

Deal Completion as a Staged Hazard: Filing-Typed Stage Conditioning, Elapsed Time, and the Limits of a Delisting Proxy

Maksim Aleksandrovich Andreev, Joseph Howden
Quark Quantitative Research

Abstract

Merger-arbitrage practice attaches a completion probability to a transaction. We argue that the quantity worth estimating attaches instead to a transaction at a stage of its regulatory process, and we estimate it. On 2,970 decided episodes assembled from typed filing sequences, the probability of completing with a delisting runs from 50.3% at the first deal-family filing to 72.9% at a preliminary merger proxy and 77.6% at a definitive one, with Wilson 95% intervals of [48.5%, 52.1%], [69.9%, 75.7%] and [75.2%, 79.9%]. The tender path sits apart, at 50.8% on 575 episodes reaching a tender offer and 76.3% on 338 reaching a target response. Every one of those rates rises in the post-2018 half of the sample, the tender stage by the largest margin, from 44.3% to 71.2%. Elapsed time within a stage carries at least as much information as the stage itself. Among the 1,759 episodes that reach a definitive proxy or a tender offer, 68.6% complete; restricted to the 847 still carrying no delisting form at day 113, the seventy-fifth percentile of the closers' duration, completion falls to 34.8%, a two-proportion z of 16.35 against a preregistered bar of 2.0. Two filing-sequence features measured in a deal's first sixty days were subsequently taken to an independent panel built to test whether the outcome window, which is anchored on each episode's own last family filing and is therefore outcome-correlated, manufactures them. One survived every control and one did not, and the surviving contrast has since been the only day-sixty conditional this programme cites as measured. All completion rates here are read through a delisting proxy and are lower bounds on true completion, tight where consideration is cash and loose where a listed successor survives; the unit that produced them is recorded as a measurement rather than a confirmed hypothesis for that reason. The severity of a break, the second term any expected-value calculation requires, remains a literature import of roughly -20% in this programme rather than a quantity it has measured, and the two attempts to measure it from free sources each sealed on a data gate before any curve was read.

1. Introduction

The elementary object in merger arbitrage is a completion probability. A position in an announced target pays the spread if the transaction closes and loses a much larger amount if it fails, so the expected value of the position is a probability multiplied by a spread, less a probability multiplied by a break loss. Practitioners and desks routinely carry a single number for the first term, applied at the class level. This programme carried 91%.

A single class-level number discards information that the public record supplies at no cost. A transaction that has filed a definitive merger proxy has cleared staff review of a preliminary one and has a signed agreement behind it; a transaction whose only filing is an initial announcement has cleared nothing. The regulatory process leaves a typed, machine-readable

trace of exactly which of those two situations obtains, and the trace is available to anyone with a network connection. Conditioning the probability on that trace is therefore free in the sense that matters, and the question is only how much it moves the estimate.

This paper reports the movement. The organising claim is that completion probability behaves as a hazard on a small state space of filing stages, in which the relevant conditioning is a pair: which state the episode occupies, and how long it has occupied it. Both components are estimated below, and the second turns out to move the estimate by more than the first. The paper also reports what conditioning of this kind cannot do, which is to price the downside, and it reports a methodological result about the outcome definition that was uncovered by taking one of the day-sixty conditionals to an independently built panel.

Contribution. Stage-conditional completion probabilities with Wilson intervals and episode counts on a filing-typed state space, estimated on 2,970 decided episodes and reported with a pre-2018 and post-2018 split. An elapsed-time conditional whose magnitude exceeds the largest stage increment, together with an account of what that conditional is and is not, since its complementary arm is degenerate by construction of the threshold. A deconfound of the day-sixty failure conditionals against an outcome-correlated observation window, which retains one of the two and withdraws the other. An explicit statement of the two terms an expected-value calculation needs and the record of which of them this programme has measured.

A companion paper places the stage chain as the topmost rung of a pre-announcement probability ladder that begins before any transaction is public. That treatment reports the stage chain as one measurement among five. The present paper takes the chain as its subject, adds the duration structure and the deconfound, and supersedes the companion on one point of fact: the companion describes a day-sixty rebase that was withdrawn on 11 August 2026 and reports it as wired.

