Econometric Detection of Collusion

Patrick Bajari
University of Minnesota
December 10, 2008

Motivation

- Goal: Identify set of firms that may be colluding within an industry
- Today: Academic empirical literature on cartel detection

Outline

1. Motivation
2. Economic theory and competitive bidding
3. Known Examples of Collusion
4. Bajari and Ye (2003)
5. Conclusions

Motivation

Caveats

1. No test for collusion is fool proof: garbage in/garbage out
2. false negatives and false positives

Benefits

1. The empirical literature suggests clear warning flags
2. Enforcement may prevent future collusion
3. Screening can be the basis for asking “hard questions”
4. Experience is cumulative and ability to detect improves over time
5. Why is no monitoring the optimal?

Examples

- **Bid-rigging**: Porter and Zona (1992, 1997), Bajari and Ye (2003), Bajari and Summers (2002), List, Milimet and Price (2004), Marx and Marshall (2008)
- **Variance Screens**: Abrantes-Metz, Froe, Geweke and Taylor (2006), Shor and Froeb (2005), FTC (2005), Nunes and Gomes (2005)
- **Others**: Crampton and Schwartz (2000), Bajari and Fox (2006), Bajari and Yeo (2008)

Competitive Bidding

- Sealed bid procurements
- Low Bidder wins
- Firms are heterogeneous
- Differentiated by:
  - location
  - backlog
  - political boundaries
  - other fixed, but persistent factors

Predictions of theory

1. Bids should be independent

- robust prediction of economic theory
- must condition on correct control variables
- requires some industry knowledge

Predictions of theory

2. Bids should reflect costs

- Ranking of bids should reflect observable cost shifters
- Comparison to control groups of other markets/firms
- Markups should “make sense”

Case Studies of Collusion

1. Failures of independence

a. High correlation of known cartel member bids

- Porter and Zona (1992, 1997)
- List, Milimet and Price (2004)

b. Perfect coordination of entry decisions

- Marx and Marshall (2008)

Case Studies of Collusion

2. Bids don't appropriately reflect costs

- Porter and Zona (1992)
- Winning Bid reflects distance/backlog
- Ranking of 2nd, 3rd and other bids unrelated to distance
- Bid levels, effects of cost covariates do not make sense relative to a control group
- Entry decisions not well explained by potential profitability

Bajari and Ye (2003)

- Bajari and Ye (2001) seal coating road repair work in Minnesota, North Dakota, and South Dakota.
- 1994-1998, 441 bids, 139 jobs, $92.8 million
- Antitrust case against several firms in the mid-1980's for bid-rigging.

Bajari and Ye (2003)

- Firm 2 received a one-year prison sentence in late 1980's.
- Firms 4 and 5 paid treble damages for bid-rigging with firm 2.
- Firms 2,4 and 5 were temporariliy banned from bidding

Data

- *BIDi,t*: The amount bid by firm *i* on project *t*.
- *ESTt*: The engineer's cost estimate for project *t*.
- *DISTi,t*: Distance from firm *i*'s headquarters to the midpoint of project *t*.
- *LDISTi,t*: *log(DISTi,t+1.0)*.
- *CAPi,t*: Used capacity measure of firm *i* at the time it bids on project *t*.
- *MAXPi,t*: Maximum percentage of free capacity of all firms on project *t*, excluding firm *i*.

Data

- *MDISTi,t*: Minimum of distances for all firms on project *t*, excluding firm *i*.
- *LMDISTi,t*: *log(MDISTi,t+1.0).*
- *CONi,t*: The proportion of work done (by dollar volume) by firm *i* in the state where project *t* is located prior to the auction.
- *DUMt*: A dummy variable for project *t*, that is, the dummy variable is equal to one if the project is *t* and is zero otherwise.
- DUMi: A dummy variable for firm *i*, that is, the dummy variable is equal to one if the bidder is firm *i* and is zero otherwise.

of 5–10 percent of their total bid, guaranteeing that they will not withdraw their bid after the public reading of all bids. After the contract is awarded, the low bidder must submit both a performance bond and a pay bond to guarantee the completion of the contract and payment of all subcontractors.

