

Competition and the Cost of U.S. Infrastructure

Lindsey Currier¹

November 2025

[Please click here for the latest version](#)

Abstract

Can limited competition in procurement auctions explain the high, and rising, price of road infrastructure in the U.S.? I assemble a new dataset covering the near-universe of state highway auctions between 2002 and 2024. I first document thin competition: one- or two-bidder auctions account for a third of awards, and this share has risen over the past decade. Using spatial variation in inter-state bidder locations, I then estimate the average causal effect of competition on prices; an additional bidder reduces prices by ten percent. To decompose bids in the data into markups and production costs, I develop a semi-parametric structural auction model that incorporates bidders' uncertainty over the number of competitors they face. I show that price increases over the past decade are primarily attributable to increasing markups, not increasing production costs. Limited competition, in turn, is consistent with patterns generated by fixed costs of entry, but not broad construction-sector fixed costs. Embedding the markup estimates in an entry model, I estimate large auction and market entry costs, consistent with an important role for procurement complexity and regulatory barriers.

¹I am grateful to Ed Glaeser, Gabriel Kreindler, Ariel Pakes, Robin Lee, and Elie Tamer for their guidance and support. I thank Shotaro Beppu, Cody Cook, Bnaya Dreyfuss, Andrew Kao, Sylvia Klosin, Aakaash Rao, Cailin Slattery, Winne Van Dijk, and participants at the Harvard IO seminar, the Urban Economics Association meeting, and the Volpe National Transportation Systems Center for helpful discussions. I acknowledge generous support of the NSF GRFP, as well as the Chae Family, the Ross Garon, and the Griffin Economics Research Fund. This draft is a work in progress and comments are welcome; all errors are my own. Department of Economics, Harvard University. Email: lcurrier@g.harvard.edu.

1 Introduction

American infrastructure is remarkably expensive. For example, the per-mile price paid by the U.S. federal government to construct transit lines can be up to double that paid by peer countries.¹ Despite extensive public attention, the underlying causes of these high prices remain a puzzle.² One longstanding view is that high prices reflect high production costs, driven for example by high labor costs or low productivity (Kessler and Katz, 1999; Goolsbee and Syverson, 2023). However, since governments typically procure infrastructure from private firms that compete for public contracts, often via auctions, high prices may also be driven by limited competition and resulting firm markups over production costs. Indeed, this view is notably present in free-response answers from transportation procurement officials in a survey by Liscow et al. (2023), where the most commonly cited cost driver is “Competition,” followed by “Materials and Labor.” Disentangling the contributions of markups versus production costs is vital to increasing the direct return on infrastructure investment, and thus realizing the well-documented benefits of better infrastructure on trade, productivity, labor markets, commuting and mobility, and consumption.³

At least three factors make it challenging to disentangle these channels in practice. First, fragmented procurement systems across the country have resulted in scarce data, such that even simple statistics — such as the number of bidders per procurement auction — are difficult to obtain. Second, infrastructure projects differ substantially in scope and complexity, making projects difficult to compare across years and across markets. Third, simple statistics do not map cleanly to conclusions about the effects of competition: estimating these effects requires understanding firms’ strategic behavior and their information sets — accounting, for example, for firms’ limited information about the number of competitors they face in any given auction. Ignoring these uncertainties biases estimates of markups.

¹The average per-km cost of U.S. projects is 33%, 103%, 200%, and 680% higher than for the U.K., Australia, Canada, and Spain, respectively, according to data compiled by the NYU Marron Institute. See Figure A.1 for a plot of the data.

²See Smith (2017) and Vartabedian (2021) for examples of news articles hypothesizing on cost drivers.

³For trade see Limão and Venables (2001), for productivity see Donaldson and Hornbeck (2016), Ghani et al. (2014), Baum-Snow (2014), Gibbons, Lyytikäinen, et al. (2019), for labor markets see Chandra and E. Thompson (2000), Michaels (2008), for commuting and mobility see Gibbons and Machin (2005), Glaeser et al. (2008), and for consumption see Gonzalez-Navarro and Quintana-Domeque (2016), among others.

In this paper, I overcome these challenges in the context of highway infrastructure auctions in the continental U.S. Nearly all major roadwork and bridges fall into this category, and this procurement method embodies the most common form of infrastructure procurement worldwide.⁴ I first assemble a novel dataset of 1.3 million project-level bids spanning two decades of auctions, and I use this data to document novel facts about costs and competition in road procurement: most importantly, that prices are rising and competition is strikingly thin and decreasing over time. I exploit spatial variation in inter-state bidder locations to estimate the average causal effect of competition on winning bids and find that an additional bidder reduces prices by ten percent. I then develop and estimate a semi-parametric structural model of auction bidding that is able to decompose bids into markups and production costs. The model accounts for firms' uncertainty over the entry decisions of other bidders and the government's reserve price, and is estimated with minimal functional form assumptions, leveraging the richness of the data. The model estimates reveal the role of markups in historical price trends, while a model counterfactual provides price reductions from hypothetical increases in participation. Estimated current and counterfactual markups additionally enable me to estimate entry costs, and thus speak not only to the *consequences*, but also the *causes*, of limited competition: I find an important role of bureaucratic costs of entering auctions and regulatory costs of entering new state markets.

I begin by assembling panel data on the near universe of auctions for state Departments of Transportation (DOTs) from 2002 to 2024. States are the principal governmental unit responsible for road construction, and the vast majority of projects are awarded through sealed-bid first price auctions. For each auction, I have a series of covariates, including a measure of size and a brief project description. Importantly, for a sub-sample of ten years, I have the bid-schedule for each project, which is the complete list of construction inputs. This data substantially extends prior sources, which in the empirical auction literature typically cover only a single state or a small handful of states.⁵

The data reveals new facts on the U.S. road procurement market, in particular rising prices and

⁴Of 40 countries surveyed by the World Bank, all primarily used design-bid-build methods with competitive bidding for traditional infrastructure investment (World Bank, 2020).

⁵Liscow et al. (2023) provide the first data spanning all states through the combination of survey and public record requests; however, the authors' bid analysis is limited to five winning bids per state. My dataset additionally captures the full distribution of bids across the continental U.S.

weakening competition. I find that over the past two decades, inflation-adjusted winning bids have more than doubled. The trend persists through a variety of sub-samples of the data, such as restricting to similarly sized projects, and increases in the price of important inputs do little to explain it. By contrast, within-state price trends from two-way-fixed-effects regressions are correlated with changes in variables related to market structure, such as the number of bidders. To my knowledge, this is the first academic study of 21st-century price growth. These results extend Brooks and Liscow (2023), who show the Interstate Highway System spending per new mile tripled in real terms in the 1960s–1980s, finding continued growth despite the shift from new construction to reconstruction.⁶

To probe the mechanisms underlying high prices, I start by providing simple measures of competition. I document that, nationwide and over time, transportation auctions have attracted only a small pool of bidders. The median auction draws just three participants, and one- or two-bidder auctions account for about a third of all projects. Since 2010, the share of auctions with 1–3 bidders increased from 47% to 63% by 2024. Declining participation occurs against a backdrop of minimal market entry and stable demand, despite the billions of dollars allocated towards transportation infrastructure in 2021 Infrastructure Investment and Jobs Act (IIJA). Together, these statistics provide a newly comprehensive picture of bidder participation in transportation procurement.

Given the challenges of quantifying the impact of the number of bidders on prices, I approach this question several times in this paper. I start by looking at gap between an the first and second lowest bids in an auction, which provides a simple measure of the loss to government if the first bidder were to drop out. I find this bid gap declines substantially and convexly as the number of bidders increases, revealing to competition hinge, intuitively, on the baseline number of bidders. However, this measure would be a naive estimate of the magnitude of those returns: it would presume which bidder is marginal, a *composition* assumption, and hold other bids fixed, a *strategy* assumption. The composition and strategy effects together determine the causal effect.

Consequently, the second contribution of this paper is to directly estimate the average causal effect of an additional bidder on auction prices. Theoretical auction models often predict large returns,

⁶Brooks and Liscow (2023) show suggestive evidence that the 20th century price growth is caused by permitting costs and citizen voice frictions, which are likely to be more binding for brand new construction than for the reconstruction projects that make up modern roadwork.

but direct empirical evidence is more limited and more mixed.^{7,8} This is at least in part due to the difficulty in finding exogenous shifts in the number of bidders in real markets. The primary identification challenge is selection into auctions, which induces a positive correlation between the number of bidders and the error term in the price equation. I address this with an instrumental variables strategy that exploits two sources of variation. I identify inter-state entry by firms based outside the market of interest; such entry is rare but lumpy, and when it occurs firms typically enter many auctions, creating plausibly exogenous and salient variation in local competition. I interact state entry with the distance to the out-of-state bidder's establishment, as distance is a strong predictor of a firm auction entry. My estimates suggest that on average an additional bidder lowers the winning bid by 10%.

I then turn to a structural auction model of bidding in order to decompose the prices paid by the government into two pieces: markups and firms' variable production costs. The model allows me to (i) estimate the heterogeneous effect of additional bidders, for example, across different baseline number of bidders; (ii) estimate whether declining competition can explain the historical rise in prices; and (iii) quantify counterfactual returns from increased bidder competition in the future.

I build on an independent private value auction model, incorporating key features of the setting. In particular, I account for the fact that while firms have information on the probability distribution of the number of bidders in an auction, they do not know the exact number.⁹ I model the government's discretion to reject bids as a secret reserve price, where again, bidders have information on the distribution but not the realization. The benefit to the researcher of the auction setting is that, under independence assumptions and information assumptions like the above, a profit maximizing firm has an optimal bid given their cost draw. I take advantage of this to identify the unobserved production costs from observed bids, with minimal functional form assumptions.

I estimate the model for all roadwork auctions in the continental U.S. between 2014 and 2023, the period for which I have the complete bid-schedule data for each project. I split the estimation by

⁷For example, Bulow and Klemperer (1996) show that under standard assumptions, increasing the number of bidders by one benefits the seller more than adopting the optimal mechanism would.

⁸For two papers on this topic, see Coviello and Mariniello (2014) and Lewis-Faupel et al. (2016).

⁹A few papers incorporate uncertainty over the number of bidders, but differ in assumptions. For instance, Li and Zheng (2009) and De Silva and Rosa (2024) model uncertainty by explicitly modeling potential entrants in mixed-strategy entry equilibria, in particular relying on parametric cost distributions.

state, year, and a covariate-based cluster, allowing trends over time and across states to be fully flexible. Joint auction estimation within samples requires accounting for project heterogeneity, which I address by “homogenizing” bids to remove the effects of project covariates, as in Haile et al. (2003). Building on the traditional linear regression method, I use machine learning on covariates and absorb residual unobserved heterogeneity with an auction random effect.

The structural estimates indicate substantial markups. The mean markup is around 0.2. This corresponds to an average per-auction margin of a bit more than \$0.7 million. A back-of-the-envelope calculation yields a total of roughly \$25 billion spent above production cost over the ten year period. While most states exhibit sizable markups, there is large dispersion across states: moving from the 25th to the 75th percentile raises the markup by 37%.

I use the model results to examine whether the rise in price between 2014 and 2023 can be explained by decreasing competition. I find that roughly one third of the 28% realized price increase over the ten year period is accounted for by the project observables, suggesting that even within project type, projects are becoming more complex. More complex projects require more inputs in total and may require novel inputs, such as environmental adjustments. Of the remaining two-thirds growth that is real price growth, costs explain little, rising only 8%. This bounds growth on all variable costs to the firm, including materials, labor, and opportunity costs. Increasing margins explain the majority of the real price growth, rising by 45% over the period. Changing competition, rather than production costs, appears to be the defining feature of infrastructure over the last decade.

How much could the government save with higher bidder participation? I solve for counterfactual prices in the model using the most recent distribution of costs but under the hypothetical guarantee of one additional bidder per auction. The counterfactual is able to account for the current distribution of number of bidders — an advantage over the reduced form analysis — and decomposes the total effect into the composition effect and the strategy effect. I find an average savings of \$350,000 per project, remarkably close to the 10% estimated in the reduced form exercise, although the identifying assumptions have no overlap. The total effect is driven primarily by the strategic effect, suggesting more productive entrants is a limited channel for price reduction.

Markups are large and persistent, yet firms do not enter auctions and drive down these variable profits. In the final section, I present evidence that fixed costs are substantial, deterring entry. I rule

out construction capital requirements, instead showing evidence of significant entry barriers at both the auction and the state market level. These barriers are consistent with procurement requirements playing a central role in limiting participation, in line with views that decades of accrued rules have led to policy congestion (Klein and D. Thompson, 2024). Entry costs also imply returns to market scale: both larger auctions and thicker demand would reduce markups.

If policymakers aim to address entry costs, it is useful to know where the bottlenecks lie. Both measuring regulation and unwinding layered, implementation-heavy rules to identify key ones are complex (Trebbi and Zhang, 2022). Instead, I estimate whether market entry or auction entry costs are larger. To do so, I combine the results from my bidding model with an entry model and, for estimation, exploit the fact that firms enter only when expected total profits are nonnegative. The model builds off of Berry (1992) and, while more stylized than the nested bidding model, yields a straightforward parameterization of auction entry costs. I find sizable bureaucratic costs of entry into auctions, with regulatory market costs of entry into state markets even larger. The results suggest policy that can lower barriers would have meaningful impacts on competition and prices.

Related Literature

This paper's main contribution is to shed light on the competitive structure of the United States highway infrastructure market and its consequences. Other work seeking to understand infrastructure costs includes Brooks and Liscow (2023) and Mehrotra et al. (2024), finding evidence that citizen voice and material input costs contributed to higher costs in the late twentieth century and in 1984–2008, respectively. Most related to this paper is Liscow et al. (2023), who combine survey data with public records requests and finds that low state capacity plays an important role in increasing project costs. In alignment with this paper, the authors also report suggestive evidence that fewer bidders results in higher prices. While the aim of this project is to quantify the competition channel, state capacity may operate in the background, affecting either firms' production costs or their entry decisions. Looking outside the U.S., Kirchberger and Beirne (2021) study a similar issue with international micro-data on input prices in the construction industry, showing that limited competition in cement generates substantial markups.

Several papers investigate distinct but intertwined questions about infrastructure. Goolsbee and Syverson (2023) study the deterioration of productivity in the U.S. construction sector from a

macro-accounting perspective, an issue that compounds rising costs. Viewed through the lens of productivity growth in their paper, the question studied here becomes why infrastructure prices have not fallen over time. Currier et al. (2023) study the quality of local road infrastructure and estimate the distribution of the costs of poor road quality across places. Kroft et al. (2025) study the relationship between construction production markups and monopsony power over construction workers; they capture the public market contribution with auction data from 28 states. Finally, Fajgelbaum et al. (2023) study the political incentives surrounding California’s high-speed rail, a famously expensive megaproject, comparing welfare quantifications from a spatial model to observed voting patterns.

To model procurement auctions, I build on a rich empirical literature. Since Paarsch (1992), economists have inferred bidders’ valuations from observed bids using the structure implied by optimal bidding strategies. Of particular relevance is the work on transportation infrastructure auctions, which includes the following. Bajari (1997) analyzes bidder asymmetry in Minnesota roadwork auctions. Hong and Shum (2002) examine winner’s curse in New Jersey procurement. Jofre-Bonet and Pesendorfer (2003) estimate a dynamic capacity-constrained model in California. De Silva, Dunne, et al. (2003) and De Silva, Kosmopoulous, et al. (2009) document incumbent–entrant asymmetry in Oklahoma, while Li and Zheng (2009) study entry in Texas mowing contracts. Krasnokutskaya (2011) account for unobserved heterogeneity in Michigan highway auctions, and Krasnokutskaya and Seim (2011) evaluate small-firm preference policies in California. Lewis and Bajari (2011) analyzes scoring auctions in California; Bhattacharya et al. (2014) quantify entry-rights effects in bridge building in Oklahoma and Texas; and Jeziorski and Krasnokutskaya (2016) study subcontracting. Balat (2017) examine ARRA impacts on capacity and prices in California. Bolotnyy and Vasserman (2023) evaluate scaling auctions for bridge maintenance in Massachusetts. De Silva and Rosa (2024) link the Great Recession’s private-sector downturn to lower Texas road prices, and Ito (2024) show that firms in Montana signal entry intent through online questions. These papers, and others, elucidate information, behavior, and optimal mechanism design in procurement auctions. I build on this work with a parsimonious model that preserves key features and limits functional-form bias. I also introduce an adjustment for bidder uncertainty over the number of competitors in a given auction. To my knowledge, no paper yet has collected national data and used the auction structure to evaluate markups in the U.S. as a whole.

2 Setting and Data

2.1 Transportation Infrastructure Market Structure

Major roads and bridges in the U.S. are the responsibility of state governments.¹⁰ In a regular year, the U.S. already allocates an annual budget of \$120 billion solely for the maintenance and development of roads and bridges.¹¹ Recent attention on infrastructure has led to additional funding, such as the \$1.2 trillion Bipartisan Infrastructure Investment and Jobs Act (IIJA) of 2021.¹² Nearly all roadwork is procured via auctions. State-managed projects procured by different means includes large projects requiring technical innovation, whose design is usually outsourced.¹³ The typical auction should be thought of as mid-sized, routine project involving limited design innovation, such as the reconstruction of a segment of highway.

The vast majority of projects are completed through the following process.¹⁴ First, project specifications are designed in-house by the state Department of Transportation (DOT). Then the contract for completing the project is awarded to a construction firm via a sealed-bid first price scaling auction. In a scaling auction, bidders are given detailed project specifications and a list of input item quantities, known as the bid-schedule. A bid consists of a price per bid-schedule item. The total price of the project is calculated, and the contract is awarded to the bidder with the lowest total.

Bolotnyy and Vasserman (2023) show that firms often bid strategically on item prices, underpricing items for which they expect the government to have overestimated quantities and overpricing those they believe were underestimated. As a result, item-level bids are an unreliable measure of input costs. In this project, I focus on the total bid, which determines firms' expected profit, and set aside the nested choice of pricing the individual items in the bid-schedule. Consequently, the auctions

¹⁰While the federal government allocates a substantial amount of funds to transportation infrastructure, these dollars are all reallocated to the states. Local governments are responsible for local roads, which account for 77% of the road system but only 13% of vehicle miles traveled (FHA, 2000).

¹¹Data from the Organization for Economic Co-operation and Development (OECD) Infrastructure Investment Survey (2024).

¹²Roads and bridges were the largest category of IIJA spending, with \$110 billion directly allocated.

¹³An example of this extreme is Boston's Big Dig project – the most expensive highway project in U.S. history – awarded through negotiation to Bechtel Corporation, a company with an annual revenue of \$16 billion and operate at a different scale than most roadwork firms.

¹⁴Alternative procurement methods account for fewer than 5% of DOT projects Liu et al. (2022).

can be treated as simple first price auctions. This allows me to focus on the role of competition in shaping total prices.

On the supply side, construction firms operate in the public sector, private sector, or both. Appendix Table A.9 profiles a random sample of five companies that compete for government contracts; their listed sub-industries span road construction, carbon capture, general building, and ship repair. All five also serve private clients. This makes sense as the private sector is larger overall: a back-of-the-envelope calculation using Kroft et al. (2020) suggests that public-sector participants account for about 22% of sales in the construction industry.¹⁵ Appendix Figure A.2 compares the distribution of sub-industries (four-digit SIC codes) for auction participants with those of all firms within two-digit SIC categories Building Construction and Heavy Construction.¹⁶ Single-Family Housing Construction alone accounts for 78% of all firms, reflecting both the dominance of this sector and the smaller scale of typical firms. Housing remains the largest category among auction participants, but the distribution is markedly flatter, with a similar number of firms in Highway and Street Construction. Firms classified under housing construction in the auction sample usually operate across several sub-industries, as illustrated by the examples in Table A.9.

Public sector construction is not equivalent to private sector jobs. For example, firms must comply with government rules and procedures, covering everything from wage rates and inspection protocols to the formatting of correspondence. These requirements are largely determined by states and vary across jurisdictions. Pre-qualification prior to bidding can be demanding: among required certifications and documents, submissions can include fully written safety programs, trainings documentation, and reference letters on past performance.¹⁷ Nearly all states also require proof of bonding capacity, a government-mandated form of insurance that is itself difficult to obtain.¹⁸ Once qualified, preparing an individual bid proposal remains itself resource-intensive: Liscow et al. (2023) report that the average submission is 164 pages of forms. If these pre-auction procedures are costly to firms, they may generate high equilibrium markups.

¹⁵Kroft et al. (2020) combine tax data on all U.S. firms with DOT auctions from 28 states. They estimate that firms participating in auctions represent 12% of sales among all firms coded as construction under NAICS.

¹⁶Data details are provided in Section 2.2 and Appendix B.2.

¹⁷For examples, see Maine and Georgia DOT prequalification documents.

¹⁸See Congressional Research Service Report “SBA Surety Bond Guarantee Program” Updated July 18, 2025 for a discussion of the difficulty in obtaining surety bonds, particularly to small businesses.

2.2 Data

Auction data My main data is state Department of Transportation procurement auction bids for all contiguous U.S. states. Observation start dates for states range from 1993 to 2007, but all are observed until 2024. I define the analysis window to begin in 2002, when 85% of states are observed.¹⁹ The data come from a construction data software company; data compilation details, state idiosyncracies, and summary statistics of variables by state are reported in Appendix B.1.

In total, I observe 370,000 auctions with 1.3 million bids. An important feature of the setting is that I observe all bids, not just winning bids, which is a common limitation of auction datasets. The data covers the allocation of nearly 1.4 trillion (real) dollars spent by the U.S. government, accounting for roughly half of total U.S. road-related infrastructure spending.²⁰ The main excluded spending is roadwork conducted by local governments, primarily for local street repairs.

The data include winning bids but exclude any changes in price due to renegotiation. To gauge the magnitude of these changes, I collect separate data on change orders and final costs for four states. Appendix Table A.1 reports mean and median cost overruns. The largest median percentage overrun is 3.25% in Colorado, while the smallest is -3.87% in New York, where projects actually come in under budget on average. Although renegotiation matters, baseline winning bids remain first-order.

The data also include firm names, which I use to match firms across states.²¹ The details of this match process are in Appendix B.2. Additional variables included in the data are the project date, the county of the project, the engineering type (defined as the input category with the largest share of spending), and “tons,” the sum of all amounts of bid-schedule items measured in tons. The latter is the best available measure of project size. I have a brief description of the project, which I use to assign each project to one of nine categories with a large language model (LLM).²² Details on the classification procedure are in Appendix B.3.

Finally, I have the full set of project bid-schedule items from 2014-2023. This is the complete list of

¹⁹When relevant, I restrict to a balanced panel or report balanced panel results in the appendix.

²⁰The data documents \$66 billion spent in 2023 via DOT auctions, while all levels of U.S. government combined spent approximately twice as much, \$127 billion, on road-related infrastructure, according to the OECD survey “Investment Spending in Transport Infrastructure”.

