© AUG 2023 | IRE Journals | Volume 7 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/ire1725645

Allocating Schedule Contingency in Multi-Site Engineering Rollouts Under Shared Delay Risk

Prince Keshi
Department of Engineering, University of Benin, Benin City, Nigeria
ORCID: https://orcid.org/0009-0006-1442-4498

Abstract- ***Multi-site engineering rollouts are reported against one programme date, yet each building is handed back to its users on its own date. This paper asks where a fixed amount of schedule contingency should sit when some delays strike one site and others, such as a supplier setback, strike several sites at once. Public records of the New York City School Construction Authority (SCA) for school security-camera installations show that lateness is strongly shared within contract bundles: the intraclass correlation of design-phase lateness is 0.93, and none of the 15 construction bundles due by December 2022 had finished within 15 days of its planned end. A Monte Carlo model of two real three-school bundles, with crew, delivery and access constraints, compared four placements of the same contingency against identical disruptions. The best placement depends on whether handover dates can still be set. In a high-dependence scenario informed by the SCA records, with 12 days of contingency, a pooled bundle reserve cut expected late handovers from 3.71 of six under equal site buffers to 2.36, but promised dates 5.5 working days later on average. Against six dates already agreed, a tight plan with an end-of-programme reserve protected them best, with 2.82 late handovers against 3.49 for pooled scheduling. Assuming independent supplier delays barely changed mean lateness but hid clustered failure: the first tested contingency keeping four or more late handovers below 10% rose from 18 days to 39 with fully shared delays. A placement guide for managers follows.***

Index Terms- ***buffer allocation, correlated risk, handover reliability, Monte Carlo simulation, multi-site programmes, public buildings, schedule contingency***

I. INTRODUCTION

A large share of public engineering work is delivered as rollouts, in which the same scope is repeated across many buildings. Security camera systems in schools, boiler replacements in housing blocks, lighting retrofits in libraries and fire alarm upgrades in clinics all follow this pattern. The owner usually manages the work as one programme with one budget and one completion date, but each building has its own users, who lose rooms, corridors or services while the work is open and who are told when they will get them back. Finishing the programme on time and handing every building back on time are therefore different outcomes, and a programme can meet its overall date while most of its buildings are returned late.

The gap between the two outcomes matters in practice. The SCA, which delivers capital improvement work in New York City public school buildings, reports the share of capital improvement projects completed on time or early in the Mayor's Management Report. The figure was 39% in fiscal year 2022 against a target of 80%, after 65%, 54%, 47% and 23% in fiscal years 2018 to 2021, and the report attributes the shortfall to schedule delays related to COVID-19 (NYC Mayor's Office of Operations, 2022). In the first four months of fiscal year 2023 the share fell to 11%, against 37% in the same period a year earlier, with schedule delays described as an ongoing effect of the COVID-19 construction pause (NYC Mayor's Office of Operations, 2023a). The indicator counts capital repair projects reaching substantial completion within 15 days of the scheduled date (NYC Mayor's Office of Operations, 2023b). Site-level reliability is evidently hard to achieve, and the delays behind it are not confined to individual projects.

Every planner holds some time contingency against this risk, and the total is often fixed by the client or by the programme envelope. Where that contingency is placed is a separate design choice. It can be spread equally as a short buffer after each site, weighted towards sites that look risky, held as a reserve shared by a group of sites, or kept as a single block at the end of the programme. The project scheduling literature has studied buffer sizing and placement for single projects in depth (Goldratt, 1997; Herroelen and Leus, 2004; Van de Vonder et al., 2005), but much less attention has been given to programmes where the protected dates belong to different sites, where one crew travels between them, and where some disruptions are common to several sites because they share a supplier, a shipment or a contract.

Shared disruptions change the logic of pooling. Pooling a reserve is attractive because independent overruns partly cancel, so a shared reserve needs less time than the sum of separate buffers for the same protection (Hopp and Spearman, 2008). When a supplier setback delays every site in a batch, overruns no longer cancel, and a planner who sizes contingency

as if site delays were independent will misjudge how often failures cluster. Correlation between schedule risks is known to widen the spread of project outcomes (Ökmen and Öztaş, 2008), but its effect on where contingency should sit in a multi-site programme has not been quantified.

This paper addresses three questions. First, for the same total contingency, which placement protects the most site handover dates, and which protects the programme completion date? Second, how does the best placement change when delays are shared across sites, and how much does assuming independence distort the decision? Third, what does each placement cost in recovery effort, measured as repeat mobilisations, temporary plant hire-days and extra vehicle travel?

The study makes three contributions. It provides evidence from public records that schedule lateness in a real multi-site rollout is strongly shared within contract bundles and across bundles started at the same time. It offers a simulation model of a real six-school rollout that enforces crew, equipment and access constraints and that varies the dependence between supplier delays while keeping each site's own delay distribution fixed, so that the effect of correlation is isolated from the effect of more risk. It turns the results into a placement guide that separates the decision about which dates to promise from the decision about when to schedule the work. Section 2 reviews related work, Section 3 presents the SCA evidence and the case, Section 4 describes the model, Section 5 reports the results, Section 6 discusses them and Section 7 sets out limitations before the conclusions.

II. RELATED WORK

A. Buffers and contingency in project schedules

Critical chain project management popularised the idea of removing safety time from individual activities and holding it as aggregated project and feeding buffers (Goldratt, 1997). Leach (1999) reviewed the case for aggregation and the main sizing rules, including the cut-and-paste rule that takes half of the removed safety time and the root-square-error rule that treats activity variations as independent. Tukel et al. (2006) proposed sizing methods that respond to resource tightness and network density and showed that the sizing rule affects both the planned duration and the chance of meeting it. These approaches concentrate protection at the end of a chain, which suits a single project with one delivery date.

Robust scheduling research asks a related question from a different angle. Herroelen and Leus (2004) distinguished solution robustness, the stability of planned activity start times, from quality robustness, the insensitivity of the project completion date. Van de Vonder et al. (2005) showed that inserting time buffers between activities trades a longer planned makespan for more stable start times, and that the best balance depends on how costly it is to change planned dates. Lambrechts et al. (2008) studied buffers under uncertain resource availability, which is relevant when a single crew can be lost for several days. In a multi-site programme the promised handover date of each site plays the role of a planned start time that others depend on: school staff plan room moves and system training around it. Solution robustness is therefore the property that site users care about, while quality robustness is the property the programme report measures.

B. Repetitive and multi-site work

Repetitive scheduling methods treat a crew moving through repeated units as the core of the plan and emphasise crew continuity, because idle crews and repeated mobilisation cost money (Harris and Ioannou, 1998). Ballard and Howell (1998) argued that buffers shield production units from upstream variation and that reliable workflow, more than float, is what protects downstream work. Schatteman et al. (2008) combined a risk register with proactive buffering in construction projects, allocating buffers according to the expected effect of identified risks. Risk-based allocation of this kind is a natural candidate for a multi-site rollout, where some buildings are known to be harder to access than others.

C. Correlated risk

Ökmen and Öztaş (2008) developed a correlated schedule risk analysis model for construction networks and showed that ignoring correlation between risk factors makes the spread of completion times look narrower than it is. Clemen and Reilly (1999) showed how copulas allow a modeller to specify dependence between uncertain quantities separately from their individual distributions, which is the property needed to test the effect of correlation without changing site-level risk. Trietsch et al. (2012) found empirical support for lognormal models of activity time, the form adopted for delay lengths in this study. Inventory theory gives the clearest statement of why correlation matters for pooling: the advantage of a pooled safety stock depends on demands being imperfectly correlated and disappears as correlation approaches one (Hopp and Spearman, 2008).

Taken together, this work suggests that pooled contingency should protect programmes well when site risks are independent, that risk-based buffers should help when risks differ between sites, and that correlation should weaken both. None of the studies compares placements against both site and programme outcomes in a programme with travel, crew and access

© AUG 2023 | IRE Journals | Volume 7 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/ire1725645

constraints, and none calibrates the strength of shared delay to observed records. That is the gap this paper addresses.

III. EMPIRICAL SETTING: THE SCA IPDVS ROLLOUT

A. Data

The SCA publishes the schedule and budget of its capital projects as an open dataset, Capital Project Schedules and Budgets, with one row per project phase (New York City School Construction Authority, 2022). This study uses the dataset as published on 8 December 2022, which holds 11,793 phase records; a copy of that version is preserved by the Internet Archive and its SHA-256 checksum is given in the data statement. Each record gives the school building, the project description, the phase (scope, design, construction and others), the phase status, the award budget and four dates. The field Project Phase Planned End Date is defined in the dataset as the date the phase was originally scheduled to be completed, so it is a true baseline date rather than a rolling forecast. Actual start and actual end dates are given where the phase has started or finished.

IPDVS projects install IP digital video surveillance systems, including cameras, cabling, network equipment and recording servers, in school buildings. They suit this study because the scope is standard across sites, the equipment comes from a small number of supply routes, and the SCA groups schools into multi-site contracts. Projects were selected whose description begins with IPDVS, optionally preceded by a fiscal year label or a Reso A funding label. This gave 597 phase records for 150 buildings, including 149 construction-phase records: 79 in progress, 69 not started and 1 complete.

The dataset has no contract identifier, so contract bundles were reconstructed. Schools whose phase shares the same actual start date and the same planned end date were treated as one bundle, which is consistent with how multi-school contracts are mobilised and scheduled. Of the 80 construction records with both dates, this rule formed 46 bundles: 21 single schools, 16 pairs and 9 bundles of three. Planned construction durations followed two templates, 269 and 299 calendar days, with a median of 299.

