

Pricing Labor Escalation in Firm-Fixed-Price Federal Services Contracts: A Seventeen-Year Backtest

Jose Aleman
University of Maryland, College Park
This version: August 29, 2026

Abstract

A contractor bidding a five-year firm-fixed-price services contract must commit, on the day of submission, to a single assumed rate of annual labor-cost escalation. A flat three percent is a common convention. I build the cost model that a price proposal requires — the FAR 15-2 rate build-up implemented as a live-formula workbook — price the same notional \$18.9 million contract in each of seventeen bid years from 2005 to 2021, and settle each bid against the wage growth that actually followed, measured by the Bureau of Labor Statistics Employment Cost Index. The three percent convention was wrong in both directions: too high in twelve of seventeen years, by up to 2.70 percent of price, and too low in five. The 2021 bid realized a 3.46 percent margin against the 6.54 percent it was priced for, keeping 53 percent of its intended profit. Judged on its average outcome the convention is close to optimal — the best fixed rate over the full history is 2.75 percent — but the risk it carries is one-sided, because price does not respond to what escalation later does. In the hottest third of simulated outcomes the fifth-percentile result is a \$174,874 loss and 8.4 percent of contracts won fail to cover cost. Escalation explains 11.5 percent of the variance in realized margin, less than direct labor rates or overhead, but it is the only input in the stack that is a forecast of an external index with no subsequent correction mechanism.

JEL classification: H57, J31, D81, E31, L33

Keywords: government procurement; firm-fixed-price contracts; labor escalation; Employment Cost Index; block bootstrap; variance decomposition; price to win

1 Introduction

On a firm-fixed-price contract the government pays the price that was bid, whatever the work turns out to cost. For a services contract, the largest cost element is labor, and labor gets more expensive over the period of performance. The bidder must therefore assume a rate at which it will escalate, apply that rate to every out-year, and live with the consequence for five years. If staff become more expensive faster than assumed, nothing recovers the difference; it comes out of fee.

A flat three percent per year is a common convention for this assumption. This paper asks what that convention actually cost, by pricing one contract repeatedly and settling each bid against realized wage growth. The exercise is a backtest in the ordinary sense: seventeen bid years, each priced with only the information available at submission, each scored against the five years of Employment Cost Index growth that followed.

Three findings follow. First, the convention erred in both directions, and the direction flipped in 2017. From 2005 through 2016 realized escalation came in below three percent in every year, so bids carried a price premium of up to 2.70 percent — not free, since a premium is what loses awards. From 2017 onward

realized escalation exceeded three percent, and margin eroded, worst in 2021. Second, on an average-outcome criterion the convention is close to right: the best fixed rate over the full history is 2.75 percent, and expected profit weighted by probability of winning moves less than one percent across every plausible choice. Third, and the reason the first two findings do not settle the question, the risk is asymmetric. Price does not depend on what escalation later does, so the same bid wins at the same rate in every state of the world; only the downside moves.

The paper also reports where escalation sits relative to other sources of margin risk, and finds that the ranking is largely a statement about assumed input uncertainty rather than a fact about the contract. That is reported as a range rather than a rank.

2 The pricing decision

A price proposal is a build-up. Direct labor dollars carry fringe, the two together carry overhead, general and administrative expense is applied to a base that includes other direct costs, and fee is applied to total cost. Each of these is a rate that the contractor negotiates and later reconciles against actual incurred cost. The escalation assumption is different in kind: it is a forecast of an index outside the firm, applied to the out-year labor rates, and there is no subsequent true-up.

The asymmetry that makes this interesting is that the bid price is fixed at submission, so the probability of winning is determined before any escalation is realized. A bidder who assumes four percent prices roughly 2.53 percent above a bidder who assumes two percent, and wins correspondingly less often, in every state of the world. Escalation realizations then move only the profit conditional on winning. The decision is therefore not a forecasting problem with a symmetric loss function; it is a trade of win probability against left-tail exposure.

I use three percent as the baseline because it is a common round number, not because it is anyone's official standard. To keep the results from resting on that choice, I also score two percent, four percent, a trailing three-year average of the index, and perfect hindsight.

