality levels associated with different quality groups, and the conditional distribution of transitory sellers' quality, $\Pr(Q_h = \bar{q}|x_h)$. We begin by summarizing the identification strategy and then providing details of the estimation methodology.

Summary of the Identification Strategy. Briefly, we exploit the permanent sellers' probability of winning conditional on relevant observable attributes to recover quality groupings from the data. Once the permanent sellers' quality group memberships are recovered, we use them to identify other model primitives. Specifically, the distributions of utility coefficients are identified through the variation in the buyers' choice sets (defined in terms of permanent sellers group memberships) and the variation in prices of permanent sellers. For this we need to consider choice sets that include at least two permanent bidders. For example, we use variation in permanent sellers' prices in the auctions that attracted two or more permanent sellers from the same quality group to identify the distribution of ϵ-components; whereas the variation in the permanent sellers' prices in the auctions which attracted permanent sellers from specific different quality groups are used to identify the distribution of α and the quality levels corresponding to different quality groups. The systematic variation in quality levels associated with variation in observable characteristics identifies utility coefficients β. It is important to emphasize that, in contrast to standard approaches to estimation of discrete choice models, we do not condition on a specific choice set in the estimation. Instead, we use moment conditions which aggregate over the choice sets that are characterized by some specific property. This feature of our methodology is motivated by the presence of transitory sellers who participate only in a very small number of auctions, and due to high variability of choice sets in general (even the part of the choice set which consists of permanent sellers) which leads to a very small number of observations available for any given choice set.

Further, for a given distribution of utility coefficients, the distribution of the payoffs associ-

[21] Online Appendix can be found at http://www.econ2.jhu.edu/people/Krasnokutskaya/Research/.

ated 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 the set of participants consists of a permanent seller and a transitory one. 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 option, we need to separately consider sets of auctions where the set of participants includes transitory sellers from different x-groups. Formal details of identification strategy can be found in Krasnokutskaya, Song, and Tang (2017a).

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 characteristics and the realized set of competitors are the same for both projects) 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 an 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:

$$r_{i,j}(b) \;=\; \Pr(i\,\text{wins}\,|\,B_i=b,\, i\in A,\, j\notin A,\, i,j\in N_{o,b};o_b). \tag{6}$$

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 while j does choose to participate in an auction of type o_b. Proposition 1 in Krasnokutskaya, Song, and Tang (2017a) establishes a pairwise ranking of bidders i and j on the basis of indexes $r_{i,j}(b)$ and $r_{j,i}(b)$. Further, Krasnokutskaya, Song, and Tang (2017b) demonstrate that if a 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. They propose 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 he 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 the strongest empirical support (in terms of average p-values). This gives the first division of $S(x)$ in 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 this approach is that it imposes in estimation transitivity of sellers' ordering which may otherwise 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.[22]

The moment conditions are primarily built around the permanent seller's winning probability given the seller attributes and quality group affiliation and some features of the set of active competitors. To be specific, let B_l be the vector of submitted bids in auction l. For each bidder j in auction l, define

$$m_{j,l} = 1\{j \text{ wins } l\} - \Pr(j \text{ wins } | B_l, I, J; o_b),$$

where $\Pr(j \,\text{wins}\, l \,|B_l, I, J; o_b)$ is the conditional winning probability of seller j in auction l with a given composition of potential sellers I, composition of the set of projects J and the type of project o_b when the vector of submitted bids is B_l. We then construct a moment condition as follows:

$$\mathbf{E}\left[\sum_j g_j(B_l, I, J, o_b) m_{j,l}\right] = 0,$$

where $g_j(B_l, I, J, o_b)$ is a function of $(B_l,\ I,\ J,\ o_b)$; and the summation inside the moment condition is over j in the set of active permanent bidders in auction l.

We use two sets of moments. The first set includes (1) moments that are based on a permanent seller's probability of winning in an auction where two or more active permanent bidders belong to the same group; (2) moments that are based on the probability of winning by a permanent seller from a group (x, q) in an auction where he competes with one or more permanent bidders belonging to a different group. As for the choice of $g_j(.)$, 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 winning seller j, B_j; the squared difference between the winning 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 winning seller

[22] The 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.

j, B_j^2; the product of the difference and the bid of winning seller j, $(B_j - B_i)B_j$; the product of winning seller j's bid and the bid of a transitory seller h, B_jB_h; the product of the square of winning seller $j's$ bid and the transitory seller h's bid, $B_j^2B_h$; as well as the product of winning seller j's bid and the transitory seller's attributes, B_jx_h; the product of the square of winning seller $j's$ bid and the transitory seller's attributes, $B_j^2x_h$. We consider a set of such moments for each buyer country group.