2. Population, Typed Stages, and the Completion Proxy

The episode population is assembled from complete per-filer submissions histories held by the Commission, harvested once for the filer set produced by a sibling unit's outcome join and cut into episodes wherever more than 365 days separate consecutive deal-family filings. The harvest yields 3,056 episodes. Of these, 86 are censored for carrying insufficient outcome runway and are excluded from every probability rather than counted as failures, leaving 2,970 decided episodes spanning 2001 to 2026.

Stages are defined by form code alone, with no recourse to filing text. The design freezes this restriction because form codes are assigned by the Commission and are immune to the extraction problems that have blocked this programme's attempts to recover consideration terms from prose.

Deal-family forms (frozen before any statistic). 425, PREM14A, DEFM14A, SC TO-T, SC 13E3, SC 14D9. The first such filing in an episode opens it and defines the first-family-filing stage. A PREM14A marks a preliminary proxy, a DEFM14A a definitive proxy, an SC TO-T a tender offer, and an SC 14D9 a target response. Episodes are separated by gaps exceeding 365 days between family filings.

Completion proxy (frozen). A Form 25 or a Form 15-12B within 548 days of the episode's last family filing. Episodes younger than 548 days are censored and excluded. Everything else is scored as not completed, which pools genuine abandonment with completion that leaves no delisting trace.

The stages form a partial order rather than a chain, since the proxy path and the tender path are alternatives and an episode may reach a definitive proxy without ever filing a tender offer. The counts reflect this: 905 episodes reach a preliminary proxy, 1,243 a definitive one, 575 a tender offer and 338 a target response, against 2,970 that reach the first family filing by definition. An episode may be counted at several stages, and the 1,759 episodes reaching either a definitive proxy or a tender offer is smaller than the sum of those two counts because some episodes reach both.

3. What a Completion Rate Means Here

The binding limitation of this paper is its outcome definition, and it belongs before the estimates rather than after them. Nothing in a filer's submissions record announces that a merger closed. What the record contains is the registration-withdrawal machinery that follows an acquisition of a listed target: a company acquired for cash stops reporting, a Form 25 or Form 15-12B appears, and the proxy fires. A company acquired for stock in a combination that leaves a listed successor may file no such form, and the proxy scores that episode as non-completion once its 548-day window elapses.

Every probability reported below is therefore a probability of close-with-delisting, which bounds true completion from below. The bound is tight where consideration is cash and loose where it is stock. The

unit that produced these estimates is recorded in the programme's register with the status of a settled measurement rather than a confirmed hypothesis, and its own results file carries the reading instruction in its note field: read these as P(close-with-delisting | stage).

Binding caveat, restated at each estimate. Every completion figure in this paper is a lower bound on true completion. It is tight for cash transactions and loose for stock-for-stock combinations whose targets' holders end up holding a listed successor. Differences between arms measured by the same proxy on the same population are far better behaved than levels, and a reader who wants one number from this paper should take a difference.

The looseness is not uniform across the sample, which is what makes the levels untrustworthy in a specific and diagnosable way rather than merely imprecise.

STAGE	N	CLOSED	P(CLOSE WITH DELISTING)	WILSON 95%	PRE-2018	FROM 2018
First family filing	2,970	1,495	50.3%	[48.5%, 52.1%]	42.4%	63.0%
Preliminary proxy	905	660	72.9%	[69.9%, 75.7%]	66.2%	80.9%
Definitive proxy	1,243	965	77.6%	[75.2%, 79.9%]	72.0%	83.9%
Tender offer	575	292	50.8%	[46.7%, 54.9%]	44.3%	71.2%
Target response	338	258	76.3%	[71.5%, 80.6%]	72.5%	84.8%

The proxy path behaves as the regulatory sequence predicts. An episode observed at its first family filing completes at 50.3%; the same episode observed after a preliminary proxy completes at 72.9%; after a definitive proxy at 77.6%. The increment from the first observation to the definitive proxy is 27.3 points, and the Wilson intervals of those two stages are separated by more than twenty points, so the ordering is not an artifact of sampling error at these counts. Against the flat 91% this programme previously applied, the measured definitive-proxy stage is thirteen points lower, and the measured first-filing stage is forty points lower.

The tender path does not behave that way, and the era split indicates why the level should be distrusted rather than interpreted. Tender completions under the proxy run at 50.8% pooled, which would place a tender offer at the same completion probability as an

Consideration mix varies by era and by transaction structure, so a comparison of levels across eras confounds a change in deal outcomes with a change in how often the proxy sees them. Section 4 exhibits a case where that confound is the more plausible reading of a large movement.

