

Dating Government Demand: A Point-in-Time, Entity-Resolved Dataset of U.S. Federal Procurement and Disclosure Records

Removing look-ahead and survivorship errors from firm-level measures of fiscal shocks

Jennie Chu

Gemral Edge · support@gemral.com

ABSTRACT

Firm-level research on government demand needs two things that raw public procurement records do not supply: the date on which a contract became public, and the listed company that held the contracting entity on that date. Hall's (1978) insight, central to the 2026 Prize in Economic Sciences, is that forward-looking agents respond only to new information, so the date information arrives defines the shock. We describe a point-in-time, entity-resolved dataset of U.S. federal contract actions and related disclosure records (850,000+ resolved events as of October 2026) built for this purpose. Each record carries its activity date, its publication date with statutory lags applied, and a ticker that traded on the publication date. We document three construction errors that affect naive joins and measure what correcting them changes: activity-date dating places defense contract information 90 days early, a current-parent join misattributes 3.3% of resolved contract dollars, and daily announcements reveal 93.5% of matched defense awards a median 90 days before the award record. Ten record types pass accuracy checks against original public documents at 95% or better. We report no return or performance results.

Keywords: fiscal shocks; government procurement; point-in-time data; look-ahead bias; survivorship bias; entity resolution; microfoundations

JEL codes: E62, H57, G14, C81

1. Introduction

On 12 October 2026 the Royal Swedish Academy of Sciences awarded the Prize in Economic Sciences to Robert E. Hall and Michael Woodford "for microeconomic foundations of macroeconomic stabilisation" (Royal Swedish Academy of Sciences, 2026). The Committee's account of the prize gives a central place to a simple idea from Hall (1978): if households base consumption on their expectations of future income, then consumption changes only when new information arrives, and because new information cannot be predicted, neither can those changes. Woodford's (2003) framework carries the same logic into price setting. Firms that cannot reprice continuously set prices looking forward, so what matters for inflation and output is what firms expect and when their expectations change.

This logic has a direct consequence for empirical work on government spending. A fiscal shock is not the moment money is obligated. It is the moment the private sector learns about it. Ramey (2011) made this point forcefully for aggregate defense spending: shocks identified from the timing of spending are largely anticipated, and once news is dated correctly the estimated effects change. A growing firm-level literature now asks the same questions of individual contractors, using contract-level records to study investment (Hebous and Zimmermann, 2021), regional multipliers (Nakamura and Steinsson, 2014), political connections (Goldman, Rocholl and So, 2013; Cohen and Malloy, 2014) and contract renegotiation (Brogaard, Denes and Duchin, 2021).

Firm-level work inherits a measurement problem the aggregate literature could partly sidestep. A single contract record must be placed in time and attached to a security. Public procurement records make both steps easy to get wrong. The date that sits on every row is the date the obligation was signed, not the date anyone outside the contracting office could see it. The name on the row is the legal entity that signed, which is often a subsidiary, and its listed parent changes over time through acquisitions, spin-offs and ticker changes. Errors in either step do not raise warnings. Every row still joins, and the result is a dataset in which information arrives early and only surviving companies receive it.

This paper describes a dataset built to remove those errors and documents how it was constructed and checked. It is a data paper, not an empirical study. We describe what the data contains, how its timing and identity fields are built, how often the corrections matter, how accurate the linked records are against original documents, and where its coverage ends. Section 2 sets out what the theory requires of a measurement. Section 3 describes three errors in naive construction and measures them. Section 4 describes the point-in-time architecture. Section 5 measures how much the corrections change, Section 6 reports accuracy checks, Section 7 outlines research designs the data supports, Section 8 gives scope and coverage notes, and Section 9 covers access.

2. What the theory asks of the data

Translating the stabilisation literature into measurement requirements yields four conditions that any firm-level series of government demand should meet.