3 The cost model

Cost elements are presented in the order FAR 15.408 Table 15-2 requires of a contractor submitting certified cost or pricing data:

direct labor + fringe + overhead + other direct costs = G&A base;
G&A base + G&A + subcontractor cost + subcontractor handling = total cost;
total cost + fee = price.

One structural choice in the build-up is worth naming. Subcontractor cost sits outside the G&A base and attracts a smaller handling rate instead of full G&A. This is the value-added base contemplated by FAR 31.203(d); the alternative charges full G&A on dollars passing straight through to a third party. Because a contractor discloses which method it uses, the choice is fixed in the model rather than exposed as an adjustable input. A unit test asserts that moving a dollar from other direct costs to subcontractor cost changes the G&A base.

The model is implemented twice. The Python pipeline computes the build-up in code; the delivered artifact is an Excel workbook with live formulas against named inputs, so that changing the named cell

Escalation_Assumed recalculates the out-year labor rates, the build-up and the price. The workbook ends with a reconciliation block carrying the same totals from the pipeline, and a test loads the file into a spreadsheet formula engine, recalculates it, and requires the two to agree to the cent. That test found a \$1.42 discrepancy the first time it ran, caused by the pipeline carrying a blended per diem rate at full floating-point precision while the workbook displayed and used two decimals.

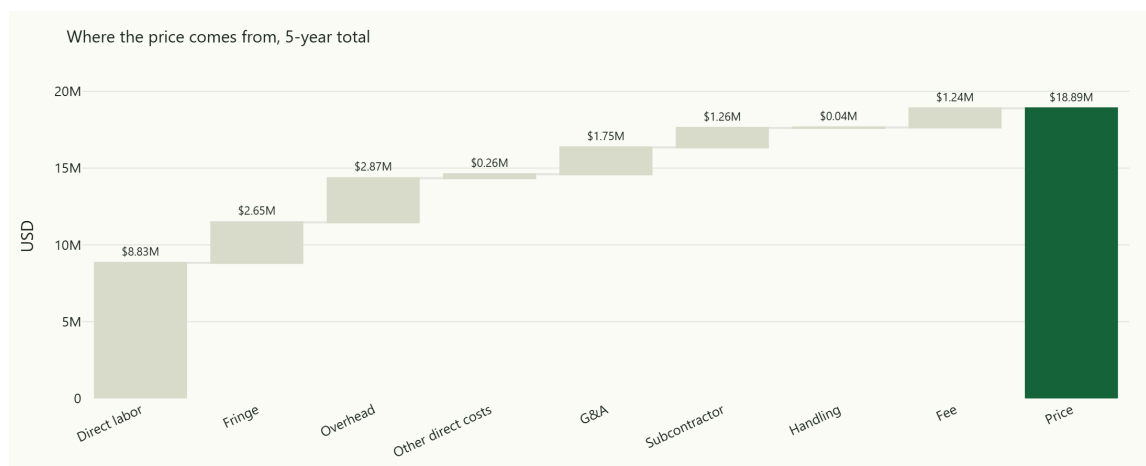


Figure 2.1. The rate build-up from direct labor to price

Notes: Cost elements in FAR 15.408 Table 15-2 order. Subcontractor dollars bypass the G&A base and take a handling rate instead, per FAR 31.203(d).

4 Data

Four public sources are pulled live, and the realized-escalation series against which every bid is settled is the Bureau of Labor Statistics Employment Cost Index.

The Employment Cost Index provides five series from 2001 to the present. Each labor category in the model is matched to a fitting index: the professional and business services wage series for senior roles, the all-industry wage series for the rest. I measure realized escalation from annual average index levels rather than point-in-time readings, which avoids making the result depend on the month in which a fiscal year is anchored.

The CPI-U is used for the non-labor portion of the contract. Travel and materials are goods, not labor, so they escalate on prices. GSA per diem rates for five cities price travel at real lodging and meal allowances, including the rule paying seventy-five percent of the meal allowance on the first and last day of a trip. USAspending supplies 2,449 professional services awards, used only to size the notional contract: the median period of performance is 5.4 years and 89 percent run three years or longer, which is why the contract is priced over five. That query sorts by award size, so it samples the large end of the market rather than the market as a whole; no finding depends on it.

