<div align="center">

Replication Instructions for

Collusion in Auctions with Constrained Bids:

Theory and Evidence from Public Procurement

</div>

<div align="center">

Sylvain Chassang   Juan Ortner*

New York University   Boston University

June 5, 2019

</div>

1   Stata files

The Stata files needed for replication are available in the stata/ folder. They include:

- Data.dta: raw auction data from the cities in the study

- gdpimport.dta: Japanese GDP data used as control in the regressions

- do_file.do: computes a variety of new variables from Data.dta

- regressions.do: runs all the regressions in the paper and online appendix

Data in Data.dta was scrapped from each of the cities' websites. The variables in this dataset are:

- bid: bid (in Yens) placed by each bidder in each auction. For some cities (i.e., Tsuchiura and Tsukubamirai), our dataset contains all bids (i.e. winning and losing) per auction. For the other cities, our dataset only contains the winning bid.

---

*Chassang: chassang@nyu.edu, Ortner: jortner@bu.edu.

- **day**: day of the month in which the auction was run

- **month**: month in which the auction was run

- **year**: year in which the auction was run

- **reserve_price**: auction's reserve price, in Yens

- **num_wins**: the number of times the bidder who submitted that bid won an auction

- **win_percentile**: for each city, we calculated the distribution of the number of wins among bidders that participate in auctions in that city. Variable **win_percentile** shows the bidder's percentile in this distribution. For cities in which the number of firms that won at least one auction is larger than 100, **win_percentile** is a number between 0 and 100. For cities in which the number of bidders that won at least one auction is smaller than 100, **win_percentile** is a number between 0 and the total number of winning bidders

- **cityXYZ**: indicator variable, which takes value of 1 if auction was run by city XYZ.

- **pid**: auction id for the city of Tsuchiura

- **winner**: indicator variable, taking value of 1 if bid was a winning bid

- **bidder_id**: bidder id for the city of Tsuchiura

- **norm_bid**: bid divided by reserve price

- **num_participated**: number of auctions in which bidder participated – available only for Tsuchiura and Tsukubamirai.

- **participation_percentile**: percentile that bidder occupies in the distribution of overall participation – available only for Tsuchiura and Tsukubamirai.

- **cartel_participation**: indicator variable, which takes a value of 1 if bidder is among the 25% of bidders who participates more frequently – available only for Tsuchiura and Tsukubamirai.

Data in **gdpimport.dta** was downloaded from the website of Japan's Cabinet Office:

http://www.esri.cao.go.jp/en/sna/sokuhou/sokuhou_top.html

2 Matlab files

The changes-in-changes estimates in Table 3 are computed using the files in the folder named athey_imbens/.The folder data/norm_bid_tsukuba_ushiku_truncated_0.80 contains the data used to run the code. The files in that folder are:

- Y00A.csv: contains auctions with winning bids above 0.8 run in control city Tsukuba prior to the policy change in Tsuchiura.

- Y00B.csv: contains auctions with winning bids above 0.8 run in control city Ushiku prior to the policy change in Tsuchiura.

- Y01A.csv: contains auctions with winning bids above 0.8 run in Tsuchiura prior to the policy, for the exact same time period as in Tsukuba (file Y00A).

- Y01A.csv: contains auctions with winning bids above 0.8 run in Tsuchiura prior to the policy, for the exact same time period as in Ushiku (file Y00B).

- Y10A.csv: contains auctions with winning bids above 0.8 run in Tsukuba in the two years after the policy change.

- Y10B.csv: contains auctions with winning bids above 0.8 run in Ushiku in the two years after the policy change.

- Y11.csv: contains auctions with winning bids above 0.8 run in Tsuchiura in the two years after the policy change.

Note that we do not merge the control data. This would bias results since the relative sample size of the pre and post period is different across control cities. Instead we separately run the algorithm of Athey and Imbens for each control city, and then average the corresponding counterfactual estimates. We report bootstrapped standard errors for our aggregated estimate. Matlab program run_athey_imbens.m runs the algorithm of Athey and Imbens for each control city, produces average estimates and bootstrapped standard errors. Results are saved in folder Output/.