Accommodating the Unobservability of Transitory Sellers' Qualities. To use these moment conditions, we need to evaluate $\Pr(j \text{ wins } |B_l, I, J; o_b)$. However, note that a buyer observes (x, q)-group affiliations of all sellers in his choice set, whereas the econometrician does not observe transitory sellers' qualities. This means that $\Pr(j \text{ wins } |B_l, I, J; o_b)$ 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 $\Pr(j \text{ wins } |B_l, I, J; Q^t = \bar{q}, o_b)$ to be the same as $\Pr(j \text{ wins } |B_l, I, J; o_b)$, except that it indicates the winning probability when the auction has participating transitory sellers with the specific quality vector $\bar{q}$. This winning probability ($\Pr(j \text{ wins } |B_l, I, J; Q^t = \bar{q}, o_b)$) reflects buyers' decisions, and is determined by the distribution of buyers' utility coefficients and their outside options. We write

$$\Pr(j \text{ wins } |B_l, I, J; o_b) = \sum_{\bar{q}} \Pr(j \text{ wins } |B_l, I, J; Q^t = \bar{q}, o_b) \Pr(Q^t = \bar{q}|B_{t,l}, I, J; o_b), \qquad (7)$$

where $B_{t,l}$ is the vector of transitory sellers' bids submitted in the auction. Note that the conditional probability of $Q^t = \bar{q}$ given $B_{t,l} = b_t$, I and J does not depend on the bids of permanent bidders.

To obtain the expression for $\Pr(j \text{ wins } |B_l, I, J; Q^t = \bar{q}, o_b)$, 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 $\Pr(Q^t = \bar{q}|b, I, J; o_b)$, because it involves the transitory sellers' behavior. Following Krasnokutskaya, Song, and Tang (2017a), we exploit the following representation of these objects:

$$\Pr(Q^t = \bar{q}|\, b_t, I, J;\, o_b) = \frac{\prod_{j \in A^t} \Pr(Q_j = \bar{q}_j) \Pr(j \in A_l|\, \tau_j, I, J) f_B(b_j|\tau_j, I_{o_b}, J;\, o_b)}{\sum_{\bar{q}'} \prod_{j \in A^t} \Pr(Q_j = \bar{q}'_j) \Pr(j \in A_l|\tau'_j, \bar{I}, J) f_B(b_j|\tau'_j, I_{o_b}, J;\, o_b)},$$

and parameterize $f_B(b_i|\, \tau_i, I_{o_b}, J;\, o_b)$ and $\Pr(j \in A|\, \tau_i,\, I, J; o_b)$ which correspond to the density of bids and the probability of participation associated with a transitory seller characterized by $\tau_i = (t, x, \bar{q}_i)$.

The identification strategy presented in the previous section permits nonparametric recovery of $\Pr(Q^t = \bar{q}|, b_t, I, J;\, o_b)$. We make use of additional restrictions in the estimation that would allow us to identify $\Pr(j \text{ is active}|\tau_j, I_N, J; o_b)$, $f_B(b_j|\tau_j, I_{o_b}, J; o_b)$, and $\Pr(Q_j = q_j|x_j)$ 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. We additionally impose optimality of the sellers' participation decisions.