²¹I assume each firm name uniquely identifies a firm. This assumption appears reasonable based on the validation exercise. The main challenge arises from cross-state formatting differences.

²²Implemented via OpenAI’s API; see <https://platform.openai.com/>.

input items, units, and quantities for every project. The items are highly specific; an example item might be “Galvanized reinforcement steel”, or “Hot-mix asphalt base course 10in”. Item identifiers and descriptions are unique to each state, allowing me high specificity in controlling for project characteristics within states over this decade of data. This project bid-schedule data is the main dataset used for the model estimation.

Additional data I obtain data on firm locations and start years from Dun and Bradstreet, a commercial provider of firm-level data and analytics. I obtain all firm from their database with SIC code 16, “Heavy Construction, Except Building Construction Contractors”, and all firms with SIC code 15, “Building Construction - General Contractors and Operative Builders.” Appendix B.2 provides details on the matching procedure. Finally, I obtain data from the Census on demographic variables across counties, and from the Bureau of Labor Statistics on price indices.

Summary statistics Table 1 reports summary statistics. Panel A summarizes the data at the auction level, starting with the project shares by the five most common engineering types. Asphalt jobs are the most common, followed by bridges and grading work. The most notable statistic is that the median number of bidders is three, a key fact discussed in Section 3. The variation is large, with a standard deviation of 2. Panel B summarizes the data at the state–year level, showing that the average state has nearly 300 auctions per year, though the distribution is skewed: on average 79 auctions account for 80 percent of annual spending. The average state–year has 145 distinct bidders and 84 distinct winners, a relatively large number given the small average number of bidders per auction.

Panel C reports bidder-level statistics. On average, firms appear for eight years, bid in 51 auctions, and win 14. Firms typically bid in three different types of auctions — for the most part they operate generally instead of being specialized. Thirty-seven percent never win at all, suggesting many test the market but do not stay. However, these firms account for only 2% of all bids, so they are not consequential in the market at large. I identify 12% of firms in the data as operating in multiple states. Because identifiers are at the state level, I conduct this match using both name matching and a large language model; detail and validation results are in Appendix B.2.²³ Among firms that ever win, the interstate share is 18%. This share is low given that, aside from the Rocky Mountains and a

²³My matching approach minimizes false positives and validation implies about a 10% under-count, so a corrected estimate is that 13% of firms are multi-state.

Variable	Statistic	Value
Panel A. Auction statistics		
Engineering Type: Asphalt	Percent	41.7%
Engineering Type: Bridge	Percent	18.9%
Engineering Type: Grading/Excavation	Percent	14%
Engineering Type: Uncategorized	Percent	13.6%
Engineering Type: Concrete Misc.	Percent	2.4%
Tons awarded	Mean	13765
Bidders per auction	Mean	3.63
Bidders per auction	Median	3
Bidders per auction	Std. dev	2.14
Panel B. State statistics		
Auctions per year	Mean	294
Auctions per 80% yearly spend	Mean	79
Bidders per year	Mean	145
Winners per year	Mean	84
Panel C. Firm statistics		
Auctions	Mean	51.37
Wins	Mean	14.12
Years active	Mean	7.54
No. types	Mean	3.01
Never win	Percent	36.9%
Multi-state	Percent	12.4%

Note: Panel A presents summary statistics at the auction level. The five types shown are the five most common auction types. Tons is the total sum of all bid-schedule items that are measured in tons. Panel B averages across each state year. Panel C presents statistics at the firm level. The last two rows gives the percent of firms satisfying the variable.

Table 1: Summary Statistics

few other ranges, there is no natural reason a state border to segment the market.

3 Empirical Facts

Infrastructure Prices Keep Rising I begin by examining levels and trends in the key variables of interest, price and competition. Starting with price, I find that over the past two decades, the winning bid for highway infrastructure roughly doubled, as shown in Figure 1. The red line shows the national median winning bid for all roadwork projects, which rose from \$1 million in 2002 to \$2.1 million in 2024, an increase of 107%.²⁴ The plot accounts for inflation by dividing by CPI, so this doubling is in real dollars. This pattern extends the increase documented by Brooks and Liscow (2023) for the late twentieth century, and shows continued price growth in the 21st century.

²⁴I define roadwork as projects classified as asphalt or concrete pavement, or with descriptions containing specified roadwork keywords (see appendix B.1).

It is possible the rise in winning bids could be caused by a compositional shift towards a different class of projects. This would constitute a rise in expense, but not a true price increase. To examine this, I plot the trend for three additional subsamples of the data, corresponding to more groupings of more homogeneous projects. The orange line shows the trend for asphalt projects, the subset of roadwork in which asphalt work is the primary expense.²⁵ The yellow line shows the trend for bridge projects, which is a separate category. Both groups have similar increases for winning bids, growing by 134% and 96%, respectively, over the full period. The final line (light blue) turns to project size, plotting the median winning bid for projects between the 45th and 55th percentile of native tons, the best proxy for size available in the data.²⁶ An advantage of tons is that it captures real differences between projects covering the same length of road with different intensities of reconstruction. A disadvantage is that it only captures inputs measured explicitly in tons. In Appendix Figure A.3 Panel (b) I report a robustness check with hand-collected data for six additional states with a converted tons measure regardless of native units.²⁷ When controlling for native tons, the price increase is larger, at 150% over the period.

²⁵For example, road widening projects in which excavation and earthwork is the primary expense are excluded.

²⁶Native tons is the total tons of all inputs that are measured in the unit tons.

²⁷The converted tons measure converts all asphalt amounts to tons from their native units. I use standard conversions between volume, weight, and area for asphalt, described in detail in Appendix B.4. Because I only have bid-schedule data starting in 2014, I hand-collect additional data covering the full period for six states. While quite noisy, the measure appears compatible with the price increase.

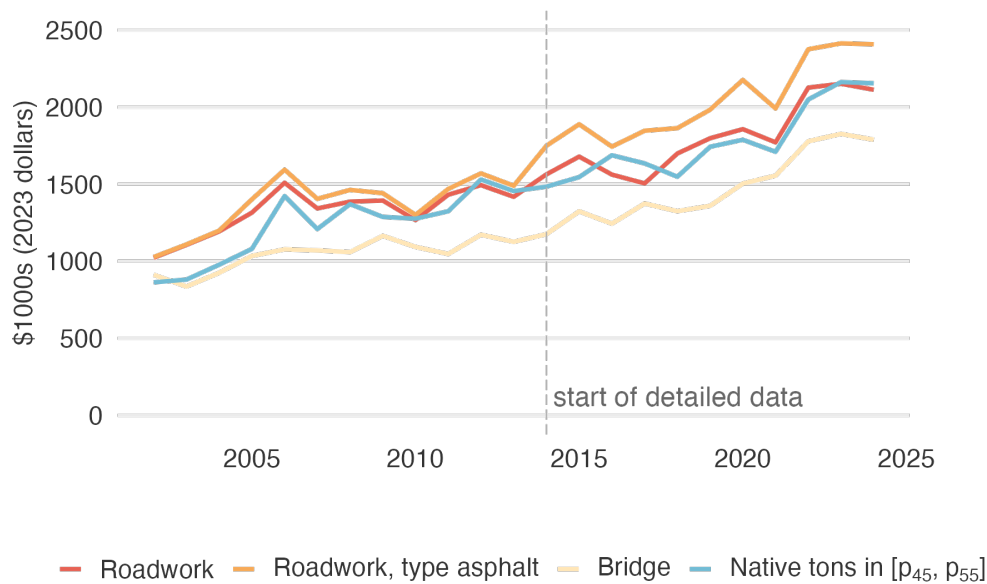


Figure 1: Winning Bids Are Rising Over Time

Note: Figure plots the median winning bid of DOT infrastructure projects between 2002 and 2024. All prices control for inflation and are in 2023 dollars. The blue lines show medians for projects in the 45th to 55th percentile of size, defined as the total tons of input items measured in tons. The percentiles are defined separately for each state. The dashed vertical line marks the beginning of the detailed data, indicating the period covered by the structural estimation later in the paper. Figure A.3 Panel (a) shows analogous trends for a balanced panel of states.

The rise in prices cannot be easily explained. Appendix Figure A.4 Panel (a) shows that the upward trend in prices is just as steep for projects in rural counties and located on flat terrain. Panel (b) shows that the construction PPI has remained flat over most of the period, and average construction wages over the period have not grown at all. Strikingly, the national unionization rate in the construction industry fell by 38% over the period, from 16.7% to 10.3%. Meanwhile, the price of crude oil rose on net since 2002, but fell substantially in 2011, completely at odds with the pattern of winning bids. Crude oil is both the main feedstock of asphalt and the main determinant of fuel and energy prices. Asphalt prices moved with crude oil through about 2010 — nearly doubling and then easing — consistent with the discussion of cost drivers in Mehrotra et al. (2024) and suggesting that asphalt likely contributed early on. Yet a back-of-the-envelope shows only about a 10% impact on total project prices, compared with the 53% increase I observe.²⁸ It cannot explain the later period, a fact

²⁸Calculation uses asphalt at a conservative 30% of pavement materials, materials at about 40% of project cost (Strassner and Moyer 2002, Table 4), and the 80% rise in asphalt prices.

corroborated by the PPI for highway and roads input goods, a new BLS series since 2015, which is very flat, except for a slight post-Covid rise.

Overall the evidence points to substantial real price increases not explained by shifts in project type, location, size, or, for the most part, input costs. I cannot rule out all production-cost drivers or changes in project complexity. Accordingly, in Section 4 I control for detailed project inputs using machine learning and then reverse engineer total production costs from observed bidding behavior, yielding an estimate that captures all cost increases borne by firms.

Price Increases across States are Correlated with Competition Measures I next turn to differences across states. Geography, precedent, and DOT institutional silos produce substantial variation in engineering design across states, confounding simple price comparisons. Unfortunately, engineer terminology in bid-schedules does not overlap across states, preventing the use of these covariates to control for design differences. Appendix Table A.3 illustrates the problem with California and Kentucky. Raw winning bids in California dwarf those in Kentucky by an average of \$9.2 million. However, the gap narrows by 82% to \$1.7 million after controlling for just a few coarse project design covariates.²⁹ To reduce the impact of omitted state-level design variables, I rely on the panel aspect of the data and focus on within-state differences. Later, I show that under plausible assumptions, markup level estimates are immune to this bias and can be compared across states.

Table 3 reports correlations between within-state changes in prices and changes in covariates of interest from two-way-fixed-effects regressions.³⁰ Notably, most of the input price, wage, and demographic covariates have no statistically significant correlation. This includes the coefficient on the state prevailing wage, a policy that has generated discourse (Kessler and Katz, 1999).³¹ This is perhaps because the Davis–Bacon Act, the federal prevailing wage law applying to all projects receiving federal funds, is the binding constraint. Regardless, while these production side variables show little correlation, there is a strong positive correlation between within-state increases in price and changes in competition-related measures. States that experienced relatively larger price increases also saw relatively larger declines in the number of active firms, the Herfindahl–Hirschman

²⁹Under the logic clarified by Oster (2019), instability of coefficients with respect to observables should raise concerns about selection on unobservables.

³⁰Appendix Table A.4 reports estimates from parallel regressions that omit state fixed effects. Because project designs differ across states, the resulting correlations are difficult to interpret.

³¹The coefficient is identified by six states that repealed their prevailing wage law over the period.

Index, and the average number of bidders per auction. These results are purely correlational, yet the magnitudes are non-trivial and motivate further investigation.

Covariate	Within-state correlation	
	Estimated coefficient	(SE)
Log avg. petrol product price	0.147	(0.267)
Log construction wage	-0.014	(0.021)
Pct unionized	0.006*	(0.003)
State prevailing wage law	-0.037	(0.062)
Log population	0.102	(0.126)
Log median HH income	-0.027	(0.025)
Log construction establishments	0.172	(0.258)
Log firms in market	-0.266***	(0.039)
Herfindahl–Hirschman Index (0–1)	0.489***	(0.162)
Log Avg. Bidders	-0.199***	(0.072)
State FE	yes	
Year FE	yes	
Observations (by row):	927, 927, 927, 815, 530, 927, 927, 927, 927	
Adj. R^2 (by row):	0.813, 0.813, 0.814, 0.813, 0.820, 0.813, 0.813, 0.822, 0.815, 0.815	

Note: Table reports estimates from a log-linear two way fixed effect regression of prices (log winning bids) on covariates. The first covariate is the log number of firms that win any auction in a given year, the second is the Herfindahl–Hirschman Index, and the final is the log number of average bidders per auction. Results are reported the unbalanced and balanced panel.

Table 3: Descriptive Evidence: Price and Competition Indicators

Auctions have Few Bidders I provide new summary statistics on competition in U.S. transportation infrastructure nationally. Figure 2 plots the share of auctions with one, two, or three bidders over time. Overall, most auctions attract few participants. Across the sample, the median is three bidders, and nearly one third of auctions have only one or two bidders. These numbers are more extreme than previous findings in the auction literature; seemingly, the states historically most willing to release data were those with above average competition, perhaps due to general higher administrative capacity.

The time trend in the number of bidders can be broken into two parts. The 2002 to 2009 period reflects the tight link between private sector and public sector construction. As the private construction sector surged with the housing boom, auction participation dropped, driving up the share of low-bidder auctions. With the onset of the Great Recession, this pattern flipped, as stimulus (and lower opportunity costs) boosted public sector construction.

In contrast, the post-2010 period is marked by a secular decline in bidder participation. In 2010, auctions with one to three bidders made up 47% of the total. By 2024, this share had risen to 63%.

Appendix Figure A.5 restricts to roadwork projects and shows a similar increase from a higher baseline: one to three bidder auctions accounted for 59% of all roadwork projects in 2010, and 72% by 2024. The share of roadwork auctions with just a single bidder increased by 40% over the period, from 13% to 18% of all roadwork projects. It is the consequences of this post-recession trend that the model section of this paper tries to understand.

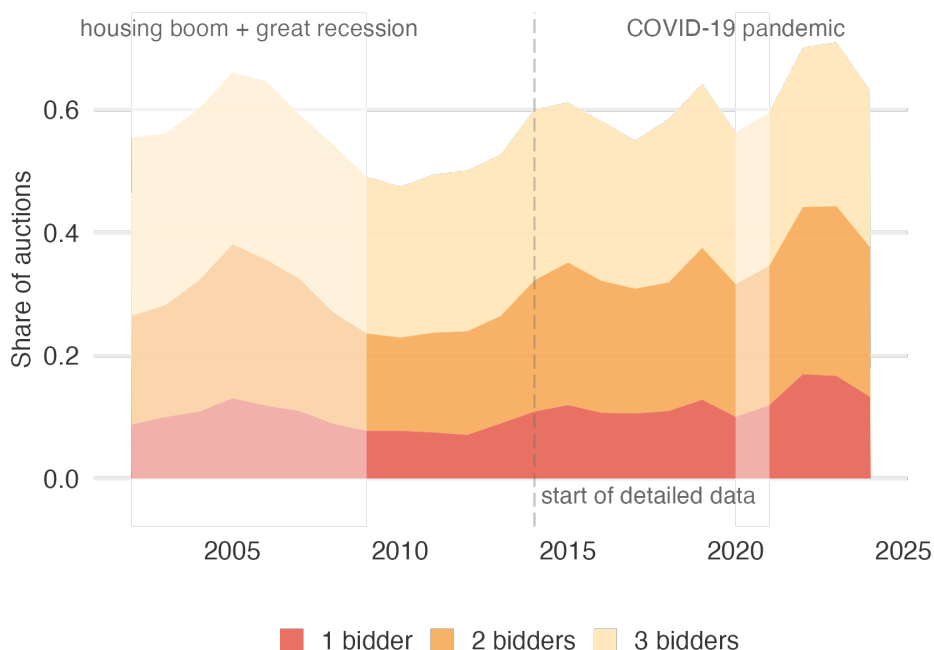


Figure 2: Macro-economic Effects and Decreasing Bidder Trend

Note: Figure plots the share of auctions by number of bidders from 2000 to 2024. The first shaded gray region marks the start of the housing boom and the end of the Great Recession, and the second marks the COVID-19 pandemic. The dashed vertical line marks the beginning of the detailed data, indicating the period covered by the structural estimation later in the paper. Appendix Figure A.5 Panels (a) and (b) plot analogous graphs for the balanced state panel and the roadwork only panel respectively.

One simple mechanism behind the post 2010 fall in bidders could be that infrastructure demand grew faster than firms could enter the market, resulting in fewer bidders per auction. This seems especially possible in the later period given the additional funds allocated to highway infrastructure by the IIJA in 2021. Indeed, entry into the market over the entire period is very limited. As shown in Appendix Table A.2, of all bids in a given year, only 2.5% are placed by a firm new to the market that year. This number increases slightly when also treating firms that cross state lines into new local markets as entrants, but only to 3.3%.³² However, the story of limited entry and rising demand

³²The rarity of entry is also inconsistent with a model of low entry barriers and high turnover, as one might expect in contestible market theory (Baumol et al., 1982).

cannot be right, because demand actually *did not* rise over the period. Appendix Figure A.6a shows that the number of projects auctioned by the government has been flat since 2010. Appendix Figure A.6b corroborates this by showing the amount of infrastructure (as measured in total tons) also did not increase over the period, although total spending did.³³ The cyclical trend in annual projects does correlate with the cyclical patterns in bidder participation, but there is no relation in the long run between demand and bidder decline.

First Bid Second Bid Gaps Decline with Bidders One simple indicator of the effect of the number of bidders on price is the gap between the winning bid and the second-lowest bid. This “bid gap” equals the loss to the government if the winning bidder disappeared and all else remained equal. Figure 3 plots the national median bid gap by number of bidders for all roadwork auctions over the sample period. Bid gaps are residualized by state and re-centered at the national average to account for omitted state-level factors that may shift both bidder counts and gap levels.

As the plot shows, the bid gap declines convexly as the number of bidders rises. Since by construction there is no such measure in one-bidder auctions, the largest estimated bid gap is for two-bidder auctions, where the median gap is nearly \$300,000 (2023 dollars). The gap falls steeply with additional entrants, leveling off around six bidders at roughly \$175,000. Appendix Figure A.7 shows the same pattern appears across different sub-samples, including asphalt projects and projects in the 45th–55th percentile of size.

³³The lack of an increase in projects built under the Biden administration may be surprising, but infrastructure output during this period remains debated. For discussion, see Zachary Liscow, “Highway investment probably didn’t go up under Biden,” *Briefing Book*, March 3, 2025.

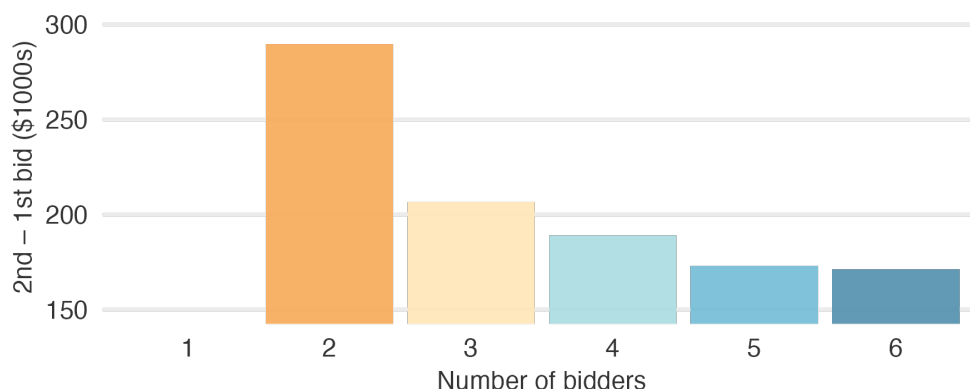


Figure 3: First Price - Second Price Bid Gap by Number of Bidders

Note: Figure plots the national median difference between the second-lowest and lowest bids by number of bidders in the auction for all roadwork auctions. To account for state omitted variables that affect both bid gaps and number of bidders, gaps are demeaned by state with the national mean added back. When taking means, 1% of the sample is trimmed due to large outliers. Appendix Figure A.7 plots analogous bid gaps for three different subcategories of auctions: bridges, asphalt only roadwork, and roadwork projects within the 45th to 55th percentile of native tons, which proxies for a constant size.

Unfortunately, the bid gap is not a measure of the price effect of a new bidder. For it to be so, one would have to make a strong assumption on the composition effect — the new bidder is the low bidder — and the strategy effect — the incumbent bidders do not adjust their bids. While the failure of these assumptions produces oppositely signed biases for the effect of a new bidder, both failures generate a downward bias for the change in the effect of a new bidder.³⁴ In other words, the decline in the bid gap as the number of bidders increases in Figure 3 is a conservative estimate the decline in the effect as the baseline number of bidders increases.

Direct Evidence that More Bidders Reduces Prices Exogenous shifts in the number of bidders are rare, which means that most of the evidence on how participation affects prices comes from theory or simulated counterfactuals. The quasi-experimental results that do exist have mixed results.³⁵ A small impact is consistent with theory if in practice the variance in costs across firms is small. In this case, any bid above cost is quickly undercut, keeping markups small even with few competitors.

³⁴If the first assumption fails, the bid gap overstates the savings to the government, and this failure is more likely when the baseline number of bidders is large. If the second fails, the bid gap understates the savings, and this failure is more likely when the baseline number of bidders is small.

³⁵For example, an Italian publicity reform was found to increase bidders by 9.3% and the winning rebate by 7% (Coviello and Mariniello, 2014). However, new e-procurement policies in India and Indonesia were found to increase participation meaningfully (increase of 0.4 bidders from a baseline of 2.9), though not statistically significantly, despite having a precise null (at least greater than -.02%) effect on prices (Lewis-Faupel et al., 2016).

Of course, theoretical predictions would also fail if firms systematically fail to bid strategically.

I directly estimate the average causal effect of the number of bidders on price, capturing the total sum of the composition and strategic effects. To identify the effect, I use plausibly exogenous variation in the number of bidders. I would like to estimate

$$\log(\text{price}_{jct}) = \alpha + \beta N_{jct} + \epsilon_{jct}, \quad (1)$$

where $\log(\text{price}_{jct})$ denotes the winning bid in auction j in year t and county c , and N_{jct} is the number of bidders in the auction. The parameter of interest is β , the average causal response to a unit increase in the number of bidders.

The primary challenge in estimating β is that N_{jct} is endogenous, since firms choose to enter auctions based on project characteristics that also affect price. This generates a positive correlation between the number of bidders and the error term, $\text{Cov}(N_{jct}, \epsilon_{jct}) > 0$. The resulting identification problem is formidable, since it requires finding an exogenous shifter of the number of bidders that does not also affect prices. A secondary challenge is that bidders do not observe the number of competitors *ex ante*. Consequently, I am interested in variation in N_{jct} that also impacts bidders' expectations, as would be consistent with a shift in the equilibrium distribution of N_{jct} , the change of interest. Unpredictable shifts, on the other hand, would likely underestimate the total effect.