R1. Information dating. The event date must be the date the information became public. In Hall's framework, a response can only follow the arrival of news, so a series dated by signing rather than publication mixes anticipated and unanticipated components and places some information before it existed. For defense contract actions the gap is statutory: these actions enter the public award records with a 90-day delay.

R2. Recipient identity at the time. The firm credited with an award must be the firm that controlled the contracting entity when the award was made. Crediting a later acquirer with a pre-acquisition award attributes demand to a firm that did not receive it.

R3. A security that traded on the event date. For any study that links government demand to market prices, financing conditions or investor attention, the identifier attached to the event must be one that traded on the date the event became public. A ticker that had already stopped trading cannot carry information to anyone.

R4. Complete history. Firms that were later acquired, merged or delisted must remain in the history under the identifiers they had at the time. Dropping them produces a sample of survivors, a bias that Shumway (1997) documented for delisting returns and that applies equally to any panel built from today's reference tables.

R2 and R3 can pull in different directions, and the case where they do is informative. When a company signs a contract and is then acquired before the contract is published, the firm that received the demand (R2) and the security on which the news can be observed (R3) are different. A dataset that keeps only one identifier must give up one of the two requirements. Section 4.3 describes how the dataset keeps both.

3. Three errors in naive construction

We call a naive join one that takes procurement records as published, uses the date on the row, and maps recipients to listed parents through a current reference table. This is the default for most researchers and most commercial tools. It fails in three ways.

3.1 The date on the record is not the date of the news

Every contract action carries an action date. For defense actions, the public record lags that date by 90 days. A study that dates a defense award by its action date places the information roughly a quarter of a year before it was available.

How much of the data does this affect? In a 3% block sample of federal contract transactions of \$1 million or more from fiscal year 2018 onward, defense actions account for about 59% of rows and 65% of dollars (measured 24 October 2026). The lag is therefore not an edge case. It applies to the majority of contract dollars in the sample.

Other disclosure types carry their own lags. Lobbying reports are filed after the quarter they describe, congressional trade disclosures arrive weeks after the trade, and insider-ownership reports have their own filing windows. In each case the activity date says what happened and only the publication date says when the market could have known.

3.2 The recipient on the record is not the firm that trades

Awards name the legal entity that signed. That entity is frequently a subsidiary, a joint venture or a holding company whose name gives no clue to its listed parent. Names also drift across records, with and without corporate suffixes, under retired brand names or under division names. Mapping such records by name to a current list of public companies either drops them or attaches them to the wrong firm.

Time compounds the problem. If a contractor wins an award and is acquired two years later, a current reference table maps the award to the acquirer, and the history now credits a firm that had no connection to the award when it was made.

3.3 Firms that disappeared are missing from the reference table

Current reference tables describe the market as it is today. Companies that were acquired or delisted are absent. Joining historical contracts to such a table silently removes them, and what remains is a history of survivors.

The defense sector illustrates the scale. Between 2018 and 2020 several large contractors left public markets through acquisition, including Rockwell Collins (acquired by United Technologies, November 2018), Orbital ATK (Northrop Grumman, June 2018), CSRA (General Dynamics, April 2018), Engility (Science Applications International, January 2019), Harris and L3 Technologies (merged as L3Harris, June 2019), and Raytheon and United Technologies (merged as RTX, April 2020). In the resolved event history these firms remain present under the tickers they carried at the time; for example, 253 events are attributed to Rockwell Collins under its own ticker, 1,115 to L3 Technologies and 4,267 to Raytheon (measured 24 October 2026). Across all record types, 4,745 distinct tickers appear in the resolved event history, including delisted, merged and renamed ones (measured 23 October 2026).

3.4 Where the errors meet: signed before a deal, published after it

The three errors interact in a specific and easily missed case. Suppose a contractor signs a defense contract action in late September and its acquisition by another listed company closes in November. The action becomes public in late December, 90 days after signing. On the date the news appears, the contractor's own ticker no longer trades.