Direct labor rates, indirect rates and fee are assumed, because real ones are proprietary. Each is named on an assumptions sheet with its basis recorded on a register. The obvious objection is that invented labor rates make the exercise vacuous. Scaling every rate by plus or minus twenty percent moves the headline result by 0.08 percentage points, because the finding concerns how rates escalate, not what they are.

One gap in the source data is recorded rather than patched. The Bureau of Labor Statistics never published an October 2025 CPI, because the lapse in appropriations that month halted price collection and the agency did not estimate the missing month afterward. The 2025 annual average therefore rests on eleven months. CPI drives only non-labor escalation, about eleven percent of the total miss, and dropping 2025 would have cost the backtest its most informative bid year, so the year is retained and the output records the month count.

5 Backtest design

For each bid year from 2005 to 2021, the same contract is priced under the strategy's escalation assumption using only information available at submission, and then re-costed using the Employment Cost Index growth that actually occurred over the following five years. Two outcomes are recorded. Margin erosion is margin as bid minus margin as realized, in percentage points; since price is fixed and cost is not, this is bid price less realized cost, over bid price. Price premium is how far the bid sits above the cheapest bid that would still have covered its own cost, and is the competitive cost of caution.

6 Results

6.1 Year by year

Table 2.1 reports every bid year. The convention was too high in twelve of seventeen years and too low in five, and the split is not random across time: the twelve run consecutively from 2005 to 2016 and the five from 2017 onward. Bidding three percent into a period of sub-three-percent wage growth carried a premium reaching 2.70 percent of price in 2008; bidding it into the 2021 wage acceleration cost 3.08 points of margin.

Table 2.1. The three percent convention settled against realized escalation, by bid year

Bid year	Assumed	Realized	Margin as bid	Margin realized	Erosion (pp)	Premium
2005	3.00%	2.70%	6.54%	6.46%	+0.08	−0.08%
2006	3.00%	2.39%	6.54%	6.99%	−0.44	+0.48%
2007	3.00%	1.95%	6.54%	8.03%	−1.49	+1.62%
2008	3.00%	1.67%	6.54%	9.00%	−2.46	+2.70%
2009	3.00%	1.74%	6.54%	8.71%	−2.17	+2.38%
2010	3.00%	1.84%	6.54%	8.49%	−1.95	+2.13%
2011	3.00%	2.00%	6.54%	8.49%	−1.95	+2.13%
2012	3.00%	2.13%	6.54%	8.45%	−1.91	+2.08%
2013	3.00%	2.30%	6.54%	8.29%	−1.74	+1.90%
2014	3.00%	2.55%	6.54%	8.04%	−1.49	+1.62%
2015	3.00%	2.73%	6.54%	7.66%	−1.12	+1.21%
2016	3.00%	2.87%	6.54%	7.27%	−0.73	+0.79%
2017	3.00%	3.24%	6.54%	6.85%	−0.31	+0.33%
2018	3.00%	3.79%	6.54%	6.26%	+0.29	−0.30%
2019	3.00%	4.20%	6.54%	5.19%	+1.35	−1.42%

Bid year	Assumed	Realized	Margin as bid	Margin realized	Erosion (pp)	Premium
2020	3.00%	4.46%	6.54%	3.82%	+2.72	-2.83%
2021	3.00%	4.32%	6.54%	3.46%	+3.08	-3.19%

Notes: Each row prices the same notional \$18.9 million, five-year, firm-fixed-price services contract in the stated bid year assuming three percent annual labor escalation, then re-costs it on realized Employment Cost Index wage growth over the following five years. Erosion is margin as bid minus margin realized, in percentage points; a positive value is lost profit. Premium is the bid price relative to the cheapest bid that would still have covered realized cost; a positive value is money left on the table. Realized escalation is the compound annual rate implied by annual average ECI levels over the period of performance.

The worst year is 2021, which realized a 3.46 percent margin against the 6.54 percent it was priced for. In dollars that is \$653,398 of an intended \$1,235,992 on an \$18.9 million contract: the contract kept 53 percent of its profit and remained profitable, which is the relevant point about how a bad escalation year actually presents.