The combined value of contracts awarded in the data set is $92.8 million.[13] Fifty-five firms in the data set won at least one job, of which only eighteen had a revenue share of contracts exceeding 1 percent.[14] Table 4 summarizes the revenue shares of these eighteen firms (identified by their ID number in the dataset) and how frequently each bid for contracts in the data set. The largest 7 firms captured 65.6% of contract revenues, led by firm 2, which alone accounted for 21.1% of contract revenues and participated in 66.9% of the auctions conducted.

Table 4
Revenue Shares and Participation by Firm

Firm ID No.	*No. of Wins*	*Avg. Bid*	*Revenue Share %*	*No. Participation*	*% Participation*
1	92	82,790	8.2	145	29.3
2	102	191,953	21.1	331	66.9
3	20	363,565	7.8	69	14.0
4	35	241,872	9.1	114	23.0
5	29	283,323	8.9	170	34.3
6	40	77,423	3.3	84	17.0
7	45	62,085	3.0	121	24.4
8	16	87,231	1.5	134	27.1
9	10	237,408	2.6	14	2.8
11	4	328,224	1.4	28	5.7
12	3	317,788	1.0	8	1.6
14	4	754,019	3.2	25	5.1
17	5	1,018,578	5.5	8	1.6
20	13	355,455	5.0	38	7.7
21	2	903,918	1.9	5	1.0
22	2	903,953	2.0	8	1.6
23	2	439,619	1.0	4	0.8
25	3	382,012	1.2	13	2.6
Total	427		87.7		

[13] The size of contracts varied greatly. Of the 495 contracts in the data set, 7 contracts were awarded for more than $1 million, 256 contracts were awarded for less than $1 million but more than $100 thousand, and 232 contracts were awarded for less than $100 thousand. A total of 98 firms bid on at least one of these 495 contracts.

[14] A firm's revenue share is defined as the total value of the firm's winning bids as a percent of the total value of winning bids for all contracts.

Table 5
Bidding Summary Statistics

Variable	*No. of Observations*	*Mean*	*Standard Deviation*	*Min. Value*	*Max. Value*
Winning Bid	441	175,000	210,000	3893	1,732,500
Winning Bid/Estimate	139	1.0031	0.1573	0.6662	1.5421
Money on the Table: 2nd Bid-1st Bid	134	15,748	19,241	209	103,481
Normalized Money on the Table: (2nd Bid-1st Bid)/ Estimate	134	0.0776	0.0888	0.0014	0.5099
Number of Bidders per Project	139	3.280	1.0357	1	6
Distance of Winning Bidder	134	188.67	141.51	0	584.2
Distance of Second Bidder	134	213.75	152.01	0	555
Capacity of Winning Bidder	131	0.3376	0.3160	0	0.9597
Capacity of Second Bidder	131	0.4326	0.3435	0	1

Another important determinant of firms' success in winning contracts is familiarity with local regulators and material suppliers. Such familiarity presumably lowers the costs of preparing a potentially successful bid and of arranging for necessary supplies. Table 6 presents, for each state and for each firm, the percentage of the firm's total dollar volume done in that state. Most firms in the data set work primarily in one state, and this pattern persists even after controlling for distance. For instance, firm 3 is located near the boundaries of Minnesota, North Dakota, and South Dakota, yet it performs over 70 percent of its dollar volume of seal coating in South Dakota. Firm 6 is located near the Minnesota-South Dakota border, but it won no contracts in South Dakota. This suggests that the concentration by state is not explained by distance alone.

1. *Bid Function Regressions*

The theoretical model of auctions in Bajari and Ye implies that a firm's bid should be a function of factors that influence both its own and its rivals' costs. Tables 5 and 6 suggest that the engineer's cost estimate, firms' distance to the project, available capacity, and previous experience in the state are important factors about which there is public information that influences costs. Bajari and Ye use regression analysis to model firms' bids as a function of the following variables.