To address these concerns, I first summarize at the county level, taking the average of log price, number of bidders, and time-varying covariates for each county-year.³⁶ County-level changes are more likely to be salient to bidders than auction level variation. I start by estimating the following within-county linear regression:

$$\log(\text{price}_{ct}) = \beta N_{ct} + \lambda x_{ct} + \gamma \tau_t + \mu_c + \epsilon_{ct}, \quad (2)$$

where τ is a time trend and μ_c is a county fixed effect. The variable x_{ct} is average native tons of inputs, my main proxy for size, described in Section 2.2. In order to identify the parameters of

³⁶I balance the panel across counties, which drops 36% of observations.

interest, this model needs to satisfy strict exogeneity.³⁷ While this specification may reduce the bias in equation 1, one is likely concerned that strict exogeneity is violated due to time-varying unobservables that are correlated with both county average price and number of bidders.

For this reason, my preferred approach uses an instrumental-variable (IV), constructed using distance between firms and auctions. Because transportation costs of equipment, materials, and labor are substantial, distance is a strong predictor of firm auction entry. Appendix Figure A.9 reports estimates from a linear probability model of auction entry on distance deciles. At distances under 3 miles, firms enter about 14% of auctions, whereas at 100 miles, the probability falls below 2%. I construct the instrument using firms for whom the given state is a secondary market to ensure their establishment location is exogenous to the affected auctions. Appendix Figure A.8 visualizes the variation used in the instrument, showing a single firm’s expansion over time. The expansion across the border is notably lumpy in time and the auctions entered are driven by distance, suggesting this instrument has a strong first stage.

Specifically, to construct the instrument, I identify each firm i ’s home state h_i as the first state they enter. For state s and year t , I define an out-of-state bidder as a firm with at least one bid in s during t whose home state is not s , i.e. if $|\text{bid}_{ist}| > 0, h_i \neq s$. I calculate the distance between i ’s location and the county of every auction in s during t . Details on the construction of the distance measure are in Appendix Section B.5. I construct a county level indicator for whether there is an out-of-state bidder within 100 miles:

$$N_{ct} = \gamma \mathbb{1}(\min_i(d_{ict}) < 100 \text{ miles}) + x_{ct} + \varepsilon_{ct}. \quad (3)$$

I choose 100 miles based on the entry and distance regression reported in Figure A.9, but report results with varying this threshold in Appendix Table A.5. For the instrument to be valid, I require (i) distance strongly predicts auction entry (relevance); (ii) firm locations are as-good-as random

³⁷Strict exogeneity requires that, conditional on the county fixed effect, the entire path of the error is independent of the entire path of the regressors: $(\epsilon_{c1}, \dots, \epsilon_{cT}) \perp (Z_{c1}, \dots, Z_{cT}) \mid \mu_c$, with $Z_{ct} = (N_{ct}, x_{ct}, \tau_t)$. I also present a similar exercise using a first-difference specification in the Appendix Table A.5. Under first-differences, unbiased estimation of β requires that $\Delta\epsilon_{c,t}$ is uncorrelated with the contemporaneous change in the independent regressors.

with respect to cross-border project proximity (random assignment); (iii) out-of-state firms affect prices only through the number of bidders (exclusion), and (iv) the instrument never reduces the number of bidders per auction (monotonicity).³⁸ I argue the second assumption is reasonable based on the observation that a firm's home state is typically its primary market (88% operate exclusively in one state). For firms that enter markets outside their home state, the average time until entry is 6 years. The geography of out-of-state auction opportunities is therefore less likely to impact firm location choices than in-state auction opportunities. Should this assumption still fail, the selection bias would shrink the magnitude of the estimated coefficient.³⁹ A potential threat to assumption (iii) would be if out-of-state bidders have systematically lower bids than state incumbents. Appendix Figure A.10 shows that out-of-state bidders actually have a slightly lower probability of winning an auction, suggesting higher bids, but the confidence intervals overlap between out-of-state and control bidders overlap.

Table 4 reports the estimated coefficients. Column (1) shows the ordinary least squares result yields a positive correlation between the number of bidders and price, consistent with selection bias. Column (2) shows that county fixed effects alone reduce the selection bias and flip the sign of the point estimate. Column (3) reports the first stage for the IV regression, showing that proximity to out-of-state bidding firms significantly increases the number of bidders in an auction. Switching from no out-of-state bidders within 100 miles to at least one raises the number of bidders on average by .15.

³⁸I follow the terminology of Imbens (2014).

³⁹An alternative reason for estimates to be biased toward zero is if panel transformations diminish the fixed-effect signal relative to measurement error, as emphasized by Griliches and Mairesse (1995).

	<i>Dependent variable:</i>					
	Log Bid	Log Bid	No. bidders	Log Bid	Log Bid	Log Bid
		(Within County)	(First Stage)	(IV)	(IV)	(IV)
No. bidders	0.071*** (0.004)	−0.014*** (0.003)		−0.125** (0.056)	−0.097* (0.050)	−0.152 (0.098)
dist. firm < 100mi			0.153*** (0.018)			
Log tons	0.096*** (0.002)	0.149*** (0.002)	−0.017*** (0.003)	0.135*** (0.002)	0.126*** (0.003)	0.135*** (0.002)
State FE			yes	yes	yes	yes
Time Trend		yes	yes	yes	yes	
County FE		yes				
Year FE						yes
Lag log bid		yes				
First stage partial F				76.26	102.79	27.48
Observations	22,050	22,050	22,050	22,050	13,461	22,050
Adjusted R ²	0.125	0.562	0.303	0.471	0.458	0.459

Note:

*p<0.1; **p<0.05; ***p<0.01

Note: Table reports coefficient estimates from equation 1 using the strategies outlined in equations 2 and 3. Column (3) includes a lagged outcome to account for dynamic bias in fixed effect panel models, which Klosin (2024) shows is greater than Nickell bias. Robustness checks to the distance threshold in the instrument are in Appendix Table A.5.

Table 4: Estimates of the Impact of Number of Bidders on Prices

Columns (4) - (6) present results from the instrument specification. All three specifications produce estimates of a similar and economically meaningful magnitude. Column (4) directly follows the specification in equation 3, while column (5) restricts the sample to counties within 50 miles of a state border and, reassuringly, estimates a similar, though slightly smaller effect, suggesting that the average causal effect of one more bidder is a roughly 10% decrease in price. To be conservative, this is my preferred estimate. Finally, column (6) replaces the time trend with year fixed effects; the estimate is not statistically significant, but the point estimate is similar and slightly larger in magnitude. Appendix Table A.5, Columns (2) to (5), vary the distance thresholds. While the smallest cutoff reduces statistical significance, overall the estimates remain consistent with an effect of approximately 10%.

These results say that, for the set of U.S. auctions whose number of bidders are affected by out-of-state entrants, on average one additional bidder reduced the government's payment by 10%. This leaves at least three questions open: (i) how do effects vary, in particular with the baseline number of bidders; (ii) can a decline competition explain the path of rising costs over time; and (iii) what are

the prospective savings if we were able to increase bidder participation in the future. To address these, I develop a model below.

4 Auction Model

I present a model of bidding in procurement auctions to decompose bids into production costs and markups. The model builds on an independent private values auction model, allows the government to reject bids through a secret reserve price, and addresses bidders' uncertainty about competitors with a probability distribution over the number of bidders. Given this setup, the structure imposed by the auction format determines firms' optimal actions, identifying the production cost with minimal functional-form assumptions, as shown in Guerre et al., 2000. The richness of my auction data enables to me preserve this functional form flexibility, using machine learning methods to control for project heterogeneity. I estimate a separate auction model by state, year, and clusters of similar auctions, in order to full capture patterns across states and over time.

4.1 Setup

I consider the problem faced by a firm bidding in an auction to win a contract to construct a road. Each project is heterogeneous and distinguished by project-specific covariates X . I omit auction subscripts to simplify notation; however, covariates, costs, and bids are always auction-specific. A firm's optimal bidding strategy trades off raising margins above its cost of constructing the road with its chances of winning the auction. I make the following assumptions.

- (A1) *Costs*. Bidder i has a cost c_i of completing the project drawn independently from a distribution with support $[\underline{c}, \bar{c}]$, and c.d.f. $F(\cdot | X)$. I assume $F(\cdot | X)$ is twice continuously differentiable on $[\underline{c}, \bar{c}]$ and its derivative is strictly positive on $[\underline{c}, \bar{c}]$, and $f(\bar{c}|X) > 0$.
- (A2) *Information I*. Bidder i does not know the total number of bidders N , but the probability distribution of N , $p_{N|X}(n) \equiv \Pr(N = n | X)$ for $n \in \{1, \dots, \bar{n}\}$ is common knowledge.
- (A3) *Information II*. Bidder i knows her own cost c_i , does not know her competitors' costs $c_{j \neq i}$, but

the distribution of all costs $F(\cdot | X)$ is common knowledge.⁴⁰

(A4) *Risk aversion.* Bidders are risk neutral and therefore maximize expected profit.

(A5) *Government reserve price.* The government sets a reserve price r for each auction such that bids exceeding r are rejected. The reserve price is secret, meaning it is unknown to bidders. Reserve r is independently drawn from the distribution $H(\cdot|X)$, where the distribution is common knowledge. Conditional on X , r is drawn independently from c_i .

Assumption (A1) rules out a common-value setting in which firms receive noisy signals about a shared underlying cost.⁴¹ Such an environment may generate the winner’s curse, where the winning firm systematically underestimates costs.⁴² As discussed in Section 2.2, empirical estimates of post-auction renegotiation are modest, with the median ranging from -4% to 3% for four sampled states. While not a test, if the winner’s curse were prevalent, one might expect substantial renegotiation to cover losses.

Assumptions (A2) and (A3) together reflect the reality that bidders have limited information on the competitors they face. While bidders may be informed about the pool of potential competitors, it is effectively impossible to know how many will ultimately bid. According to a survey by Liu et al. (2022), most DOTs publicly disclose no information on approved bidders before auctions. For the remainder that do, these lists are often large and bidders can at most form a prediction. As shown in Table 1, on average there are 145 unique bidders and 84 unique winners per state per year. Firms do not partition neatly across sub-markets, as the majority of firms bid on multiple project types and in multiple locations: the average firm bids in 11 counties; among firms with at least 10 bids, the average is 25 counties.

Assumption (A4) assumes firms maximize their expected profit. Assumption (A5) models the

⁴⁰It would be possible to add firm heterogeneity in the cost distribution. It does not matter in this case, as firms do not know who their competitors are when bidding, so their strategy is only a function of the distribution over all possible firm types.

⁴¹A1 also assumes finite (bounded) support and positive mass at the upper endpoint $\bar{c}$, which provide the initial point needed to apply the fundamental theorem of ODEs to prove a unique equilibrium bid function, as in Maskin and Riley (2000). I conjecture that the argument here follows by adapting their Proposition 2.

⁴²Alternatively, it may induce bidders to shade their bids conservatively, as Hendricks et al. (2003) provide evidence of in offshore oil lease auctions. In that setting, the value of a lease depends on the uncertain size of subsurface oil reserves. In contrast, roadwork projects have far fewer unknowns, as they are standardized and have detailed specifications provided in advance.

government's rejection of bids judged "unreasonable," i.e., bids that fail screening heuristics. Most often, these rules incorporate a percentage markup over the DOT's engineering estimate, along with other factors (see Liu et al. (2022)). Other papers have differed in their approach to modeling this feature, often treating the government as an additional bidder, or ignoring it entirely.⁴³ As more than 10% of auctions in my data have a single bidder, it is necessary to directly account for the threat of rejection.

I model the government as having a secret reserve price r with distribution $H(r) = F_b(r/1.15)$, where F_b is the distribution of bids. This assumption mimics the common screening threshold of 1.15 times the engineer's estimate. Appendix Figure A.11 shows that, across the 18 states with engineer's estimates in the data, the bid distribution closely matches the distribution of engineers' estimates. This is not surprising, as past bids are the main information engineers have with which to form their estimation. While imperfect, this assumption offers a parsimonious representation of a partly qualitative process and avoids modeling the government as a strategic bidder. It is conservative: in practice, DOTs are often hesitant to reject bids, so any bias in markup estimates is likely to be downward.

Each firm i in the auction submits a bid b_i that maximizes its expected profit. Firm i 's profit conditional on winning is simply the difference between the bid and their production cost. Firm i 's expected profit prior to the outcome of the auction is

$$\mathbb{E}\pi_i = (b_i - c_i) \Pr(b_i : b_i < \min(\{b_j : j \neq i\}, r) \mid X). \quad (4)$$

While in general no closed form solution exists for a standard first-price auction with a secret reserve price, there exists a Nash equilibrium bidding strategy β , which is symmetric and increasing in costs.⁴⁴ Consequently, the profit condition for bidder i can be written as

⁴³For example, Li and Zheng (2009), Bajari et al. (2014), and Krasnokutskaya and Seim (2011) assume no reservation price for Texas and California DOT auctions, while Bhattacharya et al. (2014) assume a publicly observed reserve price set at 1.5 times the engineer's estimate for Texas auctions.

⁴⁴I assume this is the unique equilibrium, following Maskin and Riley (2000), which provides the uniqueness proof for a similar class of auctions.

$$\mathbb{E}\pi_i = (b_i - c_i) \sum_n \left[\Pr(N = n | X) \left(1 - F\left(\beta^{-1}(b_i)|X\right) \right)^{n-1} \right] \cdot (1 - H(b_i|X)),$$

relying on the symmetric increasing bid strategy to replace the probability firm i wins with the probability all firms other than i draw costs lower than i 's cost, and to write the inverse bid function. The first order condition with respect to the bid defines the bidding strategy as a differential equation:

$$\beta(c_i) = \underbrace{c_i}_{\text{variable cost}} - \underbrace{\frac{\Pr_{\text{win}}(b_i)}{\Pr'_{\text{win}}(b_i)} \Big|_{b_i=\beta(c_i)}}_{\text{firm margin}}. \quad (5)$$

The derivation is straightforward and is in Appendix D.1 with the explicit equation. Here and throughout the paper, I define *the margin* as the difference between the bid and the marginal cost and *the markup* as the margin over cost. The margin decreases as the probability a high number of firms enter the auction grows. The amount it decreases is mediated by the underlying distribution of costs. To understand the importance of the cost distribution intuitively, consider a degenerate distribution where all firms have the same cost. In that case, the equilibrium strategy is to bid one's true cost — the margin is 0 if there are at least two bidders. On the other hand, if there is variation in cost draws, an additional bidder increases the magnitude of $\Pr'_{\text{win}}(b_i)$, pushing down the margin.

4.2 Identification and Parameterization

Semi-parametric Identification

My identification argument closely follows Guerre et al. (2000). Let $F_b(\cdot)$ denote the distribution of bids. The unique equilibrium strategy means that $F_b(b_i) = F(\beta^{-1}(b_i))$, using the equality $b_i = \beta(c_i)$. The density of bids is therefore $f_b(b_i) = f(\beta^{-1}(b_i))/\beta'(\beta^{-1}(b_i))$. Substituting these terms into the bidding equation allows for an expression in terms of observables:

$$c_i = b_i - \frac{S(b_i) (1 - H(b_i))}{S(b_i) h(b_i) - S'(b_i) (1 - H(b_i))}, \quad (6)$$

where $1 - H(b_i)$ is the reserve component, $S(b_i) \equiv \sum_n \Pr(N = n) A(b_i)^{n-1}$ is the rival response term, and $A(b_i) \equiv 1 - F_b(b_i | X, p_{N|X})$ is the bid-distribution term. The bid distribution F_b is conditional on $p_{N|X}(n)$, the probability mass function for the number of bidders, because strategies are conditional on $p_{N|X}(n)$. For comparison, the standard model that assumes the number of bidders N is known, requires that F_b is conditional on N ; the analogue here is more complicated since any shift in $p_{N|X}(n)$ changes the bid strategy.

Since bids are observed and I have assumed H is a scaled distribution of bids, observation of $p_{N|X}(n)$ renders the right-hand side of equation 6 observed, identifying costs. In practice, I will group auctions into two clusters, a “many-competitor” and a “few-competitor” group, using machine learning and their bid-schedule characteristics. I will estimate $p_{N|X}$ as the empirical distribution of N within its cluster.

Semi-parametrization Equation 6 contains high-dimensional conditional densities, the estimation of which is complicated by the curse of dimensionality. To circumnavigate this issue, I modify the approach of Haile et al. (2003) to essentially residualize out project-specific covariates from bids. My approach takes advantage of my rich covariates, machine learning, and the fact that I observe *all* bids, not just winning bids. The necessary theoretical assumption for the residualization is that costs are multiplicatively separable into two components: an auction-specific term common to all bidders, and a bidder-specific idiosyncratic term. Specifically, I assume

$$c_{ia} = \exp(X'_a \gamma) \mu_a \tilde{c}_{ia}; \quad (X_a, \mu_a, \tilde{c}_{ia}) \text{ are mutually independent,}$$

where μ_a is an auction-specific cost shifter. The benefit of this assumption is that it directly implies

$$b_{ia} = \exp(X'_a \gamma) \mu_a \tilde{b}_{ia}, \quad (7)$$

where $\tilde{b}_{ia}$ is the bid corresponding to an auction with $c_{ia} = \tilde{c}_{ia}$, i.e. the cost for bidder i in a hypothetical auction with covariates $\exp(X'_a\gamma)\mu_a = 1$. The proof is in Appendix D.2. Note that I have already assumed the government reserve follows a scaled distribution of the bid; for statement 7 to hold, I require the less restrictive assumption that whatever distribution r follows, it scales with the same $\exp(X'_a\gamma)\mu_a$ factor. This assumption, of course, is still restrictive in that it assumes the government's reserve fully adjusts in response to project cost shifters. While this is what most DOTs attempt to do, they may suffer from incomplete information.

The fact that bids are homogeneous of degree one in costs means that $\tilde{b}_{it}$ from regression (7) can be interpreted as bids for hypothetical homogeneous projects. This individual component of the bid is thus comparable across auctions and suitable for estimating the distributions and densities in equation (6). The resulting estimator is semi-parametric: I have assumed that the distribution of bids belongs to a scale family, where the scale parameter depends on covariates.⁴⁵ Estimated margins for the homogenized projects can be scaled back to true bid magnitudes by respectively multiplying each by $\exp(X'_a\gamma)\mu_a$. A useful implication is that the markup for the homogenized auction is the same as for the original auction.

4.3 Estimation

Step 1: Bid homogenization I estimate $\ln \tilde{b}_{ia}$ as the residuals from the regression

$$\ln b_{ia} = X'_a\beta + \alpha_a + \epsilon_{ia}, \quad (8)$$

where X_a as a vector of covariates at the project level, including (i) fixed effects for the type of engineering of the project and its AI-derived classification; (ii) ruggedness and rurality measures; (iii) the logarithm of the total quantities of material, summed separately by measurement unit: tons, cubic yards, linear feet, square yards and counts of "each"; and (iv) the logarithm of the number of distinct material items listed in the bid schedule.

I use the equivalence between machine learning and hierarchical bayesian methods to take advantage of my covariates and to estimate the auction random effect, α_a , which controls for remaining

⁴⁵Note that the probability distribution of N does not need to be homogeneous across projects for estimation.

unobserved heterogeneity across projects. I estimate regression 8 with maximum likelihood, treating α_a as a random effect and applying a ridge penalty on the remaining coefficients to guard against over-fitting.⁴⁶ Estimation is done separately for each state, allowing the covariate–cost relationship to be state-specific. I include year fixed effects but do not residualize by them.⁴⁷ I then normalize every homogenized bid to a benchmark asphalt resurfacing and repair project—at national median ruggedness and rurality—using each state’s predicted material quantities for the national median asphalt project size (Appendix B.4). This procedure enables cross-state comparison despite differing measurement conventions across state DOTs. For skeptical readers, note that most results do not rely on level comparisons across states.

Step 2: Cluster auctions To estimate $p_{N|X}(n)$, I cluster auctions into “many-competitor” and “few-competitor” groups, and estimate $\hat{p}_{N|X}(n)$ as the empirical distribution within each (cluster, state, year) tuple. Specifically, I apply a gradient tree boosting algorithm to predict the number of bidders in each auction using the full set of bid-schedule items.⁴⁸ Auctions are then divided into these clusters at the median predicted bidder count. The estimated $\hat{P}$ should reflect the probability according to bidder beliefs; although these beliefs are unobserved, Appendix Table A.14, columns (4) and (6), shows the predicted bidder count correlates more strongly with homogenized bids than the actual bidder count, suggesting the prediction aligns closely with bidders’ expectations. Figure A.15 shows the distribution of the actual number of bidders across the clusters; the distributions are distinct but overlap in the right tail of the “few-competitor” group and the left tail of the “many-competitor” group.

Robustness Appendix Table A.13 and Figure A.13 present a summary table and accompanying violin plots for regression 8 and two alternative specifications for bid homogenization. The first alternative is a parsimonious baseline that retains only the project classification, engineering type,

⁴⁶An L2 (ridge) penalty on a coefficient block is equivalent to modeling that blocks as random effects if the Langange multiplier satisfies $\lambda = \frac{\sigma_\epsilon^2}{\tau^2}$ where σ_ϵ^2 is the error variance and τ^2 is the variance of the prior. See Hastie et al. (2009) for details.

⁴⁷Because I will estimate the auction model separately by year, the cost distribution need not be homogenized over time, so I leave time out of X_a . The coefficients γ therefore absorb the effect of the project covariates in X_a , removing only the covariate-driven component of time variation but still allowing costs for a fixed project to shift over time.

⁴⁸The advantage of using a machine learning algorithm here, relative to bid homogenization, is threefold. First, since I no longer have multiple observations per auction, I cannot estimate auction-specific fixed effects. Second, the comparability of covariates across states is no longer necessary. Third, the relationship between auction entry decisions and inputs is likely more complex than the relationship between construction costs and inputs.

county ruggedness and rurality, and total tons. The second augments the baseline with exact quantity measures for the individual bid-schedule items, conditional on items having passing a prespecified threshold frequency; for example, “wet-weather pavement-marking tape” is included when it is sufficiently common. Because the resulting set of items is extensive, I estimate these regressors under a moderate ridge penalty. Full estimation details are provided in the appendix.

The table shows that the preferred specification achieves the lowest cross-validated root-mean-squared error on every reported metric, and the violin plots confirm that, while the three residual distributions are broadly similar, the preferred model yields the most concentrated distribution.

Appendix Table A.14 compares predictions of the number of bidders from the gradient tree boosting algorithm described above to those from a penalized multinomial logit model. Columns (1) and (2) show that gradient boosting achieves substantially higher R^2 values in explaining both the actual number of bidders and the resulting bids. Figure A.14 visualizes the bidder count predictions.

5 Estimates

5.1 Estimates of Costs and Markups

Table 5 reports summary statistics of estimates for costs and margins over roadwork auctions in the continental U.S. from 2014 to 2023. Reported costs and margins are rescaled back from the homogenized bids to match the actual bid levels, reversing the homogenization procedure in Section 4.3. Although estimation uses all bids, I report results across winning bids only to represent realized prices. The first panel reports auction-level summaries over all winning bids; the second panel reports the interquartile range across state means; the final panel groups projects by terciles of estimated costs and reports results by tercile.

