

A Readiness-Gated Management System for Manufacturing Ramp-Up

Nikhil Sharma

nikhilsharma8877@gmail.com

Research Article

Keywords: manufacturing ramp-up, production management, production stabilization, readiness gate, Monte Carlo simulation, digital manufacturing, quality management

Posted Date: August 27th, 2026

DOI: <https://doi.org/10.21203/rs.3.rs-10720064/v1>

License:  This work is licensed under a Creative Commons Attribution 4.0 International License.
[Read Full License](#)

Additional Declarations: No competing interests reported.

Abstract

Manufacturing ramp-up requires teams to increase output while processes, equipment, workforce capability, and production data are still stabilizing. Fixed calendar milestones can allow permanent balancing, automation, and process changes before normal production capability is understood. Temporary labor, inspection, or engineering support can further hide whether the system is truly ready. This paper develops a readiness-gated management system for deciding when a high-volume assembly area should move from stabilization to broader improvement.

The framework uses two phases. Phase I controls critical risks, stabilizes operations, builds workforce capability, improves data reliability, and makes temporary support visible. Phase II uses this stronger baseline to remove verified losses, improve standard work and process capability, and allow statistical and digital tools to guide more important decisions. Readiness is reviewed by a cross-functional team across six areas: safety and quality containment, throughput repeatability, equipment reliability, workforce readiness, data reliability, and change-control stability. Critical safety, quality, regulatory, or data problems can block transition. Based on the remaining evidence, the team may authorize, limit, defer, or reverse wider process and digital changes. Planned milestones therefore become evidence-based reviews rather than automatic phase changes.

A simulation-based stress test compared fixed-time and readiness-based decisions across 25,000 synthetic launches. For the same launches, the baseline gate lowered the average modeled chance of performance declining after transition from 18.2% under a fixed mid-ramp transition to 15.7%. It achieved this improvement by adjusting transition timing to the condition of each launch. A later fixed transition lowered the modeled risk further, to 14.4%, but delayed every launch, including those ready earlier. Stricter gates also lowered risk but increased the number of launches waiting for approval. Additional tests showed that the gate added value when readiness helped predict later performance and offered no advantage when readiness provided no useful information. The simulation therefore shows how the framework can balance the risk of moving too early against the cost of waiting too long.

The main contribution is a practical management system that defines what evidence should be reviewed, who should participate, and how improvement authority should increase as production becomes more stable and understandable. It can complement APQP, PPAP, Control Plans, and safe-launch practices by connecting their evidence to a clear operating decision. Field studies are required to establish suitable measures, thresholds, and production benefits.

1. Introduction

Manufacturing ramp-up links product and process development with stable full-scale production. Teams must increase output while they continue to adjust work content, equipment settings, material flow, and workforce routines. As a result, the launch may face technical and organizational problems at the same time [1]-[6].

Performance pressure begins immediately. Planned volume, takt time, quality, and cost targets are expected to be met quickly. However, the production baseline may still be changing. Temporary workarounds can protect output during this period. Teams may repeatedly adjust equipment or add inspectors, technicians, and engineers. These actions may be necessary, but they can hide what the process can achieve under normal conditions.

Lean methods are important from the first production build. Clear standards, visual controls, quality at the source, and frontline problem-solving practices should be established early [7]-[10]. Problems arise when teams use unstable loss patterns to make permanent balancing, automation, or waste-reduction decisions. Temporary launch conditions may then become part of the long-term production system. Teams should therefore match the timing and use of these methods to the condition of the launch [11].

Digital technologies create a similar challenge. Decision-support systems, virtual commissioning, digital twins, and analytics can provide faster feedback during ramp-up [12]-[16]. However, these tools need clear event definitions, matching system configurations, and reliable data. Research shows that technology and operating practices should be developed together [17]-[22]. Installing technology alone does not improve production decisions.

The key management question is not whether teams should use Lean or digital tools early. They should. The real question is how much decision-making power these tools should have at each stage of ramp-up. A dashboard can show production conditions before managers use it to set automatic priorities. A digital twin can help test possible causes before it is allowed to guide process settings. Early standard work can be updated frequently before teams use it for detailed line balancing.

This paper addresses that decision. A readiness-gated management system is proposed for high-volume assembly. The system uses two operating phases. A cross-functional evidence review connects them. Transition is approved when the production system, workforce, and information are ready. Time alone does not determine the decision.

The paper provides both a practical framework and an analytical test. The framework explains what teams should stabilize, what evidence they should review, who should participate, and when broader improvement decisions should be allowed. No company production data are used, so the framework remains open to further testing. A synthetic simulation tests how the decision rules behave under different conditions. It examines the balance between moving too early and waiting too long, as well as the conditions under which the approach may lose value. The simulation does not claim to prove results for a real factory or predict performance at a specific manufacturing site.

2. Conceptual Framework Development

This study used a concept-based approach to build the framework. It did not claim to conduct a complete systematic review. Instead, it combined several areas of research around one management question: when should a launch move from stabilization to intensive improvement?

The research base covered production ramp-up, learning curves, Lean production, statistical quality, digital manufacturing, production readiness, and launch management. Older sources were kept when they introduced ideas that remain useful over time. Recent studies were added to cover Industry 4.0, digital twins, and current ramp-up needs.

The framework was developed in three steps. First, recurring launch problems were identified. These included changing work content, incomplete learning, repeated equipment failures, gaps in workforce skills, unreliable data, and engineering changes that affected one another. Second, the required management responses were grouped by their main purpose. These purposes included containment, visibility, learning, standardization, improvement, and long-term control. Third, each response was placed either before or after the readiness review. This placement depended on how stable and well-understood the production system needed to be before the response could be used reliably.

The author developed the six readiness dimensions, critical veto conditions, and transition rules. Existing research helped guide their design. However, they have not yet been validated and are not presented as standards that apply to every production system. Table 1 shows how the recurring ramp-up problems relate to the proposed framework responses.

A readiness dimension was included when it served one of two purposes. It either captured a separate source of uncertainty that could make production results difficult to understand, or it represented a capability needed for a reliable transition decision. The six dimensions are therefore a starting point rather than a complete list. Supplier and material readiness, ergonomics, logistics, and process capability were not created as separate gate dimensions. Instead, the framework treats them as supporting conditions or evidence within one or more of the six dimensions. This arrangement reflects a design choice by the author. Independent experts must still assess whether the dimensions are complete and whether their boundaries are appropriate.

Table 1

Traceability from recurring ramp-up problem classes to readiness dimensions, cross-cutting framework responses and principal supporting literature.

Observed ramp-up problem	Management requirement	Framework response	Principal supporting literature
Critical safety, regulatory, or critical-quality exposure during early production	Temporary protection must be noncompensatory, owned, independently verified, and closed only against explicit exit evidence.	Safety and quality containment; critical-veto logic	[36], [38], [39], [51]
Unstable work content and mixed disturbances	A provisional baseline is needed. Deviations must remain visible.	Minimum viable standardization; throughput repeatability	[1], [2], [4], [6]
Incomplete learning and weak problem reporting	Short feedback cycles and protected escalation are required.	Rapid feedback; workforce readiness	[2], [24], [25], [26]
Recurring equipment faults and variable recovery	Failure patterns and recovery methods must be made repeatable.	Equipment reliability; Phase II loss removal	[1], [4], [6], [49]
Unreliable events, labels or denominators	Measurement and classification integrity must be established.	Data reliability; statistical-quality role	[12], [15], [37], [47], [56]
Interacting engineering changes	Configuration traceability and change sequencing are required.	Controlled change; change-control stability	[1], [4], [6], [38]
Digital model mismatch or configuration drift	Digital decision authority must remain proportional to data integrity, configuration synchronization, and demonstrated model-to-plant agreement.	Cross-cutting application of data reliability and change-control stability; phase-aligned digital authority	[27], [29], [30], [33]

Digital model mismatch is not treated as a seventh readiness dimension. It is governed principally through data reliability and change-control stability. The permitted level of digital decision authority is adjusted according to the credibility of the model and its inputs.