A naive join handles this case in one of two wrong ways. Dated by signing and mapped through a current table, it credits the acquirer with an award made before the acquisition, three months before anyone could see it. Dated correctly but mapped to the signer's own historical ticker, it attaches the news to a security that no longer exists, so the event drops out of any price-linked study.

We measured how often this case occurs. In the resolved contract history, 659 contract actions were signed before an acquisition of the signing company closed and published after it (measured 24 October 2026). Of these, 651 roll to a listed acquirer: 361 to RTX, 124 to L3Harris, 49 to United Technologies, 49 to General Dynamics, 43 to Northrop Grumman and 25 to Science Applications International. The remaining 8 belong to a company whose acquirer was privately held and therefore carry no ticker. The dataset handles each of these as described in Section 4.3: the signing company is retained as the recipient, and the ticker is the one that traded on the publication date.

A related case arises when the signing company stays independent but changes its ticker between signing and publication, as when United Technologies became RTX or Harris became L3Harris.

Another 254 contract actions fall in this case; each carries the ticker that traded on its publication date.

These numbers are small relative to the full history, but these are the events a naive construction misdates or loses, and they are concentrated among the largest defense contractors in exactly the years when consolidation was most active.

4. Point-in-time architecture

4.1 Coverage

The dataset links the following public record types to companies and tickers: federal contract transactions of \$1 million or more (fiscal year 2018 onward), daily defense contract announcements (2015 onward), pre-award notices naming the supplier, lobbying filings with lobbyists and prior government positions, congressional trade disclosures with committee membership, insider trades together with officers, directors and major holders, subcontractor links, corporate political action committee contributions (2016 onward), bid protests (filed September 2018 onward), and weapon programs with manufacturers. Sources are described by category: U.S. federal procurement and contract-award public records; U.S. congressional financial disclosure filings; U.S. federal lobbying disclosure filings; U.S. federal campaign-finance disclosure filings; U.S. securities-regulator public filings (company reports and insider-ownership reports); and U.S. federal arms-transfer public notices.

As of October 2026 the resolved event history holds 850,000+ events, built from 650,000+ federal contract transactions, 50,000+ congressional trades and the other record types listed above. All counts are measured estimates dated October 2026; the live figures are published with each delivery.

4.2 Three timestamps per record

Every record carries three times:

- `source_published_at`: when the record became publicly available. Point-in-time views never show a record before this time.
- `first_seen_at`: when the dataset first stored the record.
- `is_backfill`: true when the record was loaded historically rather than captured live, so users can test historical and live-captured data separately.

For historical federal contract actions, `source_published_at` is assigned by rule: the action date plus 90 days for defense actions, reflecting the statutory delay, and the action date plus 3 days for other agencies. Actions reported late may have become public after that date, so the rule gives the earliest plausible publication date; researchers who want a stricter bound can add a buffer. For other record types the publication time is taken from the source document. Where a source gives no publication time at all, the dataset assigns an estimate that can only be late, never early.

4.3 Two identities per record

Each event carries two identities.

- **Recipient** (`direct_entity_id`): the legal entity named in the record, linked to the parent that controlled it on the activity date. For contract actions the activity date is the action date, so an award signed before an acquisition stays with the original company.
- **Security** (`ticker_asof`, `listed_entity_id`): the listed company and ticker that traded on the publication date. Parent links and ticker histories carry validity dates, so the security is resolved as of that date.

In most records the two coincide. They differ in the cases described in Section 3.4, where the signing company is retained as the recipient and the security is the listed acquirer whose shares traded when the news appeared. Where no listed security traded on the publication date, `ticker_asof` is empty. Every step from recipient to security is written to `relationship_path` with the date used for each step, so the resolution of any row can be audited.

An automated check runs over the full event table and fails if any event carries a ticker that did not trade on its publication date, or a ticker symbol other than the one in use on that date. On 24 October 2026 it found none across 903,168 rows.