4. Stage-Conditional Completion

The estimates appear in Table 1. Each row reports the episodes reaching that stage, the count completing under the proxy, the resulting proportion with its Wilson 95% interval, and the same proportion computed separately on episodes before and from 2018.

unadvanced announcement. Splitting the sample, the tender stage moves from 44.3% before 2018 to 71.2% from 2018, a movement of 26.9 points and the largest of any stage in the table. The alternative reading, that tender offers genuinely completed at 44% for a decade and a half and then at 71%, requires a change in transaction outcomes of a magnitude that no other stage exhibits and that the practitioner literature does not record. A change over the same period in how reliably a completed tender leaves a delisting trace in this population accounts for the pattern more economically. The target-response stage, which is the same transactions seen through the target's own filing, sits at 76.3% and moves by 12.3 points across the split, which is in line with the proxy path and is the more plausible reading of the tender path's underlying completion.

Every stage rises across the 2018 split, by between 11.9 points at the definitive proxy and 26.9 at the tender offer. A uniform rise of that shape is what a change in proxy sensitivity would produce, and it is also what a genuine secular improvement in completion would produce. The sample cannot separate those, and the paper does not claim to. The operational consequence is that a reader applying

these numbers to a current transaction should prefer the post-2018 column, and should treat even that as a lower bound.

5. Elapsed Time Within a Stage

A stage is a state, and a state has an occupancy time. Table 2 reports the duration distributions the same episode set supports, in days.

TRANSITION	P10	P50	P75	P90	N
First filing to definitive proxy	11	44	91	160	1,243
Definitive proxy to close	31	51	111	245	948
Tender offer to close	29	36	129	476	290
First filing to close	38	120	225	429	1,495

The distributions are long-tailed in the way that makes a median an inadequate summary. A transaction reaching a definitive proxy closes in a median of 51 days, and one episode in ten of those that close takes more than 245. A tender offer closes in a median of 36 days, the fastest transition in the table, and its ninetieth percentile is 476 days. The gap between the median and the upper decile is where the information sits.

The unit's single falsifiable rider was frozen on that gap. Among the 1,759 episodes reaching a definitive proxy or a tender offer, completion under the proxy

runs at 68.6%. Restricting to the 847 that were still carrying no delisting form at day 113, the seventy-fifth percentile of the closers' duration, completion falls to 34.8%. The two-proportion statistic is $z = 16.35$ against a preregistered bar of 2.0, and the rider cleared. The magnitude of the drop, 33.8 points, exceeds the 27.3-point increment that separates the first-filing stage from the definitive proxy. Elapsed time is the stronger of the two conditioning axes on this sample.