2.1 Theoretical Mechanism and Contingency Logic

The proposed readiness gate is mainly explained by organizational information-processing theory [52]–[54]. This theory states that an organization must match its need for information with its ability to collect, combine, and understand that information. During ramp-up, both uncertainty and ambiguity increase. Uncertainty means that teams do not have enough knowledge about the process or the likely results of a decision. Ambiguity means that the same production loss may have several reasonable explanations [54].

Before teams make permanent improvements, they need a production baseline that is stable enough to understand. Recurring losses must be separated from one-time launch problems. When the amount or complexity of evidence exceeds the team's ability to evaluate it, poor decisions can follow. Process settings, work standards, or digital recommendations may then be based on temporary or incorrectly classified conditions.

Phase I helps restore the balance between the information needed and the team's ability to process it. Controlled changes and temporary standards reduce the number of possible explanations that teams must examine at the same time. Measurement checks, clear event definitions, product and process version history, and correct time order improve the quality of the evidence. Fast feedback and cross-functional review help teams combine information from different departments. Trained employees can also identify problems and respond correctly.

Each readiness dimension supports this process in a different way. Safety and quality containment prevent unacceptable results. Throughput and equipment data show which operating patterns continue to occur. Workforce readiness helps employees detect, report, and respond to problems across the production system. Reliable data allow teams to classify events and connect results with possible causes. Stable change control helps teams understand cause-and-effect relationships.

The readiness review does not require all uncertainty to be removed. Instead, it asks whether the team can collect and understand enough evidence to manage what remains unknown. The amount and type of evidence depend on product complexity, automation, regulation, product mix, process-transfer maturity, and the number of ongoing engineering changes [53]. More difficult launches may require longer periods of normal operation, stronger tracking of product and process versions, and more independent review. The immediate goal is to improve decision quality and reduce the chance that performance will fall after transition. It is not simply to maximize output as quickly as possible. Organizational learning is strengthened when production experience is turned into knowledge that teams can keep and use [55].

3. Readiness-Gated Management System

The proposed management system is shown in Fig. 1. Phase I is focused on stabilization and observability. The readiness review is used to assess whether a reliable improvement baseline has been created. Phase II is focused on improvement, maturation and sustainment. The phases can overlap by area. The distinction concerns the dominant management logic and the authority assigned to permanent optimization.

3.1 Key Definitions and Level of Application

Operational readiness. The demonstrated ability of a production area to generate safe, repeatable and interpretable performance under representative conditions.

Representative operation. Operation with the planned product mix, staffing pattern, shift conditions and normal support structure. Exceptional launch support is recorded separately.

Extraordinary support. Temporary labor, inspection, engineering attention, bypass activity or manual intervention that would not be part of the intended steady-state system.

Observability. The ability to detect, classify and trace important production events. Reliable configurations, timestamps, denominators and ownership are included.

Optimization authority. The degree to which an improvement method, analytical model or digital recommendation is allowed to change standards, resources, parameters or production decisions.

The framework can be applied to a line, production area or process family. A plant-wide application may be too broad. A station-level application may be too narrow. The unit should be chosen before indicators are defined. Connected upstream and downstream effects must then be reviewed.

The six dimensions are treated as determining components of operational readiness. They jointly constitute the readiness judgment. They are not assumed to be interchangeable measures of one hidden characteristic. Safety, regulatory, critical-quality and material data-integrity conditions are assessed as non-compensatory vetoes. The remaining evidence is interpreted together through recorded cross-functional judgment.

The dimension boundaries are defined as follows. Throughput repeatability concerns recurring output behavior under representative demand, staffing and product mix. Equipment reliability concerns asset failure and recovery behavior. Data reliability concerns whether event records represent the physical process accurately. Change-control stability concerns whether the current physical and documented configuration remains interpretable. Workforce readiness concerns independent execution and abnormality response, rather than training attendance. Safety and quality containment concerns unacceptable consequence conditions and the effectiveness of temporary controls.

Where one event affects several dimensions, evidence should be assigned according to its primary causal domain and cross-referenced as a consequence elsewhere. For example, an equipment failure belongs primarily to equipment reliability. Its output effect may also be reflected under throughput repeatability, but it should not be counted again as an independent cause.

3.2 Phase I: Stabilization and Observability

Critical-risk containment. Safety, regulatory and critical-quality conditions are treated as non-negotiable. Additional inspection or temporary verification may be used. Each temporary control is assigned an owner, a review date and an exit condition. Otherwise, temporary containment can become permanent waste.

Minimum viable standardization. The safest known sequence is documented. Critical parameters, material identity, quality points and escalation triggers are included. The standard is controlled, but rapid

revision is permitted. Detailed balancing is delayed when work content is still changing.

Rapid feedback and problem visibility. Abnormalities are reviewed through short cycles. Defect and downtime categories are defined. Ownership is assigned. Hidden support is separated from normal operation. Problem reporting is protected because learning is reduced when employees fear blame [23]-[26].

Workforce qualification. Training attendance is not treated as competence. Independent execution, abnormality recognition and correct escalation are demonstrated under representative conditions. Qualification is made visible by operation and shift.

Controlled change. Engineering changes are prioritized through a cross-functional process. Affected standards, training, tooling and digital configurations are identified. Validation evidence is defined before release. Several interacting changes are avoided when causal learning would be lost.

Digital observability. Digital systems are used to strengthen traceability, genealogy, parameter history, event coding and revision control. Advanced analytics may be tested. Decision authority remains limited by data completeness and configuration validity.

3.3 Statistical Quality Under Unstable Conditions

Statistics still play an important role in Phase I. During this phase, teams create the conditions needed for reliable analysis. They define each measure clearly and standardize defect and downtime categories. Product and process configurations must also be preserved in the data record.

Teams should match their measurement checks to the type of evidence and the risk of making a wrong decision. Numerical measurements may require calibration, resolution checks, and repeatability and reproducibility studies. The review should also cover discrimination, bias, linearity, and stability. These factors must be considered in relation to process variation, tolerance, and the intended decision [56].

Attribute data, such as pass/fail results or defect categories, also require clear rules. Within-appraiser and between-appraiser agreement should be evaluated. Analysts can report percent agreement together with a measure that accounts for agreement by chance. Cohen's kappa may be used for two raters, while Fleiss' kappa may be used for several raters [58], [59]. Analysts should also report whether some categories appear much more often than others. Category imbalance can affect agreement measures. The framework does not set one acceptance limit for every situation. Teams should choose the study design and acceptance criteria based on the harm caused by an incorrect classification. They should also control any rules used to reclassify data.

Before comparing rates, teams must confirm that the denominators are consistent. Time order should be preserved. Records should allow users to trace the product variant, process revision, shift, operating condition, and use of extra support. Teams should not assume that data collected under temporary containment represent normal production.

Teams should select control charts based on the data type, subgroup structure, denominator behavior, event frequency, product mix, and dependence over time [37], [57]. X-bar and R or X-bar and S charts should be used only when teams can form rational subgroups. Individuals and moving-range charts may be suitable for separate continuous measurements. The choice among p, np, c, and u charts depends on sample size and opportunity counts. Some data require other methods. Examples include rare events, greater-than-expected variation, measurements related over time, and mixed-product data. These cases may require specialized charts, model adjustments, or separate analysis by product or operating condition. Teams should use defined rules to identify runs, shifts, trends, cycles, and other nonrandom patterns. Visual judgment alone is not enough.

Statistical stability and meeting specifications are not the same. A stable process may still fail to meet requirements. A process may also meet requirements for a short time while remaining unstable. Capability indices should therefore be interpreted only after the process and measurement system become stable enough for the intended decision [36], [37].

Teams should investigate known launch disturbances as possible special causes. They should not simply include these disturbances in a long-term average. Designed experiments can still support ramp-up decisions. However, teams must consider process changes, time trends, and control runs when the baseline continues to change [37].

Control limits in Phase I should be treated as temporary. Teams should investigate a launch event as a possible special cause. However, the event should not be classified as a special cause only because it occurred during ramp-up [37]. Before removing observations or changing control limits, teams should connect the event to a recorded change, disturbance, configuration, or operating condition. A confirmed change in the process or measurement system may create a new baseline. In that case, analysts should restart the analysis window or divide the data into separate periods. They should retain product mix, shift, configuration, containment status, and extra support as factors for separate analysis. Teams should not create the appearance of stability by combining data from different conditions.