Estimated margins are economically large. The mean margin is approximately \$0.7 million against a mean cost of \$2.5 million. Correspondingly, the mean markup over all winning bids is .22, meaning for the average project the government pays 22% above the production cost.⁴⁹ There is a wide range

⁴⁹In the broader markup literature, this estimate is quite similar to the benchmark in De Loecker and Warzynski (2012) – though this paper uses a vastly different methodology – who report a median markup of .2 for U.S. public firms, stable over time, though with a large rise in the right tail since 1980.

in the estimated markup across projects, with the 25th percentile markup equal to .1 and the 75th percentile triple that, at .34. The median markup is slightly lower than the mean, at .17, reflecting a long right tail. To address concerns that extreme tails are due to estimation error, all means are estimated with a 1% sample trim; Appendix A.6 shows that the mean markup is stable under larger sample trims.⁵⁰

Variable	Mean	Median	25 th pctl	75 th pctl
<i>Auction-level</i>				
Markup	0.22	0.17	0.10	0.34
Cost	2477	1251	401	3210
Margin	693	339	147	795
<i>State-level (means)</i>				
Markup			0.19	0.26
Cost			2097	4100
Margin			588	879
<i>By project cost</i>				
<i>Small</i>				
Markup	0.62	0.28	0.15	0.67
Cost	436	416	259	608
Margin	222	118	64	257
<i>Medium</i>				
Markup	0.26	0.17	0.11	0.30
Cost	1606	1537	1158	2028
Margin	405	268	167	474
<i>Large</i>				
Markup	0.18	0.13	0.09	0.20
Cost	7013	5129	3608	8431
Margin	1161	784	467	1436

Note: Table reports the mean and interquartile range of winning bids for roadwork projects from 2014–2023. All monetary values are in thousands of 2023 USD. The bottom and top 1% of observations are trimmed to remove extreme tails. Appendix Table A.6 reports means under alternative trim percents of 0%, 1%, 3%, and 5% trim; the markup stabilizes after the 1%. For the third panel, only projects with estimated winning costs greater than \$.1 million are used, dropping the smallest 10% of auctions.

Table 5: Summary Statistics of Estimates

Some state procurement markets appear noticeably more compelling to bidders. Variation in the markup across states is substantial. The 25th-percentile state has a mean markup of 0.19, whereas the 75th-percentile state has a mean of 0.26. The second row in this panel shows that costs also differ markedly across states, nearly doubling from the 25th to 75th percentile state. This variation

⁵⁰ A small number (less than 3%) of estimated costs are negative. These arise when extremely low bids can only be rationalized by negative costs, which may reflect genuinely negative costs (e.g. if a firm needs to use sitting materials) or dynamic incentives to win, violations of the model such as non-risk-neutrality, or failure to control for some project heterogeneity. This project pushes hardest on the third concern using rich controls, machine learning, and random effects, though bias may persist.

reflects a combination of differences in project design and differences in the cost of a given project holding design fixed. Unfortunately, because detailed project covariates are defined *within* state, distinguishing between these two channels is complex. However, as discussed in Section 4.2, under the assumption of multiplicative project covariates, the markup is scale-invariant and therefore directly comparable across states, regardless of the baseline cost differences.

High-markup states are concentrated in the Midwest and the South, as shown in the map in Appendix Figure A.17. Kentucky stands out with a median estimated markup of 0.55, reflecting a remarkably low average number of bidders per auction (1.6 in the estimation sample). Large markets such as California and Texas exhibit lower markups. Rhode Island is an outlier, achieving low markups despite its small market size. However, this likely reflects unusually high cross-border participation, since the state has the highest share of multistate firms active within its borders (74% over the period).

The third panel of Table 5 shows that markups decline with project size. The median falls from 0.28 in the smallest tercile to 0.13 in the largest. In auctions, markups adjust with both the distribution of costs and the number of bidders. In the setting at hand, an additive increase in costs (a location shift of the cost distribution) that leaves its shape and scale unchanged keeps the margin constant and lowers the markup. By contrast, if costs increase with a multiplicative scaling of the cost distribution, the margin increases while the markup is constant. Empirically, the dispersion of both costs and prices appears to increase as mean increases, which would suggest larger projects see higher margins and variable profits, and even higher markups if the effect is large enough. At the same time, bidder counts are not fixed, and in fact increase with project size from 2.69 to 3.18, which in turn pushes markups down. On the whole, the net effect of these two forces results in declining markups. The model so far is agnostic about why certain auctions draw a higher numbers of bidders; Section 7 takes up that question.

Figure 4 plots estimated markups against the realized number of bidders, paralleling the bid gap curve in Figure 3. The median markup drops by about one quarter from one to four bidders. Because bidders observe only the distribution of bidder counts rather than the realization, the pattern is generated by between state–year–cluster variation. The figure reports up to five bidders; beyond this point, most auctions are classified into the high-competitor cluster, mechanically restricting the

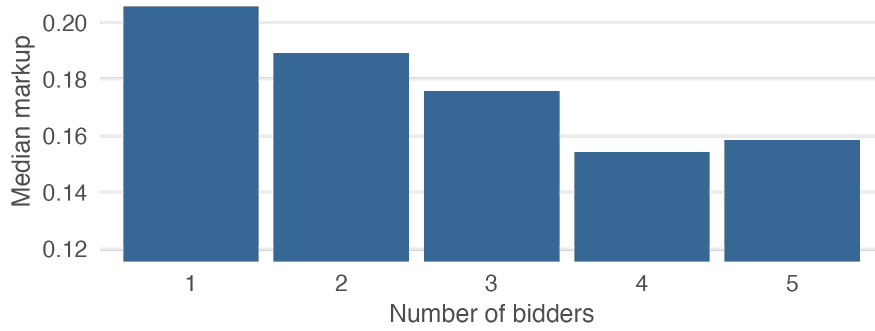


Figure 4: Estimated Markups by Realized Number of Bidders

Note: Figure plots the median estimated markup by actual number of bidders across all winning bids for roadwork projects from 2014-2023. Markups are computed from estimated costs and margins; observations with non-positive costs/markups are excluded. Bars reflect unweighted medians across auctions (not value-weighted).

incremental effects.⁵¹

5.2 Decomposing price growth over time

Using the estimates of costs and markups for every auction between 2014 and 2023, I decompose the price trend over this period into its component parts. Because estimation is split by year (and state cluster), the time trend of estimated costs and margins is allowed to be fully flexible. I first show the difference between the actual mean price and the mean of the homogenized projects, which are residualized by covariates, as discussed in Section 4.3. The resulting gap between the actual and homogenized price shows the change due to changing project inputs for roadwork, which I term project complexity. The gap would not capture changes in input prices, but such a change would appear in the estimated production cost. The mean estimated production cost and the mean estimated margin together sum to equal the mean homogenized bid each year.

Figure 5 visualizes the results of the time trend decomposition. The two lines show that detailed project covariates can account for a piece of the price growth. The red dashed line shows the mean actual bid, which rose 28% over the ten-year period. As in Figure 1, this is controlling for

⁵¹The slight uptick at five bidders may reflect coarse clustering; finer clusters trades off increasing the similarity of projects estimated together with a loss in precision. The upward turn at five bidders could also arise from genuine variation, e.g. if five-bidder auctions have higher cost variance which generates larger markups.

inflation.⁵² The yellow line reports results for homogenized price, which in turn rose 18% over the period. Consequently, project complexity alone can explain 36% of rising prices. What does increasing project complexity look like? Appendix Figure A.16 plots trends for inputs with the highest growth for Nebraska, a state with particularly high price growth over the period. The highest growth inputs include a combination of new environmental adjustments, technical items such as membrane waterproofing, and training. Additionally, the average number of inputs per project in Nebraska increased by 19% from 54 to 64.

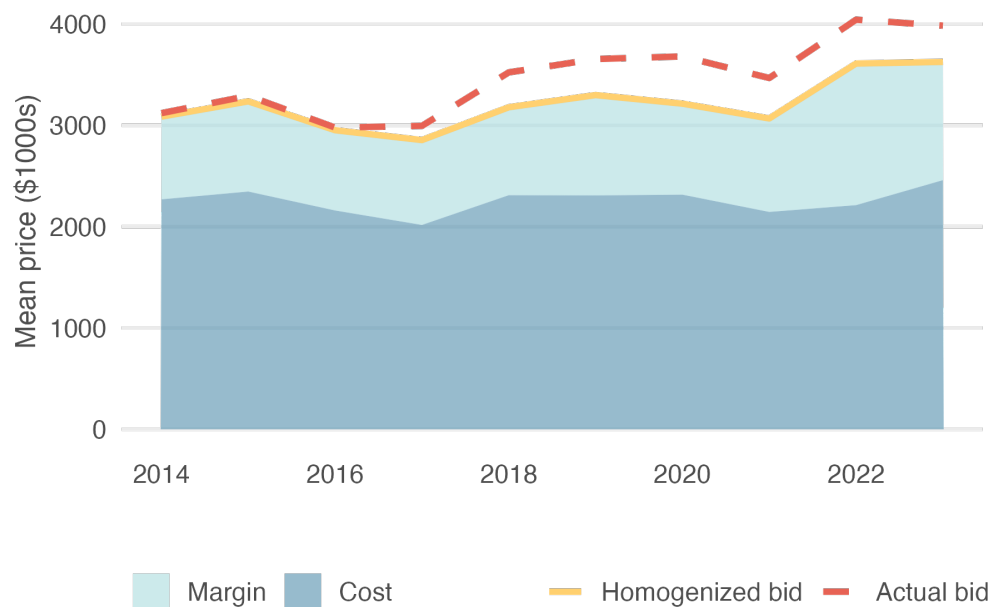


Figure 5: Decomposition of Price over Time

Note: Figure decomposes the average winning bid between 2014 and 2023 into changes due to project characteristics, estimated production costs, and estimated margins. The red dashed line shows the mean actual bid, while the yellow line shows the mean homogenized bid. Both series are expressed in real 2023 dollars and exclude the top and bottom 1% of bids. The dark shaded region represents the production cost and the light shaded region the markup.

The majority of the increase in prices paid over the period appears to be genuine price growth, with 64% of the growth persisting after project homogenization. The shaded blue regions in Figure 5 separate this price growth into production cost growth and margin growth. The dark blue area denotes mean total cost. Over the full period, estimated costs rose minimally – 8% – while estimated

⁵²The red dashed line mirrors the red line in Figure 1, with small differences arising from dropping the top and bottom 1% of projects by bid amount prior to homogenization and from using the mean here, rather than the median, to facilitate the decomposition.

margins rose 44%, explaining two thirds of the real rise in price.

The absence of cost growth is notable because the cost estimated in the model represents *all variable costs of production*, which together form the basis of a firm's bid. In particular, this includes labor costs, which are often discussed as a potential driver of high U.S. infrastructure expenses.⁵³ It also includes the prices of physical inputs, and also permitting costs, time delays, and the firm's opportunity cost. Over the decade since 2014, no portion of production costs appears to have dramatically worsened, though prices may well be above the level they could reach if such factors were addressed. Yet, margins, which at the start account for a substantial share of price, are growing.

6 Counterfactuals

How much would the government save if bidders faced more competition? I estimate the prices the government would pay if each auction had one additional bidder. Unlike the previous reduced-form exercise, the model allows me to account for the current distribution of bidders. As Figure 4 demonstrates, the effect of one more bidder depends on the current baseline. Holding bidders' cost fixed, I estimate counterfactual bidding strategies under the hypothetical guarantee of one additional bidder in every auction. Specifically, I shift the bidder-count distribution up by one – i.e., the probability mass on N bidders is reassigned to $N + 1$. The model makes explicit how bidders' beliefs and strategies adjust, allowing for the decomposition of the total price effect into the bidders' strategic response and the composition effect generated by the expanded bidder pool.

Estimation Having estimated costs under the observed distribution of bidders, $p_{N|X}$, I now hold these costs fixed and derive counterfactual bids under an alternative bidder distribution, $p_{N|X}^{cf}$. The bidding strategy comes from equation 5, restated fully here:

$$\beta = c_i + \frac{\sum_n \Pr(N = n|X) (1 - F(\beta^{-1}))^{n-1} (1 - H(\beta))}{\sum_n \Pr(N = n|X) (1 - F(\beta^{-1}))^{n-2} \left((n-1) \frac{f(\beta^{-1})}{\beta'} (1 - H(\beta)) + (1 - F(\beta^{-1})) h(\beta) \right)},$$

⁵³For example, see Goldwyn et al. (2023).

where I have omitted conditioning covariates for legibility. As before, β is the bid strategy, c_i is bidder i 's random cost draw, F is the distribution of costs, H is the distribution of the secret reserve price, and F is the probability distribution of the number of bidders.

The secret government reserve price complicates estimation relative to other empirical auction settings. If there were no secret reserve price, the equilibrium bid solves a straightforward ordinary differential equation with solution

$$\beta = \mathbb{E} \left[\min \{C_j : j \neq i\} \mid \min \{C_j : j \neq i\} > c_i \right] = \frac{1}{G(c_i)} \int_{c_i}^{\bar{c}} tg(t)dt,$$

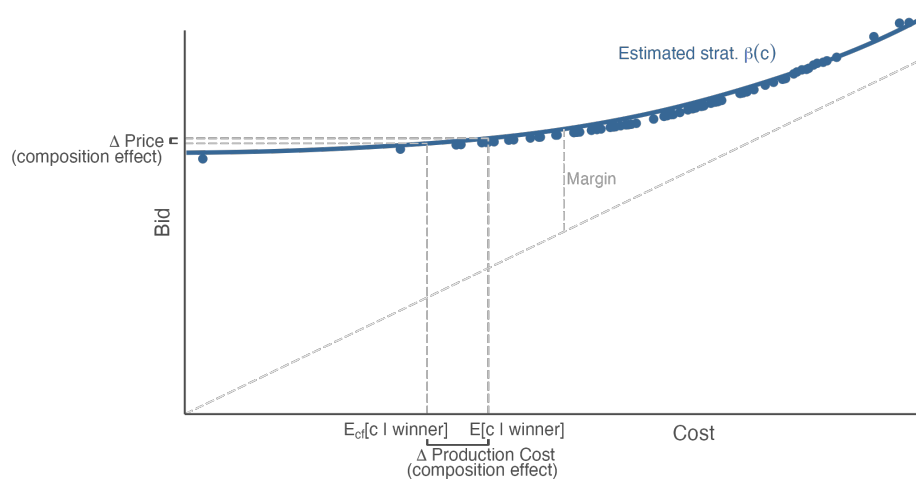
where $G(c_i) \equiv \sum_{n=1}^N \Pr(N = n)(1 - F(c_i))^{n-1}$. Under this bid strategy, counterfactuals could be computed straightforwardly by integrating over the mixture distribution $G(c_i)$.

With the secret reserve price, the ordinary differential equation cannot be solved analytically. Therefore, I instead estimate the bid function numerically from an initial boundary condition. Intuitively, the solver expresses β' directly as a function of β , iteratively tracing out the equilibrium bid curve from this initial point. The estimating equation and implementation details given in Appendix [D.3](#).

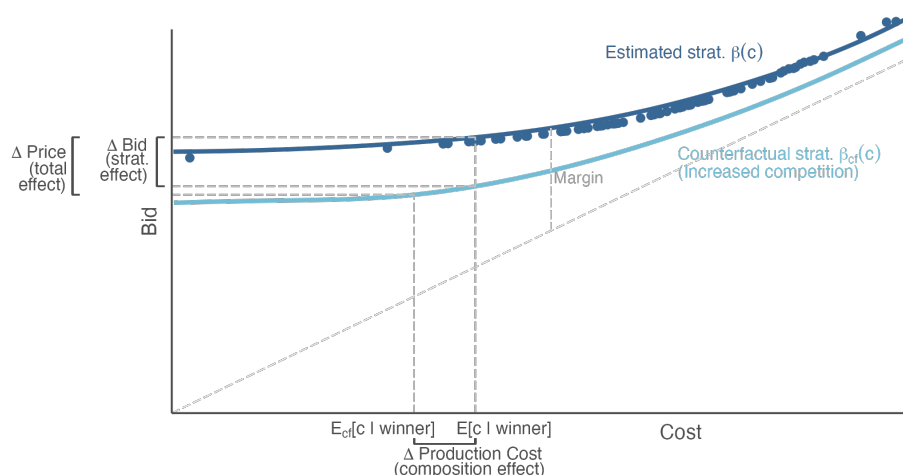
Results An additional bidder lowers prices first by threatening incumbents' win probabilities, and second, by potentially having the lowest cost draw. Figure 6 illustrates these two mechanisms for an example state. The shift from the estimated strategy curve $\beta(c)$ to counterfactual strategy curve $\beta_{cf}(c)$ reflects more aggressive bidding – the strategic effect. The shift is larger for bidder i when the probability distribution of the second-lowest cost, conditional on i winning, changes more with entry. This varies with both the initial number of bidders and the cost distribution. Holding the winner's cost constant at the expected minimum cost draw, the distance between the curves isolates portion of the price decrease due to the strategic response.

The second channel is the extensive margin. A bigger pool of bidders means that there may be a new bidder who is simply cheaper. In this exercise, I draw new bidders from the same distribution, though endogenous entry would likely imply a right-shifted distribution. Still, the graph illustrates why most of the price drop likely comes from strategy, not the extensive margin. Because $\beta(c)$ flattens in the right tail, much of the gain from a low-cost draw is captured by the firm. This is

intuitive: firms in the right tail of the cost distribution anticipate little competition. The full effect, ΔPrice , includes both the strategy effect and the pool effect.



(a) Pool Effect



(b) Total Effect

Figure 6: Decomposition of Effect of Increased Competition

Note: Figure shows estimated bidding strategies under the actual and counterfactual bidder distributions for Alabama (2023, cluster 1). Points are observed bids plotted against estimated costs. The gap between the curves reflects the strategic adjustment holding cost fixed. The total price effect combines this with the shift left in expect cost of a new lowest draw.

I estimate effects for all states using the 2023 cost distribution and baseline bidder counts. Table 7 reports results for the strategic effect alone and the total effect for winning bids. On average, a guaranteed additional bidder yields just over \$350,000 in savings per auction. Over a baseline

average winning bid of \$3.7 million in 2023, the savings represent a 10% fall in prices. This finding is surprising aligned with the quasi-experimental results, given the two exercises use entirely disparate identification strategies and sources of variation. Comparing the last two rows of the table shows that, as in Figure 6, most of price effect in the counterfactuals comes from the strategic margin, with an estimated \$300,000 in savings.

Variable	All winning bids				
	Auction-level			State-level (means)	
	Mean	25 th pctl	75 th pctl	25 th pctl	75 th pctl
<i>Baseline</i>					
Markup	0.37	0.12	0.44	0.17	0.51
<i>Counterfactual</i>					
<i>Strategic effect</i>					
Markup	0.19	0.06	0.26	0.07	0.34
Δ winning bid (\$1000s)	-286	-348	-55	-392	-184
<i>Total effect</i>					
Markup	0.33	0.08	0.39	0.18	0.47
Δ winning bid (\$1000s)	-358	-427	-62	-494	-265

Note: Table reports the mean and interquartile range of all winning bids for roadwork projects in 2023 under baseline and counterfactual distributions of the number of bidders. Cost figures are expressed in thousands of 2023 USD. The strategic effect adjusts only incumbent bidders' strategies, while the total effect incorporates the probability that the additional bidder may win. When computing means, the lowest and highest 1% of observations by cost are trimmed..

Table 7: Summary of Counterfactual Competition and Prices

The complexity of the interaction between the cost distribution and bidder counts is shown in the changing markup. The strategy effect alone produces a lower markup than the total effect, although both are lower than the baseline. This is due to the that firms that draw a particularly low cost know they have good odds of winning even with a higher markup.

7 Entry Barriers

In the preceding sections, this paper documented large, pervasive markups in highway infrastructure generated by few-bidder auctions. In this section, I turn from the consequences of low bidder participation to the causes. Identifying sources of limited competition matters in particular for policy, since remedies depend on why entry is scarce.

One common explanation for limited entry in the face of large variable profits is large fixed

costs. There are many such fixed costs that might impact the infrastructure market. For example, construction machinery is a substantial capital investment. Alternatively, there might be fixed costs that are more intangible, such as the costs of complying with regulatory and administrative requirements. At the market level, working with the government requires meeting a number of pre-qualification criteria, which can include extensive paperwork, such as financial audits, proof of safety procedures, and proof of bonding capacity, a federally mandated insurance for government work. Firms in the public sector may pay additional legal fees, or additional wages to legal staff. Firms also face fixed costs at the auction level, including the administrative cost of preparing bid proposals, and even the cost of obtaining the necessary information about the project to make a bid.⁵⁴

I begin by examining evidence of barriers to entry at three different levels: the construction industry, the infrastructure market, and individual auctions. First, despite the capital-intensity of the construction industry, construction capital appears unlikely to be pivotal. I find that infrastructure market participants are typically established construction firms well before their first government contract. Based on the subset of Heavy Construction and Building Construction firms with reliable start-year data, I find a median age of the firm at first bid is 17 years (Appendix Figure A.12 shows the full distribution). It's also worth noting that a sizable portion of the industry leases machinery, for whom equipment costs might manifest as variable costs, depending on lease length.⁵⁵ Regardless, it appears that most construction firm fixed costs such as building and machinery investments were sunk prior to entry into the infrastructure market.

I now turn to auction- and market-level entry barriers.⁵⁶ As detailed in Bresnahan and Reiss (1991), fixed costs generate economies of scale, so in a larger market firms can operate under lower unit variable profits. Consequently, in equilibrium, a larger scale means that the market can support more firms and results in each firm having a smaller variable return on a unit of production. I apply that logic at both the market and the auction level. My measures of scale are project size, now measured

⁵⁴Liscow et al. (2023) report that a one-standard-deviation increase in bid-document length is correlated with 16% higher costs, while a one-standard-deviation increase in outreach is correlated with an 18% lower cost.

⁵⁵Machinery rental is the third largest category of intermediate inputs in the "Transportation structures and highways and streets" industry according to the BEA Input-Output Use table (following fabricated structural product manufacturing and ready-mix concrete manufacturing).

⁵⁶By "entry barriers" I mean the classic definition of Stigler (1968): a barrier to entry is a cost of producing (at some or every rate of output) that must be borne by firms seeking to enter an industry but not by firms already in the industry.

by estimated cost, and state market size, measured by the number of auctions. I report the effect of scale on the number of bidders, which is a sufficient statistic for firm variable profit, all else equal. Table 8 reports results. Column (1) shows that doubling project cost increases bidders per auction by 0.05. Column (2) shows that doubling the number of auctions within 100 miles increases bidders per auction by 0.22—about 7% relative to the median of three bidders. Columns (3)–(4) push this logic further by looking across state lines. Remarkably, doubling the number of nearby auctions across state borders has no effect on bidders per auction. This implies that market fixed costs scale with state boundaries. Crossing into another state requires incurring a new entry cost. In short, the fixed costs at infrastructure market level are triggered anew with each new state, suggesting they are the result of regulatory entry barriers.