5.2 Supply Side Estimation

Recovering the Distributions of Project 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:

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

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

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

Both $\widehat{\Pr}(i\,\text{wins}|\,b,\ \tau,\ I_{o_b},\ m_{o_b};\ o_b)$ and $\frac{\partial}{\partial b}\widehat{\Pr}(i\,\text{wins}|b,\ \tau,\ I_{o_b},\ m_{o_b};\ o_b)$ can be expressed from the model once the primitives on the demand side are recovered.[23] 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, all 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}. It is worth emphasizing that in a model where sellers consider one auction at a time, the dependency of these objects on the set of available projects does not arise. However, the estimates could be significantly biased if such dependency is not taken into account in the setting where multiple projects are available for bid. Indeed, consider two markets that are characterized by the same set of potential sellers I_{o_b} but in one market a large number of projects is offered whereas in the second market the number of projects is small. Then, in the first market, the expected number of participants in any individual auction would be small and thus the distribution of submitted bids would be located around high bid values. In contrast, in the second market, the expected number of participants in the individual auctions would be large and thus the distribution of bids would be located around small bid values. If we

[23]We implement kernel smoothing over I_{o_b} and m_{o_b} when estimating $\hat{F}^{B}_{(p,x,q);o_b}(.|I_{o_b}, m_{o_b})$ or $\widehat{\Pr}(j\text{ wins }|\,b,\ I_{o_b}, m_{o_b})$, $\frac{\partial}{\partial b}\widehat{\Pr}(j\text{ wins }|b,\ I_{o_b}, m_{o_b})$ and similarly $\widehat{\Pr}(i\,\text{is active}|\tau, I_{o_b}, m_{o_b})$ below.

fail to take the number of projects into account we would misinterpret such variation in bids as being generated by the variation in costs, which would lead us to overestimate the variance of the distributions of costs.

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 $\{\tilde{\nu}_{i,t,1}, .., \tilde{\nu}_{i,t,O_b}\}$ are independent from each other and across sellers and periods and distributed according to the exponential distributions with parameters λ_0 and $\{\tilde{\lambda}_1, ..., \tilde{\lambda}_{O_b}\}$ respectively.[24]

Heuristically, the joint distribution of $\{\nu_{i,t,0} + \tilde{\nu}_{i,t,1}, \nu_{i,t,0} + \tilde{\nu}_{i,t,2}, ..., \nu_{i,t,0} + \tilde{\nu}_{i,t,O_b}\}$ 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 a seller with $\tau = (p,\, x,\, q)$ entering an auction of type o_b is given by

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

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 $\{\nu_{i,t,0}, \tilde{\nu}_{i,t,1}, \tilde{\nu}_{i,t,2}, ..., \tilde{\nu}_{i,t,O_b}\}$ are also identified.[25]

Let us use $\bar{\pi}_\tau(o_b)$ to denote $\bar{\pi}_\tau(\bar{I}_{N,t}, \, \bar{J}_{M,t}; o_b)$, and $\Delta_{p,o_b}\bar{\pi}_\tau$ to denote $\bar{\pi}_\tau(\bar{I}_{N,t}, \, \bar{J}_{M,t}; p) - \bar{\pi}_\tau(\bar{I}_{N,t}, \, \bar{J}_{M,t}; o_b)$ for brevity. Then, under the assumption that components $\nu_{i,t,0}$ and $\tilde{\nu}_{i,t,o_b}$ are exponentially distributed the equation above can be re-written as below.

$$\Pr(i \in N_{o_b} | \, \tau, \bar{I}_{N,t}, \, \bar{J}_{M,t}; o_b) = \frac{-\lambda_0 \tilde{\lambda}_{o_b}}{\sum_{p=1}^{O_b} \tilde{\lambda}_p} \exp(-\sum_{p=1}^{O_b} \tilde{\lambda}_p \Delta_{p,o_b} \bar{\pi}_\tau) \Big(\exp(-\sum_{p=1}^{O_b} \tilde{\lambda}_p \bar{\pi}_\tau(o_b)) \times \tag{8}$$
$$\frac{\exp((\sum_{p=1}^{O_b} \tilde{\lambda}_p - \lambda_0)\bar{\pi}_\tau(o_b)) - 1}{\sum_{p=1}^{O_b} \tilde{\lambda}_p - \lambda_0} - \exp(-\sum_{p=1}^{O_b} \tilde{\lambda}_p \max(0, \max_p(-\Delta_{p,o_b}\bar{\pi}_\tau))) \frac{\exp(-\lambda_0 \bar{\pi}_\tau(o_b)) - 1}{\lambda_0} \Big).$$

The probability on the left-hand-side of this equation can be computed from the data, whereas the expression on the right-hand-side can 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 $\lambda_0, \, \tilde{\lambda}_1 \, \tilde{\lambda}_2, .., \tilde{\lambda}_{O_b}$ by minimizing the distance between the left-hand-

[24] Such entry cost structure is similar to the one used in Berry (1992).