Phase II gives statistical improvement methods greater decision-making authority. Teams can apply DMAIC, statistical process control, capability analysis, and designed experiments to recurring losses that have been confirmed. These methods are not delayed because they lack value. Their findings become more useful when the process and measurement conditions support reliable conclusions.

The framework does not set one MSA design, control chart, capability limit, or required number of observations for every situation. Teams should make these choices based on the measured characteristic, sampling method, production setting, decision risk, and consequences of error.

3.4 Cross-Functional Readiness Review

The readiness review asks one practical question: is the system sufficiently stable and observable for intensive improvement to produce durable results? A formula is not required. Evidence is reviewed by dimension. Critical vetoes are considered first.

Safety, regulatory, critical-quality and data-integrity conditions are non-compensatory. A strong throughput result cannot offset an unresolved critical defect. High equipment availability cannot compensate for unreliable product genealogy. When a veto remains open, the transition is not authorized.

After vetoes are cleared, the remaining evidence is reviewed across representative operating periods. A scorecard may be used. It is optional. The final decision remains cross-functional and recorded. Missing or disputed evidence is not estimated merely to complete a dashboard.

Table 2

Operational definitions, illustrative evidence and accountable functions for the six readiness dimensions.

Dimension	Illustrative indicators	Evidence expected before transition	Primary owner
Safety and quality containment	Critical incidents; severe defect recurrence; containment escape; special-characteristic control	No open critical condition; temporary controls have owners and exit criteria	Quality
Throughput repeatability	Output attainment; schedule adherence; cycle-time pattern; extraordinary support	Representative shifts operate within the agreed range without hidden support	Production / Industrial Engineering
Equipment reliability	Recurring failure frequency; restoration method; availability pattern	Dominant failures are understood; recovery is repeatable; ownership is assigned	Maintenance / Engineering
Workforce readiness	Independent qualification; abnormality recognition; staffing coverage	Required operations are independently covered across relevant shifts	Production / Training
Data reliability	Event completeness; code consistency; configuration match; missing records	Loss and defect data support credible prioritization and traceability	Quality / IT / Production
Change-control stability	Open changes; validation status; documentation synchronization; configuration mismatch	Change load can be absorbed without losing cause-and-effect visibility	Engineering / Program Management

3.5 Phase II: Improvement and Maturation

Verified loss removal. Improvement resources are directed toward recurring losses that have been confirmed by evidence. Pareto analysis and structured root-cause methods are used. The most visible problem of the day is not automatically treated as the largest system loss.

Refined standard work and balancing. Provisional standards are converted into validated methods. Timing, ergonomic needs, quality points and training content are integrated. Balancing is tightened as cycle-time distributions become more stable.

Statistical improvement and control. DMAIC, SPC, capability analysis and designed experiments are applied to recurring losses. Control plans are updated. Regression signals are defined. Improvement is closed only after the new condition is sustained.

Robust error-proofing and automation. Poka-yoke and automation are designed around verified failure mechanisms. Detection, response, override and recovery functions are validated. Automation is managed as a production-system change, not as an isolated installation.

Validated digital decision support. Digital twins and analytical models may be used for scenario evaluation, scheduling, maintenance and parameter decisions. Verification and validation continue after transition [27]-[35]. Passage through the gate does not make a model permanently valid.

Sustainment. Temporary controls are removed after evidence is recorded. Standards, training and control plans are synchronized. Ownership is transferred into daily management. Performance regression is monitored.

Table 3
Phase-specific management logic.

Design element	Phase I: stabilization and observability	Phase II: improvement and maturation
Primary question	Can the system operate safely, repeatably and visibly?	Which verified losses should be removed or optimized?
Management emphasis	Containment, qualification, feedback, repeatability and data integrity	Capability, productivity, robustness, cost and sustainment
Standardization	Provisional, risk-focused and rapidly revisable	Validated, trained, balanced and scalable
Lean role	Visual abnormality control, basic problem solving and minimum viable standard work	Loss-prioritized Kaizen, refined standard work, flow and balancing
Statistical role	Operational definitions, measurement adequacy and evidence of stability	DMAIC, SPC, capability analysis, DOE and sustained control
Digital role	Capture, traceability, revision control and human-reviewed diagnosis	Prediction, simulation, optimization and controlled automation
Change governance	Limit interacting changes and protect causal learning	Execute a prioritized roadmap and institutionalize gains

3.6 Phase-Aligned Digital Support

The same technology may serve different purposes as maturity changes. A manufacturing execution system may first be used to establish genealogy. It may later support statistical monitoring. A digital twin may first be used for virtual commissioning or hypothesis generation. Optimization authority can be

increased after model-to-plant agreement has been demonstrated. This progression is consistent with the reported need for verification, validation and governance in digital manufacturing [27]-[35].

Table 4
Phase-aligned use of digital production capabilities.

Capability	Phase I use	Main readiness risk	Phase II use
MES and genealogy	Capture execution, configuration, operator and defect records	Missing events; inconsistent codes; late corrections	Trend analysis, capability evidence and control-plan support
Digital work instructions	Controlled revision and point-of-use access	Version mismatch; temporary methods outside control	Stable standard work, training integration and change traceability
Andon and escalation systems	Short feedback loops and response timestamps	Alarm inflation; unclear closure categories	Recurrence analysis and response optimization
Digital twin / simulation	Virtual commissioning, sensitivity analysis and hypothesis generation	Model fidelity and synchronization uncertainty	Validated scheduling, capacity and parameter decisions
Machine learning / AI	Anomaly screening with human review	Sparse labels; drift; unstable process states	Predictive quality, maintenance and prescriptive support
Dashboards and analytics	Visibility of missing data, disturbances and learning progress	False precision; unstable denominator	Loss prioritization, control and sustainment

4. Practical Implementation Protocol

4. Practical Implementation Protocol

Implementation should begin before start of production. The gate is designed with the launch plan. It is not added after problems appear. The unit of application, review frequency, veto conditions and evidence owners are defined first.

4.1 Governance Roles and Review Cycle

Table 5

Practical governance matrix for the readiness review. A standard decision record should be completed at each formal readiness review. The blank structure in Table 6 preserves the evidence basis, dissent, conditions and limits of the authorization. It does not prescribe universal numerical thresholds.

Function	Evidence owned	Typical review frequency	Decision responsibility
Production	Output, staffing, representative-shift performance, support hours	Daily review; weekly gate summary	Confirm normal operating performance and unresolved execution barriers
Quality	Critical conditions, escape risk, containment, defect recurrence	Each shift and daily	Maintain veto status; verify closure evidence
Maintenance	Recurring failures, recovery method, temporary technical support	Daily and weekly	Confirm equipment repeatability and restoration standards
Engineering	Open changes, validation, process configuration and work-content changes	Weekly or after major change	Protect configuration integrity and causal visibility
Industrial Engineering	Cycle-time behavior, workload, flow and balancing assumptions	Weekly	Determine whether timing data can support permanent balancing
Training / Supervision	Independent qualification and abnormality recognition	Weekly and after standard revision	Confirm workforce coverage by operation and shift
IT / Data owners	Event completeness, code control, interfaces and synchronization	Daily quality check; weekly review	Confirm that information is current and representative
Launch leadership	Conflicting evidence, extraordinary support and transition decision	Formal gate review	Record the decision, rationale, conditions and next review date

Table 6

Blank decision record for documenting readiness evidence, vetoes, extraordinary support, authorization boundaries, reversion triggers and dissent.

Decision-record field	Required entry
Unit of application	Line, production area, or process family
Review date and evidence period	Dates covered by the evidence and the formal review date
Product and configuration	Product mix, process revision, software state, and relevant change status
Representative-operation basis	Staffing, shifts, demand, support structure, and operating assumptions
Open vetoes	Condition, owner, closure requirement, verification evidence, and independent challenger
Extraordinary support	Temporary labor, inspection, engineering attention, bypasses, or manual intervention
Dimension evidence	Source, owner, period, limitations, missing evidence, and unresolved disagreement
Decision	Authorize, defer, limit, or return to Phase I
Approved authority	Methods, parameters, standards, decisions, or areas included in the authorization
Conditions and reversion triggers	Actions required before broader authorization and events that invalidate the baseline
Next review	Date and event-based trigger
Approvers and challengers	Participating functions, final approver, and recorded dissent

4.2 Decision Sequence

First, the planned transition date is treated as a review point rather than an automatic phase change. Second, critical vetoes are examined. Third, evidence is reviewed across the six dimensions. Fourth, extraordinary support is separated from normal performance. Fifth, conflicts are discussed across functions. Finally, the transition is authorized, deferred or limited to a defined area.