	<i>Dependent variable:</i>			
	(Project scale)	Number of bidders		
		(Market scale)		
	(1)	(2)	(3)	(4)
Log cost	0.047*** (0.007)	0.051*** (0.007)	0.050*** (0.007)	0.050*** (0.007)
Log auctions < 100 mi		0.224*** (0.017)		
Log in-state auctions < 100 mi			0.239*** (0.015)	0.214*** (0.017)
Log out-of-state auctions < 100 mi			-0.010* (0.006)	-0.004 (0.006)
Log mi to border				0.005*** (0.002)
Rurality	-0.415*** (0.018)	-0.343*** (0.019)	-0.346*** (0.018)	-0.353*** (0.018)
Ruggedness	-0.004*** (0.0004)	-0.004*** (0.0004)	-0.004*** (0.0004)	-0.004*** (0.0004)
Tons (1000s)	-0.002*** (0.0003)	-0.002*** (0.0003)	-0.002*** (0.0003)	-0.002*** (0.0003)
Type FE	yes	yes	yes	yes
Class FE	yes	yes	yes	yes
year FE	yes	yes	yes	yes
state FE	yes	yes	yes	yes
Observations	36,657	36,657	36,657	36,657
Adjusted R ²	0.224	0.227	0.230	0.230

Note:

*p<0.1; **p<0.05; ***p<0.01

Note: Table reports estimates from regressions of the number of bidders per auction on measures of market scale. The first column looks at project scale, measured by project cost, and the last three columns look at market scale, measured by the number of auctions within 100 miles of a given auction. The 100 mile threshold is chosen based on the entry-by-distance results in Appendix Figure A.9.

Table 8: Market Scale and Fixed Costs

7.1 Auction Entry Model

I am interested in understanding sources of barriers to entry. Measuring regulation, let alone less formal beaurocratic practices, is a complicated task (Trebbi and Zhang, 2022). I take a different approach in this section and estimate whether market entry or auction entry costs are larger based on observed firm behavior. To do so, I combine the results from my bidding model with an auction entry model and, for estimation, exploit the fact that firms enter auctions only when expected total profits are nonnegative. The model builds off of Berry (1992). Imposing zero profits at the market entry margin provides a rough answer to whether market entry or auction-level barriers bind more.

I start by embedding the estimated actual and counterfactual margins from Section 4.1 in a threshold-crossing entry framework, following Berry (1992). Firms draw heterogeneous auction entry costs from a common distribution and decide whether to enter before observing variable production cost. The assumptions are considerable, but permit identification of fixed costs from the observed number of entrants and estimates of variable profits. A benefit of auction setting is that I have already carefully estimated variable profits, avoiding the linear in observables profit specification in Bresnahan and Reiss (1991).

I assume the following sequence of events. First, firms in the market draw an auction entry cost. Second, firms make an entry decision based on their fixed cost and expected variable profit, conditional on covariates. I assume the firm with the lowest entry costs enter first to ensure a unique equilibrium. Finally, firms draw their production cost and make their bid. At step one, the expected total auction profit is given by

$$\mathbb{E}\Pi_{ia} = \overbrace{V_a}^{\text{variable profit}} - \overbrace{\phi_{ia}}^{\text{entry cost}},$$

where variable profit is the expected auction margin $V_a = \mathbb{E}[(b_{ia} - c_{ia}) \Pr(\text{win} \mid p_N(n))]$. As before, $p_N(n)$ is the probability mass function of the number of bidders in the auction, b_{ia} is firm i 's bid and c_{ia} is firm i 's cost. The firm thus conditions on the distribution of its number of rival bidders, rather than on the exact value. The expectation at this stage is also over the firm's own cost draw.

I parameterize auction entry costs as

$$\phi_{ia} \equiv \kappa + Z_{ia}\alpha + \rho\mu_{ao} + \sigma\mu_{ia},$$

where Z_{ai} are demeaned firm characteristics, specifically years of experience and number of state markets. The variables μ_{ao} and μ_{ia} indicate a common auction shock and a firm auction shock, respectively.

The likelihood of observing N_a firms in auction a is

$$\Pr(N_a = N) = \underbrace{\Pr(|\{i : \mathbb{E}\Pi_{ia}(p_N(n)) \geq 0\}| \geq N)}_{\text{at least } N \text{ profitable under } p_N(n)} - \underbrace{\Pr(|\{i : \mathbb{E}\Pi_{ia}(p_N(n+1)) \geq 0\}| \geq N+1)}_{\text{at least } N+1 \text{ profitable under } p_N(n+1)}.$$

The first (second) equation counts the number of firms with positive expected profits in an auction with the covariates of a and N ($N+1$) bidders in the $p_N(n)$ ($p_N(n+1)$) equilibrium distribution. In other words, the likelihood is given by the probability that at least N_a firms could be profitable in the observed equilibrium minus the probability that $N_a + 1$ would be profitable if firms shifted to an equilibrium with one bidder per auction. I assume μ_{ao} and μ_{ia} are distributed standard normal with zero covariance. The parameters of interest are $\theta = (\kappa, \alpha, \rho, \sigma)$. This likelihood can essentially be estimated similarly to an ordered probit, but with simulated firm shocks, following Berry (1992).

There are two advantages to this entry model. The first is that, unlike class threshold-crossing entry models, I do not need to estimate variable profits on market observables. Instead, I can simply plug in my previously estimated margins and counterfactual margins. Second, the parameters are identified in levels, due to variable profits V_a fixing the scale in dollars. This is in contrast to standard discrete choice models, which are usually only identified up to scale.

I estimate my main results for each state using a Markov chain Monte Carlo (MCMC) procedure. For illustration, Table 9 reports full results for Illinois; the procedure scales cleanly to all states. The results suggest sizable entry costs, with a mean posterior estimate of \$72,000 for the first bidder and up to \$446,000 for the fourth bidder. These estimates bundle all fixed costs at the auction-entry stage, including, for example, engineer hours for plans, legal review of contract terms, bid bonds and prequalification work, DBE documentation, and site visits. The estimates may also reflect capacity constraints: for firms near capacity, some fixed entry tasks can become more expensive, for example if they hire outside estimators or counsel. However, for the lowest-cost bidders actually observed to enter, capacity is least likely to bind.

Parameter	Estimate	95% CI
Expected cost (1st entrant)	-72.6	[-83.4, -62.8]
Expected cost (2nd entrant)	-183.7	[-211.6, -161.5]
Expected cost (3rd entrant)	-312.8	[-360.7, -272.3]
Expected cost (4th entrant)	-446.3	[-515.4, -386.0]
α_1 (experience)	0.5	[-5.1, 6.1]
α_2 (# states)	0.2	[-5.3, 5.5]
ρ	121.1	[23.0, 226.9]
σ	1242.9	[941.8, 1599.2]

Note: Table reports estimates posterior means and 95% credible intervals from the state-year structural model of bidding for Illinois in 2023. Parameters include the fixed entry cost (κ), cost-shifter coefficients (α_1 , α_2), the auction-level common shock (ρ), and idiosyncratic cost dispersion (σ).

Table 9: Posterior Estimates

I compare auction entry costs to market entry costs by imposing a zero-profit condition at the state–market level. This assumption is arguably reasonable given the large number of active firms; as shown in Table 1, the average state–year includes 84 unique winning firms. Consequently, an entry model in which firms draw heterogeneous entry costs would yield tight bounds on the difference between fixed costs and profits.⁵⁷ I therefore conduct a back-of-the-envelope calculation. I take the average per-auction fixed cost for the first three bidders, weighting by the mean number of bidders. I multiply by the average number of auctions a firm enters per year to obtain annual auction-entry costs. This implies that auction-entry costs account for about 25% of total fixed costs, leaving about 75% attributable to market-entry costs. The calculation builds on a restrictive model, as I have assumed a Normal distribution for entry costs and further, no selection on variable production costs. However, it cleanly separates auction- from market-entry costs. Despite greater attention to auction-entry frictions in the empirical auctions literature, the evidence here points to larger market-entry costs and suggests they are the more important margin.

⁵⁷The market entry condition is $V(y_m + 1) < F_m \leq V(y_m)$, where y_m denotes the observed number of active firms in market m , $V(y_m)$ is the variable profit per firm at that number of firms, and F_m is the fixed cost of market entry. If $\frac{\partial V(y_m)}{\partial y_m} < 0$ and $\frac{\partial^2 V(y_m)}{\partial y_m^2} > 0$, then as y_m increases the bounds tighten, and for sufficiently large y_m we have $v(y_m) \approx F_m$.

8 Conclusion

U.S. transportation demand is growing while the physical system is aging, making it increasingly important to understand how to reduce the costs of reconstruction and upgrades. The lessons learned may also be relevant for other types of public infrastructure, such as the electricity grid, which is lagging far behind goals. This paper focuses on one long-hypothesized but understudied channel: competition in the market for infrastructure construction. In particular, I examine the role auction competition has played in raising prices above firms' costs of production, focusing on the direct impact of the number of auction participants on bids. Yet market structure may further affect prices in the long run by shaping incentives for innovation — my results may be an underestimate of the long-run consequences of limited competition.

This paper documents that weak competition is a defining feature of U.S. infrastructure procurement. Across the country, auctions routinely attract only a handful of bidders, and estimated markups are large and have been rising since 2014. These markups account for more of the price growth over the last decade than rising production costs do. I show that securing one more bidder per auction would yield substantial returns to the government, a result that I establish with both a quasi-experimental design and a structural model with separate identification strategies. Yet, entry remains scarce. Entry patterns are consistent with the existence of sizable fixed costs generated by regulatory burdens rather than physical investments. These entry costs deter participation in both individual auctions and in the market itself; they appear largest at the state-market level. While reducing regulation and bureaucracy would reduce prices, in some cases doing so may be difficult or undesirable, as regulations have benefits. In such a case, increasing the scale of either auctions or market size would grow variable profits relative to fixed entry costs and reduce markups. These results motivate further research on barriers to entry and on other policy approaches, from informational outreach to interstate coordination, to improve procurement outcomes.

References

- Bajari, P.** (1997). “The First Price Auction with Asymmetric Bidders: Theory and Applications”. *Ph.D. Dissertation, University of Minnesota*.
- Bajari, P., S. Houghton, and S. Tadelis** (Apr. 2014). “Bidding for Incomplete Contracts: An Empirical Analysis of Adaptation Costs”. en. *American Economic Review* 104.4, pp. 1288–1319.
- Balat, J.** (2017). “Highway Procurement and the Stimulus Package: Identification and Estimation of Dynamic Auctions with Unobserved Heterogeneity”. *working paper*.
- Baum-Snow, N.** (2014). “Urban Transport Expansions and Changes in the Spatial Structure of U.S. Cities: Implications for Productivity and Welfare”. en. *The Review of Economics and Statistics*.
- Baumol, W. J., J. C. Panzar, and R. D. Willig** (Jan. 1982). *Contestable Markets and the Theory of Industry Structure*. en. Harcourt College Pub.
- Berry, S.** (1992). “Estimation of a Model of Entry in the Airline Industry”. *Econometrica* 60.4, pp. 889–917.
- Bhattacharya, V., J. W. Roberts, and A. Sweeting** (2014). “Regulating bidder participation in auctions”. en. *The RAND Journal of Economics*.
- Bolotnyy, V. and S. Vasserman** (2023). “Scaling Auctions as Insurance: A Case Study in Infrastructure Procurement”. en. *Econometrica* 91.4.
- Bresnahan, T. F. and P. C. Reiss** (Oct. 1991). “Entry and Competition in Concentrated Markets”. en. *Journal of Political Economy*.
- Brooks, L. and Z. Liscow** (Apr. 2023). “Infrastructure Costs”. en. *American Economic Journal: Applied Economics* 15.2, pp. 1–30.
- Bulow, J. and P. Klemperer** (1996). “Auctions vs. Negotiations”. en. *American Economic Review* 86.
- Chandra, A. and E. Thompson** (2000). “Does public infrastructure affect economic activity?: Evidence from the rural interstate highway system”. *Regional Science and Urban Economics* 30.4.
- Coviello, D. and M. Mariniello** (2014). “Publicity requirements in public procurement: Evidence from a regression discontinuity design”. en. *Journal of Public Economics* 109.C, pp. 76–100.

- Currier, L., E. L. Glaeser, and G. E. Kreindler** (Dec. 2023). *Infrastructure Inequality: Who Pays the Cost of Road Roughness?* en. Tech. rep. w31981. National Bureau of Economic Research.
- De Loecker, J. and F. Warzynski** (May 2012). “Markups and Firm-Level Export Status”. en. *American Economic Review* 102.6, pp. 2437–71.
- De Silva, D. G., G. Kosmopoulous, and C. Lamarche** (Feb. 2009). “The effect of information on the bidding and survival of entrants in procurement auctions”. en-US. *Journal of Public Economics* 93.1-2. Publisher: North-Holland, pp. 56–72.
- De Silva, D. G., T. Dunne, and G. Kosmopoulou** (2003). “An Empirical Analysis of Entrant and Incumbent Bidding in Road Construction Auctions”. en. *The Journal of Industrial Economics*.
- De Silva, D. G. and B. Rosa** (Dec. 2024). *Highway Procurement During the Great Recession and Stimulus*. en. Tech. rep. w33299. National Bureau of Economic Research.
- Donaldson, D. and R. Hornbeck** (2016). “Railroads and American Economic Growth: A “Market Access” Approach”. en. *The Quarterly Journal of Economics*.
- Fajgelbaum, P. D., C. Gaubert, N. Gorton, E. Morales, and E. Schaal** (July 2023). *Political Preferences and Transport Infrastructure: Evidence from California’s High-Speed Rail*. en. Tech. rep. w31438. National Bureau of Economic Research.
- FHA** (2000). *Our Nations Highways: Selected Facts and Figures*. Tech. rep. The Federal Highway Administration, US Department of Transportation.
- Ghani, E., A. G. Goswami, and W. R. Kerr** (2014). “Highway to Success: The Impact of the Golden Quadrilateral Project for the Location and Performance of Indian Manufacturing”. en. *The Economic Journal*.
- Gibbons, S., T. Lyytikäinen, H. G. Overman, and R. Sanchis-Guarner** (2019). “New road infrastructure: The effects on firms”. en. *Journal of Urban Economics* 110.C, pp. 35–50.
- Gibbons, S. and S. Machin** (2005). “Valuing rail access using transport innovations”. en. *Journal of Urban Economics* 57.1, pp. 148–169.
- Glaeser, E., M. Kahn, and J. Rappoport** (Jan. 2008). “Why do the poor live in cities? The role of public transportation”. en-US. *Journal of Urban Economics* 63.1, pp. 1–24.
- Goldwyn, E., A. Levy, E. Ensari, and M. Chitti** (2023). *Transit Costs Project: Understanding Transit Infrastructure Costs in American Cities*. Tech. rep. Marron Institute of Urban Management: New York University.

- Gonzalez-Navarro, M. and C. Quintana-Domeque** (2016). “Paving Streets for the Poor: Experimental Analysis of Infrastructure Effects”. en. *The Review of Economics and Statistics*.
- Goolsbee, A. and C. Syverson** (Jan. 2023). *The Strange and Awful Path of Productivity in the U.S. Construction Sector*. en. Tech. rep. w30845. National Bureau of Economic Research.
- Griliches, Z. and J. Mairesse** (1995). “Production Functions: The Search for Identification”. *Econometrics and Economic Theory in the 20th Century*. Chapter 6. Cambridge University Press.
- Guerre, E., I. Perrigne, and Q. Vuong** (2000). “Optimal Nonparametric Estimation of First-Price Auctions”. en. *Econometrica* 68.3.
- Haile, P., H. Hong, and M. Shum** (2003). “Nonparametric Tests for Common Values at First-Price Sealed-Bid Auctions by”.
- Hastie, T., J. Friedman, and R. Tibshirani** (2009). *The Elements of Statistical Learning*. en. 2nd ed. Springer Series in Statistics.
- Hendricks, K., J. Pinkse, and R. H. Porter** (2003). “Empirical Implications of Equilibrium Bidding in First-Price, Symmetric, Common Value Auctions”. en. *The Review of Economic Studies*.
- Hong, H. and M. Shum** (2002). “Increasing Competition and the Winner’s Curse: Evidence from Procurement”. en. *The Review of Economic Studies* 69.4.
- Imbens, G. W.** (Mar. 2014). “Instrumental Variables: An Econometrician’s Perspective”. en. *Statistical Science* Vol. 29.No. 3.
- Ito, Y.** (2024). “Entry Deterrence in Procurement Auctions”. *working paper*.
- Jeziorski, P. and E. Krasnokutskaya** (2016). “Dynamic auction environment with subcontracting”. en. *RAND Journal of Economics* 47.4, pp. 751–791.
- Jofre-Bonet, M. and M. Pesendorfer** (2003). “Estimation of a Dynamic Auction Game”. en. *Econometrica* 71.5.
- Kessler, D. P. and L. F. Katz** (Dec. 1999). *Prevailing Wage Laws and Construction Labor Markets*. en. SSRN Scholarly Paper. Rochester, NY.
- Kirchberger, M. and K. Beirne** (June 2021). “Concrete Thinking About Development”. en. *Trinity Economics Papers*.
- Klein, E. and D. Thompson** (2024). *Abundance*. en. Avid Reader Press.
- Klosin, S.** (2024). “Dynamic Biases of Static Panel Data Estimators”.

- Krasnokutskaya, E.** (2011). “Identification and Estimation of Auction Models with Unobserved Heterogeneity”. en. *The Review of Economic Studies* 78.1.
- Krasnokutskaya, E. and K. Seim** (Oct. 2011). “Bid Preference Programs and Participation in Highway Procurement Auctions”. en. *American Economic Review* 101.6, pp. 2653–86.
- Kroft, K., Y. Luo, M. Mogstad, and B. Setzler** (June 2020). *Imperfect Competition and Rents in Labor and Product Markets: The Case of the Construction Industry*. en. Tech. rep. w27325. National Bureau of Economic Research.
- (2025). “Imperfect Competition and Rents in Labor and Product Markets: The Case of the Construction Industry”. en. *American Economic Review*.
- Lewis, G. and P. Bajari** (2011). “Procurement Contracting With Time Incentives: Theory and Evidence *”. en. *The Quarterly Journal of Economics* 126.3.
- Lewis-Faupel, S., Y. Neggers, B. A. Olken, and R. Pande** (Aug. 2016). “Can Electronic Procurement Improve Infrastructure Provision? Evidence from Public Works in India and Indonesia”. en. *American Economic Journal: Economic Policy* 8.3, pp. 258–83.
- Li, T. and X. Zheng** (2009). “Entry and Competition Effects in First-Price Auctions: Theory and Evidence from Procurement Auctions”. en. *The Review of Economic Studies* 76.4.
- Limão, N. and A. J. Venables** (2001). “Infrastructure, Geographical Disadvantage, Transport Costs, and Trade”. en. *The World Bank Economic Review*.
- Liscow, Z., W. Nober, and C. Slattery** (Sept. 2023). *Procurement and Infrastructure Costs*. Working Paper.
- Liu, H., V. Kwigizile, W.-C. Huang, and W. M. University** (2022). “Competitive Bidding in Construction Contracting”. en. *Michigan Department of Transportation*.
- Maskin, E. and J. Riley** (2000). “Uniqueness of Equilibrium in Sealed High-Bid Auctions”. en. *The Review of Economic Studies*, 67.
- Mehrotra, N., M. Turner, and J. Uribe** (2024). “Does the US have an infrastructure cost problem? Evidence from the interstate highway system”. *Journal of Urban Economics* 143.C.
- Michaels, G.** (2008). “The Effect of Trade on the Demand for Skill: Evidence from the Interstate Highway System”. *The Review of Economics and Statistics* 90.4, pp. 683–701.
- Oster, E.** (Apr. 2019). “Unobservable Selection and Coefficient Stability: Theory and Evidence”. EN. *Journal of Business & Economic Statistics*.

- Paarsch, H.** (1992). “Deciding between the common and private value paradigms in empirical models of auctions”. *Journal of Econometrics* 51.1-2. Publisher: Elsevier, pp. 191–215.
- Smith, N.** (May 2017). “The U.S. Has Forgotten How to Do Infrastructure”. en. *Bloomberg.com*.
- Stigler, G.** (1968). *The Organization of Industry*. en. R.D. Irwin, Inc., located in Homewood, Illinois.
- Trebbi, F. and M. B. Zhang** (Nov. 2022). *The Cost of Regulatory Compliance in the United States*. en. Tech. rep. w30691. National Bureau of Economic Research.
- Vartabedian, R.** (Nov. 2021). “Years of Delays, Billions in Overruns: The Dismal History of Big Infrastructure”. en. *New York Times*.
- World Bank** (2020). *Benchmarking Infrastructure Development 2020 : Assessing Regulatory Quality to Prepare, Procure, and Manage PPPs and Traditional Public Investment in Infrastructure Projects*. Tech. rep. World Bank: World Bank.

A Supplemental Figures and Tables

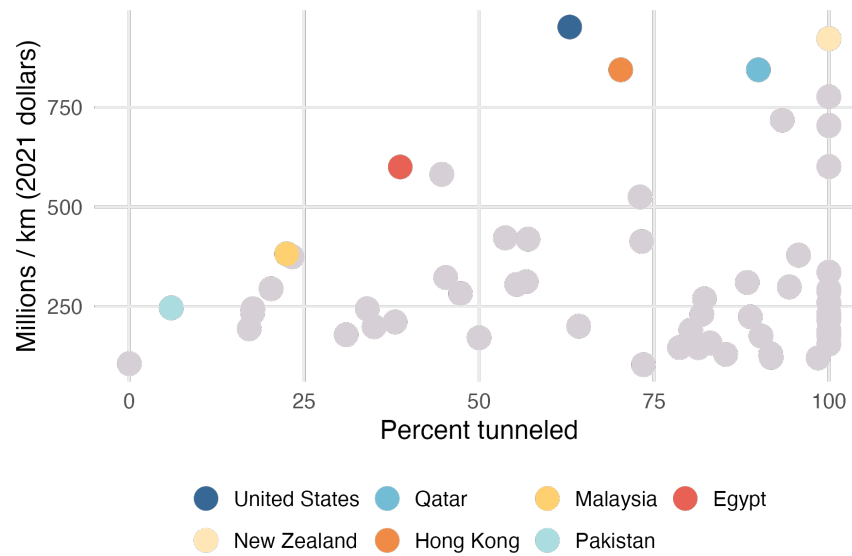


Figure A.1: Mean spending per kilometer over complexity (mean percent of project tunneled)

Note: Figure plots the mean spending per kilometer over mean percent of project tunneled for 59 countries. Country averages are taken over projects, so that each project is weighted equally. When weighting each project by length, the U.S. falls to 9th most costly. Data is from the Marron Institute at NYU.