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

side and right-hand-side expressions in (8) for different seller groups and for several values of $(I_N,\, J_M)$.

It is important to note that the potential non-uniqueness of the equilibrium in participation strategies does not impact our estimation of the demand-side parameters or of the distributions of the projects' costs. However, in order to recover the distribution of participation costs we need to assume that the same equilibrium is played in all markets characterized by the same value of $(I_N,\, J_M)$. This assumption is standard in the literature.

6 Estimation Results

This section presents the estimation results. 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 the willingness to pay for quality, as well as the estimated distributions of sellers' costs. We conclude with a 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 $x \in \bar{x}_1, \cdot, \cdot, \cdot\bar{x}_P$.

Implementation Details. We focus on the set of projects associated with graphics-related programming (media, computer games, computer-generated animation) of medium ($200-$500) and medium-large ($500-$800) size. 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.[26]

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 according to their average scores: less than 9.7 (low), at or above 9.7 and below 9.9 (medium), and at or above 9.9 (high). This definition allocates sellers approximately 30%, 30%, and 40% across the three groups.[27]

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.

[26] To minimize ambiguity of this definition we discard the first three years of the market operation. The majority of permanent sellers in our data are past their six months cut-off in this dataset. Those permanent sellers who reach the cut-off during the remaining three years constitute less than 2% of the total number of permanent sellers.

[27] 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 of ten). The estimation results for such specifications are very similar to those reported here though they are less precise.

We implement this classification procedure using only the sample of the US-based buyers, because the classification procedure requires that a large number of observations should be available for each pair of permanent sellers and for projects from a given buyer country and with a specific vector of project characteristics. This condition is not satisfied for other buyer countries (or 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 below that this approach nevertheless allows us to achieve good fit to the data.

The Results of Classification Analysis. The classification of permanent sellers from the sample of US-based projects has been implemented in Krasnokutskaya, Song, and Tang (2017a). We summarize the main finding here. The reader is referred to Krasnokutskaya, Song, and Tang (2017a) for details and the summary of the robustness checks performed by the authors.

The results of the classification step for the sellers from North America, Eastern Europe and South and East Asia country groups are presented in the Online Appendix (Table 1). 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 tends to be larger than 1 for all country groups and all levels of the reputation scores. The latter 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 the 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. This comparison is deferred to the next section.

6.2 Results of Estimation: Demand Side Parameters

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

Parameterization. We assume that the 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 the 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 a buyer-country specific mean and the standard deviation σ_α.[28] Since we estimate the standard

[28]Strictly 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.

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 hence we 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.

We assume that transitory and permanent sellers' bid distributions are well approximated by normal distributions $N(\mu_{B^t}, \sigma^2_{B^t})$ and $N(\mu_{B^p}, \sigma^2_{B^p})$,[29] 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.[30]

Results of Estimation. Tables 2 and 3 in the Online Appendix report the estimated parameters for three specifications.[31] 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 specifications 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 the seller's country group.

Our results confirm the importance of the unobserved quality component highlighted in Krasnokutskaya, Song, and Tang (2017a). The estimates for quality levels indicate that two or three distinct quality levels are associated with each value of the sellers' observable characteristics and the distributions of qualities vary with seller's observable characteristics. Buyers are willing to pay a substantial premium for improvement in the quality: for example, the average buyer would be prepared to pay a premium of 60% of the project size to move from the lowest to the highest quality level of the medium-score Eastern European seller. The same buyer world 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 the stochastic components appear

[29] See the comment for the distribution of α above.

[30] 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.

[31] The estimates for the auxiliary objects such as the distribution of bids and participation frequencies for various buyer-seller country group pairs are available from the authors upon request.

to be of a moderate size, which indicates that buyers' choices are driven by the variation in observable covariates and quality to an 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 the 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 buyers from the United Kingdom appear to have the lowest outside option whereas South and East Asian buyers appear to have the highest outside option.