A partial transition can be used. One process family may enter Phase II while another remains under Phase I governance. The downstream effect must be reviewed. Local optimization is not approved when it shifts instability elsewhere.

The decision record in Table 6 is mandatory for a formal gate review. Representativeness is confirmed by the participating functions and the assumptions are recorded. Veto closure is verified by the evidence owner and challenged by an independent function. Missing evidence is listed as unresolved; it is not

estimated. Partial authorization is bounded by area, method and decision authority. A return to Phase I controls is required when a major product, process or data-system change invalidates the baseline.

4.3 Illustrative Decision Vignette

Consider a hypothetical mixed-model assembly launch. Management has scheduled a move from launch governance to intensive optimization after a fixed number of weeks. Output has improved by the planned review. Several engineering changes are still open. Workforce independence is uneven across shifts. Defect and downtime codes are also being corrected. No critical safety condition is open.

A calendar-based decision would allow the full optimization roadmap to begin because the date has been reached. The readiness-gated system would retain Phase I governance. The decision would not be based on a low aggregate score. It would be based on unresolved evidence concerning change interaction, workforce capability and data reliability.

At a later review, configurations and work instructions have been reconciled. Training status is current. Representative shifts can operate without extraordinary support. Recurring losses can be separated from temporary launch disturbances. The broader Phase II roadmap may then be authorized. The decision is recorded by the cross-functional team.

The vignette shows how the information used in a transition decision is changed. It does not prove that readiness-gated governance is superior. That claim requires longitudinal or comparative evidence.

Table 7
Difference between calendar-led and readiness-gated transition logic.

Decision feature	Calendar-led transition	Readiness-gated transition
Trigger	Scheduled date or nominal milestone	Demonstrated operational and information readiness
Treatment of support	May be embedded in reported output	Recorded separately from normal operation
Treatment of critical risk	Managed within launch controls	Explicit veto before transition
Treatment of data quality	Often assumed after system activation	Reviewed as a readiness dimension
Transition scope	Often plant-wide or program-wide	May be limited by area or process family
Decision output	Phase change	Authorize, defer, limit or reverse improvement authority

5. Simulation-Based Policy Stress Test

5.1 Purpose and Validation Boundary

A computer simulation with random variation tested how the proposed transition rules behave under uncertain production conditions. Its main purpose was to determine whether a readiness gate can select a reasonable transition time when the available evidence is incomplete or noisy. The experiment did not model a specific factory or estimate the actual improvement that a company might achieve. The model used no company production data to define its structure, set its inputs, or calculate its outcomes. The experiment therefore serves as a policy stress test rather than proof of real-world performance. This use follows accepted simulation principles. A model can help explain how decision rules work and where they may fail even before it has been tested in a real production system [61].

The experiment also avoided a circular comparison by separating the information used to make decisions from the information used to judge the results. The calendar and readiness policies could observe only the planned timing, noisy readiness measures, and visible temporary support. The post-transition regression result was generated from hidden readiness conditions, risk factors that varied across launches, and combined effects among those factors. Neither policy could observe this hidden information. The gate was therefore not given direct access to the outcome measure used to judge its performance.

The experiment tests whether the policy can make reasonable decisions with limited information. It does not test whether the six readiness dimensions cause better results in real factories. The model assumes that hidden readiness gaps can increase the risk of performance declining after transition. Given that assumption, the simulation asks whether a policy can choose a suitable transition time by using only noisy evidence that may be affected by temporary support. If readiness does not affect later performance in a real production system, the simulation cannot establish that the gate will be effective. Independent field evidence would then be required.

5.2 Stochastic Launch Model

Each synthetic launch contained six latent readiness states $x(i,t)$ in the interval $[0,1]$, corresponding to the six proposed dimensions. States began at heterogeneous immature levels and generally matured over 36 review periods. Maturation slowed with launch complexity and unresolved disruption. Random shocks temporarily reduced one to three dimensions. Extraordinary support began at a launch-specific level and decayed over time.

Observed readiness $y(i,t)$ differed from latent readiness because of measurement noise. The observation error decreased as data reliability improved. Temporary support was also allowed to mask observed safety/quality, throughput and equipment performance to different degrees. In compact form, observed readiness was represented as

$$y(i,t) = \text{clip}[x(i,t) + \text{measurement error} + \text{support-masking effect}, 0, 1].$$

This structure operationalizes the paper's distinction between true operating maturity and a launch that appears capable because of temporary intervention.

The latent-observed separation represents a management problem rather than only a computational device. During ramp-up, apparent performance may reflect measurement error, temporary engineering support, containment or other extraordinary intervention. A decision rule based on visible performance can therefore authorize too early when temporary support masks immaturity; conversely, a rule that demands excessively persistent evidence can authorize too late when measurements remain noisy. The experiment deliberately permits both errors.

Post-transition regression risk was generated separately from the decision rules. The hidden risk function increased with latent readiness deficits, the interaction between data and change-control deficits, the weakest readiness dimension, extraordinary-support dependence, launch complexity and unresolved disruption. Dimension weights varied across launches. The resulting probability is a synthetic model output, not an empirical probability of a plant event. The outcome equation was intentionally unavailable to the transition policies.

5.3 Policy Definitions, Experimental Design and Reproducibility

Five environmental inputs were varied across broad synthetic ranges: complexity, baseline learning rate, disturbance rate, observation noise and support-masking strength. One hundred input combinations were generated with Latin hypercube sampling to spread design points across the multidimensional input space [60]. Each combination was replicated 250 times, producing 25,000 synthetic launches. A fixed master random seed was used. Common random numbers were retained within each launch so that policies faced the same underlying trajectory and period-specific stochastic outcome draws, reducing Monte Carlo noise in paired comparisons.

Table 8
Synthetic experimental design and decision-policy settings.

Element	Synthetic setting	Purpose
Launch complexity	0.00–1.00	Varies maturation difficulty and disturbance exposure
Baseline learning rate	0.09–0.17 per review period	Varies speed of latent readiness maturation
Base disturbance rate	0.04–0.20 per review period	Varies recurring launch shocks
Observation noise	0.025–0.120	Varies disagreement between latent and observed readiness
Support masking	0.05–0.80	Varies how strongly temporary support can improve apparent performance
Design size	100 Latin-hypercube settings x 250 replications	25,000 synthetic launches
Calendar policies	Transition at periods 10, 14, 18 or 22	Tests early-to-late fixed milestones
Gate-lenient	Lower thresholds; support ≤ 0.35 ; one qualifying period	Tests early authorization tendency
Gate-baseline	Dimension thresholds 0.62–0.72; support ≤ 0.30 ; two consecutive periods	Primary readiness-gate setting
Gate-strict	Thresholds + 0.04; support ≤ 0.22 ; three consecutive periods	Tests over-conservatism and deferral risk

The numerical thresholds are experimental settings, not recommended industrial limits. Three gate settings were tested specifically to prevent the conclusion from depending on one preferred threshold choice. Four calendar milestones were included because any fixed date can look favorable if it is chosen after observing the simulated trajectories.

The 36-period horizon and all numerical policy thresholds are dimensionless experimental constructs. A review period is not intended to represent a day, week or month. Thirty-six periods were selected to accommodate early, intermediate and late fixed-transition policies (periods 10–22) while preserving a further observation window for adaptive authorization and deferral. The threshold vector similarly instantiates the noncompensatory gate logic rather than a transferable industrial standard. Lenient, baseline and strict variants were included to expose sensitivity to policy conservatism rather than to identify a universally optimal threshold set.

5.4 Results and Robustness

The fixed calendar policies displayed the expected risk-delay trade-off. Mean modeled regression risk decreased from 0.409 at calendar period 10 to 0.189 at period 22. The baseline readiness gate authorized 84.3% of launches within the 36-period horizon. Among authorized launches, its mean transition period was 18.37, mean modeled regression risk was 0.157 and the simulated regression-event rate was 0.155. The lenient gate authorized earlier but at higher modeled risk. The strict gate reduced modeled risk further but deferred 27.4% of launches. Table 9 summarizes the policy outcomes.