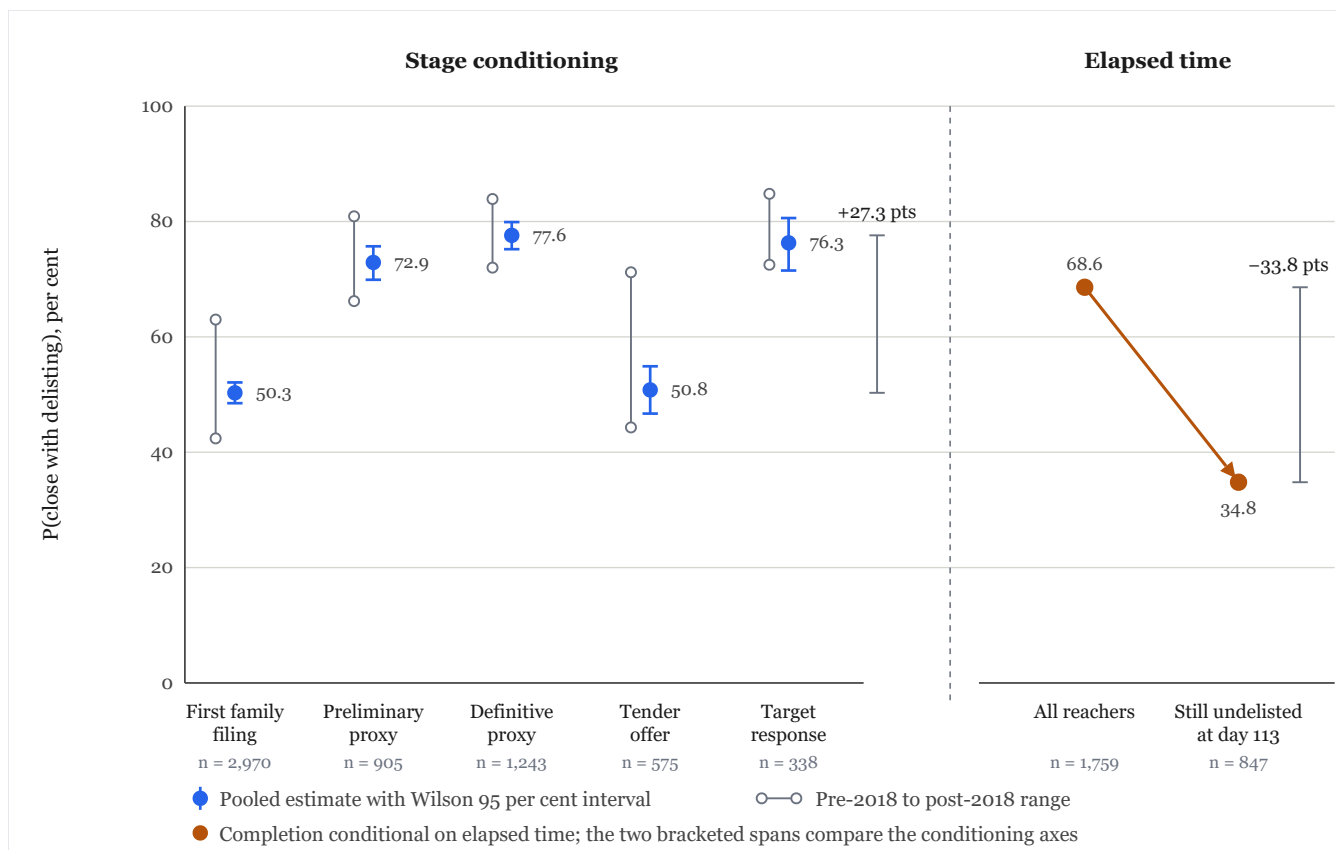


Figure 1. The two conditioning axes on a common scale. Left panel: probability of completing with a delisting at each of five filing-typed stages (blue, with Wilson 95% intervals), each accompanied by the range from its pre-2018 to its post-2018 estimate (grey, open markers). Right panel: the elapsed-time conditional among the 1,759 episodes reaching a definitive proxy or a tender offer (amber). The bracketed spans are derived differences of the plotted values and are the comparison the figure exists to make: advancing from the first family filing to a definitive proxy is worth 27.3 points, while conditioning on an episode still carrying no delisting form at day 113 is worth -33.8. Two features of the left panel resist a table. The stage axis is non-monotone, with the tender stage at 50.8% falling below both proxy stages; and the era ranges are wide and uniformly upward, the tender stage moving 26.9 points across the split against 11.9 at the definitive proxy, which is the pattern that makes the tender level a statement about the proxy rather than about tender offers. Every level plotted is a lower bound on true completion in the sense of Section 3. Data: the stage-chain results file (August 2026); episode counts are printed beneath each category.

The elapsed-time conditional requires a precise statement of what it is, and the register's reported form is the correct one. The comparison is the overdue arm against the pooled rate among all reachers, and it is stated that way because the complementary arm is degenerate by construction. An episode that is not overdue at day 113 is, by the definition of the threshold, one that already carries a delisting form, and therefore completes with probability one. The reported counts confirm the identity arithmetically: the 847 overdue episodes at 34.8% contribute approximately 295 completions, the pooled 1,759 at 68.6% contribute approximately 1,207, and the difference of 912 matches the 912 non-overdue episodes exactly. That derivation is ours rather than

the unit's, and we report it because it changes how the rider should be read. The rider is a survival statement: 48.2% of advanced-stage episodes are still undelisted at day 113, and among those the eventual completion probability is 34.8%. It is not a covariate measured at day sixty and followed forward, and treating it as one would overstate what it establishes.

Read as a survival statement the rider is still the more useful of the paper's two conditioning results for a position already held, because it updates continuously as a transaction ages, whereas the stage indicator updates only when a form is filed. It is also the conditional this programme wired first, replacing a hand-set fifteen-point penalty for overdue transactions with the measured 34.8%.