4.4 Entity resolution

Recipients are resolved through registry identifiers first and through dated parent and subsidiary links second. Name matching alone never attaches a ticker. Tickers carry validity windows, so renamed, merged, spun-off and reused symbols are handled by date; a symbol reused by an unrelated company after a delisting is not joined to the earlier company's history. About 49.3% of fiscal year 2018+ contract transaction dollars are linked to a listed ticker (measured 23 October 2026). Most of the remainder is awarded to private or unlisted companies, which no join can map to an equity.

4.5 Revisions

Amendments do not overwrite history silently. When a source record changes, the previous version is written to a revision log with its hash and the time the change was observed. Silent revisions of historical records can change research results, as Ljungqvist, Malloy and Marston (2009) showed for analyst recommendation histories, so the dataset keeps the earlier versions. The log started in October 2026 (see Section 8).

5. Three measurements of what the corrections change

This section reports three measurements taken on the dataset itself. Each compares the point-in-time construction with the naive one on the same records, so that a reader can judge how much the corrections matter before running any model. None of them involves prices or returns.

5.1 Table 1: How early naive dating places the information

Dating a record by its activity date instead of its publication date places the information early. Table 1 measures by how much, record type by record type.

Table 1. Look-ahead introduced by activity-date dating (measured 24 October 2026).

Record type	Records measured	Activity date to publication
Defense contract actions	59% of contract rows, 65% of dollars (3% block sample, FY2018+)	90 days (statutory delay)
Congressional trade disclosures	63,122 trades since 2018	median 27 days; 90th percentile 63 days
Lobbying reports	8,975 quarterly filings since 2018	median 20 days after quarter end; 90th percentile 21 days
Insider trades	35,226 trades since 2018	median 1 day; 90th percentile 4 days

The defense lag dominates in both size and reach. It is three times the median congressional lag and applies to the majority of contract dollars.

5.2 Table 2: How much a current reference table misattributes

Mapping each contract recipient through today's parent table, instead of the parent that held it on the publication date, credits part of the history to the wrong company. Table 2 measures the share among federal contract transactions whose recipient is resolved to an entity.

Table 2. Misattribution from a current-parent join, federal contract transactions (measured 24 October 2026).

Measure	Transactions	Dollars
Resolved contract transactions	197,201	\$2,706 billion
Credited to a different ticker under today's parents	8,568	\$89.6 billion (3.3%)
Reach a listed company only through a dated subsidiary link	13,465	\$191.9 billion (7.1% of ticker-linked dollars)
Signed before an acquisition closed, published after it (Section 3.4)	659	carry the ticker trading on publication day
Signer changed ticker between signing and publication	254	carry the ticker in use on publication day

The third row is a measure of how hard the data is to rebuild. For \$191.9 billion of contract dollars the name on the award is not a listed company, so matching names against a list of public companies cannot find the link. It is recovered only through parent and subsidiary relationships that carry dates.

5.3 Table 3: Information that appears before the award record

Some public records reveal an award before it reaches the award records. Daily defense contract announcements are the clearest case. We matched announcements to the base award record of the same contract by contract identifier.

Table 3. Lead of daily defense contract announcements over award records, FY2018+ (measured 24 October 2026).

Measure	Value
Contracts with both an announcement and an award record	9,158
Dollars announced	\$1.07 trillion
Announced before the award record became public	93.5%
Lead, median	90 days
Lead, 10th / 25th / 75th / 90th percentile	85 / 90 / 90 / 91 days
Announcement date relative to signing date, median	0 days (same day)

A series built from award records alone learns about these contracts about a quarter later than a series that also reads the announcements. Because the dataset links both record types to the same company and ticker on their own publication dates, a researcher can date each contract by the first public record that revealed it.

6. Accuracy checks

A record type is released only after a random sample of up to 200 linked rows is checked against the original public documents for three fields: entity, ticker and date. The threshold is 95%, and samples must be no older than 30 days. As of 23 October 2026, ten of ten record types pass; Table 4 reports the samples.