Table 9

Synthetic policy outcomes across 25,000 launches. Gate risk and regression values are calculated among launches authorized within the 36-period horizon. All numerical risks are outputs of the synthetic model and must not be interpreted as empirical plant probabilities.

Policy	Mean transition period	Deferral rate	Mean modeled risk	Simulated regression rate
Calendar 10	10.00	0.0%	0.409	0.410
Calendar 14	14.00	0.0%	0.297	0.298
Calendar 18	18.00	0.0%	0.230	0.229
Calendar 22	22.00	0.0%	0.189	0.188
Gate lenient	14.44	5.5%	0.257	0.259
Gate baseline	18.37	15.7%	0.157	0.155
Gate strict	21.34	27.4%	0.113	0.111

Paired comparisons show why the result should not be framed as universal superiority. On the 21,071 launches authorized by the baseline gate, its mean modeled risk was 0.157. The same launches had mean risks of 0.247, 0.182 and 0.144 under fixed calendar transitions at periods 14, 18 and 22, respectively. Thus, compared with period 18, the adaptive gate reduced mean modeled risk by 0.025, but compared with the later period-22 policy it increased mean risk by 0.013. A sufficiently late calendar transition can therefore reduce risk, but it imposes the same delay on easy and difficult launches.

A second robustness check used the first period at which hidden modeled regression risk fell below 0.20 as a diagnostic low-risk benchmark. This benchmark was not available to the policies and is not proposed as an industrial threshold. Among launches that reached the benchmark within the horizon, the baseline gate authorized before the benchmark in 16.8% of cases and had a mean absolute timing error of 3.46 periods. The corresponding early-authorization rates for calendar periods 14, 18 and 22 were 53.0%, 33.3% and 20.1%, with mean absolute timing errors of 5.54, 5.94 and 7.69 periods. The baseline gate also deferred 7.8% of launches after the synthetic benchmark had already been reached. This false-deferral result is important: the gate can be too conservative when evidence remains noisy or heterogeneous.

Scenario-level sensitivity analysis reinforced the contingency interpretation. Launch complexity was strongly associated with baseline-gate deferral (Spearman $\rho = 0.809$) and later authorization ($\rho = 0.795$). Faster learning reduced deferral ($\rho = -0.412$) and transition time ($\rho = -0.535$). Disturbance rate had a smaller positive association with deferral ($\rho = 0.211$). These associations describe the synthetic experiment; they are not empirical causal estimates.

Two counter-results are especially important for interpretation. First, the later period-22 calendar policy produced lower mean modeled risk than the baseline gate among baseline-authorized launches, showing that lower risk can be purchased through uniform delay. Second, the baseline gate falsely deferred some launches after the hidden diagnostic benchmark had already been reached, showing that imperfect evidence can make adaptive governance overly conservative. These unfavorable outcomes are retained because they define the boundary of the contribution: readiness gating is proposed as an adaptive risk-timing mechanism, not as a universally dominant transition rule.

5.5 Horizon and Alternative Outcome-Model Sensitivity

Two additional robustness tests were conducted to challenge the two most direct model-form criticisms. First, the same 25,000 synthetic launch trajectories were evaluated with finite analysis horizons of 30, 36, 42 and 48 review periods; the latent and observed trajectories and hidden reference outcome path were held fixed so that only the reporting cutoff changed. Second, the baseline gate at the 36-period horizon was rescored under five hidden outcome structures: the reference heterogeneous model, a sparse-dimension model in which only two or three randomly selected readiness dimensions affected outcome, a weak-readiness model in which contextual terms dominated, an interaction-dominant model, and a placebo-null model in which outcome risk was constant across periods and independent of readiness, support and transition timing. The robustness run used the same 100 Latin-hypercube environmental settings, 250 replications per setting and an independent fixed seed.

The independent 36-period robustness run closely reproduced the primary baseline result: 15.4% deferral, a mean authorization period of 18.44 and mean modeled risk of 0.158 among authorized launches, compared with 15.7%, 18.37 and 0.157 in the primary run. Horizon extension reduced the fraction classified as deferred because late-authorizing launches were given additional time, but mean modeled risk among authorized launches changed by less than 0.002 across the 30–48 period range. Deferral is therefore horizon-dependent, as expected, whereas the risk estimate for authorized launches was comparatively insensitive to the cutoff.

Table 10

Horizon sensitivity of the baseline readiness gate in an independent 25,000-launch robustness run. Positive Delta(C18-Gate) values indicate lower mean modeled risk for the gate than for a fixed period-18 transition on the same authorized launches. All values are synthetic.

Horizon	Authorized launches	Deferral (%)	Mean gate period	Mean gate risk	Delta(C18-Gate)
30	19,431	22.28	17.13	0.1582	+ 0.0077
36	21,144	15.42	18.44	0.1575	+ 0.0260
42	22,243	11.03	19.48	0.1572	+ 0.0387
48	22,966	8.14	20.29	0.1568	+ 0.0473

The hidden outcome equation was then made progressively less favorable to the framework. Relative to calendar period 18, the baseline gate retained a positive mean paired risk difference under the reference (+ 0.0260), sparse-dimension (+ 0.0106), weak-readiness (+ 0.0123) and interaction-dominant (+ 0.0185) models. The magnitude declined when readiness carried less or more selective outcome information. Under the placebo-null model, every paired policy difference was exactly zero because outcome risk was constant over time for each launch. This null result is a diagnostic safeguard: the gate receives no mechanical risk advantage from its timing rule when readiness is removed from the outcome mechanism.

Table 11

Alternative hidden-outcome-model sensitivity on the 21,144 launches authorized by the baseline gate within 36 periods. Delta values are calendar mean modeled risk minus gate mean modeled risk; positive values favor the gate within that model. Absolute risk levels should not be compared across model forms.

Hidden outcome model	Gate mean risk	Delta(C14-Gate)	Delta(C18-Gate)	Delta(C22-Gate)
Reference	0.1575	+ 0.0917	+ 0.0260	-0.0126
Sparse-dimension	0.1409	+ 0.0374	+ 0.0106	-0.0060
Weak-readiness	0.1948	+ 0.0289	+ 0.0123	+ 0.0013
Interaction-dominant	0.1267	+ 0.0528	+ 0.0185	-0.0020
Placebo-null	0.2114	0.0000	0.0000	0.0000

These tests sharpen rather than expand the claim. The readiness gate is not supported because it always wins; it does not. A late calendar policy can still produce lower modeled risk, the reported deferral percentage depends on the analysis horizon, and the gate loses all modeled advantage when the assumed readiness-outcome relationship is removed. The computational contribution is therefore conditional: if readiness contains meaningful but imperfectly observed information about transition vulnerability, an adaptive gate can improve the risk-timing decision relative to selected fixed milestones.

Whether that premise holds, and which dimensions and thresholds matter in a real production system, remains an empirical question.

6. Discussion

The simulation provides a structured check of how the proposed decision rules behave. It shows that the readiness gate creates a real trade-off between reducing risk and avoiding unnecessary delay. The gate does not produce a guaranteed advantage in every situation. Under difficult simulated conditions, high-risk early transitions were reduced when the gate adjusted approval timing to each launch. However, settings that were too strict caused avoidable delays and left more launches waiting for approval.

The practical lesson is that launch teams should adapt the gate to their operating conditions. Evidence requirements, readiness limits, and review frequency should be adjusted for product complexity, automation, workforce experience, production changes, and other launch conditions. Teams should not convert readiness into one fixed score for every launch.

The goal is not to make the gate as strict as possible. If the requirements are too lenient, temporary stability may be mistaken for true production readiness. If they are too strict, launch controls may continue after the production system has become ready for broader improvement. The main rule is simple: managers should allow greater decision authority as the evidence becomes more reliable, the production system becomes more stable, and dependence on temporary support decreases.