6. Whether the Observation Window Manufactures the Day-Sixty Conditionals

A separate unit measured nine filing-sequence features over each deal’s first sixty days, with each feature’s direction frozen before the data were seen, against the same delisting-proxy outcome on the 2,625 episodes still alive at day sixty. Two features cleared a bar of $|z| \geq 3$ in their frozen direction. A proxy-path episode with no definitive merger proxy on file by day sixty fails at 62.24% on 1,671 exposed episodes against 45.60% on 954 unexposed, $z = 8.27$. A tender offer still unresolved at day sixty fails at 73.18% on 343 exposed against 53.64% on 2,282 unexposed, $z = 6.80$. The family-wise control was 500 label permutations within announcement-year strata, whose maximum absolute z reached 3.77 at the ninety-fifth percentile, so the stronger of the two exceeded the permutation ceiling by more than a factor of two.

Both were wired, and one of them was then withdrawn. The reason is a defect in the outcome definition that the hazard framing makes visible. The

outcome window runs 548 days from the episode’s own last deal-family filing, so its length is a function of the episode’s filing behaviour, and filing behaviour is what the features measure. In survival terms the censoring time depends on the covariate, which is the classical condition under which a contrast can be produced by the labelling rather than by the process.

A deconfound unit was registered to test exactly this. It built an independent panel offline from a local copy of the Commission’s full-index files covering 4,569 filer identifiers, 865,854 form rows and the period from 1 January 2020 to 7 August 2026, and imported the parent unit’s episode, feature and label code verbatim so that the data source was the only difference. Four arms were frozen: the parent label as a reconciliation gate, a primary arm replacing the endogenous window with a fixed 548 days from episode start, a 730-day sensitivity arm, and an arm restricted to terminal episodes. A Mantel-Haenszel arm matched on observation-window quintiles was included at an external reviewer’s suggestion.

ARM	WINDOW DEFINITION	Z, SLOW PROXY PATH	Z, TENDER UNRESOLVED
Reconciliation	Parent label, full panel	3.12	12.25
Reconciliation, restricted	Parent label on the primary arm’s rows	1.44	11.76
Primary	Fixed 548 days from episode start	3.24	11.33
Sensitivity	Fixed 730 days from episode start	2.32	11.16
Terminal episodes only	Fragment control	2.54	12.05
Mantel-Haenszel	Matched on window-length quintiles	9.70	11.63

The tender conditional survived every arm, with statistics between 11.16 and 12.25 and a contrast of 35.8 points in the primary arm. The label mechanics run against it: among exposed episodes that close, 11.0% do so after day 548, against 15.2% among unexposed, so the window rule if anything suppresses the contrast. The wiring stands.

The slow-proxy conditional did not survive. Its primary arm passed at 3.24 against a bar of 3.0 while both remaining controls failed, at 2.32 and 2.54, and the frozen ladder required all three. The marginal primary pass is accounted for arithmetically by the

channel the sensitivity arm was frozen to bound: 19.9% of exposed episodes that close do so after day 548, against 5.9% of unexposed, a differential mislabelling of 14.1 points against a measured contrast of 6.75. Two further numbers sit below the bar without having been ladder inputs. The parent label applied to the primary arm’s own rows gives 1.44, and the reconciliation arm’s 3.12 falls short of that arm’s own family-wise permutation ceiling of 3.69. The parent’s 8.27 replicates at 3.12, and the day-sixty close rate of 37.8% that had been wired does not replicate at all, the independent panel giving 48.6% in the primary arm and 53.5% in the sensitivity arm.

The duration diagnostics explain the mechanism directly. Exposed episodes that close take a median of 217 days and an upper-decile 888; unexposed episodes that close take a median of 75 and an upper-decile 314. A feature defined as the absence of a definitive proxy by day sixty selects transactions that are slow, and a fixed-length window scores slow completions as failures. The feature and the label share a cause.

The externally proposed Mantel-Haenszel arm passes for the slow-proxy feature at 9.70 and rescues nothing, which is itself a methodological finding worth recording. Window length equals 548 days plus the episode's family-filing span, so it is a post-treatment variable and matching on it conditions on a consequence of the exposure. The common-effect assumption also fails visibly: the exposed and unexposed failure rates run 76.3% against 30.1% in the shortest-span stratum and reverse to 24.9% against 34.0% in the longest. A statistic computed under an assumption its own strata contradict is not evidence for the effect it reports.