B. Measures

Lateness was measured as the actual end date minus the originally planned end date, in calendar days, and for phases still open at the snapshot as the snapshot date minus the planned end date, which is a lower bound on eventual lateness. A bundle was counted as on time if every school in it was complete within 15 days of the planned end, which mirrors the SCA reporting threshold. The degree to which lateness is shared within bundles was measured by the one-way random-effects intraclass correlation coefficient (ICC) (McGraw and Wong, 1996), which is the share of the variance in site lateness that lies between bundles rather than within them. Single-school bundles contribute to the between-bundle mean square only, and unequal bundle sizes are handled with the adjusted mean bundle size of the one-way analysis of variance. Confidence intervals came from 2,000 bootstrap samples in which whole bundles were resampled. Because only one construction phase had been completed at the snapshot, the ICC was estimated for the scope and design phases, using completed phases in bundles reconstructed in the same way.

C. Findings

Table 1 summarises the evidence. Fifteen construction bundles, covering 25 schools, had a planned end on or before 8 December 2022. Only one of them had been completed, 834 days after its planned end, and the other 14 were still in progress between 37 and 976 days after their planned end, with a median of 259 days (Figure 1a). None of the 15 met the 15-day threshold, and none would have met a 30-day threshold. Twelve of the overdue bundles had started between May and August 2021 and were all between 187 and 301 days overdue, which suggests that the lateness was driven by conditions common to the whole wave rather than by the individual schools. This matches the authority-wide pattern of delays attributed to the COVID-19 construction pause (NYC Mayor's Office of Operations, 2023a).

Table 1. Schedule evidence from SCA IPDVS records (dataset as published on 8 December 2022)

Measure	Value	Basis
IPDVS phase records / buildings	597 / 150	Description filter
Construction phase: in progress / not started / complete	79 / 69 / 1	Phase status
Construction records with start and planned end; bundles	80; 46	Shared actual start and planned end
Bundle sizes: one / two / three schools	21 / 16 / 9	Reconstructed bundles
Median planned construction duration	299 days	80 records
Bundles due by snapshot (schools)	15 (25)	Planned end on or before 8 Dec 2022
Due bundles complete within 15 days of planned end	0 of 15	All schools in bundle complete

© AUG 2023 | IRE Journals | Volume 7 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/ire1725645

Measure	Value	Basis
Open due bundles: days past planned end	Median 259 (37 to 976)	Lower bound on lateness
ICC of lateness, scope phase	0.59 (0.16 to 0.90)	136 sites, 103 bundles (22 multi-site)
ICC of lateness, design phase	0.93 (0.86 to 0.97)	108 sites, 76 bundles (17 multi-site)
Design phases complete within 15 days of planned end	78%	108 completed phases

Note: ICC values are one-way random-effects estimates with 95% bootstrap intervals from 2,000 resamples of bundles. Only one construction phase was complete, so no construction-phase ICC could be estimated. Source: New York City School Construction Authority (2022), author's analysis.

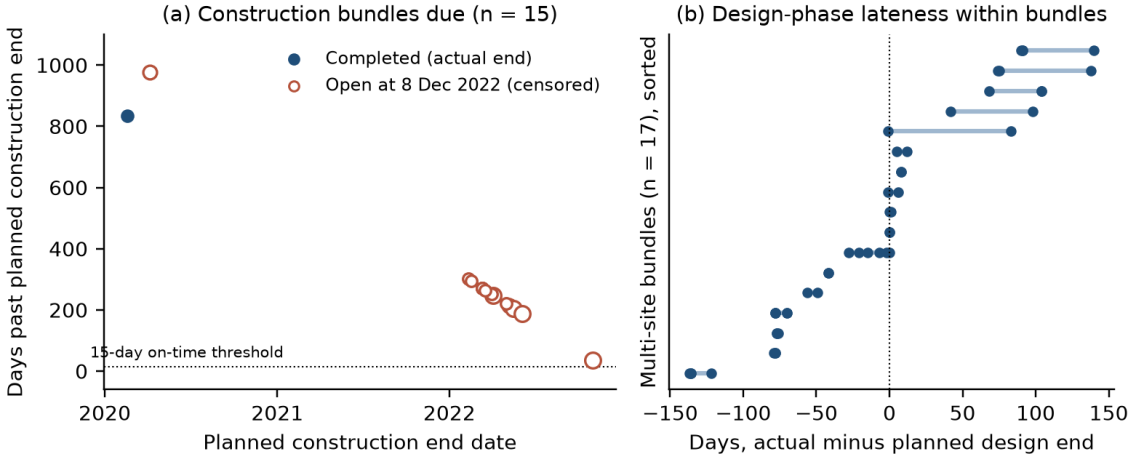


Figure 1. (a) IPDVS construction bundles with a planned end on or before 8 December 2022, plotted as days past the originally planned end; the filled marker is the one completed bundle, hollow markers are bundles still open at the snapshot and marker area is proportional to the number of schools. (b) Design-phase lateness of the schools in each of the 17 multi-site bundles with completed design, one row per bundle; most bundles lie on a single point or a short line. Source: New York City School Construction Authority (2022), author's analysis.

Lateness is strongly shared within bundles. In the design phase, which has the largest sample of completed multi-site bundles, the ICC is 0.93, and the within-bundle standard deviation of lateness is 18.9 days, whereas mean lateness spans about 290 days across the 76 bundles and about 240 days across the 17 multi-site bundles shown in Figure 1b. Because bundles were reconstructed from shared dates, part of this similarity could reflect shared scheduling or recording practice rather than shared delay. Two checks limit that concern: grouping on the actual start date alone gives the same 76 bundles, and the ICC of actual design duration, which does not use the planned end date, is also 0.93 (95% CI 0.86 to 0.97). Common administrative recording of completion dates within a contract cannot be ruled out with these data. Design phases were more often early than late, with a median of 22 days early, so the high ICC reflects bundles moving together rather than a general tendency to overrun. The scope phase shows a weaker but still substantial ICC of 0.59, with a wide interval. The design-phase value is used as the calibration target for shared delay, because it is the most precise estimate available and it describes phases that were completed. Construction lateness cannot yet be decomposed in the same way, but the clustering of overdue bundles from the 2021 wave in Figure 1a points in the same direction.

D. The case rollout

The model is built around two real bundles from the same dataset, both from the May 2022 construction wave and both on the 299-day template. The Bronx bundle covers three schools in District 12 whose construction started on 16 May 2022 with a planned end of 11 March 2023. The Manhattan bundle covers three schools in Districts 1 and 2 whose construction started on 24 May 2022 with a planned end of 19 March 2023. Both were in progress at the snapshot date. Table 2 lists the schools with their SCA award budgets. Four of the six buildings house more than one school, so each handover affects several school communities. School coordinates came from the Department of Education school point locations file

© AUG 2023 | IRE Journals | Volume 7 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/ire1725645

(NYC Department of Education, 2020). Road distances from a base at the SCA headquarters in Long Island City were approximated by multiplying straight-line distance by a circuity factor of 1.2, following the urban detour factors reported by Ballou et al. (2002).

Table 2. Case rollout: two three-school IPDVS bundles from the May 2022 construction wave

Code	School building	Bundle	Budget (USD)	Planned install / commission (working days)	Road km from base	Local risk class
X050	Fairmont Neighborhood School, 1550 Vyse Avenue	Bronx	274,326	9 / 4	12.9	Normal
X167	P.S. 214, 1970 West Farms Road	Bronx	308,370	11 / 4	13.8	High
X193	P.S. 211, 1919 Prospect Avenue	Bronx	302,384	10 / 4	14.0	Normal
M066	P.S. 151 Yorkville Community School, 421 East 88th Street	Manhattan	120,678	4 / 2	4.5	Normal
M440	Hudson High School of Learning Technologies, 351 West 18th Street	Manhattan	198,231	7 / 3	6.7	Normal
M025	Lower East Side Preparatory High School, 145 Stanton Street	Manhattan	201,929	7 / 3	6.0	High

Note: Bronx bundle: construction start 16 May 2022, planned end 11 March 2023. Manhattan bundle: start 24 May 2022, planned end 19 March 2023. Budgets are SCA award base budgets (New York City School Construction Authority, 2022). Planned durations are scaled from budgets and the local risk class is an assumption (Section 4.1); neither is observed in the data. Road distance is straight-line distance from SCA headquarters multiplied by 1.2.

The two bundles are treated as one rollout served by a single commissioning team, as would be the case if one systems integrator commissioned both contracts. This is an assumption: the dataset does not name contractors. It is tested in Section 5.5 by giving each bundle its own team. The two bundles also form the two shipment batches and the two groups for pooling. The case is used prospectively relative to the data. Its baseline dates and site attributes are real, but its outcome was not in the records analysed, so the simulation describes the closing stage that lay ahead of it at the snapshot rather than reconstructing a known result. The six schools' completed design phases are among the 108 used to estimate the design-phase ICC, but their construction outcomes, which the model describes, were not in the records analysed, so the model is not calibrated and tested on the same outcomes. The recorded completions of the two bundles will provide a later check on the model.