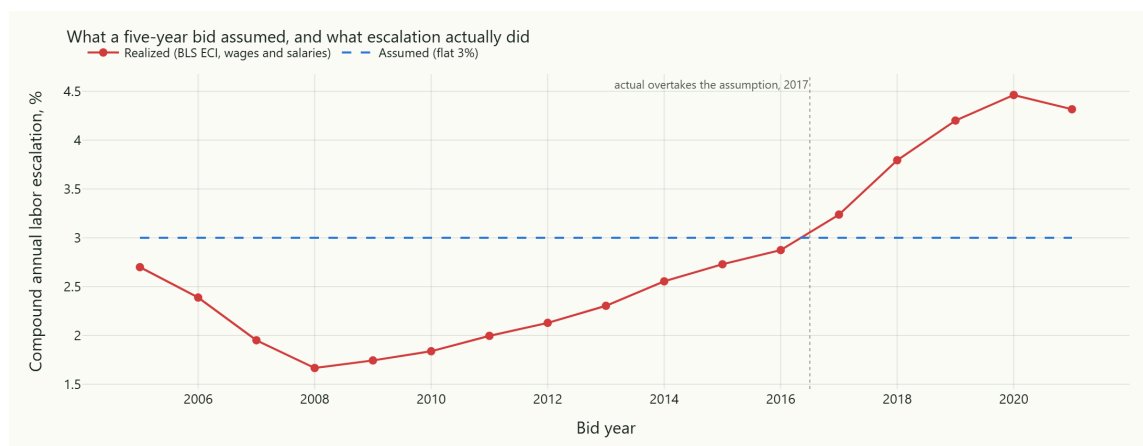


Figure 2.2. Assumed versus realized escalation, bid years 2005–2021

Notes: The three percent line against realized Employment Cost Index wage growth over each bid’s five-year period of performance. The convention sat above realized growth for twelve consecutive years and below it for the last five.

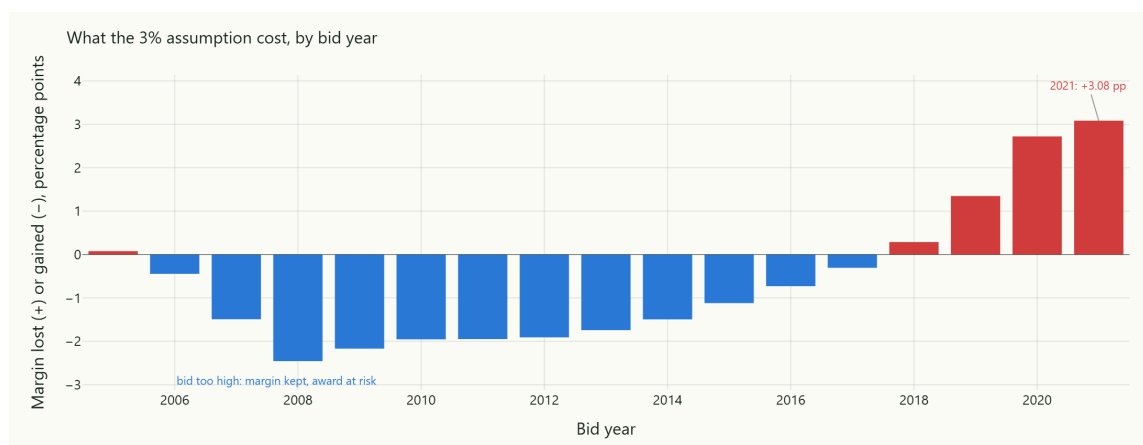


Figure 2.3. Margin as bid against margin realized, by bid year

Notes: The vertical distance between the two series is margin erosion. Negative erosion, in the years to 2016, means the contract earned more than it was priced for — which is to say it was priced too high to be competitive.

6.2 Average outcome against tail outcome

Scored on its average, the convention is close to optimal. Across every escalation rate a pricer would plausibly choose, probability of winning multiplied by expected profit varies by less than one percent, and the best fixed rate over the full seventeen-year history is 2.75 percent. A decision rule that maximizes expected value would leave the assumption alone.

That conclusion does not survive attention to the tail, for a structural reason. Because price is set at submission and escalation is realized afterward, the win rate of a three percent bid is identical in calm and turbulent escalation regimes; only the profit conditional on winning changes. Conditioning simulated outcomes on the realized escalation environment makes the point: in the calmest third of outcomes the fifth-percentile result is a profit of \$372,530, while in the hottest third it is a loss of \$174,874, and 8.4 percent of contracts won fail to cover their cost. The distribution does not widen symmetrically; its left tail extends while the rest of it stays put.