6.1 Contribution to Production Engineering

The framework provides a management system rather than a new Lean tool. Basic Lean practices remain important from the beginning of ramp-up. However, their main purpose changes as the production system becomes more stable. Early Lean practices are used to identify problems, control risks, and make abnormal conditions visible. Later practices are used to remove confirmed losses and maintain a stable production method.

The second contribution is treating the ability to see and understand production conditions as an important production capability. Output, quality, and equipment availability remain important measures. However, these measures alone do not show whether teams understand the causes behind the results. Teams also need complete event records, consistent definitions, and the ability to trace product and process versions. Temporary labor, inspection, engineering help, and other launch support must also remain visible.

The third contribution is the gradual increase of improvement authority. Analytical and digital tools can support the launch from an early stage. However, managers should not give their recommendations the same decision-making power at every stage. Greater authority is given after teams confirm that the production baseline is stable and clear enough to support reliable decisions.

This approach supports earlier research that connects Lean practices with digital technologies [17]-[22], [40]-[45]. It also applies research on readiness and decision management [46]-[50] to one specific ramp-up question: when should a production area move from stabilization to broader improvement?

6.2 Relationship to Existing Launch Systems

APQP, PPAP, Control Plans, and safe-launch practices remain necessary. They manage required launch documents, validation evidence, product and process risks, customer approval, release requirements, and temporary containment [38], [39], [51]. The proposed framework does not replace any of these systems.

Instead, the framework addresses a different decision. It asks when production evidence is strong enough to support broader improvement and greater decision authority. Required launch documents may be completed while the production baseline is still changing or difficult to understand. In other cases, one production area may become ready for broader improvement before the overall program reaches its planned milestone.

Teams can add the readiness review to their existing launch-management process. It serves as an additional operating decision rather than a separate launch system. Duplicate work should be avoided. Teams should reuse evidence that already comes from quality planning, maintenance, training, engineering, and production systems.

Table 12

explains the different decisions made by the existing launch systems and the proposed readiness review. APQP, PPAP, Control Plans, and safe launch continue to manage their established requirements. The paper claims a narrower contribution: deciding whether production operations and their supporting information are ready for greater improvement authority.

System	Primary purpose	Typical decision object	Relationship to the proposed framework
APQP [38]	Plan and govern product and process development, risk, change, metrics and launch deliverables.	Completion and approval of planning and validation requirements.	Remains the overarching launch-assurance system.
PPAP [51]	Demonstrate that design-record and specification requirements are understood and can be met by the production process.	Customer production-part approval.	Provides approval evidence but does not, by itself, determine when optimization authority should increase.
Control Plan [39]	Define and maintain controls for product and process characteristics across production phases.	Approval and maintenance of control methods.	Provides control and evidence inputs used by the readiness review.
Safe launch [39]	Apply heightened temporary controls during early production and define their exit.	Entry, maintenance and exit from additional launch controls.	Overlaps with Phase I containment but does not cover every readiness dimension.
Proposed readiness review	Assess whether the operating baseline is sufficiently safe, repeatable and interpretable.	Authorize, defer, limit or reverse increased improvement and digital decision authority.	Operates within, rather than instead of, existing launch systems.

Table 12. Relationship of the readiness review to APQP, PPAP, Control Plan and safe-launch governance.

6.3 Risks and Boundary Conditions

A readiness gate can be gamed. Indicators may be manipulated. Critical conditions may be reported late. A threshold may become a bureaucratic approval exercise. These risks are reduced when definitions are transparent, evidence is independently challenged and the predictive value of the indicators is reviewed.

Excessive phase separation is another risk. Real ramp-ups contain local pockets of maturity. A plant-wide gate may be too coarse. An unrestricted station-level gate may encourage suboptimization. The unit of application should therefore be selected with system flow in mind.

The framework is less distinctive in copy-exactly transfers and low-change processes. It may be more valuable when complexity, product mix, automation, workforce change or engineering-change load is high. Regulated production may also require permanent controls before start of production. The dimensions and evidence would need to be adapted.

Phase I should not be used to excuse weak performance. Temporary controls require closure plans. Provisional standards require version control. Below-target performance should be accompanied by visible learning and declining recurrence. Inefficiency should not be relabeled as learning without evidence.

7. Research Propositions and Field-Validation Agenda

The management system remains conceptually developed and has only been stress-tested with synthetic data. The computational experiment evaluates internal policy behavior; it does not test the following causal propositions in real production. The propositions therefore remain a field-validation agenda and should not be interpreted as empirical results of the present study.

The propositions are stated at the production-area or process-family level and are intended for longitudinal or comparative testing. A transition is treated as readiness-conditioned when documented veto review, representative-operation evidence, extraordinary-support disclosure and bounded authorization are completed before broader optimization authority is granted. Post-transition regression may be operationalized through reopened control-plan actions, reintroduced temporary containment, recurrence of previously closed loss categories, reversal of validated standard work, or loss of previously demonstrated process performance within a predefined observation period.

Table 13

Testable propositions, units of analysis, observation periods and candidate operational measures for future validation.

Proposition	Testable formulation and candidate operationalization
P1: Readiness-based sequencing	At the production-area level, readiness-conditioned transitions will be followed by a lower 90-day rate of control-plan reopening and recurrence of previously closed loss categories than matched transitions triggered primarily by a scheduled date. Candidate operationalization: Matched ramp-ups or difference-in-differences. Candidate measures: control-plan reopenings, reintroduced containment, recurrence of closed loss codes, and reversions to earlier standards during the first 90 days after transition.
P2: Measurement and data reliability	The positive association between post-transition improvement intensity and 90-day sustained production performance will be stronger when baseline measurement-system adequacy, classification agreement, denominator completeness, configuration traceability and time-order integrity are higher. Candidate operationalization: Multilevel moderated analysis across lines or launches. Improvement intensity may be represented by verified improvement hours, completed DMAIC actions or controlled process changes. Sustained performance may include first-pass yield, cycle-time attainment, downtime recurrence and reopened controls.
P3: Controlled change and learning	Within a ramp-up area, a higher weekly count of concurrent, insufficiently validated engineering changes will be associated with a longer time to confirm stable cause-and-effect relationships for recurring losses. Candidate operationalization: Event-history or survival analysis. Candidate measures: concurrent unvalidated changes per week and elapsed time from the first recurring-loss signal to independently verified causal closure, with complexity, staffing and automation treated as controls.
P4: Workforce qualification and reporting	Ramp-up areas with higher independent-qualification coverage and shorter abnormality-reporting latency will identify recurring loss mechanisms sooner. The association is expected to be stronger when both conditions are present. Candidate operationalization: Multilevel time-to-event analysis. Candidate measures: qualified-operation coverage by shift, event-to-report time, escalation completeness and time to a verified recurring-loss mechanism.
P5: Phase-aligned digital authority	The realized benefit of predictive or prescriptive digital recommendations will be greater when model-to-plant validation and configuration synchronization are demonstrated before decision authority is increased. The benefit is expected to weaken when process state or product mix departs from the validated baseline. Candidate operationalization: Interrupted time-series, stepped-wedge or controlled before-and-after design. Candidate measures: prediction error, recommendation acceptance and override rates, realized versus predicted benefit, decision reversals and performance after out-of-baseline states.

External validation is now the first priority. Several ramp-ups should be observed from pre-production through stable series operation. Readiness evidence, support resources, improvement timing and later regressions should be captured. Empirical work should test whether the synthetic risk-timing behavior

observed here persists when indicators, disturbances and outcomes are measured in operating production systems.

Comparative case studies could examine calendar-led and readiness-led governance. Product complexity, automation, workforce experience, supplier readiness and planned volume should be matched. Outcomes should extend beyond nominal output. Sustained capability, defect recurrence, standard revision, downtime concentration and support hours should also be considered.

Content validity should also be assessed. Production, quality, maintenance, engineering, training and digital-manufacturing experts could review the dimensions and veto conditions. A favorable opinion alone would not prove effectiveness. It would show whether the proposed architecture is understandable and complete enough for field testing.

8. Limitations

No empirical production dataset was used. The illustrative vignette demonstrates decision logic only. The computational experiment uses entirely synthetic trajectories and parameter ranges; it is not calibrated to a specific company, product, process or launch. Horizon and alternative outcome-model sensitivity tests reduce dependence on a single cutoff and a single hidden risk equation, but they do not convert the computational study into field validation. Real operating data remain necessary to estimate construct validity, thresholds, effect sizes and causal performance consequences.