The consequence, frozen at preregistration and applied on settlement, is that the slow-proxy day-sixty re-base was downgraded to a directional flag, the 37.8% figure may no longer be cited as a measured number, and the short-side arming path may cite only the tender conditional and the day-113 rider of Section 5. The programme's live surface now carries that language in place of the withdrawn rate. We record the episode because a conditional that cleared a preregistered bar, a permutation control and a settlement review was still produced in part by its own label, and the instrument that detected this was an independently built panel rather than a stronger statistic on the same one.

One further reading of the same table deserves recording. A third feature, the volume of additional soliciting material filed in the first sixty days, carried the largest effect in the parent study at $z = -14.73$, with high-volume episodes failing at 33.10% against 65.00% for the rest, and its sign is opposite to the direction frozen for it. The frozen settlement rules admit a feature only in its declared direction, so the finding was reported without being adopted, and it

replicates in the independent panel at -9.58 . A registration and a forward test would be required to use it.

7. Arrival Non-Stationarity and the Era Split

A hazard model on a stage space presumes that episodes entering the space are drawn from a stable process. A separate unit tested that presumption directly by fitting a self-exciting point process to deal announcement times, against a homogeneous Poisson alternative, on 1,868 business-day-timed events. Clustering is present in the pooled sample: the likelihood-ratio statistic against Poisson is 19.1 at $p = 7.35 \times 10^{-4}$, the branching ratio is 0.421 with a background rate of 0.558 events per business day and an excitation half-life of 13.1 business days, and the fitted branching ratio exceeds the ninety-fifth percentile of 100 within-weekday shuffles at 0.212, so day-of-week seasonality does not account for it.

The design's stability bar failed. Fitted on halves, the branching ratio runs 0.532 in the earlier half and 0.042 in the later, the latter being indistinguishable from a Poisson process. The unit's own results file records its overall pass flag as false. The consequence for arrival-intensity modelling was to close it: a state variable present in one half of a sample and absent in the other cannot be conditioned upon.

The consequence for the present paper is narrower and is a caution rather than a result. The stage estimates of Section 4 pool 2001 to 2026, and the arrival process generating those episodes is demonstrably different across the sample. That is a second and independent reason, alongside the consideration-mix argument of Section 3, to read the era split in Table 1 as informative about sample composition and to prefer the post-2018 column for current use.

8. The Term This Programme Has Not Measured

An expected-value calculation on an announced-deal position requires two quantities. The first is the completion probability, which Sections 4 and 5 estimate under a stated proxy. The second is the loss

conditional on a break, and this programme does not have a measurement of it. The desk that consumes these probabilities carries a break move of 20% as a constant, annotated in the source as a literature prior of the Mitchell-Pulvino class with an uncertainty of roughly ten points, and it is labelled as a literature prior on the surface that displays it. Nothing in this paper improves it.

Two units were registered to replace that import with a measurement, and each sealed on a gate before any curve was read. The first attempted to calibrate completion against entry spread on a reconstructed cash-deal panel, with a differential-coverage gate frozen in advance as the selection-bias killer, since a failed transaction remains priceable while a completed one delists. Its blind twenty-sample validation failed at 8 of 20, the failures being junk tickers produced by a cover-page extraction bridge along with wrong-quantity and wrong-era prices, and it was independently terminal on population: the eligible set was 83 episodes containing 13 failures, so the failure arm could not have reached the frozen citable bar of 20 per bucket at any price coverage.

The second unit answered each of those walls, keyed identity from the Commission rather than from a text bridge and priced both arms from a single delisted-capable instrument, and reached the pricing stage with its upstream work intact: 132 candidates, 46 verified cash reconstructions, 41 outcomes resolved from the panel's own delisting filings and 37 of 41 tickers recovered. It sealed on coverage, at 11 usable rows against a floor of 40 and 8.3% coverage against a floor of 60%. The binding constraint has a date: the price vendor's monthly cap of 500 unique symbols was exhausted mid-harvest, and the unit's own results file records that a single repricing pass after the monthly reset completes the remaining names. The differential-coverage gate that killed the predecessor passed here, at 14.4 points against a ceiling of 20.

The structural finding that emerged from the second attempt bears on any replication of this class of work. Both free machine-readable maps from filer identifier to ticker, the Commission's company file and the per-filer submissions ticker field, remove a company's symbol when it delists. A panel keyed on either map

therefore self-selects toward transactions that failed, which is the differential-coverage failure relocated from the price layer to the identity layer, where a coverage gate on prices would not see it. The company's own filing cover page retains the symbol and recovered the missing names.