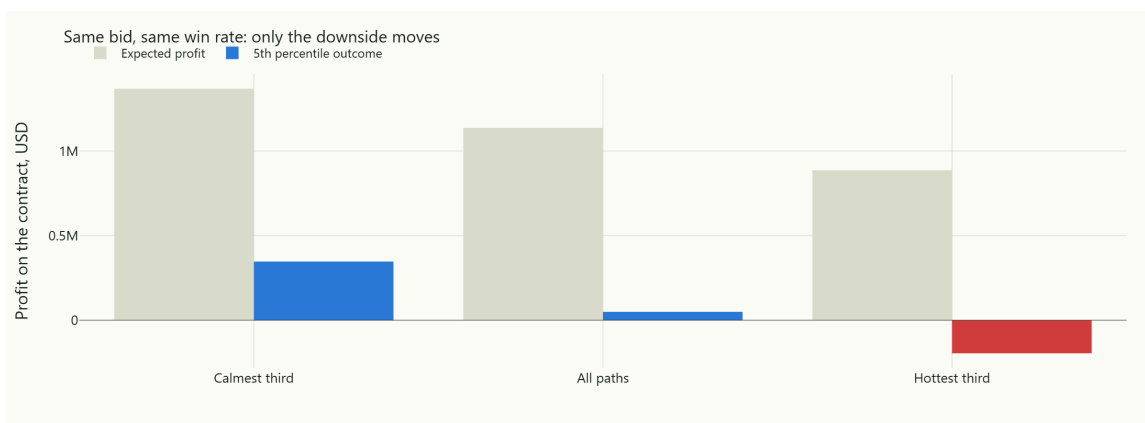


Figure 2.4. Same bid, same win rate, only the downside moves

Notes: Simulated outcomes for the three percent bid, partitioned by realized escalation environment. Win probability is invariant across the panels because price is fixed at submission; the fifth-percentile outcome is not.

6.3 Alternative strategies

No alternative dominates. Table 2.2 scores five strategies on mean erosion, worst single-year erosion, and mean price premium.

Table 2.2. Escalation strategies compared over seventeen bid years

Strategy	Mean erosion (pp)	Worst erosion (pp)	Mean price premium
Flat 4 percent	−2.28	+1.34	+2.53%
Flat 3 percent (the convention)	−0.60	+3.08	+0.68%
Perfect hindsight	−0.21	+0.80	+0.22%
Trailing three-year index	+0.38	+3.13	−0.38%
Flat 2 percent	+1.10	+4.85	−1.13%

Notes: Mean erosion averages margin erosion across the seventeen bid years; a negative value means the strategy on average over-recovered, which is the same thing as having priced high. Worst erosion is the single worst bid year. Mean price premium is the average distance above the cheapest bid that would have covered realized cost. Perfect hindsight uses the realized index and is

included as an upper bound on what any escalation rule could achieve; it still leaves 0.80 points of worst-case erosion, because the contract's staffing mix does not track any single index.

The trailing three-year index prices about a point lower than the convention, which is worth real win probability, but gives up about a point of margin and records the worst single year of any strategy tested. That is the characteristic failure of a backward-looking rule: it reacts slowest exactly at a turn, which is when the assumption matters. Flat two percent is worse on both measures. Flat four percent protects margin at the cost of pricing 2.53 percent high, which on a competitive procurement is a decision not to win.