Table 6
Concentration of Firm Activity by State

Firm ID No.	*MN. Concentration*	*ND. Concentration*	*SD. Concentration*
1	1	0	0
2	0.2781	0.7218	0
3	0	0.2377	0.7623
4	0	1	0
5	0.1246	0.5338	0.3414
6	0.8195	0.1804	0
7	0.9572	0.0427	0
8	0.7290	0.2709	0
11	0	0	1
14	0	1	0
20	0	1	0

- $BID_{i,t}$: The amount bid by firm i on project t.
- EST_t: The engineer's cost estimate for project t.
- $DIST_{i,t}$: Distance from firm i's headquarters to the midpoint of project t.
- $LDIST_{i,t}$: $log(DIST_{i,t}+1.0)$.
- $CAP_{i,t}$: Used capacity measure of firm i at the time it bids on project t.
- $MAXP_{i,t}$: Maximum percentage of free capacity of all firms on project t, excluding firm i.
- $MDIST_{i,t}$: Minimum of distances for all firms on project t, excluding firm i.
- $LMDIST_{i,t}$: $log(MDIST_{i,t}+1.0)$.
- $CON_{i,t}$: The proportion of work done (by dollar volume) by firm i in the state where project t is located prior to the auction.
- DUM_t: A dummy variable for project t, that is, the dummy variable is equal to one if the project is t and is zero otherwise.
- DUM_i: A dummy variable for firm i, that is, the dummy variable is equal to one if the bidder is firm i and is zero otherwise.

Bajari and Ye begin to explore the empirical relationship between costs and bids by estimating the following regression using ordinary least squares and information in the database for the bids of all firms on all contracts.

$$\frac{BID_{i,t}}{EST_t} = \beta_0 + \beta_1 LDIST_{i,t} + \beta_2 CAP_{i,t} + \beta_3 MAXP_{i,t} + \beta_4 LMDIST_{i,t} + \beta_5 CON_{i,t} + \varepsilon_{i,t}\,. \tag{1}$$

The coefficients, β_1, β_2, and so forth, measure the average effect on bids (relative to the engineering estimate of cost) of each factor.[16] The residual, $\varepsilon_{i,t}$, captures the variation in bids not explained by the factors included in the regression. Some of this unexplained variation in bids is due to private cost information not observable to either the econometrician or other firms.

In equation (1), there are too few data points to include all the variables that measure factors affecting the costs of all competing firms. The example in Part II.C. suggests that the lowest-cost rival will play in the most important role in determining the market prices. Using advanced computational techniques, Bajari and Ye find that this is also true in more general models of bidding. Therefore, they include the maximum free capacity among all rival firms and the distance of the closest rival firm as controls. Their estimates for this regression model are displayed in Table 7; t-statistics are contained in parentheses.[17]

All of the estimated coefficients in Table 7 have intuitively plausible signs. The coefficient on $LDIST_{i,t}$ is positive—on average, the greater the distance of the firm to a project, the higher its bid—which makes sense because increasing distance from the job increases transportation costs. The coefficient on $CAP_{i,t}$ is also positive. As a firm utilizes more of its available capacity, the opportunity cost of winning a job today increases

Table 7
Estimated Bid Functions

Variable	*Coefficient*
Constant	.68 (5.95)
$LDIST_{i,t}$	.040 (3.45)
$CAP_{i,t}$	.17 (8.51)
$MAXP_{i,t}$	.026 (.71)
$LMDIST_{i,t}$	.24 (1.81)
$CON_{i,t}$	–.59 (–1.87)
Sample Size	450
R^2	.85

[16] Bajari and Ye divide by the engineer's estimate to correct the regression for what econometricians refer to as the heteroskedasticity problem.