The hidden regression-risk function shares the framework's general premise that unresolved readiness deficits, support dependence and interacting uncertainty can increase post-transition instability. This structural relationship is necessary for a policy stress test but prevents the simulation from serving as independent empirical proof. Separation of observed policy inputs from latent outcome information, heterogeneous risk weights, Latin-hypercube sampling and threshold variation reduce tautology risk but do not eliminate model-form dependence.

The modeled regression probabilities, the 0.20 diagnostic benchmark, gate thresholds and review-period horizons are synthetic experimental constructs. They are not recommended operating limits. The horizon test showed that the reported deferral rate changes materially with the finite cutoff, even though mean modeled risk among authorized launches remained nearly unchanged from 30 to 48 periods. Alternative hidden outcome structures also changed the magnitude of the gate advantage, and the advantage disappeared under the placebo-null outcome model. These are useful model-form checks, but they do not validate the assumed latent readiness-outcome relationship.

The literature was synthesized through a concept-driven approach. It was not selected through a complete systematic-review protocol. Relevant studies may therefore have been omitted.

The six dimensions are provisional formative components. Their boundaries may overlap. Indicator definitions, review periods and local adaptations require validation.

The framework assumes that reasonably credible evidence can be obtained. Political reporting, weak instrumentation or fragmented systems may limit that assumption. Independent observation may then be required.

The system simplifies spatial and temporal differences. Stations, shifts, suppliers and product variants may reach readiness at different times. Further research should examine the appropriate level of aggregation.

9. Conclusions

A readiness-gated management system has been proposed for manufacturing ramp-up. The system separates stabilization-oriented governance from improvement-oriented maturation. Lean, statistical and digital practices are used in both phases. Their purpose and authority are changed as evidence quality improves.

Six dimensions are reviewed: safety and quality containment, throughput repeatability, equipment reliability, workforce readiness, data reliability and change-control stability. Critical vetoes are considered first. Cross-functional evidence is then reviewed under representative operating conditions.

The practical message is direct. A planned date should be used as a review point, not as automatic proof of readiness. Temporary launch support should be made visible. Improvement should be intensified when recurring losses can be interpreted without embedding unstable conditions into permanent standards or digital decisions.

The framework is intended to complement established launch systems. A synthetic policy stress test, horizon sensitivity analysis and alternative hidden-outcome checks indicate that adaptive readiness gating can improve the alignment between transition timing and modeled post-transition vulnerability when readiness contains meaningful outcome information, while excessive conservatism can create delay and deferral. The advantage is conditional rather than universal and disappears when readiness is deliberately removed from the synthetic outcome mechanism. Field validation through expert review, longitudinal launches and comparative cases is therefore still required.

Declarations

Competing interests.

No competing financial or non-financial interests are declared.

Declarations of generative AI use. During preparation of this manuscript, the author used OpenAI ChatGPT to assist with simulation-design code drafting. The tool was not used to generate, infer or analyze empirical organizational production data. The author is responsible for reviewing the model

assumptions, executing and verifying the code, checking all sources and calculations, and approving the final arguments and conclusions.

Funding.

No external funding was received.

Author Contribution

Nikhil Sharma is the sole writer of the paper.

Acknowledgement

Pre-Print of the work exists on Research Square

Data availability.

No empirical organizational dataset was generated or analyzed. The computational experiment uses only synthetic data generated from the parameter ranges and stochastic rules reported in the manuscript. Reproducible source code and synthetic output summaries are available as supplementary material.

Ethics. No human-participant study or proprietary organizational study was conducted.

Confidentiality. The manuscript is based on generalized manufacturing concepts and published literature. No confidential information from a specific organization is disclosed.

References

1. Surbier L, Alpan G, Blanco E (2014) A comparative study on production ramp-up: state-of-the-art and new challenges. *Prod Plann Control* 25(15):1264–1286. 10.1080/09537287.2013.817624
2. Terwiesch C, Bohn RE (2001) Learning and process improvement during production ramp-up. *Int J Prod Econ* 70(1):1–19. 10.1016/S0925-5273(00)00045-1
3. Almgren H (2000) Pilot production and manufacturing start-up: the case of Volvo S80. *Int J Prod Res* 38:4577–4588. 10.1080/00207540050205316
4. Schuh G, Gartzen T, Wagner J (2015) Complexity-oriented ramp-up of assembly systems. *CIRP J Manufact Sci Technol* 10:1–15. 10.1016/j.cirpj.2015.05.007
5. Bultó R, Viles E, Mateo R (2018) Overview of ramp-up curves: a literature review and new challenges, *Proceedings of the Institution of Mechanical Engineers, Part B: Journal of Engineering Manufacture*,

- vol. 232, no. 5, pp. 755–765. 10.1177/0954405416654416
6. Verhaelen B, Martin M, Peukert S, Lanza G (2023) Practice-oriented methodology for increasing production ramp-up efficiency in global production networks of SME. *Prod Eng Res Devel* 17:145–177. 10.1007/s11740-022-01154-7
 7. Ohno T (1988) *Toyota Production System: Beyond Large-Scale Production*. Productivity, Cambridge, MA
 8. Womack JP, Jones DT (2003) *Lean Thinking: Banish Waste and Create Wealth in Your Corporation*, revised ed. Free, New York, NY
 9. Liker JK (2004) *The Toyota Way: 14 Management Principles from the World's Greatest Manufacturer*. McGraw-Hill, New York, NY
 10. Dennis P (2015) *Lean Production Simplified*, 3rd edn. Productivity, Boca Raton, FL
 11. Netland TH (2016) Critical success factors for implementing lean production: the effect of contingencies, *International Journal of Production Research*, vol. 54, no. 8, pp. 2433–2448. 10.1080/00207543.2015.1096976
 12. Doltsinis S, Ferreira P, Mabkhot MM, Lohse N (2020) A decision support system for rapid ramp-up of Industry 4.0-enabled production systems, *Computers in Industry*, vol. 116, art. 103190. 10.1016/j.compind.2020.103190
 13. Barbieri G et al (2021) A virtual commissioning based methodology to integrate digital twins into manufacturing systems. *Prod Eng Res Devel* 15:397–412. 10.1007/s11740-021-01037-3
 14. Hussein K et al (2022) Development of a digital twin for improved ramp-up processes in the context of Li-ion battery cell stack formation. *Procedia CIRP* 106:27–32. 10.1016/j.procir.2022.02.150
 15. Stang J, Kroeger S, Zaeh MF (2023) Semi-structured expert interview-based requirements elicitation for a digitalized production ramp-up. *Procedia CIRP* 120:279–284. 10.1016/j.procir.2023.08.050
 16. Haller J, Kaven L, Göppert A, Schmitt RH (2026) Industry 4.0 advancements in discrete production ramp-ups: a systematic literature review. *J Intell Manuf* 37(6):2563–2583. 10.1007/s10845-025-02656-8
 17. Buer S-V, Strandhagen JO, Chan FTS (2018) The link between Industry 4.0 and lean manufacturing: mapping current research and establishing a research agenda. *Int J Prod Res* 56(8):2924–2940. 10.1080/00207543.2018.1442945
 18. Rosin F, Forget P, Lamouri S, Pellerin R (2020) Impacts of Industry 4.0 technologies on Lean principles. *Int J Prod Res* 58(6):1644–1661. 10.1080/00207543.2019.1672902
 19. Ciano MP, Dallasega P, Orzes G, Rossi T (2021) One-to-one relationships between Industry 4.0 technologies and Lean Production techniques: a multiple case study. *Int J Prod Res* 59(5):1386–1410. 10.1080/00207543.2020.1821119
 20. Schumacher S, Hall R, Bildstein A, Bauernhansl T (2023) Lean Production Systems 4.0: systematic literature review and field study on the digital transformation of lean methods and tools. *Int J Prod Res* 61(24):8751–8773. 10.1080/00207543.2022.2159562