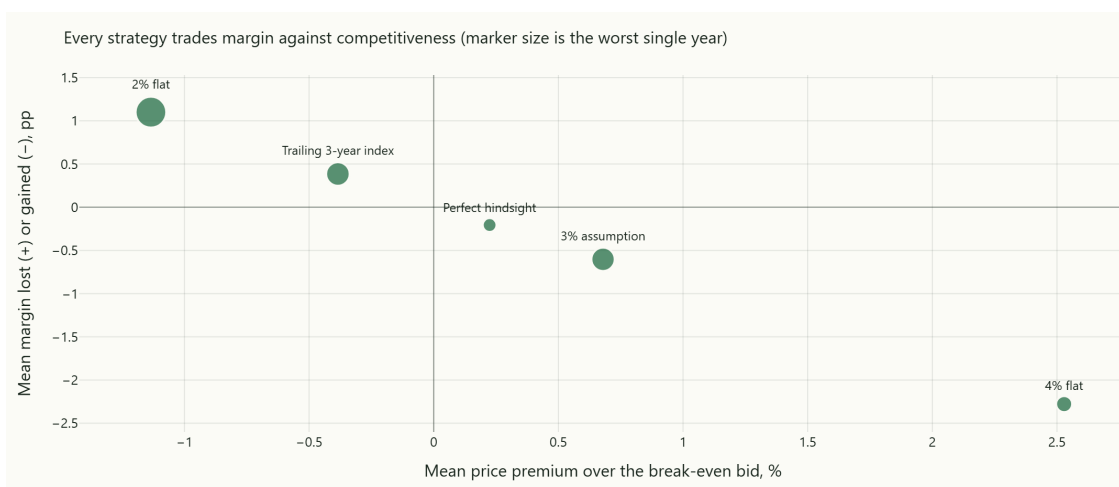


Figure 2.5. Every strategy trades margin protection against competitiveness

Notes: Mean margin erosion against mean price premium. Perfect hindsight sits nearest the origin and is not available to a bidder; the remaining strategies trade along a line.

6.4 Where margin risk actually comes from

Escalation is not the largest source of variance in realized margin. Table 2.3 reports first-order variance indices from a Monte Carlo over the model's inputs.

Table 2.3. First-order variance decomposition of realized margin

Input	First-order index	Share of explained variance
Direct labor rate realization	0.583	60.8%
Overhead rate	0.139	14.5%
Escalation (labor and ODC)	0.115	12.0%
Fringe rate	0.080	8.3%
G&A rate	0.043	4.4%

Notes: First-order indices are estimated by binning Monte Carlo draws on each input and measuring how far the conditional mean outcome moves across bins, with a correction for the scatter that bin means exhibit under sampling alone. Without that correction the shares summed above one and the residual went negative, which is how the bias was detected. Indices do not sum to one because interaction terms are not attributed.

The ranking is not robust, and reporting it as though it were would overstate what the model knows. Tightening the assumed uncertainty on the internal rates raises escalation's share to 32.2 percent; widening it drops the share to 5.7 percent. Since those uncertainty ranges are my assumptions rather than measurements,

the honest report is the range and not the rank.

What does not move under that sensitivity is the qualitative difference between the inputs. Fringe, overhead, G&A and direct labor rates come out of a contractor's own accounting system and are reconciled against actual incurred cost after the fact. Escalation is a forecast of an external index applied to a fixed price, and nothing reconciles it. Two inputs with similar variance shares are not equally worth managing when one of them has a correction mechanism and the other does not.

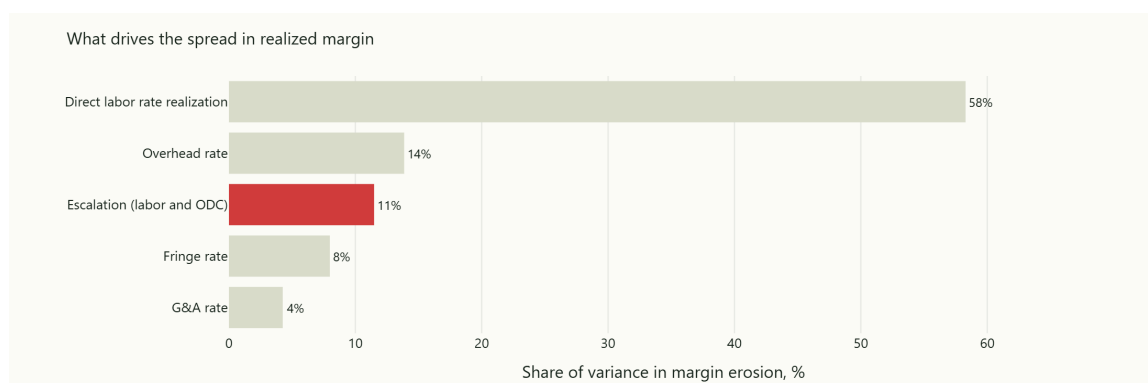


Figure 2.6. First-order variance shares and their sensitivity to assumed input uncertainty

Notes: Bars are the shares in Table 2.3; the ranges show how each share moves when the assumed uncertainty on the internal rates is tightened or widened.