IV. SIMULATION MODEL

A. Scope and time base

The model represents the closing installation, commissioning and handover stage of the rollout in working days. This is the stage in which contingency decisions bite: equipment is arriving, each school must release rooms for installation and testing, and the handover dates that school staff plan around are fixed. Day 0 is the start of the stage and the first commissioning slot opens on day 25. Planned commissioning durations were set as max(2, round($3f_i$)) and installation durations as max(4, round($8f_i$)), where f_i is the site budget divided by the mean budget. This gives commissioning slots of 4, 4, 4, 2, 3 and 3 days (20 in total) and installation spells of 9, 11, 10, 4, 7 and 7 days. The largest scope in each bundle, X167 in the Bronx and M025 in Manhattan, was designated high local risk because larger installations involve more rooms and more access negotiation. These durations and risk classes are assumptions; Section 5.7 tests the effect of making local risk uniform.

B. Constraints

Three kinds of constraint are enforced. For people, one commissioning team serves all six sites and can work at only one site on any day; a variant with one team per bundle is tested in Section 5.5. Installation is carried out by site crews that are not treated as a binding constraint. For equipment, each site's shipment is planned to arrive five working days before installation starts, installation is planned to finish three working days before the commissioning slot opens,

and installation cannot start until the day after the shipment arrives. For site access, commissioning cannot begin before the date booked with the school, a site is closed to the team during any random access loss, and a site that fails its acceptance test cannot be worked until rectification is complete. In the base case the team cannot start a site early even if it is ready; Section 5.7 relaxes this.

C. Uncertainty

Each simulated rollout samples one set of disruptions (Table 3, with full detail in Appendix Table A1). Supplier delay follows a hurdle model: with probability p a site's shipment is late, and if late the delay in working days is the ceiling of a lognormal draw with median m and log-scale standard deviation σ, capped at 40 days. Dependence between sites enters through a Gaussian copula (Clemen and Reilly, 1999). For site i in shipment batch g, a latent score is formed as

$$u_i = \sqrt{\lambda}\left(\sqrt{\kappa}\, Z_0 + \sqrt{1-\kappa}\, Z_g\right) + \sqrt{1-\lambda}\, \varepsilon_i \quad (1)$$

where Z_0 is a programme-wide shock, Z_g a shock to the batch and ε_i a site-specific term, all independent standard normals. The shipment is late when $\Phi(u_i)$ exceeds 1 - p, and the delay length is read from the lognormal at the corresponding conditional quantile. A common supplier setback is therefore drawn once per run and applied to every site in the batch it affects. The parameter λ sets the latent correlation between sites in the same batch, with κ = 0.5 of that correlation shared across the whole programme, reflecting the wave-wide lateness seen in Figure 1a. Because each u_i remains standard normal, every site has the same delay distribution whatever the value of λ, so changing λ changes how delays cluster without changing how much delay there is.

Local disruptions are independent across sites: a lognormal multiplier on installation duration, a chance of a one to three day installation access loss, variation in commissioning duration, a chance of a commissioning access loss (higher at high-risk sites), and a chance of a failed acceptance test that rises with scope and triggers two to six days of rectification and a one-day retest. The team itself can be lost for two to five days at a random point in the plan window. All policies are evaluated against the same sampled events, so no policy sees a different supplier setback from another.

Table 3. Main uncertainty parameters and their status

Element	Mixed-risk regime	SCA-pattern regime	Status
Probability a shipment is late, p	0.30	0.60	Assumed; SCA regime tuned
Median delay if late, m (working days)	4	14	Assumed; SCA regime tuned
Lognormal σ of supplier delay	0.8	0.9	Assumed
Within-batch latent correlation, λ	0.9	1.0	SCA regime calibrated to ICC
Programme-wide share of common shock, κ	0.5	0.5	Assumed
Commissioning access loss, normal / high sites	0.15 / 0.55	0.03 / 0.10	Assumed
Test failure probability	0.12 to 0.22	0.12 to 0.22	Assumed, rises with scope
Team outage probability (length)	0.20 (2 to 5 days)	0.20 (2 to 5 days)	Assumed
Simulated within-bundle ICC of site lateness	0.27 / 0.47	0.92 / 0.92	Model output
Observed within-bundle ICC (design phase)	0.93	0.93	Observed (Table 1)

Note: Simulated ICC pairs are Bronx / Manhattan. Observed quantities come from the SCA records; assumed quantities are modelling choices and are varied in the sensitivity analysis.

D. Two risk regimes

Because activity-level delay data are not published, two regimes are used. The mixed-risk regime has modest supplier risk and higher local access risk, so local events dominate; its simulated within-bundle ICC of site lateness is 0.27 in the Bronx and 0.47 in Manhattan. The SCA-pattern regime raises supplier risk and lowers local access risk until the simulated ICC reaches 0.92 in both bundles, close to the observed design-phase value of 0.93. This value is

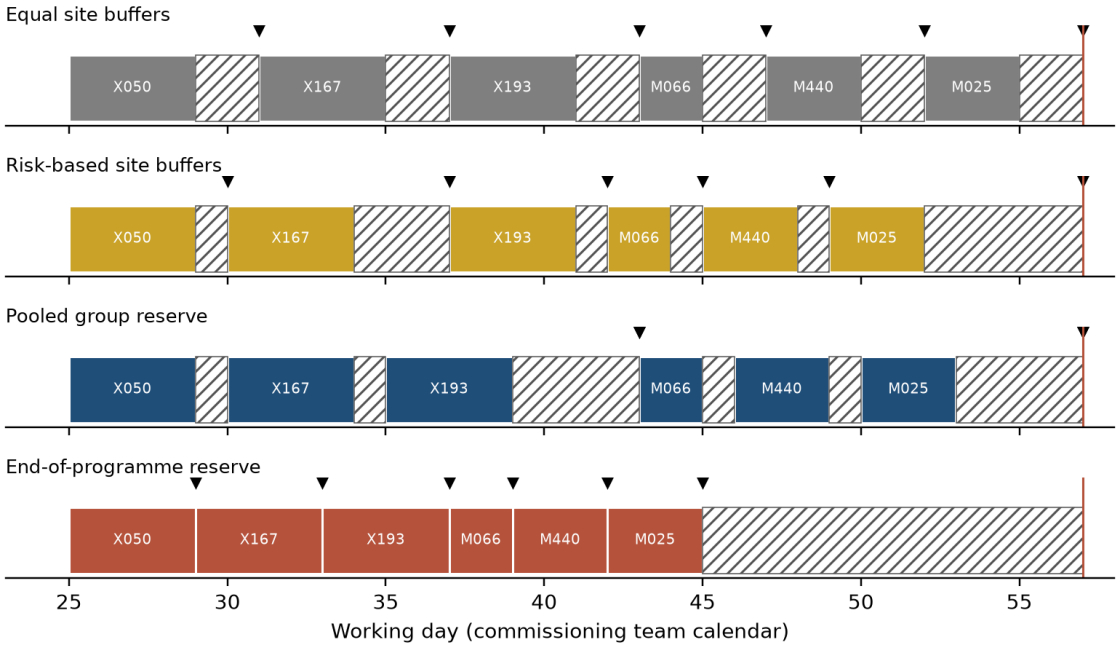
© AUG 2023 | IRE Journals | Volume 7 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/ire1725645

reached at λ = 1, the upper limit of the copula, so the SCA-pattern regime represents the strongest within-batch dependence the model can express. It is therefore best read as a high-dependence scenario informed by the design records, not as an estimate of dependence in construction, and Section 5.3 reports how the results change at intermediate dependence. The calibration transfers the dependence structure of SCA lateness, not its magnitude. SCA lateness is measured in months over a construction phase of about ten months, whereas the model's closing stage spans about twelve working weeks, so the absolute delays in the model should not be read as estimates of SCA delays. Calibrating to an observed pattern rather than to point values follows the logic of pattern-oriented modelling (Grimm et al., 2005). Keeping the observed quantity (the ICC) separate from the assumed quantities (the activity-level parameters) is deliberate, and Table 3 labels each.

E. Contingency placement policies

Every policy receives the same total contingency C in working days and the same programme deadline P = 25 + 20 + C, which is day 57 when C = 12. Sites are sequenced X050, X167, X193, M066, M440, M025, so each bundle is completed before the team crosses to the other. The policies differ only in where the contingency sits and therefore in the access dates booked with each school and the handover dates promised to them.

Under equal site buffers, C is divided equally across the six sites and placed after each commissioning slot; the promised handover is the end of the site's buffer. Under risk-based site buffers, C is allocated by a greedy heuristic: one day at a time, each day goes to the site where an extra day of buffer most raises the probability that the site absorbs its own local overrun. The aim is to reduce the expected number of sites whose local overrun exceeds their buffer; the heuristic is not guaranteed to find the best allocation. The overrun distributions are estimated from 20,000 pre-simulated runs with a separate random seed. Only local risks enter this score, which mirrors how risk registers usually assess each building on its own. With C = 12 it gives buffers of 1, 3, 1, 1, 1 and 5 days in the mixed-risk regime and 2, 4, 3, 1, 1 and 1 days in the SCA-pattern regime. Under the pooled group reserve, half of C is placed as equal site buffers and half as a reserve window after each bundle, and every school in a bundle is promised the end of its bundle window, as a sectional completion. This mirrors the SCA practice of scheduling work at bundle level. Under the end-of-programme reserve, sites are planned back to back, each school is promised the end of its own slot, and all of C sits after the last site. Figure 2 shows the four plans for C = 12.



Filled: planned commissioning slot. Hatched: uncommitted contingency. Triangle: promised handover. Red line: programme deadline (day 57).

Figure 2. The four contingency placements for the case rollout with C = 12 working days (risk-based allocation shown for the mixed-risk regime). Filled bars are planned commissioning slots, hatched bars are uncommitted contingency, triangles mark promised handover dates and the red line is the programme deadline (day 57).