[17] One can reject the hypothesis that the coefficient equals zero and the factor had no effect on bids with at least 90% confidence if the t-statistic exceeds 1.65 and with at least 95% confidence if the t-statistic exceeds 1.96.

Tables

- Engineers cost estimate explains large fraction of bids
- Ratio of bid to cost estimate is one
- Ranking of distances and capacities reflects costs
- Firm size skewed
- Political boundaries are important
- Regression coefficients sensible
- Bids well explained by cost controls

Testing

- $BID_{i,t}/EST_t = \beta_{0i} + \beta_{1,i}\ LDIST_{i,t} + \beta_{2i}CAP_{it} + \beta_{3i}MAXP_{it} + \beta_{4i}LDIST_{it} + \beta_{5i}CON_{it} + \varepsilon_{it}$ for 11 largest firms in the industry
- Include project and firm fixed effects.
- Restrict regression coefficients to be the same for “fringe” firms.

Testing

- They find, using standard confidence levels, that the bids are conditionally independent for all but four pairs of firms: (firm 1, firm 2), (firm 2, firm 4), (firm 5, firm 14), and (firm 6, firm 7).
- Only (firm 2, firm 4) bid against each other often.

Testing

- Do costs make sense?
- Test if $\beta i,k = \beta j,k$ for all pairs i,j.
- All firms, except for the pair firm 2 and firm 5 pass the exchangeability test.
- Bajari and Ye conduct 46 separate tests
- Forty-one of these tests fail to reject the hypothesis of competitive bidding at conventional levels of statistical significance.

Case Studies.

- The bidding patterns of firm pairs 2 and 4 and 2 and 5, however, do not appear to be consistent with those predicted by the Bajari and Ye model of competitive bidding.
- It is worth noting, as mentioned above, that the owners of these three firms were previously sanctioned for bid-rigging.

Estimate Markups

- Use econometric methods to estimate markups
- The industry insiders gave us a distribution over markups.
- Estimate markups for 3 competing models:

-no collusion

-(2,4) collude

-(2,5) collude

- No collusion model fits best

Table 8
Distribution of Percentage Markups of Winning Bids Implied by Models A, B, and C

Percentile	*A: Competitive Model*	*B: Cartel (2,4)*	*C: Cartel (2,5)*
10	1.229	1.273	1.14
20	1.597	1.818	1.82
30	2.077	2.422	2.56
40	2.536	3.201	3.43
50	3.329	4.126	4.47
60	4.227	5.434	5.84
70	5.692	7.54	9.30
80	10.0	16.21	17.56
90	23.81	33.54	58.26

have fifty years of experience in the industry. Since many firms do not survive in this industry, Bajari and Ye argue that the market survival of these sources indicates they probably have good insights into the overall distribution of costs and markups.[19] The two industry insiders had quite similar views, and Bajari and Ye averaged their beliefs to get the following distribution over markups.

Table 9
Beliefs of Industry Sources About the Distribution of Markups

Percentile	*Percentage Markup*
25th Percentile	3%
50th Percentile	5%
75th Percentile	7%
99th Percentile	15%

In the final stage of the statistical horse race, Bajari and Ye used methods from Bayesian statistics to calculate the probabilities that models

[19] Bajari and Ye report that the industry sources, in their own language, seemed to understand many of the intricacies of the Bajari and Ye model of competitive bidding. In fact, they found that the sources had insights into the competitive bidding model garnered through experience in the real world that had escaped the attention of leading economic theorists with whom they had spoken! The bidders had also spent a great deal of time thinking about their competitors' costs and monitoring the prices of inputs because this is an important strategic variable in determining their optimal bid. One of the industry experts with whom they spoke said, "I think some guys in the industry spend more time thinking about their competitors' costs than their own costs!"

Conclusions

- Academic work on testing for collusion, particularly bid-rigging
- This literature suggests a large number of warning flags and potential diagnostics
- Known examples of cartels fail to submit independent bids
- Bids also fail to make sense with cost controls
- Despite the limitations of any testing, why is doing nothing the optimal solution?