Simulation of Public Procurement Networks

Maximiliano González

BUSN36917 - Networks and Markets

The University of Chicago

March 8, 2021

Context and Project Objectives : Public Procurement

- More than 500 government units in Chile procure different types of buildings and construction-related services.
- Most are awarded through competitive bidding.
- Each contractor can bid to more than one government unit, each government and each unit has multiple contractors $\Longrightarrow$ a natural bipartite network.

1. Project Objective

Develop a predictive model of outcomes in a networked market of first price, sealed bid auctions.

2. Project Objetive

Simulate how entry and entry creation would impact welfare and prices.

Literature Review

The project draws from three strands of literature:

- Simulation of bipartite networks: (Guillaume and Latapy, 2006), (Newman, Strogatz, and Watts, 2001).
- Auction models that deal with endogenous/exogenous entry: (Athey, Levin, and Seira, n.d.), (Bajari and Hortaçsu, 2003), (Li and Zheng, 2009).
- Influence of market structure in competition, networked markets : (Kranton and Minehart, 2001) (Bimpikis, Ehsani, and Ilkılıç, 2019).

Gaps to adress: auction models without "intermediate" set-up, empirical work, bridge between "economic" and physics/computer science approach.

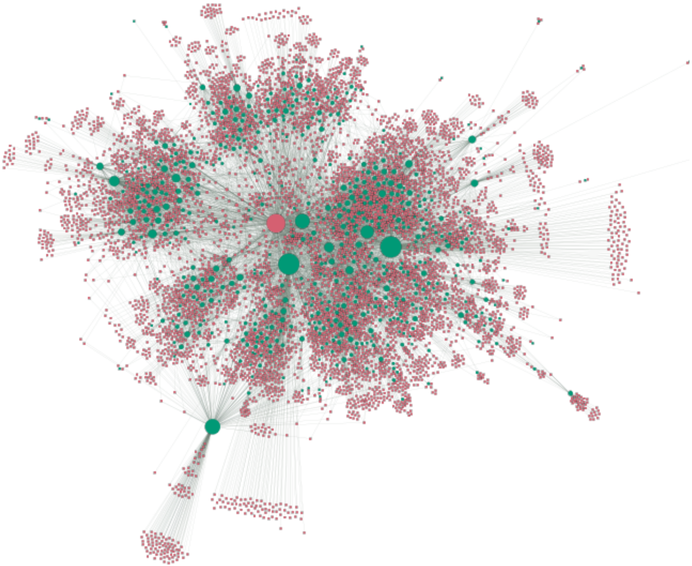
Data

- Data consists in first price, sealed bid auctions for construction goods and services procured in Chile from 2012 to 2016.
- The government provides an estimate of the project, so all bids are standarized by diving on the estimate. Only single item auctions were selected.
- A firm f and a government unit g have a link between them if f bid on any of the auctions developed by g during the timeframe selected.
- The variables employed in the project are the government unit who auctioned the item, the firm who bid for it, and the amount of the submitted bid.

Period	Contracts	Firms	Gov.units	Bid(mean)	Bid(sd)	Participants(mean)	Participants(sd)
Period 1	13500	6640	593	0.852	0.27	2.56	2.01
Period 2	14500	7190	638	0.834	0.24	2.72	2.19

Table: Sample Descriptive Statistics

Network



Bipartite projection

- To model bipartite networks it is usually convenient to employ the top and bottom projections:

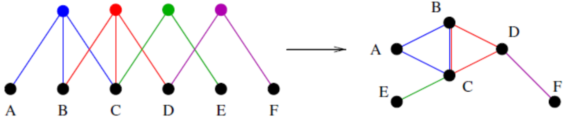


Figure: A bipartite network and its bottom projection. If two nodes share a common neighbor in the left they have a link in the network on the right (picture: Guillaume and Latapy, 2006).

Structure of the Project: three models in two stages

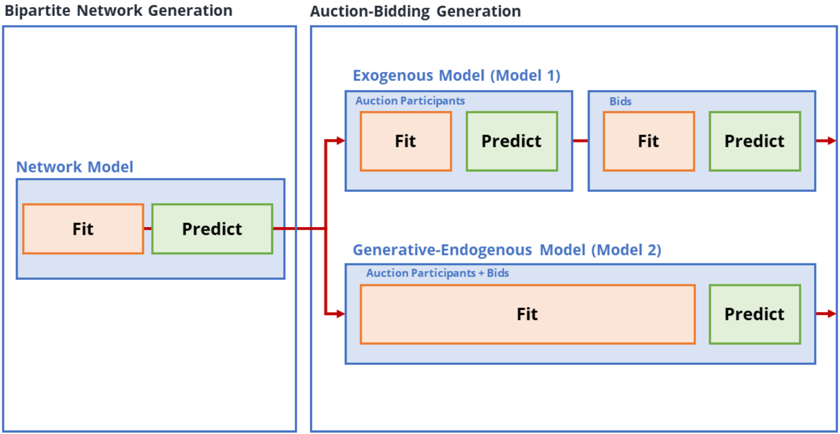


Figure: Structure of the simulation procedure

I. Bipartite Network Generation

1. Fit
   - Fitted a negative binomial for the top distribution and a lognormal distribution for the bottom distribution.
   - Higher order statistics need to be checked.
2. Predict
   - The process requires i) randomly drawing from the fitted distributions and ii) creating an edge list compatible with both of the random draws from the distributions.
   - Adapted algorithm from (Guillaume and Latapy, 2006), in two stages (small bias).