Table 4. Precision of linked records against original public documents (samples dated 23 October 2026).

Record type	Correct / checked	Precision
Company officers	198 / 200	99.0%
Board directors	195 / 199	98.0%
Major shareholders (insiders)	195 / 197	99.0%
Subcontractor links	192 / 200	96.0%
Congressional trades	168 / 172	97.7%
Insider trades	200 / 200	100%
Lobbying filings	192 / 200	96.0%
Federal contract transactions	193 / 196	98.5%
Pre-award notices	192 / 198	97.0%
Bid protests	194 / 200	97.0%

For congressional trades, 28 of the 200 sampled records could not be checked automatically because the source site blocks automated access; these are excluded from the denominator and reported separately.

Accuracy checks of this kind measure whether the dataset faithfully reflects the public record. They do not measure whether the public record is itself complete.

Expanded accuracy check (24 October 2026). Every record type is checked against the original public documents before it ships. In the latest graded batch, all 13 record types scored between 95.6% and 100% precision, with at least 500 graded rows per type across batches (site permits: the full set), Wilson 95% intervals, and a blind second grader on at least 10% of rows. Cumulative figures are lower for four types (contracts 91.2%, pre-award notices 76.4%, subcontractor links 94.0%, bid protests 92.5%) because they include batches graded before the entity resolver was fixed.

Attribution coverage. Edge attributes 822 power-plant status events and 807 foreign arms-sale notices to the listed parent company valid on the day each record became public. Exact matching on the raw operator or contractor name attributes 0 and 6 of them. This measures coverage, not return predictability.

7. Research designs the data supports

The dataset is built to support designs in which the timing of information and the identity of the recipient both matter. We list several without reporting estimates, since our purpose here is to describe the measurement, not to test hypotheses.

1. **News-dated firm-level demand shocks.** Following Ramey (2011), construct firm-level government demand shocks dated by publication rather than signing, and compare responses

under the two datings. The difference is a direct measure of how much signing-date construction contaminates estimates with anticipated or not-yet-public information.

2. **Financing frictions and investment.** Hebous and Zimmermann (2021) find that a dollar of federal purchases raises capital investment of financially constrained firms by 10 to 13 cents over four quarters, with no effect for unconstrained firms. The dataset allows the same design with publication-dated contracts and with acquired firms retained, so researchers can check how sensitive such estimates are to timing and survivorship.
3. **Regional transmission.** Following Nakamura and Steinsson (2014), aggregate publication-dated awards by place of performance and recipient to study regional responses, keeping recipients that were later absorbed into other firms.
4. **Political economy of procurement.** Link awards, bid protests, lobbying filings and congressional disclosures through dated entity relationships to study how connections relate to the allocation and renegotiation of contracts (Goldman, Rocholl and So, 2013; Cohen and Malloy, 2014; Brogaard, Denes and Duchin, 2021), with every relationship valid only for the period the public record supports.
5. **Information frictions.** Woodford's framework, and the broader literature on limited attention (Sims, 2003), suggests that the gap between when information is public and when it is reflected in decisions is itself an object of study. Publication timestamps, first-capture timestamps and the revision log allow researchers to measure that gap directly.

In each case the dataset supplies the dated, entity-resolved inputs. Outcome variables, such as financial statements, prices or regional statistics, come from the researcher's own sources.

8. Scope and coverage notes

The following notes describe the scope of the data so that researchers can design studies around it.

- **Listed coverage of contract dollars.** About 49.3% of fiscal year 2018+ contract transaction dollars are linked to a listed ticker (measured 23 October 2026). The rest goes to private or unlisted awardees or is not yet attributed. Federal spending does not map neatly onto public equities.
- **Fiscal year 2026 is incomplete.** Defense contract actions from July to September 2026 are not yet public because of the 90-day delay, and the current fiscal year 2026 load will be refreshed when the next archive becomes available.
- **Publication dates for historical contract actions are rule-assigned** (Section 4.2). Late-reported actions may have become public after the assigned date.
- **Delisted prices.** Roughly half of events on delisted tickers have historical prices available to us. Studies that require prices for delisted firms should treat those firms with care, because missing prices on failed firms can bias results upward.