F. Daily dispatch

Each day the team takes decisions using only what is known that day. If the site where it is working can still be worked, it stays. Otherwise it moves to the workable site with the earliest promised handover date, a due-date rule. A site is workable if its booked access date has arrived, installation is finished, it is not in an access loss and it is not awaiting rectification. The team does not know future deliveries, failures or access losses. If no site is workable the team waits; after two consecutive idle days it demobilises, and returning to a site it has already visited counts as a repeat mobilisation with a round trip of twice the road distance.

G. Outcome measures

Site reliability is measured by the expected number of the six schools handed over after their promised date, the probability that all six are on time and the probability that four or more are late, which captures clustered failure. Programme reliability is the probability that the last site is complete by P, with the 90th-percentile programme tardiness as a measure of the upper tail. Recovery effort is measured by repeat mobilisations and by team days worked beyond P. Contingency use is measured by idle team days within the plan window and by blocked days, when at least one site is open but none can be worked, which captures reserve time that exists but cannot be used. Temporary plant hire-days are the days between installation start and commissioning completion summed over sites, representing access towers, lifts and site storage held while work is open. Extra vehicle-km are the round trips caused by repeat mobilisations; emissions are derived from them with the 2023 UK government conversion factor for an average diesel van up to 3.5 tonnes, 0.23128 kg CO_2e per km (Department for Energy Security and Net Zero, 2023), used as a proxy because no equivalent New York factor for this vehicle class was found. Because the pooled policy promises some schools later than the other policies, each policy is also described by its mean promise deferral: the average number of working days between a school's promised date and the date its slot would end in a plan with no contingency.

H. Experimental design

Six experiments were run. E1 compares the four policies at C = 12 and C = 24 in both regimes. E2 varies λ from 0 to 1 with fixed marginals. E3 varies C from 0 to 48 days in steps of three, with independent (λ = 0) and shared delays. E4 varies the share of contingency held in the pooled windows and the grouping of sites (two bundle groups of three, or three pairs). E5 is a one-at-a-time sensitivity analysis over twelve cases. E6 compares one commissioning team with one team per bundle. All policies within an experiment are run against the same sampled disruptions (common random numbers), so differences between policies are not confounded by sampling noise (Law, 2015). E1, E2, E4 and E6 use 20,000 runs per configuration, and E3 and E5 use 10,000. Running means stabilised within about 2,000 runs (Appendix Figure A1), and the standard error of expected late handovers is at most 0.015 in E1. The experiments were run with a discrete-time implementation of the model written in Python, using the NumPy and Numba libraries. A spreadsheet version for Microsoft Excel, with its random draws, inputs and per-run outputs supplied as CSV files, accompanies the paper so that the main comparison can be inspected and rerun without programming; Appendix B describes it and shows that on identical draws it reproduces the expected number of late handovers to within 0.1 for every policy in both regimes.

V. RESULTS

A. The same contingency placed four ways

Table 4 reports the main comparison at C = 12. The pooled group reserve protects the most handover dates in both regimes. In the mixed-risk regime it gives 0.82 late handovers on average against 2.42 for equal site buffers, and a 0.45 chance that all six schools are handed back on time against 0.09. In the SCA-pattern regime the pooled reserve gives 2.36 late handovers against 3.71, and the chance of four or more late handovers falls from 0.55 to 0.31. Risk-based buffers behave differently in the two regimes. In the SCA-pattern regime they are indistinguishable from equal buffers on site reliability (3.70 against 3.71) and worse on programme reliability (0.24 against 0.33). In the mixed-risk regime the allocation puts five days behind the last site, M025, so it acts partly as an end reserve: programme reliability rises to 0.71 but late handovers rise to 2.64.

Table 4. Policy outcomes with C = 12 working days (programme deadline day 57)

Measure	Equal site buffers	Risk-based site buffers	Pooled group reserve	End reserve
Mixed-risk regime				
Expected late handovers (of 6)	2.42	2.64	0.82	4.87

Measure	Equal site buffers	Risk-based site buffers	Pooled group reserve	End reserve
P(all six on time)	0.09	0.10	0.45	0.01
P(four or more late)	0.23	0.32	0.02	0.87
P(programme on time)	0.51	0.71	0.65	0.93
90th-percentile programme tardiness (days)	6	4	4	0
Repeat mobilisations	1.39	1.42	1.42	1.45
Plant hire-days (all sites)	98.7	99.8	99.1	107.4
Idle / blocked team days	10.8 / 6.0	8.9 / 5.9	9.5 / 5.7	4.8 / 4.0
Mean promise deferral (days)	7.0	5.8	12.5	0.0
SCA-pattern regime				
Expected late handovers (of 6)	3.71	3.70	2.36	5.22
P(all six on time)	0.07	0.08	0.27	0.01
P(four or more late)	0.55	0.55	0.31	0.90
P(programme on time)	0.33	0.24	0.39	0.57
90th-percentile programme tardiness (days)	28	29	27	22
Repeat mobilisations	1.01	1.01	1.04	1.07
Plant hire-days (all sites)	92.8	92.4	94.0	100.1
Idle / blocked team days	15.3 / 12.0	16.1 / 12.1	14.3 / 11.6	10.7 / 9.9
Mean promise deferral (days)	7.0	8.3	12.5	0.0

Note: 20,000 runs per policy with common random numbers. Standard errors of expected late handovers are between 0.007 and 0.015.

The end-of-programme reserve is the mirror image. It gives the highest chance of finishing the programme on time, 0.93 in the mixed-risk regime and 0.57 in the SCA-pattern regime, and the smallest 90th-percentile programme tardiness, but because every school is promised the end of a tight slot, about 4.9 to 5.2 of the six handovers are late. Doubling contingency to C = 24 makes no difference to that count (4.86 and 5.21), because the extra days sit after the last site where no school can draw on them. By contrast, doubling C under the pooled reserve reduces late handovers to 0.18 in the mixed-risk regime and 1.57 in the SCA-pattern regime, and raises its programme on-time probability to 0.93 and 0.61. Site reliability and programme reliability therefore call for different placements, and the choice between them is a choice about which outcome the owner is accountable for.

The idle and blocked days show how the contingency is used. Under equal buffers the team is idle for about eleven days within the plan window in the mixed-risk regime, and on six of those days at least one site is open but cannot be worked because it is waiting for equipment, rectification or access. A buffer attached to a site that runs to plan is simply lost, while a site in difficulty can use only its own two days. The pooled reserve lowers idle and blocked days because a bundle window can be drawn on by whichever school in the bundle needs it. In the SCA-pattern regime blocked days are high under every policy, about 10 to 12, because a late shared shipment holds every school in the bundle at once, and no placement can create workable sites where none exist.

B. Equal contingency, unequal promises

The pooled policy has a cost that the late-handover count hides: some schools are promised later. With C = 12 its mean promise deferral is 12.5 days against 7 for equal buffers, because the first schools in each bundle are promised the bundle window end rather than their own slot end. A fair comparison therefore asks how many late handovers each policy gives for the same average deferral. Figure 3 traces this frontier by varying C. At every level of deferral the pooled reserve gives fewer late handovers than either form of site buffer. In the mixed-risk regime the pooled reserve at C = 12 (deferral 12.5 days) gives 0.82 late handovers, whereas equal buffers at C = 24 (deferral 14 days) still give 1.36. In the SCA-pattern regime the corresponding figures are 2.36 for pooling and 3.00 for equal buffers at the larger contingency. Pooling is therefore not just buying reliability with later promises; it uses the promise time more efficiently.

© AUG 2023 | IRE Journals | Volume 7 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/ire1725645

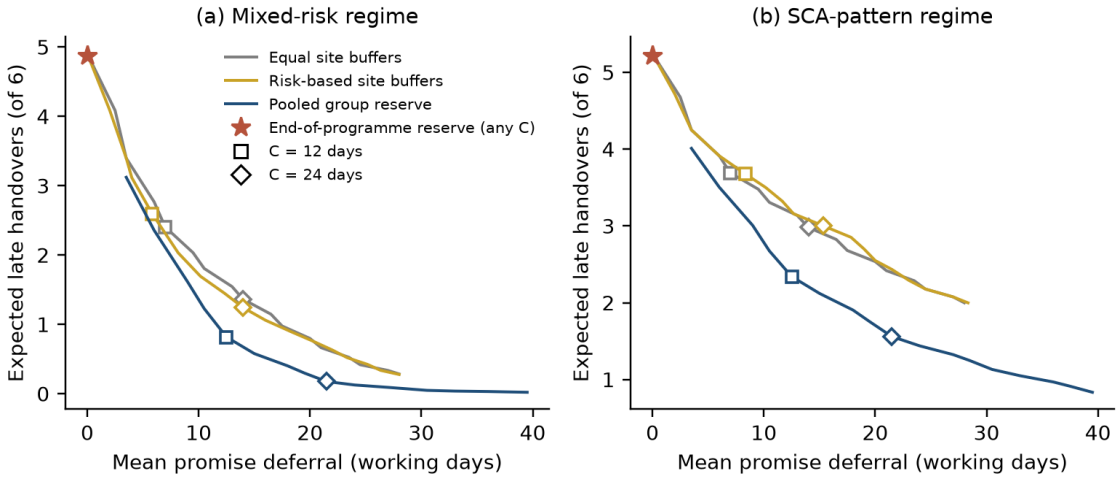


Figure 3. Expected late handovers against mean promise deferral as total contingency C varies from 0 to 48 working days. Squares mark C = 12 and diamonds C = 24. The end-of-programme reserve defers no promises and stays at the star for every C. Results from 10,000 runs per point.