The findings on the outside options are consistent with statistics on the rates charged by software developers around the world reported in various media sources[32] which indicate that the programmers in the United States and in the United Kingdom are paid almost ten times more than the programmers in South and East Asia or 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 the outside option of Eastern European or Asian buyers should be more similar in value to that delivered through an online market than the outside option for the US, the UK and Western European buyers. It is more difficult to evaluate price sensitivity findings. The data on average household incomes across countries are widely available and indicate that an individual's purchasing power in the US, Australia, the UK and Western Europe is much higher than the purchasing power of individuals from Southern Europe, South and East Asia or Eastern Europe. It is unclear, however, how well these numbers reflect the paying power of the technologically-savvy small businesses in these countries.

6.3 Results of Estimation: Supply Side Parameters

Table 4 reports the estimated means and standard deviations of the project-cost distributions. Our modeling approach maintains 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 the Asian buyer group and those associated with other buyer groups is borderline rejected (with a 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 latter two groups have comparable costs. Interestingly, the costs are U-shaped in quality, i.e., across different seller country and score groups the lowest

[32]See, for example, http://www.bloomberg.com/visual-data/best-and-worst//highest-paid-software-engineers-countries.

Table 4: Distribution of Project Costs

S:	Low	Low	Low	Medium	Medium	Medium	High	High	High
Q:	1	2	3	1	2	3	1	2	3
				North American Sellers					
μ_C	0.92 (0.056)		0.83 (0.023)	0.91 (0.056)		0.89 (0.26)	0.98 (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	0.82 (0.044)	0.71 (0.032)		0.97 (0.031)	0.69 (0.025)	0.88 (0.036)	1.04 (0.043)	0.72 (0.033)	0.83 (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	0.81 (0.034)	0.67 (0.023)	0.86 (0.034)	0.91 (0.033)	0.65 (0.023)	0.78 (0.041)		0.67 (0.023)	0.85 (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 a bootstrap procedure. In the table 'S' denotes a reputation score group whereas 'Q' refers to a quality group.

quality sellers appear to have high costs whereas the costs of medium quality seller are generally lower than the costs of high quality sellers. This illustrates the heterogeneity of sellers who participate in this online market.

Table 5: 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 5 reports the estimates of the distribution of entry costs. These estimates correspond to the specification where the distribution of the $\nu_{0,j,t}$ is seller-country specific whereas the distribution of vector $\tilde{\nu}_{j,t} = \{\tilde{\nu}_{1,j,t}, ..., \tilde{\nu}_{O_b,j,t}\}$ remains the same across seller countries.[33] 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 have generally lower costs (relative to the sellers from other countries). Further, the entry cost 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 4 in the Online Appendix 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 to illustrate the magnitude of entry costs incurred in this market. As the table indicates the entry costs constitute on average about 7% of the project costs.

Lastly, we comment on the limitations of the analysis presented in this section. In this analysis, we take the 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 permanent sellers. 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 have a large impact on the average reputation score once a seller has completed ten or more projects. Indeed, we observe in the data that a bad score is not associated with a statistically significant change in the probability of winning or an alteration in the bid of an established seller.

6.4 Model Fit

Tables 6 and 7 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.

We summarize the fit of the 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 are recorded in Table 6. This table demonstrates that the 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 the UK and Western European auctions and overpredicts

[33]In this market sellers usually submit bids for multiple projects. To accommodate this data regularity, we assume that a permanent seller present in a market in a given week wishes to bid for at most three projects; corresponding upper bound on the number of projects for a transitory seller is equal to two projects. These assumptions reflect regularities observed in the data. We further assume that a seller makes independent entry decision for each of the 'potential submissions' and these decisions are based on independent draws of entry costs. We have also experimented with the following assumptions on entry costs for different 'submissions': (a) holding $\nu_{0,i,t}$ fixed and independently re-drawing $\tilde{\nu}_{i,t}$, or (b) holding $\tilde{\nu}_{i,t}$ fixed and independently re-drawing $\nu_{0,i,t}$. We find that results remain reasonably robust to these assumptions.