Consequence for the expected-value product. A reader combining the probabilities of Sections 4 and 5 with a break loss is combining a measured lower bound with an imported constant. The completion term is conservative and the loss term is neither conservative nor measured. This programme does not present the product as an estimate, and the registered hypothesis governing the merger-arbitrage spread has not settled.

9. What Is Wired, and the Forward Status

Three consequences were frozen before settlement and applied on the day each unit settled. Deal cards on the programme's internal surface gained typed stage detection and display the measured stage-conditional probability in place of the flat 91% wherever a stage is detected, with the flat prior retained as the fallback where it is not. The overdue trigger became the measured 34.8% conditional on the day-113 threshold, replacing a hand-set penalty. The day-sixty tender conditional re-bases an unresolved tender offer to a 26.8% completion probability, and the day-sixty slow-proxy conditional does not re-base anything, having been downgraded to a directional flag on 11 August 2026 with its withdrawn rate named in the flag text.

The immediate effect of the stage wiring was that seven of eleven then-open transactions moved off the flat 91% prior to stage-conditional values of 78% and 76%. Precedence among the conditionals was frozen: the day-113 rider outranks the day-sixty features, which outrank the stage base rate.

None of this rests on out-of-sample evidence. Every estimate in this paper is an in-sample measurement on historical filings. The registered hypothesis that would supply forward confirmation, which governs whether the announced-deal spread earns a positive risk-adjusted return net of break losses, remains unsettled with no verdict date, and its quorum is stated in

resolved deals rather than in calendar time because the claim is per-deal. No capital is committed at any point during that accrual. The honest description of the present status is that a set of hand-set numbers has been replaced by measured ones, with an interval on each and a named proxy through which each must be read, and that the forward record which would license using them is outstanding.

10. Limits

Four limits bound what this paper establishes, and they are of different kinds.

The outcome proxy bounds the levels. Every completion rate is a rate of close-with-delisting, tight where consideration is cash and loose where a listed successor survives, and the looseness varies with sample composition so that cross-era comparisons of levels are unreliable. This is why the unit settled as a measurement rather than as a confirmed hypothesis. Contrasts measured by the same proxy on the same population are the citable objects.

The observation window bounds which conditionals may be cited. Section 6 establishes that a window anchored on an episode's own last family filing can produce a contrast on a feature that measures filing behaviour, and that one of the two day-sixty conditionals was so produced. The deconfound panel covers 2020 to 2026 only, so it bounds the parent's later years and does not speak to its earlier ones.

The severity term is absent. Section 8 records that the break loss remains a literature import rather than a measurement, and that both attempts to measure it from free sources sealed on gates. A measurement-limited verdict of that kind establishes what bound the measurement and establishes nothing about the quantity. Neither unit produced a spread-to-outcome curve, and neither should be read as weak evidence for one.

Forward validation is absent. The estimates are in-sample, the hypothesis that would test them forward has not settled, and this paper licenses no change to a decision rule beyond the changes already recorded at each unit's settlement.

11. Reproducibility

Each unit is a command-line program with modes that run in a fixed order: a harvest that queries the Commission's endpoints at a 1.2-second throttle and caches every response to disk, a selftest that must pass on planted fixtures before scoring is permitted, and a scoring pass that reads only cached material and writes a results file. No network request is made during scoring, so a replicator who has completed a harvest re-derives every statistic reported here without returning to the Commission. The stage unit's selftest plants a synthetic sequence set with known transitions, known duration percentiles and a known overdue effect, all of which the scorer must recover, together with a null plant whose overdue effect must fail the rider bar. The deconfound unit's selftest additionally plants a pure window artifact, which its ladder must correctly call definitional, and verifies that a genuine effect survives the same ladder.

A selftest that passed and a defect that survived it.

The stage unit's censoring branch compared a string to a date. The selftest could not detect it because every planted episode in the fixture completed, so the censoring branch never executed. The programme's register records this as the third instance of the same lesson and now requires that fixtures exercise every branch they are meant to protect. The defect was corrected before the statistics reported here were computed.

Each unit froze a single probability that its primary bar would clear, scored at settlement as a Brier contribution. The stage chain froze 0.60 against a rider that cleared, the failure-predictor unit 0.55 against a rider that cleared, and the deconfound 0.70 against a primary that passed. The programme's intent is that after enough such units the calibration curve replaces argument about confidence with arithmetic.