A second check fixes the promises instead. If the six handover dates had already been agreed with the schools on the equal-buffer dates and each policy only changed how the team's work is scheduled, the end reserve would protect the most of those fixed dates: 1.28 late handovers in the mixed-risk regime against 1.97 for pooled scheduling and 2.42 for equal buffers, and 2.82 against 3.49 and 3.71 in the SCA-pattern regime. Working to a tight internal sequence and holding the slack centrally means that the slack sits in front of each fixed date rather than behind it. The two views are consistent. Placement of contingency is really two decisions, which dates to promise and when to plan the work, and the best answer to the first is not the best answer to the second.

C. *Shared delays and the cost of assuming independence*

Varying λ with fixed marginals changes the mean number of late handovers very little. Under the pooled reserve in the SCA-pattern regime, expected late handovers are 2.41 with independent supplier delays and 2.36 with fully shared delays; under equal buffers they are 3.84 and 3.71. The single commissioning team links the sites: a delay at one site pushes the team's later work, so late handovers already cluster through the team even when supplier delays are independent. What shared supplier delays change is the shape of the distribution (Figure 4). Outcomes become polarised: rollouts in which the common supplier is on time go well everywhere, and rollouts in which it is late go badly everywhere. The standard deviation of the late count under pooling rises from 1.20 to 2.10.

© AUG 2023 | IRE Journals | Volume 7 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/ire1725645

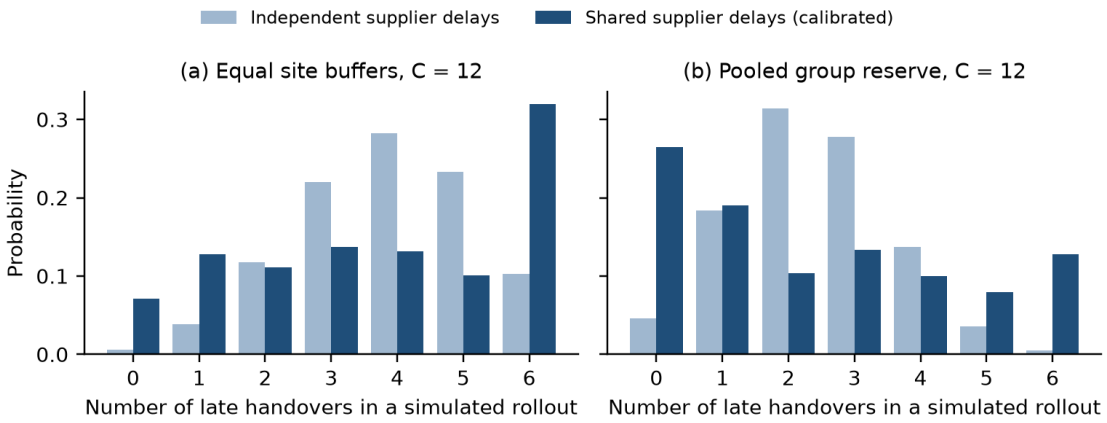


Figure 4. Distribution of the number of late handovers with C = 12 in the SCA-pattern regime, with independent supplier delays (λ = 0) and fully shared supplier delays, the high-dependence scenario informed by the SCA records (λ = 1). Marginal delay distributions are identical in both cases. 20,000 runs per bar set.

Table 5 puts numbers on the distortion. Under the pooled reserve, assuming independence puts the probability of four or more late handovers at 0.18 when the shared-delay value is 0.31, a factor of 1.7. It puts the probability of a clean rollout at 0.05 when the shared-delay value is 0.27, and the probability that all six schools are late at 0.005 when the shared-delay value is 0.13. Under equal buffers the independent model overstates the chance of four or more late handovers slightly (0.62 against 0.55) but understates the chance that all six are late by a factor of about 3 (0.10 against 0.32). The direction of the error therefore depends on which tail is examined. What is consistent is that the independent model makes both extremes look rare: it underestimates how often everything goes well and how often everything goes wrong, and the second of these is the outcome that attracts public attention.

Table 5. Effect of assuming independent supplier delays (SCA-pattern regime)

Measure	Policy	Independent (λ = 0)	Shared (λ = 1)	Ratio or gap
Expected late handovers, C = 12	Pooled group reserve	2.41	2.36	0.98
SD of late handovers, C = 12	Pooled group reserve	1.20	2.10	1.8
P(four or more late), C = 12	Pooled group reserve	0.179	0.307	1.7
P(no late handovers), C = 12	Pooled group reserve	0.046	0.265	5.8
P(all six late), C = 12	Pooled group reserve	0.005	0.128	about 25
P(four or more late), C = 12	Equal site buffers	0.618	0.553	0.89
P(all six late), C = 12	Equal site buffers	0.103	0.320	3.1
P(programme on time), C = 12	Pooled group reserve	0.113	0.394	3.5
First tested C with P(four or more late) ≤ 0.10	Pooled group reserve	18	39	+21 days
First tested C with P(all six on time) ≥ 0.50	Pooled group reserve	more than 48	27	reversed
First tested C with P(programme on time) ≥ 0.90	End reserve	33	36	+3 days
First tested C with P(programme on time) ≥ 0.80	End reserve	30	27	-3 days

© AUG 2023 | IRE Journals | Volume 7 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/ire1725645

Note: Contingency values are the first point on a three-day grid (10,000 runs per point) that meets the target; Section 5.3 gives Monte Carlo intervals around each crossing. For P(four or more late) ≤ 0.10 the first tested values at λ = 0.25, 0.5 and 0.75 are 24, 30 and 36 days. In the mixed-risk regime the distortion is small: the pooled requirement for P(four or more late) ≤ 0.05 is 9 days under both assumptions.

The distortion matters most when contingency is sized to a risk target. Under the pooled reserve, the first tested contingency that keeps the chance of four or more late handovers at or below 10% is 18 days with independent supplier delays and 39 days with fully shared delays (Figure 5a). These are grid values from a three-day grid with 10,000 runs per point, not exact thresholds. With independent delays the probability is 0.132 (95% interval 0.125 to 0.139) at 15 days and 0.095 (0.090 to 0.101) at 18 days; with shared delays it is 0.113 (0.107 to 0.120) at 36 days and 0.098 (0.092 to 0.104) at 39 days. At intermediate dependence the first tested value rises steadily: 24, 30 and 36 days at λ = 0.25, 0.5 and 0.75. A planner using the independent model would therefore hold roughly half the contingency that the high-dependence scenario calls for, and at the 5% level the shared-delay requirement lies beyond the 48 days tested while the independent model asks for 24. The error runs the other way for targets on the good tail: to give a 50% chance that every school is handed back on time, the shared-delay model needs 27 days, while the independent model never reaches that level within 48 days. For the programme date under the end reserve the two models are close. At 33 days the probability of meeting the deadline is 0.904 (0.898 to 0.910) with independent delays and 0.898 (0.892 to 0.904) with shared delays, so the gap between the first tested values for a 0.90 target (33 and 36 days) is within sampling error; for a 0.80 target they are 30 and 27 days (Figure 5b), because shared delays concentrate overruns in fewer runs and leave more runs comfortably early. Independence is therefore not a conservative simplification in either direction; the size and sign of the error depend on the target. The ranking of the four policies did not change with λ: across the full contingency grid at λ = 0, 0.25, 0.5, 0.75 and 1, the pooled reserve gave the fewest late handovers and the end reserve the highest programme on-time probability at every C above zero.

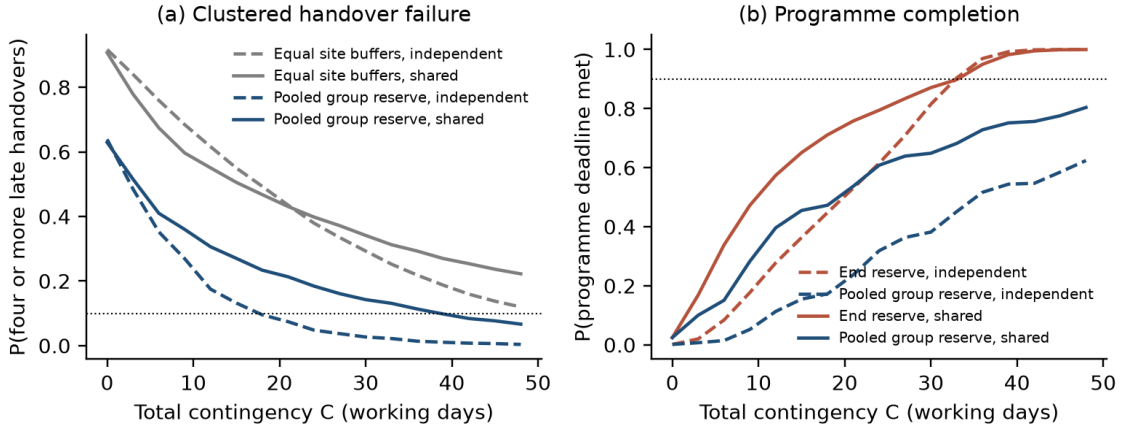


Figure 5. Contingency needed to meet two reliability targets in the SCA-pattern regime under independent and shared supplier delays. (a) Probability of four or more late handovers, with a 10% target shown dotted. (b) Probability that the programme deadline is met, with a 90% target shown dotted. 10,000 runs per point.