Table 6: Participation Fit

Buyer Country	Data: Northern America	Data: Eastern Europe	Data: SE Asia	Model: Northern America	Model: Eastern Europe	Model: 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.

their participation in Australian auctions. Additionally, the model slightly overpredicts the participation of Eastern European sellers 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 in estimation we attempt to rationalize the sellers' observed participation decisions through the realized subsequent sellers' profitability in different types of auctions. Specifically, we explicitly model and fit to the data sellers' endogenous pricing decisions and buyers' allocation decisions. These are used to predict expected profitability for a given type of seller from participation in a given type of auction under existing market conditions (i.e., given the set of potential sellers present in the market and the set of projects available for bid). Further, the model primitives which are used to construct expected profit conditional on participation are estimated separately and are not adjusted to facilitate the fit to participation patterns and they do not include bilateral components.[34] Additionally, an important indicator of a good fit of the model to the data is the fact that the empirical participation probabilities are close to the theoretical participation probabilities under the plausible values of estimated parameters. Indeed, we might expect to find a set of bilateral parameters characterizing participation costs (one parameter for every buyer-seller country pair) which reconcile the empirical and theoretical participation probabilities if these parameters are allowed to take arbitrary numerical values. However, in our model these parameters have a structural interpretation which means that they have to take reasonable values for us to consider the model a success. To begin with, they should be non-

[34]Such parameterization is more parsimonious relative to a typical single auction entry model where the parameters of the distribution of participation costs are chosen to reconcile the observed probability of entry and observed profitability for a given type of bidder and a given type of project.

negative but also they should not be too large relative to other profit components, e.g. costs of completing the project. Not only do we find this to be the case in our estimation, our model is much more restrictive in that it does not allow for bilateral entry costs.

The top panel of Table 7 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 the allocation of projects towards these sellers for all country groups except North America. Note that Table 7 reflects the fit in terms of the *ex-ante* market shares corresponding to buyer-seller country group pairs. Thus, it indicates that the model is capable of correctly reproducing the allocative decisions conditional on participation and prices, as well as the pricing and participation decisions that generate the choice sets on which allocation decisions are based.

Table 7: 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 the 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 investigates the clustering patterns generated by the 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 the model reproduces the 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 lower rate 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 results indicate that our model does 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:

$$R = \sqrt{\frac{1}{M_B}\sum_{c_2}\frac{1}{M_S}\sum_{c_1}(\frac{s_{c_1|c_2}}{s_{c_1}} - 1)^2} \tag{9}$$

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 the conditional and the marginal frequencies, respectively, of different seller-country groups.[35] 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 in which buyers are presented with 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 the model under the counterfactual scenario outlined above using an extension of the computational method proposed by Marshall, Meurer,

[35]This index is a statistic which is commonly used to test for conditional independence of random variables. If the frequency of sellers' participation does not vary with (is independent of) the buyer country then $s_{c_1|c_2} = s_{c_1}$ for all c_1 and c_2 and thus the index R is equal to zero, whereas a deviation from zero points towards statistical dependence. We use this statistic to construct a measure (index) of dependence. However, in our setting, the number of observations used to construct this index is so large that any reasonable positive value is indicative of a statistically significant dependence.

Table 8: Conditional Market Shares: Clustering Under Non-Strategic Participation

Buyer country	Benchmark Model Northern America	Benchmark Model Eastern Europe	Benchmark Model SE Asia	Random Participation Northern America	Random Participation Eastern Europe	Random Participation 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 a simulation analysis which studies clustering under non-strategic participation. The bidding strategies are re-computed to account for non-strategic participation.

Richard, and Stromquist (1994). Our computational algorithm is summarized in the Appendix. In this exercise we hold the global market conditions – the 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 8 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; between European buyers and Eastern European sellers; as well as between Australasian buyers and Asian sellers is diminished. In contrast, the conditional allocation in excess of the 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 Policy Analysis

In this section we investigate the effects of policies impacting participation decisions of potential sellers in the context of markets where multiple auctions are available for bid. We focus on the policies impacting international trade since the sorting of sellers across projects most prominently occurs along this dimension.

Specifically, we consider three types of counterfactual scenarios: (1) the exclusion of low quality foreign providers, for example, through licensing; (2) the 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 the domestic price exceeds the foreign price by a specified 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 Online 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 the dimensionality of the problem.

Specifically, we first discard auctions held by South European buyers due to their small number and because the 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 (combining UK and Western and Eastern Europe), and the Pacific region (combining Australasia with South and East Asia).[36] Further, we reduce seller heterogeneity by eliminating differences in reputation scores.[37] Thus, in the end we have 9 different seller groups – 3 country groups with 3 quality groups per country group (with the exception of North America where we have only 2 quality groups) plus the residual group comprised of all other countries – and 3 buyer country groups. We additionally ignore the distinction between permanent and transitory sellers in this analysis. Instead we impose the rule that all sellers are permanent. We have experimented with a specification which

[36] We 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.