7 Inference

Confidence intervals come from a block bootstrap over the escalation history. Consecutive runs of years are resampled and reassembled into five-year paths, and all four indices are drawn from the same years, because a high-escalation year is usually followed by another one and because wages and prices rose together in 2022. Treating years as independent is not innocuous.

Table 2.4. Bootstrap block length and the width of the 95 percent interval on margin erosion

Resampling method	95% interval	Width (pp)
Single years, independent	[−1.93, +1.81]	3.73
Two-year blocks	[−2.18, +2.63]	4.81
Three-year blocks	[−2.29, +2.92]	5.21
Five-year blocks	[−2.46, +2.72]	5.18

Notes: Intervals are on margin erosion in percentage points for the three percent strategy. Escalation is serially correlated, so an independent-year resample destroys the persistence that makes a bad five-year path possible and narrows the interval by roughly a quarter. The reported results use the block resample, which makes the bid look riskier and is the correct characterization.

8 Limitations

The price-to-win curve rests on an assumed win function. No public dataset contains losing bids, so the mapping from price to probability of award is judgment. Every result that uses win probability inherits that assumption; the margin results do not, and should be read as the firmer half of the paper.

It is one contract with one labor mix. A different mix escalates differently, and even perfect hindsight on the headline index leaves up to 0.80 points of error here, because the staffing mix does not track any single series. Firm-fixed-price is also the hard case by construction: on a cost-reimbursement vehicle the government pays actual cost and fee is fixed in dollars, so an escalation miss enlarges the contract rather than compressing the margin.

Finally, the bootstrap assumes the past is representative. It resamples 2003 through 2025, a window containing exactly one inflation episode. If the escalation process is not stable over time, the intervals describe a period that may not recur.

9 Conclusion

Settled against twenty years of Employment Cost Index data, a flat three percent labor escalation assumption is a defensible average and a poor description of the risk. It was too high for twelve consecutive years, at a competitive cost of up to 2.70 percent of price, and too low for the five that followed, costing a 2021 bid 47 percent of its intended profit. Because price is committed before escalation is realized, the choice does not trade expected value against expected value; it trades win probability, which is paid with certainty, against left-tail exposure, which is not.

The practical implication is narrow. Nothing here supports replacing the convention: the best fixed rate over the full history is 2.75 percent and no alternative rule dominates on both margin and competitiveness. What the results support is treating the assumption as a risk position rather than a forecast — pricing it with an explicit view of the fifth percentile, and recognizing that escalation is the one element of the build-up that no later reconciliation corrects.

References

Federal Acquisition Regulation. 2026. 48 CFR § 15.408, Table 15-2, “Instructions for Submitting Cost/Price Proposals When Certified Cost or Pricing Data Are Required”; and 48 CFR § 31.203(d), “Indirect Costs.” Washington, DC.

Künsch, Hans R. 1989. “The Jackknife and the Bootstrap for General Stationary Observations.” *Annals of Statistics* 17 (3): 1217–1241.

Saltelli, Andrea, Paola Annoni, Ivano Azzini, Francesca Campolongo, Marco Ratto, and Stefano Tarantola. 2010. “Variance Based Sensitivity Analysis of Model Output.” *Computer Physics Communications* 181 (2): 259–270.

Sobol’, Ilya M. 2001. “Global Sensitivity Indices for Nonlinear Mathematical Models and Their Monte Carlo Estimates.” *Mathematics and Computers in Simulation* 55 (1–3): 271–280.

U.S. Bureau of Labor Statistics. 2026. Employment Cost Index and Consumer Price Index for All Urban Consumers. Washington, DC.

U.S. General Services Administration. 2026. Per Diem Rates API. Washington, DC.

Data and code availability

All sources are public and no API key is required. Realized escalation is the BLS Employment Cost Index; non-labor escalation uses the CPI-U; travel is priced on GSA per diem rates; the contract is sized from USAspending. The pipeline, the seventeen-year backtest, the block bootstrap, the Monte Carlo, the live-formula Excel pricing model, the compliance matrix and the two-page memo are available at github.com/josealemanm/federal-services-escalation-risk. An interactive dashboard runs in the browser at josealemanm.github.io/federal-services-escalation-risk.