D. How the pool should be designed

E4 varied the share of C held in the bundle windows and the size of the groups, with promises always set at group window ends. Comparing the two extremes, holding more of the contingency in the pool helped both outcomes. In the mixed-risk regime, moving from a pool holding none of C to one holding all of it reduced expected late handovers from 1.16 to 0.56 and raised the programme on-time probability from 0.51 to 0.77. In the SCA-pattern regime the same move reduced late handovers from 2.56 to 2.21 and raised the programme on-time probability from 0.33 to 0.45. Intermediate shares did not move smoothly between these end points: splitting 12 days into whole-day buffers and windows put the odd days early in the sequence and pushed later access dates back, so with a quarter or three-quarters of C pooled the programme on-time probability was below that at the neighbouring shares, although late handovers still fell steadily. Groups that match the supply batches worked better than smaller ones: with all contingency pooled, two bundle groups of three gave 0.56 late handovers in the mixed-risk regime against 1.26 for three pairs, and

2.21 against 2.83 in the SCA-pattern regime, although pairs deferred promises less (9.5 days against 12.5). With no pooled share the programme outcome reproduced the equal-buffer results exactly (0.51 and 0.33), which serves as a check on the implementation.

E. One team or two

Giving each bundle its own commissioning team removed the cross-bundle coupling but changed little for the site buffer and pooled policies. In the SCA-pattern regime late handovers fell from 3.71 to 3.51 under equal buffers and from 2.36 to 2.23 under pooling, and the probability of four or more late handovers under pooling fell from 0.31 to 0.26. Shared supplier delays still polarised outcomes, and the independent model still understated the chance of four or more late handovers under pooling (0.14 against 0.26). The second team mainly helped the end reserve, cutting its late handovers from 5.22 to 4.67 in the SCA-pattern regime and from 4.87 to 3.87 in the mixed-risk regime, and its plant hire-days in the SCA-pattern regime from 100.1 to 92.5. The finding that correlation matters is therefore not an artefact of the single-team assumption.

F. Recovery effort, plant hire and travel

Policies differed little in recovery effort. Repeat mobilisations averaged 1.39 to 1.45 per rollout in the mixed-risk regime and 1.01 to 1.07 in the SCA-pattern regime, with the end reserve slightly highest because a tight plan forces more switching between sites. The end reserve kept temporary plant on hire longest, about 7 to 9 hire-days more than equal buffers (107.4 against 98.7 and 100.1 against 92.8), because installation starts are tied to tight access dates that then slip. Extra travel from repeat mobilisations was between 20.7 and 31.1 vehicle-km per rollout, about 4.8 to 7.2 kg CO_2e, and the largest difference between any two policies in the same regime was 1.8 km, about 0.4 kg CO_2e. At this scale, emissions do not discriminate between placements.

The team's standby rule mattered more than placement. Waiting five idle days before demobilising rather than two cut repeat mobilisations under equal buffers by 29% in the mixed-risk regime (1.40 to 0.99) and by 20% in the SCA-pattern regime (1.02 to 0.82), and reduced extra travel under equal buffers in the mixed-risk regime from 29.4 to 21.9 km. Demobilising after any idle day raised repeat mobilisations to about 1.7 in the mixed-risk regime. Site and programme reliability did not change with the standby rule, because the model charges no lead time for remobilisation; this is discussed as a limitation.

G. Sensitivity

The ranking was stable across all twelve sensitivity cases in both regimes (Appendix Table A2). The pooled reserve always gave the fewest late handovers and the end reserve always gave the highest programme on-time probability. In the mixed-risk regime, halving all delays reduced late handovers under pooling to 0.18 and increasing them by half raised the figure to 1.71, while equal buffers moved from 1.30 to 3.51. Allowing the team to start a site as soon as it was ready, rather than waiting for the booked date, reduced late handovers under equal buffers in the mixed-risk regime from 2.40 to 1.87 and cut plant hire-days from 98.6 to 89.4; in the SCA-pattern regime the gain was smaller (3.69 to 3.51) because sites were usually waiting for equipment rather than for access. When two multi-school buildings, X193 and M440, were given double priority weight, risk-based allocation using the same weights reduced weighted late handovers relative to equal buffers from 3.17 to 2.92 in the mixed-risk regime and from 4.94 to 4.75 in the SCA-pattern regime.

VI. DISCUSSION

A. Two outcomes need two placements

The central result is that the best placement depends on which outcome is protected. An end-of-programme reserve is the right answer if the owner is judged only on the programme date, and the wrong one if each building has a date that matters to its users. This is the multi-site version of the distinction between quality robustness and solution robustness drawn by Herroelen and Leus (2004). Critical chain practice, which concentrates protection at the end of a chain (Goldratt, 1997; Leach, 1999), optimises the first and, applied to a rollout, will tend to produce a programme that finishes on time while most schools are handed back late. The SCA's published indicator and its bundle-level scheduling both describe projects and contracts rather than the order in which schools within a contract get their rooms back, so site-level lateness is largely invisible in official reporting.

B. Why pooling within groups works

Pooling works here for a reason that does not depend on independence. In a rollout served by one team, an overrun at one site propagates to the sites after it, so separate site buffers tend to be in the wrong place: the site that runs to plan does not need its buffer, and the site in difficulty and those behind it need more than they have. A group window placed after the sites that share a team and a shipment absorbs the delay wherever it arises within the group. This is close to the argument of Ballard and Howell (1998) that buffers should shield the work flow rather than individual

activities. The idle and blocked days in Table 4 show the same effect: pooled time is idle less often and blocked less often. Groups should follow the structure of shared risk; bundle groups that matched the shipment batches beat smaller groups even though smaller groups defer promises less.

C. What correlation changes

Shared delay did not change which placement is best, which is reassuring for practice. It changed how often the extreme outcomes happen and therefore how much contingency a target requires. Sizing rules that combine site variations by root-sum-square, as in some buffer sizing methods reviewed by Leach (1999), assume independence and will misstate the risk of clustered failure, consistent with the warning by Ökmen and Öztaş (2008) for single projects. In this case the mean was almost unaffected, so a planner who looks only at expected lateness would see no reason to model dependence. The error appears in the tails: under pooling the chance that every school is late was about 25 times higher with shared delays, and the first tested contingency keeping four or more late handovers below 10% rose from 18 to 39 days. A pooled reserve should be sized against the delay that the group's common supplier can cause, not against the spread of independent site delays.

D. The limits of risk-based site buffers

Risk-based buffers gave no consistent benefit. They assess each site on its own, as a risk register does (Schatteman et al., 2008), and so they miss the two dominant sources of lateness in this setting: the cascade through the shared team and the shared supplier. Their effect also depends on where the risky site falls in the sequence. In the mixed-risk regime the riskiest local site was last, so the allocation behaved like a partial end reserve and traded site reliability for programme reliability; in the SCA-pattern regime the protection moved forward and programme reliability fell. Their one clear use is when sites differ in importance, for example a building shared by several schools or one with an examination period. With priority weights the risk-based allocation reduced weighted lateness, so it is better seen as a way of expressing priorities than as a way of handling uncertainty. Combining priority-weighted site buffers with a pooled structure is a natural next step, but it was not tested here.

E. A contingency placement guide

Table 6 condenses the results into guidance. The guide is conditional on the setting studied here, a modest number of sites served by a common team and supply route, and it should be read alongside the limitations in Section 7.

Table 6. Contingency placement guide for multi-site rollouts

Situation	Recommended placement	Evidence in this study	Watch for
Each building has a date its users plan around	Pool contingency in windows after groups of sites that share a team and a supply route; promise group window ends	Fewest late handovers at every level of promise deferral (Figure 3)	Later promises for early sites in each group; agree sectional completion dates
Only the programme completion date is contractually binding	Hold contingency at the end of the programme	Highest programme on-time probability (Table 4)	Most site handovers late; do not publish slot ends as promises
Handover dates already fixed with users	Plan the work tight and hold the slack centrally in front of the fixed dates	Fewest late handovers against fixed dates (Section 5.2)	Plant on hire longer; earlier access needed
Sites share a supplier, shipment or contract	Size the pool to the common delay; test with a correlated model	Independence understated both tails; first tested C for a 10% clustered-failure target rose from 18 to 39 days (Table 5)	Root-sum-square sizing rules; judging risk by the mean alone
Some sites matter more than others	Candidate for testing: priority-weighted buffers at those sites inside a pooled structure	Priority-weighted site buffers lowered weighted lateness against equal buffers (Section 5.7); the pooled hybrid is untested	Effect depends on sequence position
Repeat mobilisation or travel is a concern	Change the standby rule rather than the placement	Five-day standby cut repeat mobilisations by 20% to 29% (Section 5.6)	Idle crew cost; travel differences between placements are small

Situation	Recommended placement	Evidence in this study	Watch for
Sites become ready before booked access dates	Negotiate flexible early access	About a fifth fewer late handovers under site buffers when local risk dominates (Section 5.7)	School disruption from earlier works

Note: Evidence refers to the six-school case and the two risk regimes studied here.

For owners such as the SCA, two practical implications follow. Recording completion at school level within multi-school contracts would make site reliability measurable alongside the project-level indicator. Sectional completion dates for groups of schools, set at the end of a group reserve, would give users dates that are realistic and give the contractor room to absorb shared setbacks without breaking every promise at once.