[37] We 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 a different reputation score group.

Table 9: Participation and Allocation

Buyer Country	Participation Seller Country North America	Eastern Europe	South and East Asia	Allocation Seller Country 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. For a given specification (horizontal panel) each entry in the left half of the panel reflects the conditional probability that a bid is submitted by a seller from a given country group conditional on this bid being submitted to an auction conducted by a buyer from a given buyer country group; similarly, each entry in the right half of the panel reflects the conditional probability of an auction being won by a seller from a given country group conditional on this auction being conducted by a buyer from a given buyer country group. The numbers reported in the table are simulated averages across sample time periods.

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.[38]

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 the rule that the number of projects by buyer country and the number of potential entrants by seller group for each time period should be as observed in the data. Table 9 reports the distribution of the number of bids (participation behavior) and the allocation of projects across seller countries conditional on the buyer country (see the benchmark case).[39] The table replicates the

[38]In 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 a non-stable type.

[39]Specifically, each entry in the top left panel reflects the conditional probability that a bid is submitted by

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 the buyer country. Further, Table 10, which reports the probabilities of participation (that is the probability of submitting a bid to a given buyer country group) for different seller groups, demonstrates that the majority of bids for each seller group are 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.

Table 10: Probabilities of Entry

		Benchmark			Quota		
		Buyer Country			Buyer Country		
Seller Country	Quality	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 computed for the benchmark and counterfactual settings. The entries in a given row reflect the probabilities with which a seller from a given country group (summarized by this row) enters one of the auctions conducted by buyers from the country group corresponding to a given column conditional on being present in the market.

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 the 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 Tables 9 and 11 which report conditional market shares (participation and allocation), and the welfare measures, respectively.

Quality Restriction on Participation. First, we consider a setting where the restriction on trade excludes foreign sellers with the quality levels below the lowest quality level of North

a seller from a given country group conditional on this bid being submitted to an auction conducted by a buyer from a given buyer country group; correspondingly, each entry in the lower left panel reports the conditional probability of an auction being won by a seller from a given country group conditional on this auction being conducted by a buyer from a given buyer country group.

American sellers. Under this scenario, the lowest quality Eastern European sellers are prevented from participation in North American auctions. This restriction has a 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 European sellers 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 the excluded group. The changes in welfare for all market participants except for low quality European sellers are rather small.

Quota on Foreign Participation. Next, we consider the case when the availability of foreign sellers is restricted without regard to quality. In this experiment, foreign sellers have to obtain permission in order to submit a bid in a North American auction. The number of permissions granted is set to be 30% below the number of foreign bids submitted to the North American auctions in the unrestricted setting.[40,41] Under this quota restriction, the number of bids submitted to 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 reduction in 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 increase. North American buyers are hurt by this increase. Interestingly, this effect is mitigated by the change in the mix of sellers participating in the North American market. Specifically, higher quality foreign sellers who are inherently more competitive (they are more attractive to buyers and have lower costs) shift their participation away from North American market at a lower rate relative to the low quality foreign sellers. Similar regularity also holds in the case of North American sellers with the higher quality group gaining more in terms of market share relative to the low quality group. As a result, the mix of sellers participating in the market shifts towards higher quality. These effects can be observed in Table 10 which reports the probabilities of participation across seller groups under a participation quota and in Table 11 which reports welfare changes resulting from a quota restriction both under fixed participation (sorting is not allowed) and when participation is allowed to adjust.[42] As the results indicate, under fixed participation the welfare loss to buyers is substantial (-12%) and the overall welfare is reduced (-5%). The adjustment in participation lowers the negative impact on buyers (to -3%

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

[41]The 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.

[42]Specifically, when computing an equilibrium without sorting we impose that the relative frequencies of participation in different types of auctions for a given seller type remain, as in a benchmark case. The probability of not participating is allowed to change and then all other probabilities are adjusted accordingly.