Fitting/Simulation Results

Table: *Network statistics for fitted and simulated networks*

Model	average.distance	clust.coef	density.graph
Empirical Period 1	2.15	0.806	0.0219
Fitted Period 1	2.28	0.759	0.0170
Empirical Period 2	2.22	0.808	0.0240
Predicted Period 2	2.27	0.760	0.0176

Exogenous Model

Let G be the set of top (government) nodes g and F the set of bottom (firms) nodes f. F_g is the subset of firms with links to node g and f_d is the degree of node f.

- For each g in G:
  - Draw $p \sim \mathcal{LN}(\mu, \sigma)$ and sample p firms from F_g, obtain $F'_g \subset F_g$, firms that participate in the auction.
  - For each f in F'_g:
    - Draw bid $b_f \sim \mathcal{N}(\mu_B(d,p), \sigma_B(d,p))$.

- The following parsimonius assumption is made: $\mu_B(d,p) = \beta_0 + poly(d,2) + poly(p,2) + \varepsilon$ and $\sigma_B(d,p) = \gamma_0 + \varepsilon$.
- Number of Participants' distibution was selected to be lognormal by employing the BIC criterion, comparing with Poisson, Gamma, and Negative Binomial.

Endogenous Model: Motivation

- For an item A auctioned by g, firms have private cost values drawn from a common distribution function. They bid truthfully and their strategy consists in bidding only when they expect to win.
- All firms receive the same signal z regarding the bias in the government's estimate of the project.
- A firm bids when it expects it will win, i.e. when its cost is less than the expected value of the minimum of the n competitors in the auction.
- Since the number of competitors is not known, this is an expected value.
- At the end of the round, the firm selects the most favorable market among the ones where they have preliminary bids and submits its bid there.

Endogenous Model: Algorithm

Algorithm

- For each g in G:
  - Draw common component cost: $z = \mathcal{U}(-\theta, \theta)$
  - For each f in F_g:
    - Draw cost c_f from $\mathcal{N}(1, \sigma)$.
    - If $c_f < m_f(E(A_g))$:
    - Preliminary bid $(v + z)$.
  - For each f in F:
    - Select bid f_b with the lowest cost and submit.

Random Variables:

- A_g: Number of firms bidding for A on g.
- z: bias in the government estimate.
- c_f: cost of A for firm f.
- $m_f(n)$: $E(\min_n c_{f'}), f' \neq f$
- Only two parameters!

Network Effects

1. First Model
   - The distribution of the number of participants in an auction is specific to the degree of the top node.
   - The distribution of bids depends jointly on the degree of the firm and the number of participants in the auction.
2. Second Model
   - The expected number of firms that will participate in the auction is assumed to be:

$$E(A_g) = \sum_{f \in F_G} \frac{1}{d_f}$$

   - In words, the agent estimates that a competing firm has a probability $\frac{1}{d}$ that its best chance is in the current market, and thus that it will submit a bid there ("Miopic").

- The fitted models are employed to simulate period 2 (p_2), using as features **only** the number of nodes at the top and bottom for $p2$ (so not *fully* a prediction).
- The simulation corresponds to one auction per government unit.
- The following table displays relevant result statistics:

Model	bid(mean)	bid(sd)	participants(mean)	participants(sd)	W. distance
Empirical Period 2	0.852	0.270	2.56	2.010	0.000
Exogenous	0.868	0.301	2.53	1.700	0.113
Endogenous	0.803	0.200	1.69	0.911	0.109