12. Discussion

The empirical claim of this paper is narrow and, we think, robust to its own caveats. Conditioning an announced transaction on the filing stage it has reached moves the completion estimate by 27.3 points between the first family filing and a definitive merger

proxy, and conditioning on elapsed time within the advanced stages moves it by 33.8 in the other direction. Both conditioners are readable from public form codes, at no cost, on the day they change. The levels are lower bounds and the contrasts are the citable objects, and both statements hold uniformly across the tables.

The methodological claim is the one we expect to travel further. Section 6 records a conditional that cleared a preregistered bar at $z = 8.27$, survived a 500-permutation family-wise control, passed a settlement review, was wired into a live surface, and was then found to be substantially a property of its own outcome window. What detected it was an independently constructed panel with the parent's code imported verbatim, so that the data source was the only difference; and the diagnostic that made the mechanism legible was a pair of duration distributions rather than a stronger test statistic. Where an outcome is defined by a window whose length depends on the subject's own behaviour, and the features under test measure that behaviour, the labelling is a candidate explanation for any contrast the study finds, and the ordinary apparatus of preregistration and permutation control does not address it.

What remains open is the term that would make these probabilities actionable. A completion probability without a break severity is half of an expected value, and this programme has measured the half that is bounded from below and imported the half that is unbounded. The path to closing that gap is recorded in Section 8 with a date attached to it, and until it closes the appropriate use of the tables above is to replace hand-set numbers with measured ones on a surface that displays both, which is what has been done.

References

[1] M. Mitchell & T. Pulvino. Characteristics of risk and return in risk arbitrage. *Journal of Finance*, 56(6):2135–2175, 2001. The source class for the break-severity prior of §8, which this programme carries as an imported constant of roughly 20% and has not itself measured.

[2] C. R. Harvey, Y. Liu & H. Zhu. ...and the cross-section of expected returns. *Review of Financial Studies*, 29(1):5–68, 2016. The $t > 3.0$ convention adopted as the programme's default significance bar.

[3] A. G. Hawkes. Spectra of some self-exciting and mutually exciting point processes. *Biometrika*, 58(1):83–90, 1971. The arrival model fitted in §7.

[4] E. Bacry, I. Mastromatteo & J.-F. Muzy. Hawkes processes in finance. *Market Microstructure and Liquidity*, 1(1), 2015. Estimation practice for the same.

[5] N. Mantel & W. Haenszel. Statistical aspects of the analysis of data from retrospective studies of disease. *Journal of the National Cancer Institute*, 22(4):719–748, 1959. The stratified estimator discussed and rejected in §6.

[6] E. B. Wilson. Probable inference, the law of succession, and statistical inference. *Journal of the American Statistical Association*, 22(158):209–212, 1927. The interval reported throughout.

[7] M. A. Andreev & J. Howden. A measured ladder of pre-announcement and in-process deal probabilities, reconstructed entirely from free EDGAR filings. *Quark Quantitative Research Working Paper*, 2026. Companion paper; places the stage chain of §4 as one rung of a pre-announcement ladder, and predates the deconfound settlement reported in §6.

[8] M. A. Andreev & J. Howden. The reachable boundary of free filing data for event-driven research: a catalogue of measurement limits with named defect classes. *Quark Quantitative Research Working Paper*, 2026. Catalogues the terms-extraction and delisting-and-price boundaries referenced in §8.

[9] Preregistrations for the units reported here, each frozen before its unit read any outcome statistic, held in the experiments directory of the platform repository under the register identifier of the unit it governs.

[10] Register of hypotheses and verdicts, maintained in the platform's research engine, carrying for each unit its frozen hypothesis text, terminal status, verdict, and the paths of the programs and results files behind it. Where a register entry and a results file disagree, the results file governs and the entry is a prose summary of it.

© 2026 Maksim Aleksandrovich Andreev & Joseph Howden.
All rights reserved.

Quark Quantitative Research Platform — Working Paper
Generated 2026-08-17 21:46 UTC

Every quantity reported here is recoverable at no cost from the Securities and Exchange Commission's full-text search index and per-filer submissions endpoint. Completion figures are delisting-proxy lower bounds; the break-severity term is an imported literature prior rather than a measurement; all results are in-sample and none is forward-proven. Correspondence: M. A. Andreev, Quark Quantitative Research.