Table 11: 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 the welfare effects of several counterfactual restrictions on trade in the graphic programming segment of on-line market. The numbers reported are simulated averages per time period.

only) and thus leads to an 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 the 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 cost distributions with relatively low mean and variance. Thus, in addition to delivering higher utility to buyers their increased participation also brings down prices in the equilibrium relative to the case of 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 competition at the auction level. We now consider a 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.[43] Under this policy a buyer is penalized (pays a fine equal to **x**% of the project size) if he purchases the 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 latter 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 the 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 the competitive pressure from domestic 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 cost 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 the quota, the quality mix of foreign participants shifts towards higher-quality levels. North American buyers pay lower prices relative to the case of the quota and they are presented with an improved quality mix relative to the benchmark case. As a result, the buyers' surplus and the overall welfare increase, the latter by 7%. In this case the difference between 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 profitability is substantially reduced by this policy. Additionally, even if other buyer countries impose a 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

[43] A detailed analysis of this policy can be found in McAfee and McMillan (1987) and Krasnokutskaya and Seim (2011).

sellers is very limited.[44] This experiment, of course, holds the conditions in other markets fixed, when in reality, foreign countries may retaliate through policies impacting other markets.

9 Conclusion

This paper makes a two-fold contribution to the literature. First, we develop a tractable model of an online market for services which rationalizes the sorting of heterogeneous sellers across heterogeneous projects and the realized volume of transactions. We specifically focus on the ability of our model to explain observed clustering of international trade which is a prominent feature of our data. Such an analysis has an independent substantive interest since this market is representative of the increased importance that the electronic space plays in facilitating international trade in services. Second, we investigate the importance of taking sorting (and re-sorting) into account when analyzing policies impacting participation in such markets.

The main conceptual difficulty which arises in the study of online 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 a 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. The key insight which insures a successful fit of this model to the data is that the equilibrium outcomes in a given period are determined both by the number and composition of the set of potential sellers and by the set and the composition of available projects. It is thus important to take into account that each seller's decisions reflect the fact that he and his competitors are choosing among multiple auctions when deciding where to participate and which bids to submit. In fact, we show that 70% of the observed clustering of trade is explained by sorting into participation and by pricing which takes such sorting into account.

From a policy perspective, we find that the sorting of sellers into participation should not be ignored as it may play an important role in determining policy outcomes. Specifically, the analysis of the policy restricting participation of foreign sellers in the US market which fails to take re-sorting into account finds that the welfare of North American buyers would decrease since such a policy would result in higher prices and thus lead to a reduction in buyer surplus. However, the model that takes into account participation adjustments accounts for the fact that the mix of sellers participating in US market would shift towards higher quality which mitigates the negative effect of a price increase. Such a model would predict that the overall welfare would in fact improve modestly.

[44] The results of this experiment are available from the authors upon request.

References

Ackerberg, D., K. Hirano, and Q. Shahriar (2006): "The Buy-it-Now Option, Risk Aversion and Impatientce in Empirical Model of eBay Bidding," Working Paper, University of Michigan.

Asker, J., and E. Cantillon (2008): "Properties of Scoring Auctions," *The RAND Journal of Economics*, 39(1), 69–85.

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.

Bajari, P., and G. Lewis (2011): "Procurement with Time Incentives," *Quarterly Journal of Economics*, 126(3), 1173–1211.

Balat, J., I. Brambilla, and Y. Sasaki (2016): "Heterogeneous Firms: Skilled Labor Productivity and Export Destination," Working Paper, Johns Hopkins University.

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.

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

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.

Guerre, E., I. Perrigne, and Q. Vuong (2000): "Optimal Nonparametric Estimation of First-Price Auctions," *Econometrica*, 68(3), 525–574.

Hendricks, K., and A. Sorensen (2014): "Evaluating the Role of Intermediary in Online

Auctions," Working Paper, University of Wisconsin-Madison.

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 (2017a): "The Role of Quality in Internet Service Markets," Working Paper, University of Pennsylvania.

——— (2017b): "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.

Melitz, M. J. (2003): "The Impact of Trade on Intra-Industry Reallocations and Aggregate Industry Productivity," *Econometrica*, 71(6), 1695–1725.

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.

Tadelis, S., and F. Zettlemeyer (2015): "Information Disclosure as a Matching Mechnism: Theory and Evidence from a Filed Experiment," *American Economic Review*, 105(2), 886–905.