State	Mean	Median	Mean %	Median %	Source
Massachusetts	-26,990	15,653	8.46%	1.67%	Bolotnyy and Vasserman, 1998-2015
California	457,746	58,168	4.32%	3.25%	CALTRANS, 2009–2021
Colorado	169,165	37,482	3.7%	2.1%	CDOT, 1999–2023
New York City	-13,052	-77,648	-2.14%	-3.87%	NYCDOT, 2009–2021

Note: Table shows the cost overrun (final price minus winning bid) for four states. Data for Massachusetts comes directly from Table 1 in Bolotnyy and Vasserman, 2023 and covers only bridges. The other data comes from the relevant DOTs. The sample sizes are 1651, 2256, and 1265, for California, Colorado, and New York, respectively. For the latter three states, estimates control for inflation and are in 2023 dollars .

Table A.1: Cost Overruns for Four States

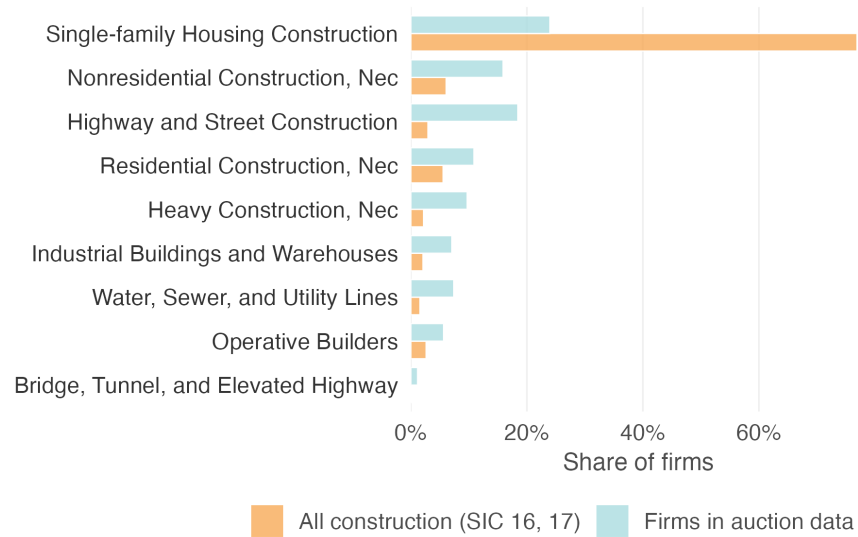
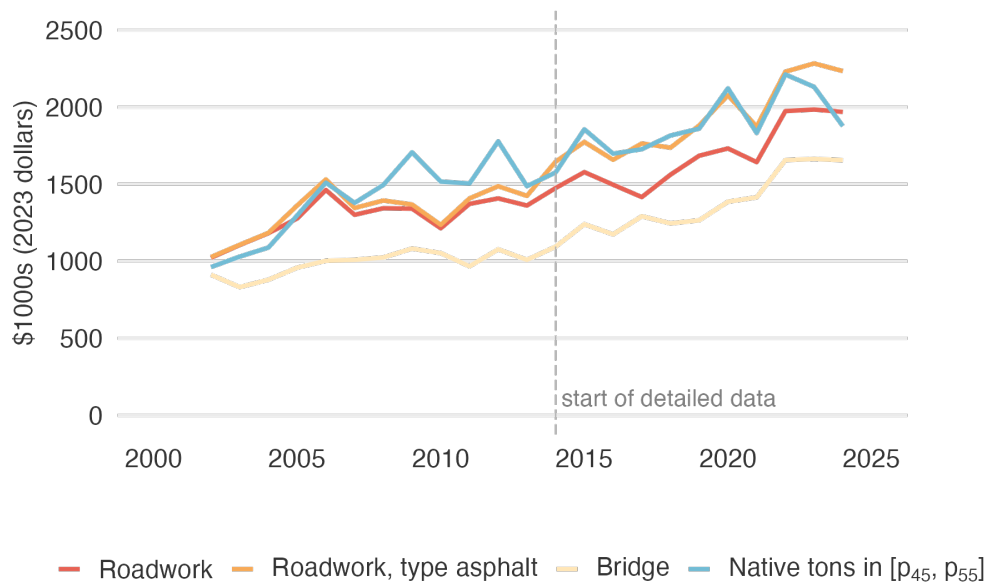
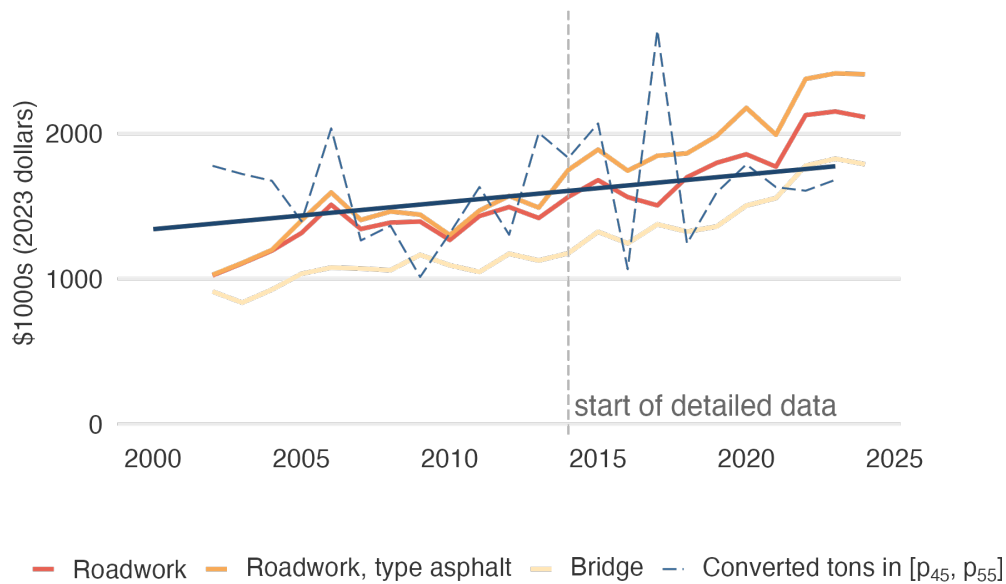


Figure A.2: Shares of Infrastructure Firms and All Firms by 4-Digit SIC Code (SIC 15–16)

Note: Figure plots the distribution of firms across four-digit SIC codes. “Firms in auction data” are those matched to the DOT auction sample by name and state, while “All firms” are Dun & Bradstreet firms in the broader SIC 15–16 construction categories. Each unique firm name is counted once. A small number of firms have multiple SIC codes; shares therefore computed by weighting each firm’s contribution equally across its distinct SIC codes.



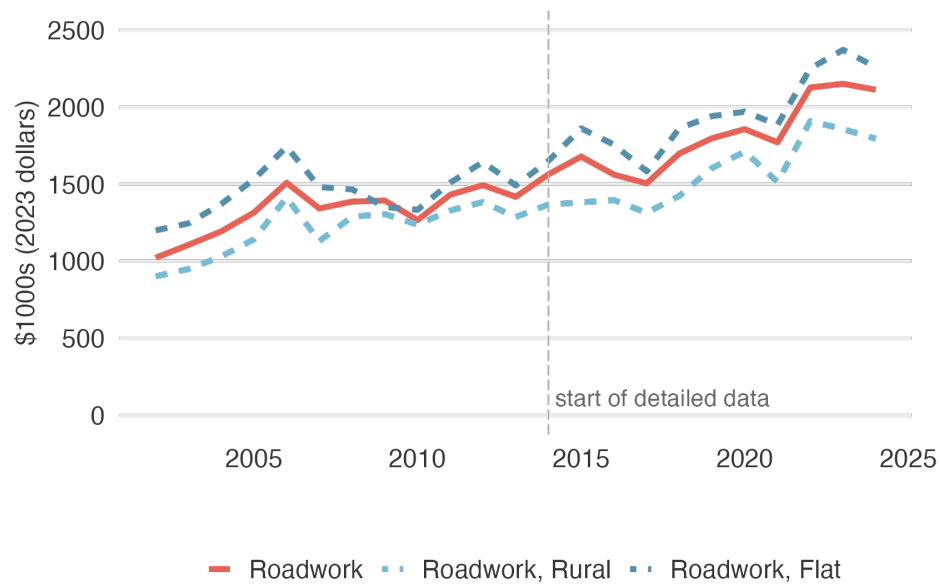
(a) Balanced Panel



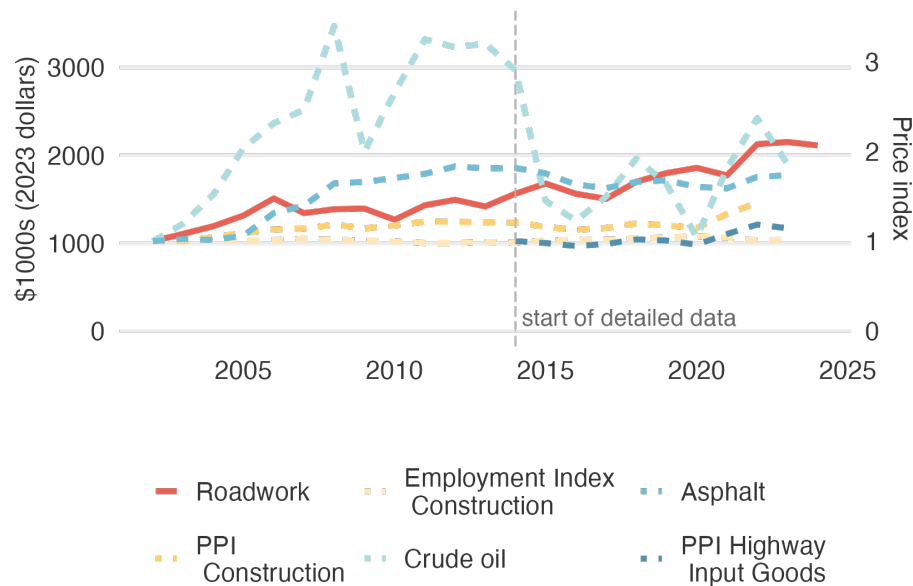
(b) Trend with Converted Tons

Figure A.3: Price of Infrastructure Over Time Robustness

Note: Panels plot the median winning bid of DOT infrastructure projects over time, as in Figure 1. All prices control for inflation and are in 2023 dollars. Panel (a) replicates Figure 1 but uses a balanced panel of states, which includes 41 states. In Panel (b), the dark dashed blue shows the same exercise for the alternative measure of size, converted tons, for the six states that have bid-schedule data since 2000. Construction of the converted tons measure is described in Appendix Section B.4. The solid dark blue line shows the jackknife bias-corrected linear trend estimated on data from 2000 to 2024 for the dark-blue series, based on leave-one-state-out resampling. Leave-one-out slopes have large variation: \$200, \$18,000 \$19,000 \$21,000 \$23,000 and \$ 64,000 per year, and with only six states the 95% jackknife confidence interval includes zero.



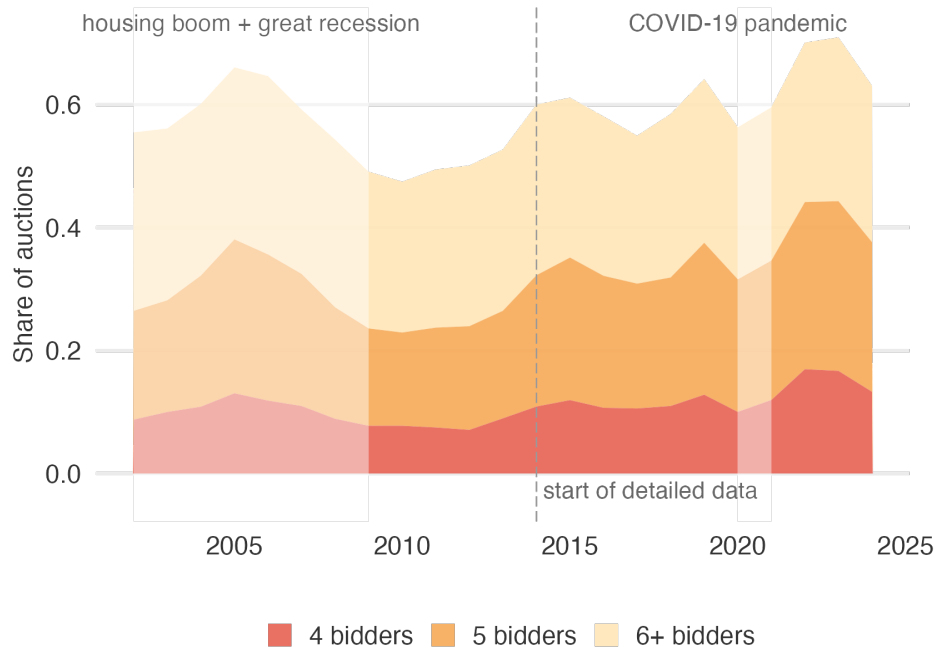
(a) Price and Location



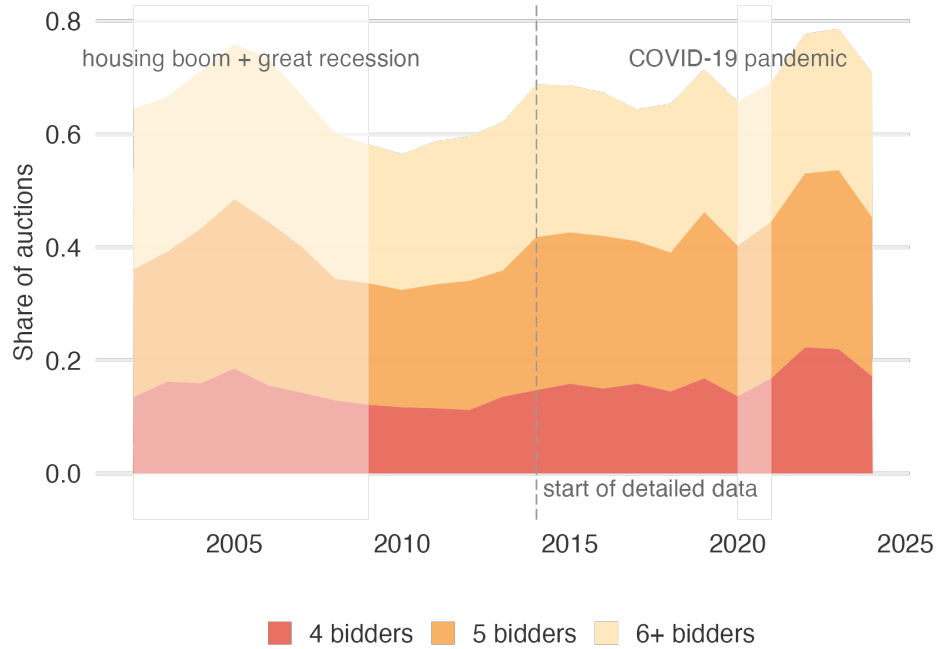
(b) Price and Input Prices

Figure A.4: Price of Infrastructure Over Time

Note: Figures plot the median winning bid for roadwork auctions over time, as in Figure 1. Figure (a) overlays the roadwork median with median prices for projects located only in rural counties and only in flat-terrain counties. Figure (b) overlays the roadwork median with three input-cost indices: (i) the Construction Producer Price Index (PPI), (ii) the Employment Cost Index (ECI) for construction labor, (iii) the national average crude oil price, the main feedstock for asphalt, and (iv) the PPI for Net Inputs to Highways and Streets, Goods. This index is only available since 2015. All four indices are also divided by the CPI for comparability.



(a) Number of Bidders Over Time, Balanced State Sample



(b) Number of Bidders of Time, Roadwork Auctions Only

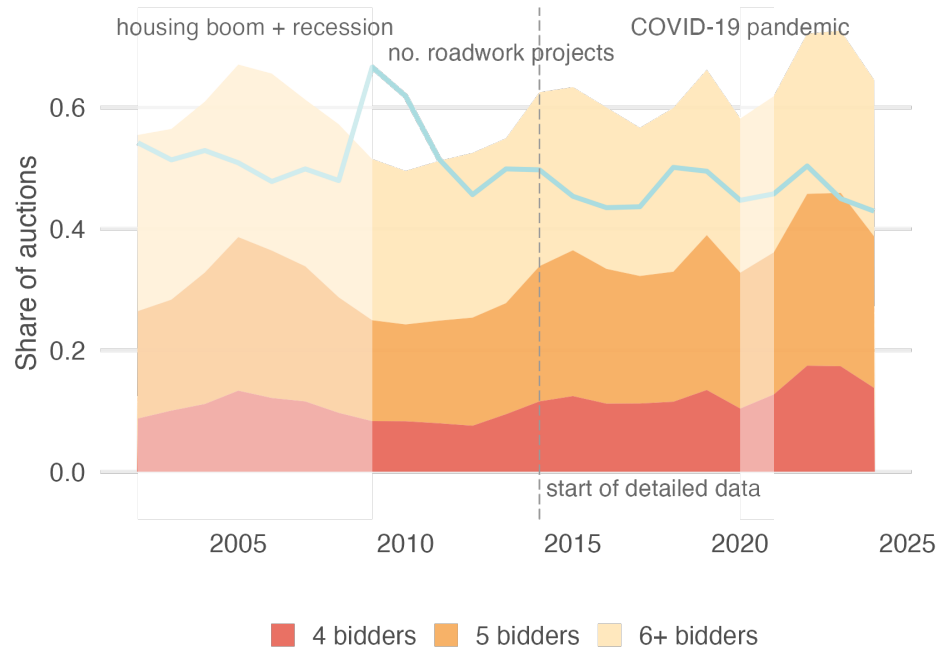
Figure A.5: Macro-economic Effects and Decreasing Bidder Trend: Robustness

Note: Figures plot the share of auctions by number of bidders over time, as in Figure 2. The dashed vertical line marks the beginning of the detailed dataset used later in the paper, indicating the period covered by the structural estimation.

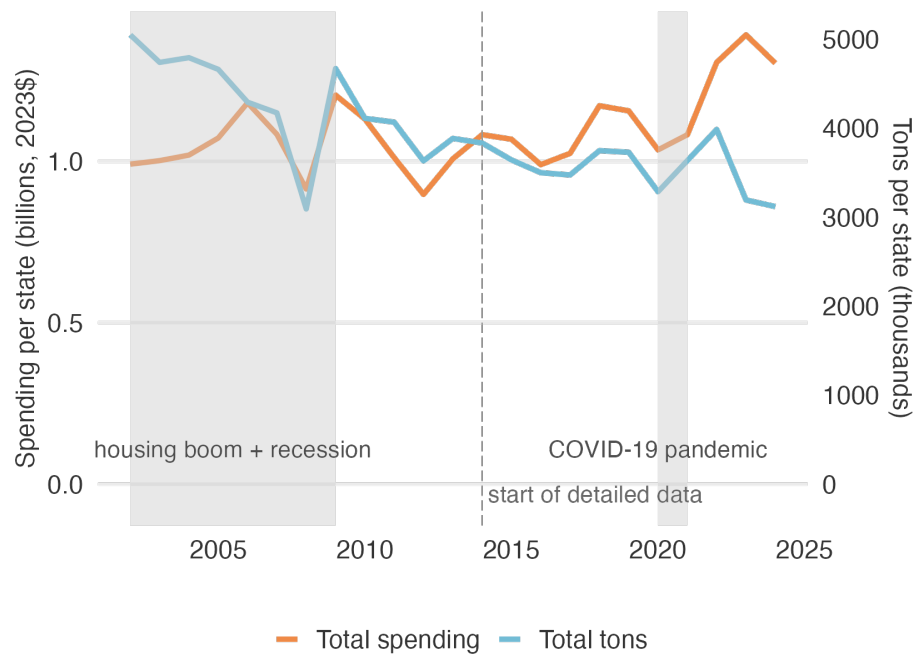
Metric	State Market Entrant	Procurement Market Entrant
Share of bids	3.3%	2.4%
Share of wins	3.0%	2.4%
Share of value	4.9%	3.8%

Note: Table reports the share of bids, wins, and total contract value captured by firms in the year they enter the market. Data begin in 2010 to include all states while allowing for at a minimum three-year lead-in period, to avoid misclassifying firms' first appearance in the data as true market entry.

Table A.2: New Entrants Account for Only A Minor Share of Market



(a) No. of Bidders Over Time with Total No. of Projects



(b) Total Spending and Total Size (Tons) Over Time

Figure A.6: Flat Infrastructure Demand Over Time, Though Rising Spending

Note: Figures (a) plot the share of auctions by number of bidders over time, as in Figure 2, along with the total number of projects each year. Figure (b) plots the total spending each year, along with the total size of projects, as measured by total tons of inputs.

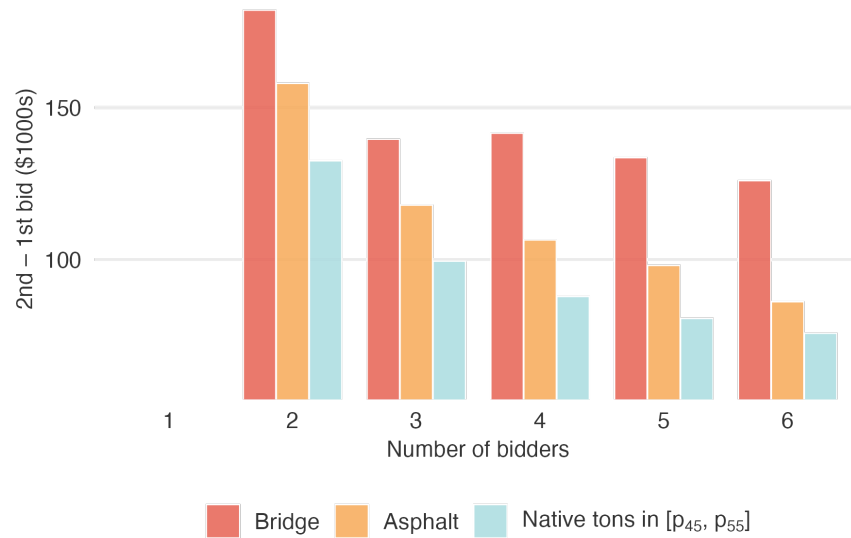


Figure A.7: 2nd - 1st Bid Gap

Note: Figure plots the national median difference between the second lowest and lowest bids by number of bidders in the auction for all bridge, asphalt roadwork, and roadwork projects within a narrow range of native tons, respectively. To account for state omitted variables that effect bid gaps and number of bidders, gaps are demeaned by state with the national mean added back. When taking means, 1% of the sample is trimmed due to large outliers.