- **Bid protests.** We hold 13,850 protests filed since September 2018, but only 25.3% (3,505) are linked to a specific contract transaction and awardee identifier. The 100-day decision deadline matches the recorded decision date exactly in 61.7% of cases; the remainder differ by weekends, holidays or extensions.
- **Revisions before October 2026.** The revision log started in October 2026. Before that, amended insider-ownership reports and amended House trade reports are kept as separate records at their own filing dates but are not yet linked to the filing they amend; amended Senate reports and corrected contract records show only the latest version.
- **Congressional trade verification.** 28 of 200 sampled congressional records could not be checked automatically (Section 6).

9. Data availability, licensing and conflict of interest

The dataset is available to research teams as Parquet files with a manifest (row counts and checksums per file), a data dictionary, a REST API and a server for AI assistants. Sample packs are available on request at support@gemral.com. Methods and known limits are documented in the public due-diligence questionnaire at gemral.com/edge-ddq.md. All records are drawn from public sources, and the dataset contains no material non-public information. Detailed source documentation is available to a licensee's compliance team under a non-disclosure agreement.

The author is the founder of Gemral Edge, which produces and licenses the dataset described here. This paper reports no return, performance or forecasting results.

References

- Brogaard, J., Denes, M., and Duchin, R. (2021). Political influence and the renegotiation of government contracts. *Review of Financial Studies*, 34(6), 3095–3137.
- Cohen, L., and Malloy, C. J. (2014). Friends in high places. *American Economic Journal: Economic Policy*, 6(3), 63–91.
- Cohen, L., Diether, K., and Malloy, C. (2013). Legislating stock prices. *Journal of Financial Economics*, 110(3), 574–595.
- Goldman, E., Rocholl, J., and So, J. (2013). Politically connected boards of directors and the allocation of procurement contracts. *Review of Finance*, 17(5), 1617–1648.
- Hall, R. E. (1978). Stochastic implications of the life cycle–permanent income hypothesis: Theory and evidence. *Journal of Political Economy*, 86(6), 971–987.
- Hall, R. E. (2009). By how much does GDP rise if the government buys more output? *Brookings Papers on Economic Activity*, Fall 2009, 183–231.
- Hall, R. E., and Mishkin, F. S. (1982). The sensitivity of consumption to transitory income: Estimates from panel data on households. *Econometrica*, 50(2), 461–481.
- Hebous, S., and Zimmermann, T. (2021). Can government demand stimulate private investment? Evidence from U.S. federal procurement. *Journal of Monetary Economics*, 118, 178–195.

- Ljungqvist, A., Malloy, C., and Marston, F. (2009). Rewriting history. *Journal of Finance*, 64(4), 1935–1960.
- Nakamura, E., and Steinsson, J. (2014). Fiscal stimulus in a monetary union: Evidence from US regions. *American Economic Review*, 104(3), 753–792.
- Ramey, V. A. (2011). Identifying government spending shocks: It's all in the timing. *Quarterly Journal of Economics*, 126(1), 1–50.
- Royal Swedish Academy of Sciences (2026). *Microfoundations deliver macro insights*. Popular science background to the Sveriges Riksbank Prize in Economic Sciences in Memory of Alfred Nobel 2026. Stockholm.
- Shumway, T. (1997). The delisting bias in CRSP data. *Journal of Finance*, 52(1), 327–340.
- Sims, C. A. (2003). Implications of rational inattention. *Journal of Monetary Economics*, 50(3), 665–690.
- Woodford, M. (2003). *Interest and Prices: Foundations of a Theory of Monetary Policy*. Princeton, NJ: Princeton University Press.