Table: Simulated Bids Statistics by model

Results II

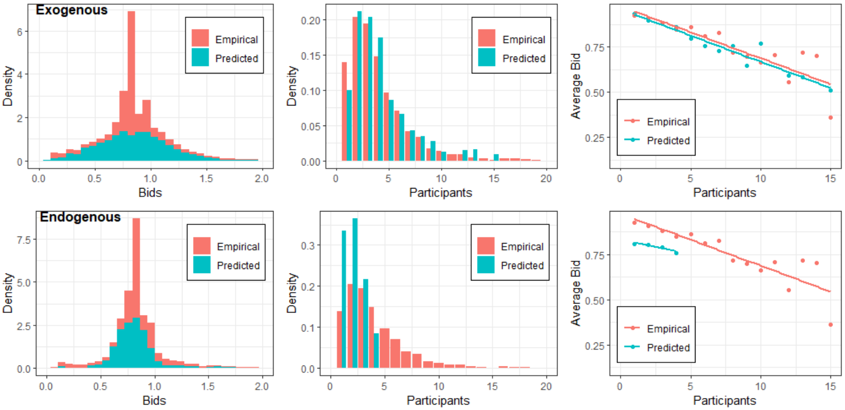


Figure: *Simulation Results comparison. Top row: exogenous model. Bottom row: endogenous model. Note that a plausible distribution of participants arises naturally in the second model.*

Counterfactuals

- We would like to know how a different configuration of top and bottom nodes would impact the efficiency of auctions.
- Entry and link creation:relevant market phenomena.
- I test the effect on bid average and number of participants of the following matrix of counterfactual scenarios:

Phenomena	Change	Type
Entry of New Firms	+10% firms	Random Degree Entry High Degree Entry Low Degree Entry
Link Creation	+10% links	Random Link Creation High Centrality Links Low Centrality Links

Counterfactual Results: Entry (10% more firms)

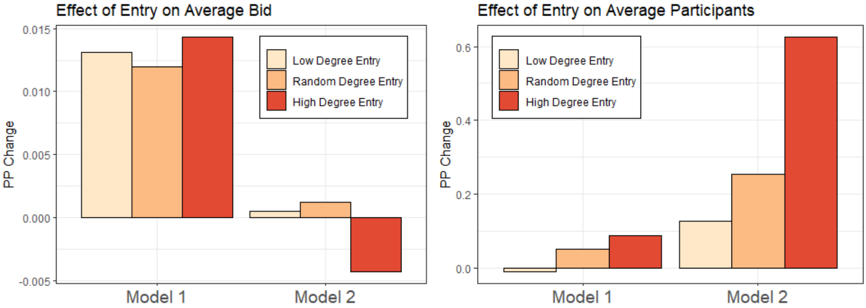


Figure: Effect of entry on bids(left) and on average participant in auctions(right). Entry in general drives up bids !

Counterfactual Results: Density (10% more edges)

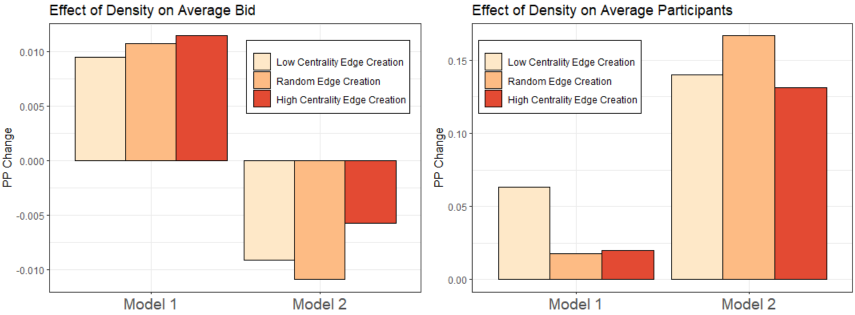


Figure: Effect of increased density on bids(left) and on average participant in auctions(right).

Discussion: computational side

- Pending: robustness and confidence of estimations employing the bootstrap or other techniques.
- Pending: examine other distributions for the bid, with fat tails (not immediately better).
- Fitting the endogenous auction model was the most computationally demanding part of the project.
- The exogenous model could be improved a lot by considering a richer set of network features and running a lasso regression in ML fashion.

Discussion: market-economic side

- Pending: winning bid analysis.
- The problem of endogenous entry in auctions seems to be difficult because it bridges two micro levels. Some partial solutions in the literature.
- The proposed model was parsimonius (only two parameters) and yet captured the high concentration of bids near the mean.
- The counterfactual analysis might might relate to two competing effects discussed in the literature: entry and competition. (Li and Zheng, 2009).
- A key feature absent which explains some problems in Model 2: heterogenous projects. A firm does not usually face the *full* set of connected firms to a government unit.

Appendix

Fitting the Endogenous Model

- Objective function: we try to minimize the distance between the observed distribution of bids and the fitted distribution of bids. Instead of summary statistics, I employ as criterion the *Wasserman distance*.
- Since the generative model performs several intermediates steps, it is unclear how to perform an optimization over the parameters via some deterministic pre-programmed method.
- Instead, I construct a $s \times t$ grid M with s values for parameter 1 and t values for parameter 2 and evaluate the objective function at each one ("MLE").
- The selected parameters are then:

$$(\mu^\star, z^\star) = \arg\min_{\mu,z} ||(B - \hat{B}(\mu, z))||_{\mathcal{W}}$$

$$(\mu, z) \in M$$