	<i>Dependent variable:</i>			
	Winning Bid			
	Road projects	Asphalt only projects	+ Size	+ Complexity
California	9,222.78*** (536.85)	6,471.71*** (520.28)	4,547.30*** (493.51)	1,733.85*** (514.73)
Log Tons of Asphalt			2,961.60*** (144.37)	2,730.33*** (233.22)
Asphalt Share of Price				-9,231.65*** (1,975.57)
Log No. items				2,169.28*** (614.53)
Constant	1,463.40*** (250.08)	1,530.78*** (247.51)	-23,295.48*** (1,231.98)	-21,997.62*** (1,879.30)
Observations	3,613	2,744	2,744	2,744
R ²	0.08	0.05	0.18	0.24

Note:

*p<0.1; **p<0.05; ***p<0.01

Note: Table shows the difference in average winning bids between California and Kentucky for projects from 2018 to 2024. The outcome is normalized by CPI and is in 2023 dollars.

Table A.3: Difference in Prices: California vs. Kentucky

Covariate	Within-state correlation	
	Estimated coefficient	(SE)
Log avg. petrol product price	0.991***	(0.286)
Log construction wage	-0.006	(0.044)
Pct unionized	-0.010***	(0.002)
State prevailing wage law	0.107**	(0.045)
Log population	0.001	(0.022)
Log median HH income	0.041*	(0.024)
Log construction establishments	1.138***	(0.077)
Log firms in market	-0.365***	(0.029)
Herfindahl–Hirschman Index (0–1)	1.445***	(0.229)
State FE	yes	
Year FE	yes	
Observations (by row):	927, 927, 927, 815, 927, 927, 927, 927	
Adj. R^2 (by row):	0.075, 0.063, 0.086, 0.050, 0.063, 0.066, 0.246, 0.206, 0.102	

Note: Table reports estimates from a log-linear regression of prices (log winning bids) on covariates, analogous to Table 3 but without state fixed effects. Given project design and engineering vary substantially by state, the correlations are difficult to interpret.

Table A.4: Price Growth and Covariates

	<i>Dependent variable:</i>		
	Δ Log Bid	Log Bid	Log Bid
	(County First-Differences)	(IV: 75 miles)	(IV: 125 miles)
No. bidders	-0.014*** (0.003)	-0.093 (0.065)	-0.215*** (0.068)
Log tons	0.133*** (0.002)	0.136*** (0.002)	0.133*** (0.002)
State FE		yes	yes
Time Trend	yes	yes	yes
Lag log bid	yes		
First-stage partial F		55.2	59.1
Observations	19,950	22,050	20,728
Adjusted R^2	0.430	0.485	0.407

Note:

* $p < 0.1$; ** $p < 0.05$; *** $p < 0.01$

Note: Table reports estimates from modified regressions from those reported in Table 4. Column shows results from a first differences model: $\Delta \log(\text{price})_{c,t} = \gamma \Delta N_{c,t} + \Delta x_{c,t} + \mu_t + \Delta \epsilon_{c,t}$, where $\Delta y_{c,t} \equiv y_{c,t} - y_{c,t-1}$. Columns 2-3 report IV results analogous to column (6) in Table 4, but adjusting the threshold distance between the auction and the firm location.

Table A.5: Estimates of the Impact of Number of Bidders on Prices

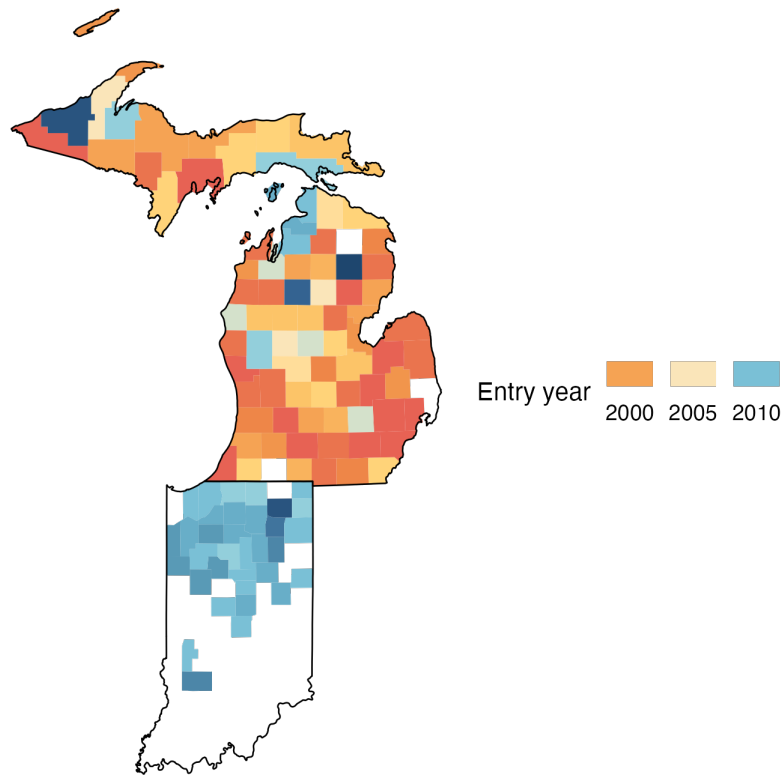


Figure A.8: Example of a firm in expanding to a new state

Note: Figure plots the expansion of one representative multi-state contractor across counties. Shaded areas indicate the first year the firm bid on a project in each county, with lighter colors corresponding to earlier years.

trim	Means				
	Cost	Margin	Cost (homogenized)	Margin (homogenized)	Markup
0%	2603	870	1493	801	1.31
1%	2477	693	1593	669	0.22
3%	2243	604	1638	576	0.22
5%	2079	554	1643	528	0.23

Note: Table reports the means of model estimates of costs, margins, and markup, by the percent of observations that are trimmed from each tail of the cost distribution. The markup stabilizes after 1% of the sample is trimming.

Table A.6: Means of Model Estimates by Tail Trim Percent

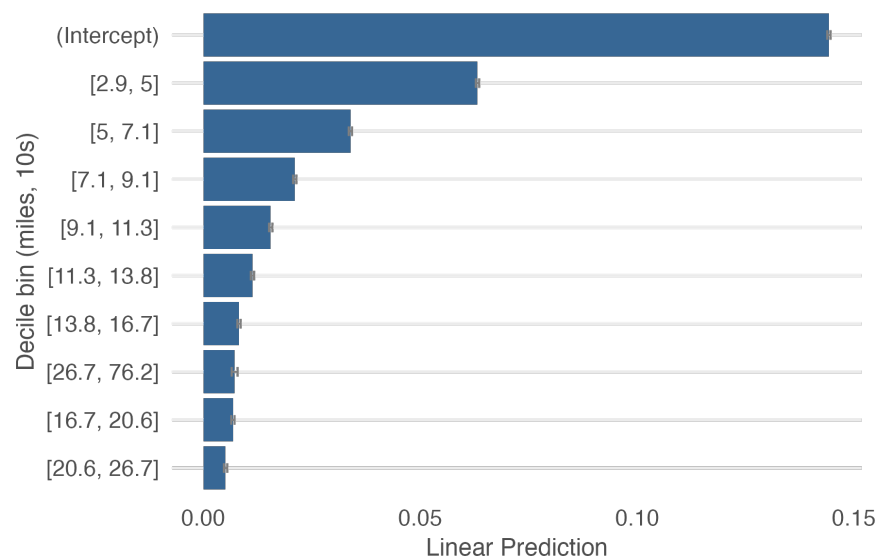


Figure A.9: Estimates from Regression of Auction Entry on Distance Bins

Note: Figure plots predicted probabilities (point estimates plus intercept) from a linear probability model of a firm entering an auction, as a function of the distance between the firm’s county and the auction’s county. Firm locations are taken from a matched subsample of Dun and Bradstreet’s heavy construction firms. The 6% of auctions lacking county identifiers—for example, those designated as “state-wide”—are dropped. An active firm is defined as one that submitted a bid in any auction during the year, or in both the preceding and following year. All auction types are included.

B Data Appendix

B.1 Auction data

For each state year, I have a file with all lettings for the year. For each letting, I extract the job id, the county, the letting date, a description of the job, the total tons of input required for the job, the engineer’s estimate when available, the bidder name and id, the bidder rank, the bid, and the project engineering “type,” a variable pre-defined by the data provider as the most expensive input category of item in the project.

States vary by their start year in the sample and their data coverage for the project description and the engineer estimate. Data is missing when either the state does either does not record it or does not share it with the data provider. Finally, while the tons variable is always present, states vary in the share of projects with inputs measured in tons; this reflects both variation in project type and in chosen units. Table A.7 below gives the min, max, 25th, 50th, and 75th percentile for each of these variables, as well as the number of auctions per state. Additionally, all states except Rhode Island



Figure A.10: Estimates from Regression on Indicator for Out-of-State Bidder

Note: Figure plots the mean share of wins for out-of-state entrants and the control group after removing state means; whiskers show 95% confidence intervals. Estimates come from OLS on the state-demeaned outcome, shifted back by the control-group mean. The left panel includes all auctions, while the right panel restricts to auctions with exactly three bidders to control for strategic effects. Home-state observations for entrants are excluded, and the unit of observation is firm by state.

provide county information.

variable	min	p25	p50	p75	max
Start Year	1993.00	1996.75	1998.00	2001.00	2007.00
Auctions per year	37.00	99.00	214.00	346.00	952.00
Description	0.61	0.99	1.00	1.00	1.00
Tons > 0	0.44	0.68	0.72	0.77	0.93
Engineer Estimate	0.00	0.00	0.00	0.68	1.00

Table A.7: Data coverage quantiles

For much of the analysis, I restrict to sample of projects identified as “roadwork”. I define these projects as projects of engineering type as asphalt or concrete pavement, or with descriptions containing specified roadwork keywords: pavement, grinding, paving, resurfacing, milling, route, or roadwork. I exclude projects of engineering type bridge. This rule is ultimately very similar to using the “roadway resurfacing and repair” LLM classification based purely on the description, described

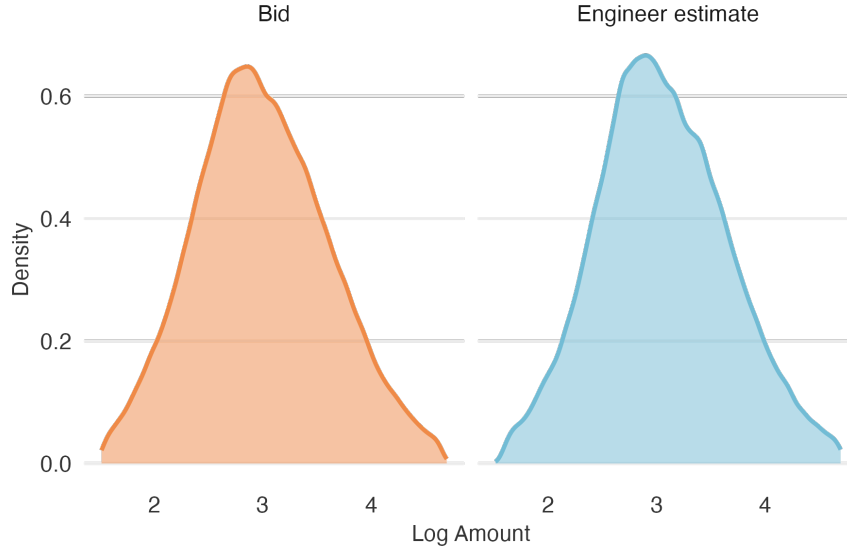


Figure A.11: Bid Density versus Engineer Estimates

Note: Figure plots kernel densities of bids and engineer estimates in log amounts (of thousands of dollars). The left panel shows bids and the right panel shows engineer estimates. The sample covers the 18 states with engineer estimates in the data. The top 1 and bottom 1 percent of bids are dropped.

in Section B.3. Roadwork auctions account for 47% of all auctions.

B.2 Firm data

The auction data include firm names and a within-state identifier, but no additional information on the firms. To illustrate firm industries, Table A.9 reports self-descriptions and listed sub-industries for a random sample of five firms.

Multi-state firms To identify multi-state firms, I use a three-stage process. First, I clean bidder names using standard procedures. Second, I build candidate pairs with fuzzy string matching. Specifically, I apply a constrained Jaro–Winkler (JW) procedure: require overall JW distance 0.20 after cleaning; require at least one shared meaningful word; require the first meaningful word to share the same initial; if that first word has ≤ 4 characters, require an exact match; if it has > 4 characters, allow a first-token JW ≤ 0.08 . This modified rule outperforms a simple JW cutoff and still shrinks the candidate set substantially. Third, I classify the remaining non-exact pairs with a large-language model (OpenAI Chat Completions, gpt-5). Each request includes the two firm names, their states, and a fixed prompt requiring a strict “Yes” or “No” on whether the names are the same company, with a confidence score. I construct firm identities by taking the transitive closure

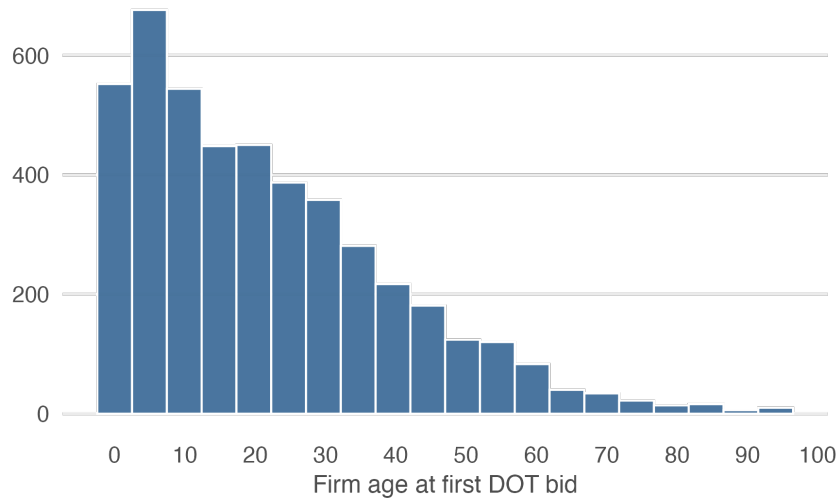


Figure A.12: Firm Age at Infrastructure Market Entry

Note: Figure plots a histogram of firm age at the time of the first infrastructure auction bid, for the 50% of auction participants matched to Dun & Bradstreet records in Heavy or Building Construction (see Appendix B.2). To reduce measurement error, the sample is restricted to firms that (i) bid in their matched state of location, (ii) have a first bid at least three years after the auction dataset’s start date for that state, and (iii) have a Dun & Bradstreet “start year” prior to the first year the database records the firm. Ages are trimmed to the 0–100 year range, with values outside this interval treated as noise, dropping about 5% of observations. The final sample covers 17% of infrastructure firms. Relative to unmatched firms, they are slightly more active bidders, participating in 28 auctions on average compared to 18, though they win slightly less often, at 23% compared to 27%.

over all positive matches, assigning each equivalence class of names a unique identifier. A firm is classified as multi-state if its identifier appears in more than one state.

Table A.8 reports validation on 100 candidate pairs, split above vs. below the 25th percentile of the model’s confidence. I determine ground truth by searching each firm online and checking websites and state records. Reassuringly, I see little evidence that extremely similar cross-state names belong to different firms. Precision is high – 90% in the higher-confidence set and 93% in the lower-confidence set – while accuracy is high in the former and lower in the latter. That’s by design: the false-positive rate stays low (3% and 5%). The benefit is that identified interstate links are reliable; the cost is a slight under-count. I quantify the under-counting with a back-of-the-envelope exercise: fuzzy matches contribute 22% of total multi-state firms; the model finds 48% of true matches in the most-uncertain quartile and 75% in the most-certain three-quarters; overall that’s 68% of true positives, implying the total multi-state count is under-estimated by about 10%.

Firm covariates To obtain firm locations and start years, I use Dun & Bradstreet, a commercial firm

Metric	High-confidence	Low-confidence
	Top 75%	Bottom 25%
Accuracy	92.0%	68.0%
Precision (PPV)	90.0%	93.3%
False positive rate	2.6%	4.8%
N	50	50

Note: Table reports validation for the LLM’s classification of candidate firm pairs as same firm vs. different firms. Pairs are pre-filtered by a Jaro–Winkler (JW) similarity threshold. Panel A randomly samples 50 pairs from the top 75% by the model’s confidence; Panel B randomly samples 50 pairs from the bottom 25%.

Table A.8: Validation of interstate firm identities

analytics provider, and pull all firms in SIC 16 (“Heavy Construction, Except Building Construction Contractors”) and SIC 15 (“Building Construction—General Contractors and Operative Builders”). I standardize names and first link exact cases. For non-exact cases, I search candidate Dun & Bradstreet firms in the same state or in contiguous states and apply a constrained Jaro–Winkler procedure: overall distance 0.06 after cleaning; require at least one shared meaningful token; require the first meaningful token to share the same initial; if that first token has 4 characters, require an exact match; if it has > 4 characters, allow a first-token Jaro–Winkler 0.07. Stronger restrictions on the first token prevent false matches between companies that are technically similar in spelling but clearly different firms, e.g. “PJ Construction Company” and “KM Construction Company”. When multiple candidates satisfy all rules, I retain the candidate with the smallest overall Jaro–Winkler distance. Exact links cover 29% of auction firms in SIC 16 and 43% pooling SIC 15 and 16; the full procedure yields a 48% match rate. A hand audit of 60 randomly sampled pairs finds 95% correct (Table A.10).

B.3 Project Description Based Classification

I classify auctions based on brief project descriptions using OpenAI’s GPT-4o-mini language model (March 2025 release) via the chat.completions.create API. The 9 categories given to OpenAI to classify on in Table A.11 below.

Firm	State	Self-description	Listed sub-industries
SCS Engineers	California	SCS is an employee-owned environmental consulting and construction firm that designs and implements sustainable environmental solutions.	Solid and hazardous waste management; Renewable energy; Remediation; Carbon capture.
Dan R. Dalton Inc.	Colorado	Dan R. Dalton Inc. is North America's heat straightening specialists for repairing damaged steel bridges and container cranes.	Railroad bridge repair; Highway bridge repair; Crane repair; Ship repair.
Cutting Edge Group LLC	New York	We specialize in civil construction and design work. We also perform mechanical and electrical work to meet our client needs.	Heavy civil; Building; Parks & recreation; Horizontal directional drilling.
Rutledge Excavating Inc	Pennsylvania	Rutledge Excavating prides itself on delivering top-notch full site development services tailored to meet your needs.	Land clearing and demolition; Concrete work; Erosion control; Stormwater management; Drainage systems; Utility installations; Road construction.
Dome Corporation of America	Washington	We do more than build spaces—we create environments where your team can grow, thrive and innovate.	Capital project feasibility study; Preconstruction; General construction; Commissioning & turnover; Modernization.

Note: Table reports names, state, self published business descriptions, and listed sub-industries for five randomly sample firms in the data. Firms without public websites were excluded.

Table A.9: Firm Profiles

SIC group	Validation	
	Share correct	N
Building construction (SIC 15)	96.7%	30
Heavy construction (SIC 16)	93.3%	30

Note: Table reports share of correct firm matches for 60 randomly sampled firms, stratified by two-digit SIC code.

Table A.10: Validation accuracy of Firm Name Match

B.4 Measuring Project Size

An accurate measure of project size is important for comparing project prices. The data includes the total tons for every project, which is a natural measure of size. However, this measure only includes items that whose unit is tons – projects may differ in which units they use for similar inputs,

Category	Definition
Roadway Surfacing and Repair	Projects involving milling, grinding, resurfacing, and overlaying of roadways, including hot mix asphalt resurfacing, rubberized asphalt overlays, and bituminous resurfacing.
Bridge Construction and Repair	Bridge replacements and rehabilitations, high-rise bridge repairs, painting, deck repairs, and waterproofing.
Culvert and Drainage	Culvert repair and replacement, stormwater pond maintenance, drainage improvements, and minor drainage structure work.
Safety and Traffic Improvements	Traffic signal upgrades, installation of cable barriers, new or replacement guardrails, traffic signs and devices, and traffic control improvements.
Grading and Earthwork	Clearing and grubbing, grading, slope protection, slide repairs, and other work altering the contour of land.
Pedestrian, Sidewalk, and Cycling Infrastructure	Construction or improvement of trails, bikeways, sidewalks, curb ramps, curb and gutter, and safety features for non-motorized users.
Environmental and Landscaping	Environmental mitigation, tree trimming, landscaping, mowing, and other activities to improve or protect the environment.
Facilities	Construction, renovation, or maintenance of buildings, welcome centers, rest areas, park-and-ride structures, weigh stations, or other facilities.
Unknown	Category not specified.

Table A.11: Project categories and definitions.

particularly across states.

I use the detailed data subsample from 2014 to 2023 to show that tons is a good proxy for size *within* state, but less so between states. I then use this data to construct a nationally comparable measure of size for projects between 2024 and 2019.

Table A.12 shows that within states, log tons of a project is highly correlated with other totals of inputs by unit, with an R^2 around .6. However, without state fixed effects, the R^2 drops substantially, in some cases close to 0. This suggests that within a state, the amount of inputs measured in tons and the amount of other inputs are complements. Across states, however, they substitute each other, suggesting states have different unit preferences for the same inputs.

I construct a cross-state measure of project size as total asphalt tons. Using the 2014–2023 item-level subsample, I restrict to projects classified as asphalt. Within those projects, I then flag the items that are truly asphalt—keywords such as asp, hma, hot mix, plant mix, bitum, superpave, or course layers (wear, intermediate, surface, leveling, binder)—while excluding removal, milling, grinding, saw-cut, and coating lines. Each flagged item is converted into short tons using standard engineering factors:

Unit	β	p	Adjusted R^2
<i>Panel A: with state fixed effects</i>			
Log Square Yards	0.2428	0.000	0.623
Log Linear Feet	0.3161	0.000	0.625
Log Each	1.6802	0.000	0.629
Log Lump Sum	1.0968	0.000	0.597
<i>Panel B: without state fixed effects</i>			
Log Square Yards	0.1259	0.000	0.018
Log Linear Feet	0.3848	0.000	0.103
Log Each	0.5115	0.000	0.008
Log Lump Sum	0.6586	0.000	0.012

Note: Table shows coefficients from regressions of the indicated size measure on log tons. Tons are measures as the sum of amounts of all inputs in the project whose unit is tons.

Table A.12: Correlation between Log Tons and Other Measures of Project Size

one ton = 1.0, a metric ton = 1.1, a hundredweight = 0.056, a pound = 0.0005, a gallon of liquid asphalt = 0.00425, a cubic yard of compacted hot-mix = 1.96, a square yard of a 1.5-inch overlay = 0.0825, and a square foot under the same assumption = 0.0092. Summing these conversions at the project level yields total asphalt tons, along with the share of each project's inputs and costs that come from asphalt items. The resulting measure is more stable across states than the raw "tons" field, which reflects only items originally recorded in tons and therefore misses variation in unit conventions.