VII. LIMITATIONS

The study has several limitations. First, the model represents the closing stage of a six-site rollout in a stylised way. Activity durations were scaled from budgets and the activity-level risk parameters were assumed, because the SCA does not publish activity data; only the dependence pattern was calibrated to observed records. The results should be read as comparisons between policies under a realistic structure of shared risk, not as forecasts of SCA delays. Second, the calibration target is the design-phase ICC, because only one IPDVS construction phase had been completed at the snapshot. Design lateness may be shared for reasons, such as common review cycles, that differ from those behind construction lateness. Bundles were also reconstructed from shared dates in the absence of a contract identifier, which may merge or split some contracts, and common recording of dates within a contract could inflate the ICC even though it is robust to the bundle definition and to a measure that does not use the planned end (Section 3.3). For these reasons the SCA-pattern regime is treated as a high-dependence scenario, and the contingency thresholds are reported across a range of dependence levels.

Third, the case is prospective relative to the data, and the two bundles are combined into one rollout under the assumption of a common commissioning team that the records cannot confirm. The model has not been validated out of sample; the recorded completions of the two bundles, planned for March 2023, offer a natural test. Fourth, the dispatch rule is a simple due-date rule with no foresight, and the policies do not re-plan when a supplier delay becomes known. An adaptive controller that reallocates the pool as information arrives could do better, and the comparison here is between static placements. Fifth, remobilisation carries a travel cost but no lead time, so the standby rule affects effort and travel but not reliability; in practice a demobilised team may not return immediately. Sixth, the emission factor is a UK government factor for an average diesel van applied to New York travel, and road distances were approximated by a circuity factor. Both are adequate for comparing policies, which differ by less than 2 km, but not for absolute carbon accounting. Finally, the study does not attach money values to late handovers, idle crew days or plant hire, so it ranks placements by reliability and effort rather than by cost.

VIII. CONCLUSION

This paper asked where a fixed amount of schedule contingency should sit in a multi-site engineering rollout when some delays affect single sites and others affect several sites together. Public SCA records for IPDVS installations in New York City schools, as published in December 2022, show that lateness is strongly shared within contract bundles and across bundles started together, and that none of the construction bundles then due had finished within 15 days of its original date. A simulation of a real six-school rollout, enforcing crew, equipment and access constraints, gave two answers that depend on whether the manager can still set the handover dates. Where each policy sets its own promises, a reserve pooled across groups of sites that share a team and a supply route protects the most handover dates for the same contingency, in the high-dependence scenario cutting expected late handovers from 3.71 to 2.36 of six, although it promises dates 5.5 working days later on average than equal buffers. Where the dates are already agreed with users, planning the work tight and holding the slack centrally protects them best, with 2.82 late handovers against 3.49 for pooled scheduling. A single end reserve also gives the best chance of meeting the programme date. Equal and risk-based site buffers performed between these options and were rarely the best choice for any outcome.

Assuming that supplier delays are independent barely changes the expected number of late handovers but misrepresents both tails: it makes clean rollouts and rollouts in which every school is late look far rarer than they are. Under the pooled reserve, the first tested contingency keeping four or more late handovers below 10% rose from 18 days with independent delays

to 24, 30, 36 and 39 days as dependence increased, while the ranking of policies did not change. Policy choices had little effect on travel and emissions; the standby rule of the team mattered more. The practical message is that programme managers should first establish whether handover dates can still be set or must be protected as agreed, choose the placement that fits that answer, group contingency around the sites that share risk, and size it with a model that lets shared delays happen together. Future work should validate the model against the case bundles once their completions are recorded, extend the analysis to adaptive reallocation of the pool, test priority-weighted buffers within a pooled structure, and test the guide on larger rollouts with several suppliers.

DATA AND CODE AVAILABILITY

The SCA dataset is publicly available from NYC Open Data. The version used in this study, as published on 8 December 2022 (SHA-256 e69fb59655d1a982db7f275fb4990a1c4879d5c8bf2fd4c6d7e0a770392370db), is preserved at https://web.archive.org/web/20221208221940/https://data.cityofnewyork.us/api/views/2xh6-psuq/rows.csv?accessType=DOWNLOAD. The accompanying reproducibility package contains the derived bundle and site files, the Python simulation code, the experiment outputs and the scripts that produce every table and figure. It also contains the Excel workbook described in Appendix B and its CSV files: the 2,000 sets of random draws, the parameters, the site data, the risk-based buffers, the plans, the per-run results and the summary measures for both regimes, together with the scripts and outputs of the additional robustness analyses reported in Sections 3.3 and 5.3. The package is provided as supplementary material with this submission and will be deposited in a public repository on acceptance.

DECLARATIONS

Funding: This research received no specific grant from any funding agency in the public, commercial or not-for-profit sectors.

Competing interests: The author declares no competing interests.

Author contribution: Prince Keshi is the sole author and carried out the conceptualisation, data curation, modelling, software, formal analysis, visualisation and writing.

Use of data: The study uses only publicly available administrative data and involved no human participants, so ethical approval was not required.

REFERENCES

[1] Ballard, G. and Howell, G. (1998) Shielding production: essential step in production control. Journal of Construction Engineering and Management, 124(1), 11-17. https://doi.org/10.1061/(ASCE)0733-9364(1998)124:1(11)

[2] Ballou, R.H., Rahardja, H. and Sakai, N. (2002) Selected country circuity factors for road travel distance estimation. Transportation Research Part A: Policy and Practice, 36(9), 843-848. https://doi.org/10.1016/S0965-8564(01)00044-1

[3] Clemen, R.T. and Reilly, T. (1999) Correlations and copulas for decision and risk analysis. Management Science, 45(2), 208-224. https://doi.org/10.1287/mnsc.45.2.208

[4] Department for Energy Security and Net Zero (2023) Greenhouse gas reporting: conversion factors 2023, flat file (updated 28 June 2023). London: DESNZ. https://www.gov.uk/government/publications/greenhouse-gas-reporting-conversion-factors-2023

[5] Goldratt, E.M. (1997) Critical Chain. Great Barrington, MA: North River Press.

[6] Grimm, V., Revilla, E., Berger, U., Jeltsch, F., Mooij, W.M., Railsback, S.F., Thulke, H.H., Weiner, J., Wiegand, T. and DeAngelis, D.L. (2005) Pattern-oriented modeling of agent-based complex systems: lessons from ecology. Science, 310(5750), 987-991. https://doi.org/10.1126/science.1116681

[7] Harris, R.B. and Ioannou, P.G. (1998) Scheduling projects with repeating activities. Journal of Construction Engineering and Management, 124(4), 269-278. https://doi.org/10.1061/(ASCE)0733-9364(1998)124:4(269)

[8] Herroelen, W. and Leus, R. (2004) Robust and reactive project scheduling: a review and classification of procedures. International Journal of Production Research, 42(8), 1599-1620. https://doi.org/10.1080/00207540310001638055

[9] Hopp, W.J. and Spearman, M.L. (2008) Factory Physics. 3rd edn. New York: McGraw-Hill.

[10] Lambrechts, O., Demeulemeester, E. and Herroelen, W. (2008) Proactive and reactive strategies for resource-constrained project scheduling with uncertain resource availabilities. Journal of Scheduling, 11(2), 121-136. https://doi.org/10.1007/s10951-007-0021-0

[11] Law, A.M. (2015) Simulation Modeling and Analysis. 5th edn. New York: McGraw-Hill.

[12] Leach, L.P. (1999) Critical chain project management improves project performance. Project Management Journal, 30(2), 39-51. https://doi.org/10.1177/875697289903000207

[13] McGraw, K.O. and Wong, S.P. (1996) Forming inferences about some intraclass correlation coefficients. Psychological Methods, 1(1), 30-46. https://doi.org/10.1037/1082-989X.1.1.30

[14] New York City School Construction Authority (2022) Capital Project Schedules and Budgets [dataset]. NYC Open Data, version published 8 December 2022. https://data.cityofnewyork.us/Housing-Development/Capital-Project-Schedules-and-Budgets/2xh6-psuq

[15] NYC Department of Education (2020) 2019-2020 School Point Locations [dataset]. NYC Open Data. https://data.cityofnewyork.us/Education/2019-2020-School-Point-Locations/a3nt-yts4

[16] NYC Mayor's Office of Operations (2022) Mayor's Management Report, Fiscal 2022: School Construction Authority. New York: City of New York. https://www.nyc.gov/assets/operations/downloads/pdf/mmr2022/sca.pdf

[17] NYC Mayor's Office of Operations (2023a) Preliminary Mayor's Management Report, Fiscal 2023: School Construction Authority. New York: City of New York. https://www.nyc.gov/assets/operations/downloads/pdf/pmmr2023/sca.pdf

[18] NYC Mayor's Office of Operations (2023b) Preliminary Mayor's Management Report, Fiscal 2023: School Construction Authority indicator definitions. New York: City of New York. https://www.nyc.gov/assets/operations/downloads/pdf/pmmr2023/sca_idf.pdf

[19] Ökmen, Ö. and Öztaş, A. (2008) Construction project network evaluation with correlated schedule risk analysis model. Journal of Construction Engineering and Management, 134(1), 49-63. https://doi.org/10.1061/(ASCE)0733-9364(2008)134:1(49)

[20] Schatteman, D., Herroelen, W., Van de Vonder, S. and Boone, A. (2008) Methodology for integrated risk management and proactive scheduling of construction projects. Journal of Construction Engineering and Management, 134(11), 885-893. https://doi.org/10.1061/(ASCE)0733-9364(2008)134:11(885)

[21] Trietsch, D., Mazmanyan, L., Gevorgyan, L. and Baker, K.R. (2012) Modeling activity times by the Parkinson distribution with a lognormal core: theory and validation. European Journal of Operational Research, 216(2), 386-396. https://doi.org/10.1016/j.ejor.2011.07.054