B.5 Measuring Distance to Firms

I construct a measure of the distance between project locations and bidding firms for the instrumental variable regression with first stage in equation 3. I use the auction data and the procedure described in Section B.2 to identify multi-state firms. I assign each firm's home state h as the first state in which it appears. I define "state entrants" for state s and year y as firms active in a state other than their home state that year: i is an "entrant" if $|\text{bid}_{i, st}| > 0$, $h_i \neq s$. I obtain firms' county locations from the matched sample with Dun & Bradstreet data (see Section 2.2). I use only matched firms in the Heavy Construction industry for this exercise, dropping Building Construction Firms, as the latter is a much larger industry and therefore has a higher probability of duplicate firms names. This match covers 60% of state entrants. I drop firms that "enter" in the first year of data for each state. Finally, I clean the sample by restricting the sample to firms with locations in the identified home

state and where said home state is a neighbor to the target state. Finally, I remove firms that ever establish a location within the target state. The final match covers 13% of entrants.

Each project is assigned a location based on either the centroid of its county or, if multiple counties are involved, the centroid of those counties. I exclude the 6% of projects referenced only by “District,” “various,” or “statewide.” The distance between a firm and a project is defined as the distance from the project location to the firm’s nearest establishment. Distance is calculated as the miles between the county centroid of firm location and the county centroid of the auction locatio

C Model estimation and details

C.1 Robustness of homogenization regression specification

Model	Median RMSE	IQR RMSE	Pooled RMSE
Baseline	0.250	0.092	0.271
+ unit covariates	0.223	0.065	0.220
+ item covariates	0.232	0.078	0.244

Table A.13: Cross-validated root mean squared error (RMSE) for three regression specifications. Median and inter-quartile range (IQR) are calculated across states; the pooled RMSE weights observations equally. The baseline specification includes fixed effects for engineering type and AI-derived project class, county-level terrain ruggedness and rurality, and log total inputs measured in tons. The “+ unit covariates” specification (preferred, corresponding to regression 8) adds log total material quantities aggregated by measurement unit—tons, cubic yards, linear feet, square yards, and counts of “each”—and log the number of distinct bid-schedule items; all physical quantities are converted to common units (e.g., pounds to tons). The “+ item covariates” specification further includes individual item quantities for every item used in at least 25% of projects, along with log the number of distinct items and the count of “rare” items below this threshold. Item effects are regularized with ridge penalties of $\lambda = 0.5$ on the item coefficients. All specifications ridge penalties of $\lambda = 1$ on the auction-level intercept, and 0 on the baseline covariates.

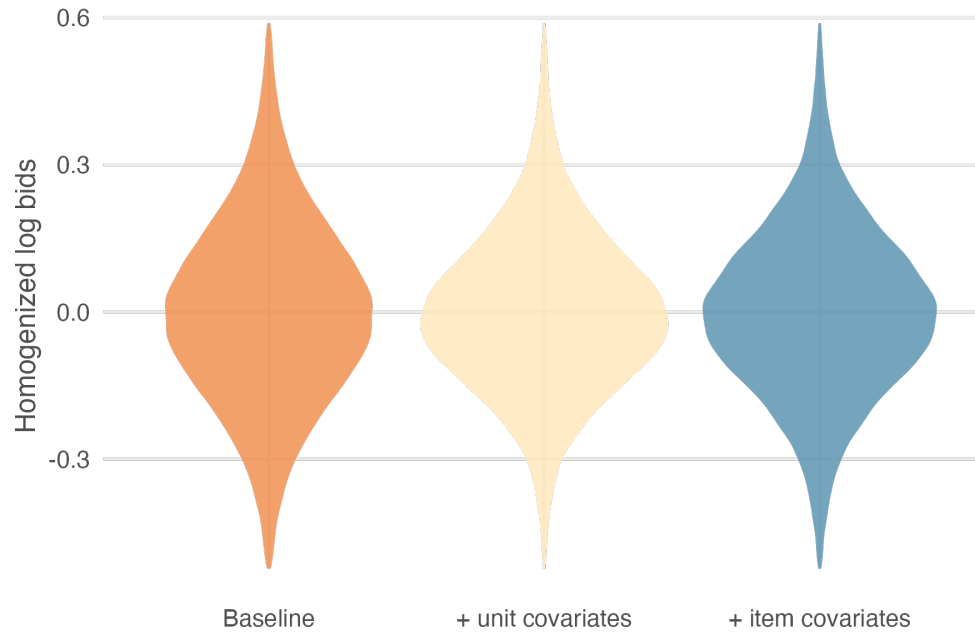


Figure A.13: Distribution of homogenized log bids by specification

Note: Figure plots violin densities of state-demeaned log bids for the three regression specifications defined in Table A.13. Observations beyond the 0.5th and 99.5th percentiles of the residual distribution are dropped before plotting.

C.2 Robustness of number of bidder prediction

	<i>Dependent variable:</i>					
	N		log bid	log homogenized bid		
	(1)	(2)	(3)	(4)	(5)	(6)
E[N], multinomial	1.639*** (0.015)				-0.032*** (0.002)	
E[N], xgboost		1.239*** (0.004)				-0.020*** (0.001)
N			-0.0005 (0.003)	-0.010*** (0.001)		
state FE	yes	yes	yes	yes	yes	yes
Projected R-squared	0.201	0.719	0	0.009	0.004	0.01
Observations	47,431	47,431	46,952	46,952	44,793	44,793

Note:

*p<0.1; **p<0.05; ***p<0.01

Table A.14: Columns (1) and (2) show the correlation between the predicted number of bidders from two models and the actual number of bidders. Columns (3) through (6) show the correlations between the number of bidders—either actual or predicted—and both actual and homogenized bids. All models include the same covariates: fixed effects for engineering type and AI-derived project class, county-level terrain ruggedness and rurality, total input quantities measured in tons, individual quantities for all bid items present in at least 25% of projects, the log number of distinct items, and the count of "rare" items below this frequency threshold. The multinomial model is a multinomial logistic regression with a ridge penalty. The XGBoost model is a gradient-boosted decision-tree model.

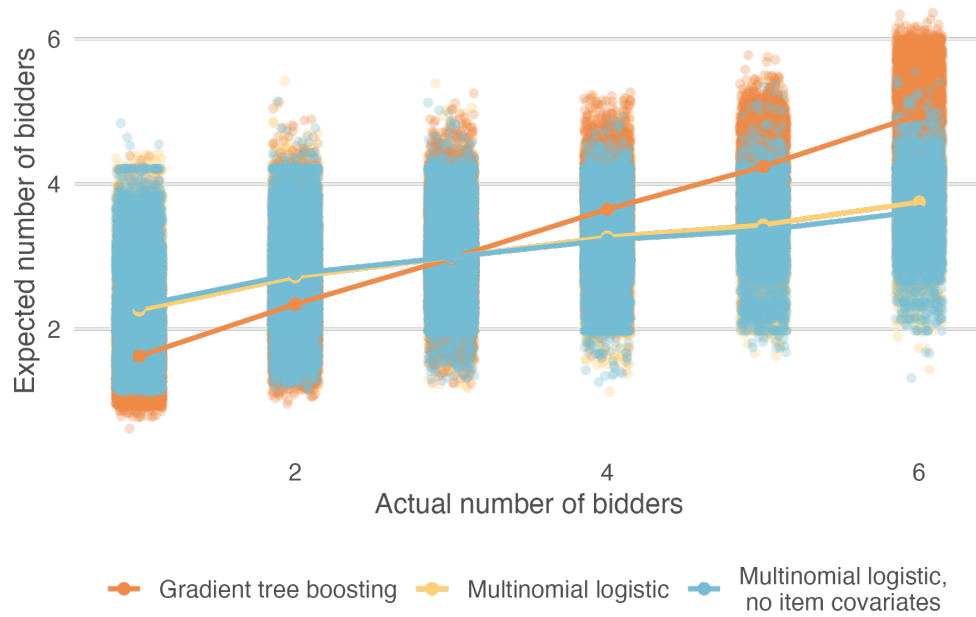


Figure A.14: Number of bidder prediction by different models

Note: This figure compares actual and predicted bidder counts using the models shown in Table A.14. It also includes predictions from a third model—identical to the multinomial logistic model but excluding bid schedule item covariates—to illustrate that omitting these covariates has only a minor impact on predictions. In contrast, switching from the penalized multinomial logistic model to the XGBoost model substantially alters predicted bidder counts.

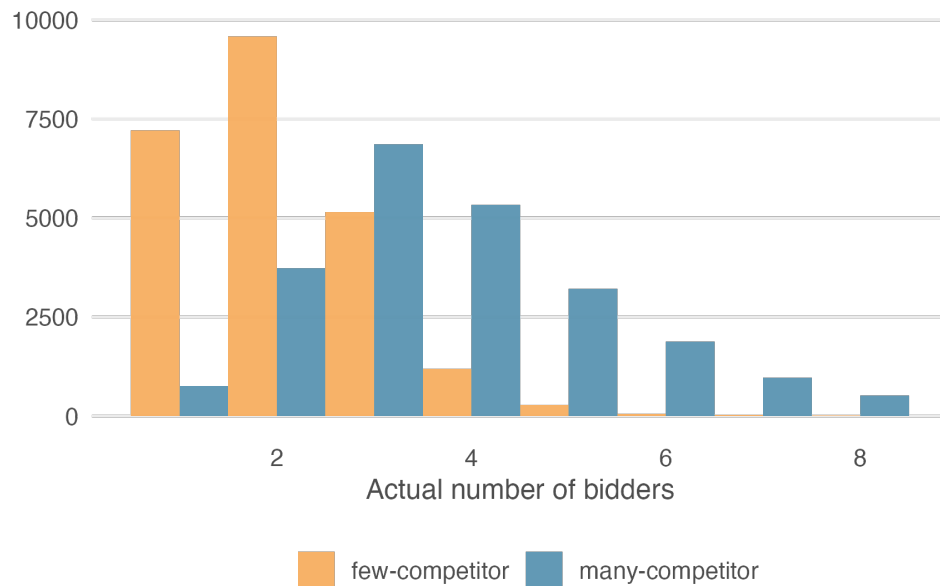


Figure A.15: Auction clusters based on predicted number of bidders

Note: Figure shows the distribution of auctions across the actual number of bidders – “few-competitor” and “many-competitor” – by the predicted bidder count clusters. Clusters are based on a median split of the predicted number of bidders from the gradient tree boosting (XGBoost) algorithm. Details on the algorithm are given in Table A.14.

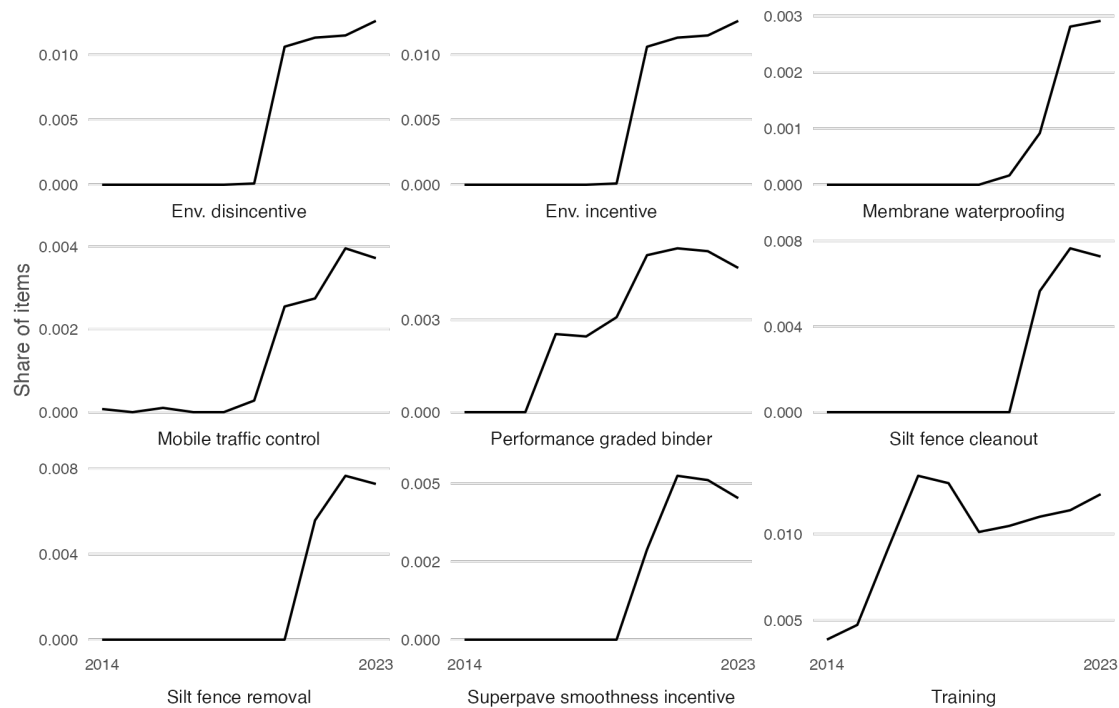


Figure A.16: Most Rapidly Growing Input Items, Nebraska 2014–2023

Note: Figure shows trends in Nebraska pay items, 2014–2023. Each panel plots the yearly share of items for the items with the highest share growth over the period. Shares are defined as the item’s count divided by the total items that year. Mobilization is omitted as a category, as that represents a lump sum amount and is not an input used to control for project characteristics. Over the full period, the average unique items per project rose by 19% from 54 to 64.

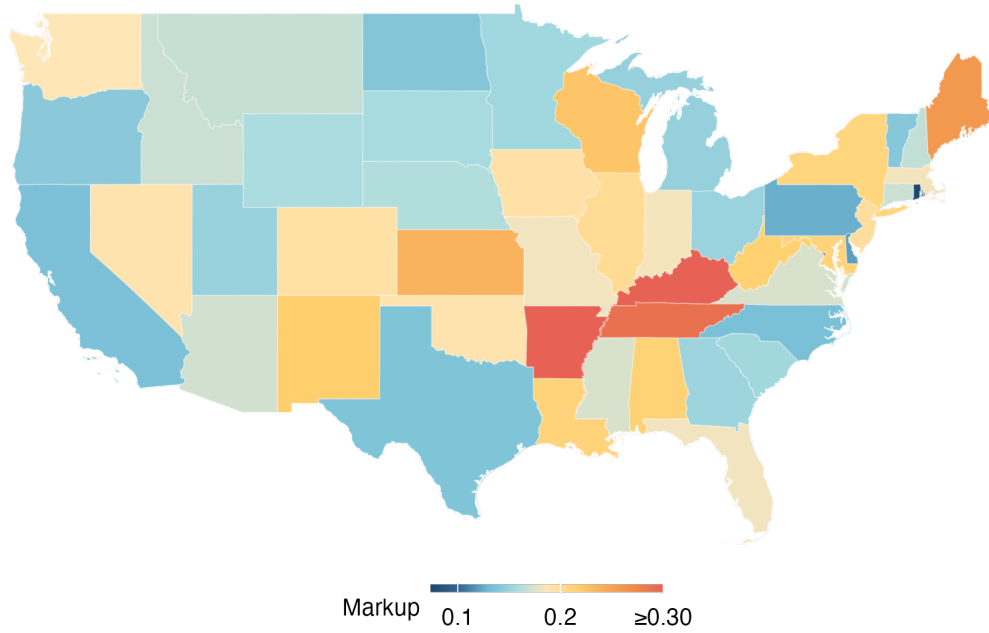


Figure A.17: Median Markup Across States

Note: Figure shows the median winning-bid markup by state from 2019 to 2023. Each state's color represents the median markup among winning bids on all roadwork projects over the period. Projects with negative cost estimates are dropped from the sample. Projects with a markup above .3 are grouped together in the highest-color category. The highest median markup estimate is for Kentucky, with an estimate of .54.

D Technical Appendix

D.1 Optimal bidding equation

The first order equation for the bidding firm is i

$$\begin{aligned}
 \frac{\partial \mathbb{E}\Pi}{\partial \beta} = & (\beta - c_i) \left[\sum_n \Pr(N = n|X) (n-1) (1 - F(\beta^{-1}))^{n-2} \left(-f(\beta^{-1}) \frac{1}{\beta'} \right) (1 - H(\beta|X)) \right. \\
 & + \sum_n \Pr(N = n|X) (1 - F(\beta^{-1}))^{n-1} (-h(\beta|X)) \Big] + \\
 & + \sum_n \Pr(N = n) (1 - F(\beta^{-1}))^{n-1} (1 - H(\beta|X)) = 0.
 \end{aligned}$$

Solving for the bid strategy gives

$$\beta(c_i) = c_i + \frac{\sum_n \Pr(N = n|X) (1 - F(\beta^{-1}|X))^{n-1} (1 - H(\beta|X))}{\sum_n \Pr(N = n|X) (1 - F(\beta^{-1}|X))^{n-2} \left((n-1) \frac{f(\beta^{-1}|X)}{\beta'} (1 - H(\beta|X)) + (1 - F(\beta^{-1}|X)) h(\beta|X) \right)}$$

D.2 Proof of equation 7

The proof closely follows the proof of Krasnokutskaya, 2011 Proposition 1. I show that the proposed bidding function satisfies the conditions for equilibrium. Note that I have already assumed equilibrium uniqueness.

Suppose $z_i = c_i \cdot y$ where $y \in [\underline{y}, \bar{y}]$ is a common multiplicative shock observed by all bidders in the auction. Let $\alpha(c_i)$ denote the symmetric equilibrium bidding strategy in the benchmark case $y = 1$, with distribution F of c_i , density f , entry probabilities $p_N(n)$ and government reserve $r \sim H(a)$.

The first order condition for the $y = 1$ auction over is

$$a_i = c_i + \frac{\sum_n P(n) (1 - F(\alpha^{-1}(a_i)))^{n-1} (1 - H(a_i))}{\sum_n P(n) \left[(n-1) (1 - F(\alpha^{-1}(a_i)))^{n-2} f(\alpha^{-1}(a_i)) \frac{1}{\alpha'(\alpha^{-1}(a_i))} (1 - H(a_i)) + (1 - F(\alpha^{-1}(a_i)))^{n-1} h(a_i) \right]}$$

Now consider a general auction with cost $z_i = c_i \cdot y$. The first order condition is

$$b_i = c_i y + \frac{\sum_n P(n) (1 - G(\beta^{-1}(b_i)))^{n-1} (1 - K(b_i))}{\sum_n P(n) \left[(n-1) (1 - G(\beta^{-1}(b_i)))^{n-2} \cdot g(\beta^{-1}(b_i)) \cdot \frac{1}{\beta'(\beta^{-1}(b_i))} (1 - K(b_i)) + (1 - G(\beta^{-1}(b_i)))^{n-1} k(b_i) \right]}$$

I hypothesize that the bidder strategy is $\beta(z_i) = y \cdot \alpha\left(\frac{z_i}{y}\right)$. I now show that this satisfies the differential equation. First note that $\frac{\partial \beta(z_i)}{\partial z_i} = y \cdot \alpha'\left(\frac{z_i}{y}\right) \cdot \frac{1}{y} = \alpha'\left(\frac{z_i}{y}\right)$. Next, $\beta^{-1}(b_i) = y \cdot \alpha^{-1}\left(\frac{b_i}{y}\right)$. Finally, the distributions can be written:

$$G\left(\beta^{-1}(b_i)\right) = F\left(\alpha^{-1}\left(\frac{b_i}{y}\right)\right),$$

$$g\left(\beta^{-1}(b_i)\right) = \frac{1}{y}f\left(\alpha^{-1}\left(\frac{b_i}{y}\right)\right).$$

Plugging these definitions into the general auction FOC gives:

$$b_i = c_i y +$$

$$y \cdot \frac{\sum_n P(n) \left(1 - F\left(\alpha^{-1}\left(\frac{b_i}{y}\right)\right)\right)^{n-1} \left(1 - K\left(y \cdot \alpha\left(\frac{z_i}{y}\right)\right)\right)}{\sum_n P(n) \left[\left(n-1\right) \left(1 - F\left(\alpha^{-1}\left(\frac{b_i}{y}\right)\right)\right)^{n-2} f\left(\alpha^{-1}\left(\frac{b_i}{y}\right)\right) \frac{1}{\alpha'\left(\alpha^{-1}\left(\frac{b_i}{y}\right)\right)} \left(1 - K\left(y \cdot \alpha\left(\frac{z_i}{y}\right)\right)\right) + \left(1 - F\left(\alpha^{-1}\left(\frac{b_i}{y}\right)\right)\right)^{n-1} k\left(y \cdot \alpha\left(\frac{z_i}{y}\right)\right)\right]}$$

The final requirement for equality is on the secret reserve price, specifically that $K(y \cdot a_i) = H(a_i)$ (and consequently, $k(b_i) = \frac{1}{y}h(a_i)$). In words, when the cost and bids are shifted by y , the reserve must also shift by y , preserving scale invariance. One solution is to simply assume $r = y \cdot \rho$ where ρ corresponds to the $y = 1$ auction. In practice, I make the stronger assumption that the reserve is drawn from a distribution that is a fixed multiple of the bid distribution, so as bids scale with y the reserve distribution scales proportionally by the same constant factor.

D.3 Counterfactual estimation

Step 1: Rearrange first order equation Let $A(c) = \sum_n P(n) (1 - F(c))^{n-1}$ and $G(c) = \sum_n P(n - 1) f(c) (1 - F(c))^{n-2}$. Then equation 5 can be written

$$b - c = \frac{A(c) (1 - H(b)) \beta'}{G(c) (1 - H(b)) + A(c) h(b) \beta'},$$

$$\implies A(c) h(b) (b - c) \beta' + (1 - H(b)) (b - c) G(c) = A(c) (1 - H(b)) \beta',$$

$$\implies \beta' [A(c) h(b) (b - c) - A(c) (1 + H(b))] = -(1 - H(b)) (b - c) G(c),$$

$$\implies \beta' = \frac{(1 - H(b)) (b - c) G(c)}{A(c) [h(b) (b - c) + (1 - H(b))]},$$

$$= (b - c) \frac{G(c)}{A(c)} \frac{1 - H(b)}{1 + H(b) - h(b) (b - c)},$$

where $b = \beta(c)$. This is an explicit first order ordinary differential equation.

Step 2: Write the initial condition I assume the secret reserve price has a mass point at the right upper bound of the support $\bar{c}$, rather than a right tail past this point. Empirically, this makes no difference beyond the starting point of the ODE. At $\bar{c}$, the bid is $\bar{b} = \beta(\bar{c}) = \bar{c}$. I use this point as the initial value and iterate left along the empirical support of c to trace out the full bid function.

In a handful of (state, year, cluster) samples, the solver fails because the denominator in the ODE approaches zero at the boundary. To address this, I shift the starting point to $(\bar{c} - \epsilon, \bar{b})$, where ϵ is less than .5% of costs and $\bar{b}$ is a first order Taylor approximation of $\beta(\bar{c} - \epsilon)$, as the derivative at the boundary is zero.

Step 3: Numerically solve I estimate a separate solution curve for each (state, year, cluster) using $\beta' = (b - c) \frac{G(c)}{A(c)} \frac{1-H(b)}{1+H(b)-h(b)(b-c)}$ and a Rosenbrock solution method for stiff differential equations, provided by the Matlab function ‘ode23s’.