[22] Tukel, O.I., Rom, W.O. and Eksioglu, S.D. (2006) An investigation of buffer sizing techniques in critical chain scheduling. European Journal of Operational Research, 172(2), 401-416. https://doi.org/10.1016/j.ejor.2004.10.019

[23] Van de Vonder, S., Demeulemeester, E., Herroelen, W. and Leus, R. (2005) The use of buffers in project management: the trade-off between stability and makespan. International Journal of Production Economics, 97(2), 227-240. https://doi.org/10.1016/j.ijpe.2004.08.004

APPENDIX A. MODEL PARAMETERS, CONVERGENCE AND SENSITIVITY

Table A1. Full parameter list

Parameter	Value	Status
Working-day calendar; first commissioning slot	Day 25	Assumed
Planned commissioning days, K_i	max(2, round($3f_i$)) = 4, 4, 4, 2, 3, 3	Scaled from SCA budgets
Planned installation days, I_i	max(4, round($8f_i$)) = 9, 11, 10, 4, 7, 7	Scaled from SCA budgets

© AUG 2023 | IRE Journals | Volume 7 Issue 2 | ISSN: 2456-8880
DOI: https://doi.org/10.64388/ire1725645

Parameter	Value	Status
Delivery lead before installation; installation lag before commissioning	5 days; 3 days	Assumed
Supplier delay model	Hurdle lognormal, cap 40 days, Gaussian copula	Assumed form
Installation duration multiplier	Lognormal, log-mean 0.05, log-sd 0.15	Assumed
Installation access loss	Probability 0.20, 1 to 3 days	Assumed
Commissioning duration variation	Lognormal, log-sd 0.20	Assumed
Commissioning access loss length	1 to 3 days (normal), 2 to 5 days (high)	Assumed
Test failure probability	$0.12 + 0.12(f_i - 0.5)$	Assumed

Parameter	Value	Status
Rectification and retest	2 to 6 days plus 1 day	Assumed
Team outage	Probability 0.20, 2 to 5 days	Assumed
Standby before demobilising	2 idle days	Assumed
Road distances	Straight line × 1.2: 4.5 to 14.0 km	Derived from DOE locations
Emission factor	0.23128 kg CO_2e per km	DESNZ (2023)
Risk-based allocation at C = 12	Mixed: 1, 3, 1, 1, 1, 5; SCA-pattern: 2, 4, 3, 1, 1, 1 days	Model output
Runs per configuration	20,000 (E1, E2, E4, E6); 10,000 (E3, E5)	Convergence checked

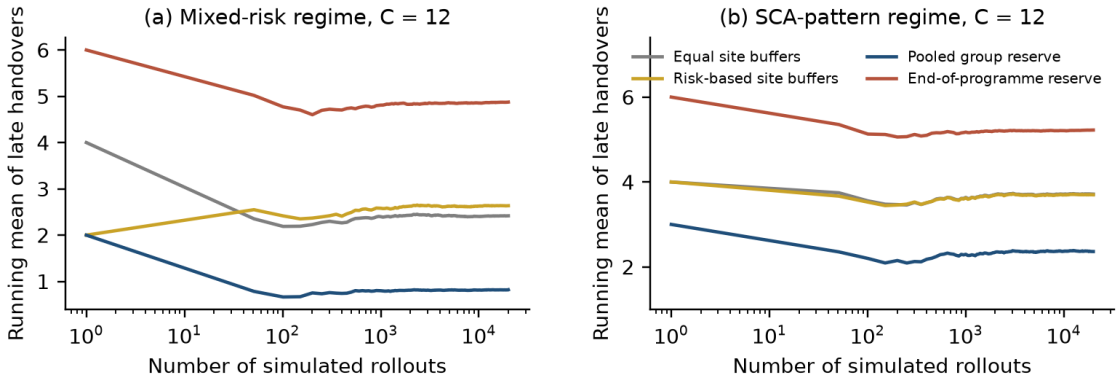


Figure A1. Running mean of late handovers by number of simulated rollouts at C = 12. Estimates are stable well before the 20,000 runs used.

Table A2. Sensitivity of expected late handovers and programme on-time probability, SCA-pattern regime, C = 12

Case	Late: equal	Late: risk-based	Late: pooled	Late: end	Programme: pooled	Programme: end
Base case	3.69	3.68	2.34	5.22	0.40	0.57
All delays × 0.5	2.47	2.48	1.05	4.53	0.70	0.83
All delays × 1.5	4.41	4.44	3.23	5.45	0.23	0.38
Supplier probability × 0.5	2.62	2.59	1.38	4.83	0.60	0.76
Supplier probability × 1.5	4.73	4.72	3.30	5.61	0.22	0.40
No team outage	3.60	3.59	2.28	5.18	0.41	0.59
Team outage probability 0.4	3.78	3.77	2.40	5.26	0.38	0.56
Standby 0 days	3.69	3.68	2.34	5.22	0.40	0.57
Standby 5 days	3.69	3.68	2.34	5.22	0.40	0.57
Uniform local risk	3.66	3.65	2.32	5.20	0.41	0.57
Early access allowed	3.51	3.49	2.28	4.73	0.42	0.58

Case	Late: equal	Late: risk-based	Late: pooled	Late: end	Programme: pooled	Programme: end
Priority weights 2 for X193 and M440	4.94	4.75	3.26	7.02	0.40	0.57

Note: 10,000 runs per case with common random numbers, so base values differ slightly from Table 4. With priority weights the late count is weighted. The mixed-risk regime gave the same ranking in every case.

APPENDIX B. SPREADSHEET IMPLEMENTATION

A spreadsheet version of the model is provided so that readers can follow each calculation and rerun the main comparison in Microsoft Excel without programming. It holds 2,000 simulated rollouts drawn with a fixed seed, stored as values so that every policy faces the same draws. The Inputs sheet holds the regime, the total contingency C, the pooled share and every parameter in Table A1, with an optional override of λ for testing the independence assumption. The Plans sheet computes buffers, booked commissioning dates and promised handover dates for the four policies, and the Events sheet converts the draws into supplier delays through the same Gaussian copula, together with installation spells, access losses, test failures, rectification and team outage. The four simulation sheets hold one row per rollout, and the Summary sheet reports the outcome measures of Section 4.7 with the distribution of late handovers. All results are live formulas, so a change to any input recalculates the whole sample.

The daily dispatch loop of Section 4.6 cannot be written cleanly in worksheet formulas, so the spreadsheet uses an event-step form of the same rule. The team serves the schools one at a time; at each step it moves to the earliest time at which some unfinished school is workable and, among the schools workable then, chooses the one with the earliest promised date. An access loss that starts during a commissioning spell extends it, a team outage that falls inside the spell extends it by its length, and a failed acceptance test adds rectification and a one-day retest after the team has moved on. The random draws, inputs and per-run outputs are exported as CSV files. Table B1 compares the two implementations on the same 2,000 draws.

Table B1. Spreadsheet (event-step) and full (daily dispatch) implementations on identical draws, C = 12

Regime	Policy	Late, S	Late, F	P(4+ late), S	P(4+ late), F	Deadline, S	Deadline, F
Mixed-risk, λ = 0.9	Equal	2.51	2.47	0.25	0.24	0.52	0.51
Mixed-risk, λ = 0.9	Risk-based	2.66	2.65	0.33	0.32	0.72	0.71
Mixed-risk, λ = 0.9	Pooled	0.85	0.85	0.02	0.02	0.66	0.65
Mixed-risk, λ = 0.9	End	4.97	4.87	0.86	0.86	0.94	0.94
Mixed-risk, λ = 0.0	Equal	2.50	2.47	0.24	0.23	0.51	0.49
Mixed-risk, λ = 0.0	Risk-based	2.67	2.66	0.31	0.31	0.70	0.69
Mixed-risk, λ = 0.0	Pooled	0.84	0.84	0.00	0.01	0.64	0.63
Mixed-risk, λ = 0.0	End	4.97	4.87	0.86	0.87	0.92	0.93
SCA-pattern, λ = 1.0	Equal	3.75	3.76	0.57	0.57	0.33	0.32
SCA-pattern, λ = 1.0	Risk-based	3.72	3.75	0.56	0.56	0.25	0.23
SCA-pattern, λ = 1.0	Pooled	2.39	2.44	0.31	0.32	0.40	0.38
SCA-pattern, λ = 1.0	End	5.16	5.20	0.88	0.90	0.57	0.56
SCA-pattern, λ = 0.0	Equal	3.78	3.80	0.60	0.61	0.09	0.09
SCA-pattern, λ = 0.0	Risk-based	3.76	3.78	0.59	0.60	0.05	0.05
SCA-pattern, λ = 0.0	Pooled	2.32	2.35	0.16	0.17	0.13	0.12
SCA-pattern, λ = 0.0	End	4.96	4.99	0.89	0.90	0.29	0.28

Note: S = spreadsheet (event-step) implementation; F = full (daily dispatch) implementation. 2,000 runs with common random numbers, so values differ slightly from Table 4, which uses 20,000 runs. Late is the expected number of late handovers; Deadline is the probability that the programme deadline is met.

The two implementations agree to within 0.1 late handovers and 0.02 in every probability, and they give the same ranking of policies and the same effect of independence under pooling. The spreadsheet slightly overstates repeat mobilisations in the SCA-pattern regime (about 1.10 per rollout against 1.00) and differs from the full model by up to about 1.7 plant hire-days, so recovery effort, hire-days, travel and emissions in Section 5.6 are taken from the full model. All results in the main text come from the Python implementation.