21. Saad SM, Bahadori R, Bhovar C, Zhang H (2024) Industry 4.0 and Lean Manufacturing: a systematic review of the state-of-the-art literature and key recommendations for future research. *Int J Lean Six Sigma* 15(5):997–1024. 10.1108/IJLSS-02-2022-0021
22. Kassem B, Callupe M, Rossi M, Rossini M, Portioli-Staudacher A (2024) Lean 4.0: a systematic literature review on the interaction between lean production and Industry 4.0 pillars. *J Manuf Technol Manage* 35(4):821–847. 10.1108/JMTM-04-2022-0144
23. Dutton JM, Thomas A (1984) Treating progress functions as a managerial opportunity. *Acad Manage Rev* 9(2):235–247. 10.5465/AMR.1984.4277639
24. Argote L, Eppl D (1990) Learning curves in manufacturing, *Science*, vol. 247, no. 4945, pp. 920–924. 10.1126/science.247.4945.920
25. Adler PS, Clark KB (1991) Behind the learning curve: a sketch of the learning process. *Manage Sci* 37(3):267–281. 10.1287/mnsc.37.3.267
26. Edmondson AC (1999) Psychological safety and learning behavior in work teams. *Adm Sci Q* 44(2):350–383. 10.2307/2666999
27. Kritzinger W, Karner M, Traar G, Henjes J, Sihn W (2018) Digital Twin in manufacturing: a categorical literature review and classification. *IFAC-PapersOnLine* 51(11):1016–1022. 10.1016/j.ifacol.2018.08.474
28. Tao F, Zhang H, Liu A, Nee AYC (2019) Digital twin in industry: state-of-the-art. *IEEE Trans Industr Inf* 15(4):2405–2415. 10.1109/TII.2018.2873186
29. Saporiti N, Cannas VG, Pozzi R, Rossi T (2023) Challenges and countermeasures for digital twin implementation in manufacturing plants: a Delphi study, *International Journal of Production Economics*, vol. 261, art. 108888. 10.1016/j.ijpe.2023.108888
30. Bitencourt J, Wooley A, Harris G (2025) Verification and validation of digital twins: a systematic literature review for manufacturing applications. *Int J Prod Res* 63(1):342–370. 10.1080/00207543.2024.2357741
31. Psarommatis F, May G (2023) A literature review and design methodology for digital twins in the era of zero defect manufacturing. *Int J Prod Res* 61(16):5723–5743. 10.1080/00207543.2022.2101960
32. Overbeck L, Graves SC, Lanza G (2024) Development and analysis of digital twins of production systems. *Int J Prod Res* 62(10):3544–3558. 10.1080/00207543.2023.2242525
33. Wagner S et al (2024) From framework to industrial implementation: the digital twin in process planning. *J Intell Manuf* 35:3793–3813. 10.1007/s10845-023-02268-0
34. Ivanov D (2024) Conceptualisation of a seven-element digital twin framework in supply chain and operations management. *Int J Prod Res* 62:2220–2232. 10.1080/00207543.2023.2217291
35. Guo D, Mantravadi S (2025) The role of digital twins in lean supply chain management: review and research directions. *Int J Prod Res* 63(5):1851–1872. 10.1080/00207543.2024.2372655
36. Montgomery DC (2019) *Introduction to Statistical Quality Control*, 8th edn. Wiley, Hoboken, NJ

37. Heckert NA, Filliben JJ, Croarkin CM, Hembree B, Guthrie WF, Tobias P, Prinz J (2002) Handbook 151: NIST/SEMATECH e-Handbook of Statistical Methods. National Institute of Standards and Technology, Gaithersburg, MD. 10.18434/M32189
38. Automotive Industry Action Group (2024) Advanced Product Quality Planning (APQP), 3rd ed., manual code APQP-3. Southfield, MI: AIAG
39. Automotive Industry Action Group (2024) Control Plan, 1st ed., manual code CP-1. Southfield, MI: AIAG
40. Skalli D, Charkaoui A, Cherrafi A, Garza-Reyes JA, Antony J, Shokri A (2023) Industry 4.0 and Lean Six Sigma integration in manufacturing: a literature review, an integrated framework and proposed research perspectives. *Qual Manage J* 30(1):16–40. 10.1080/10686967.2022.2144784
41. Vlachos IP, Martinez Pascazzi R, Zobolas G, Repoussis P, Giannakis M (2023) Lean manufacturing systems in the area of Industry 4.0: a lean automation plan of AGVs/IoT integration. *Prod Plann Control* 34(4):345–358. 10.1080/09537287.2021.1917720
42. Ding B, Ferràs Hernández X, Agell N, Jané (2023) Combining lean and agile manufacturing competitive advantages through Industry 4.0 technologies: an integrative approach. *Prod Plann Control* 34(5):442–458. 10.1080/09537287.2021.1934587
43. Saraswat P, Agrawal R, Rane SB (2024) Technological integration of lean manufacturing with Industry 4.0 toward lean automation: insights from the systematic review and further research directions. *Benchmarking: Int J* 32(6):1909–1941. 10.1108/BIJ-05-2023-0316
44. Sanders A, Elangeswaran C, Wulfsberg JP (2016) Industry 4.0 implies lean manufacturing: research activities in Industry 4.0 function as enablers for lean manufacturing. *J Ind Eng Manag* 9(3):811–833. 10.3926/jiem.1940
45. Tortorella GL, Fettermann D (2018) Implementation of Industry 4.0 and lean production in Brazilian manufacturing companies. *Int J Prod Res* 56(8):2975–2987. 10.1080/00207543.2017.1391420
46. Wagire AA, Joshi R, Rathore APS, Jain R (2021) *Prod Plann Control* 32(8):603–622. 10.1080/09537287.2020.1744763. Development of maturity model for assessing the implementation of Industry 4.0: learning from theory and practice,
47. Saunila M, Ukko J, Nasiri M, Garengo P (2024) The role of digital governance in the integration of performance measurement systems uses and Industry 4.0 maturity. *Int J Comput Integr Manuf* 37(7):809–822. 10.1080/0951192X.2023.2257918
48. Salimbeni S, Redchuk A, Rousserie H (2023) Quality 4.0: technologies and readiness factors in the entire value flow life cycle, *Production & Manufacturing Research*, vol. 11, no. 1, art. 2238797. 10.1080/21693277.2023.2238797
49. Haller J, Kumar B, Göppert A, Schmitt RH (2024) A system dynamics stability model for discrete production ramp-up, *Systems*, vol. 12, no. 12, art. 575. 10.3390/systems12120575
50. Melnychuk O, Baum J, Göppert A, Schmitt RH, Tolio T (2025) Pit-Stop Manufacturing: decision support for complexity and uncertainty management in production ramp-up planning, *Systems*, vol. 13, no. 5, art. 393. 10.3390/systems13050393

51. Automotive Industry Action Group (2009) Production Part Approval Process (PPAP), 4th ed., 2nd printing, manual code PPAP-4. Southfield, MI: AIAG
52. Galbraith JR (1974) Organization Design: An Information Processing View, *Interfaces*, vol. 4, no. 3, pp. 28–36. 10.1287/inte.4.3.28
53. Tushman ML, Nadler DA (1978) Information Processing as an Integrating Concept in Organizational Design. *Acad Manage Rev* 3(3):613–624. 10.5465/AMR.1978.4305791
54. Daft RL, Lengel RH (1986) Organizational Information Requirements, Media Richness and Structural Design. *Manage Sci* 32(5):554–571. 10.1287/mnsc.32.5.554
55. Argote L, Miron-Spektor E (2011) Organizational Learning: From Experience to Knowledge, *Organization Science*, vol. 22, no. 5, pp. 1123–1137. 10.1287/orsc.1100.0621
56. Automotive Industry Action Group (2010) Measurement Systems Analysis Reference Manual, 4th ed., manual code MSA-4. Southfield, MI: AIAG
57. International Organization for Standardization (2023) ISO 7870-2:2023, Control Charts-Part 2: Shewhart Control Charts. ISO, Geneva, Switzerland
58. Cohen J (1960) A Coefficient of Agreement for Nominal Scales. *Educ Psychol Meas* 20(1):37–46. 10.1177/001316446002000104
59. Fleiss JL (1971) Measuring Nominal Scale Agreement among Many Raters. *Psychol Bull* 76(5):378–382. 10.1037/h0031619
60. McKay MD, Beckman RJ, Conover WJ (1979) A comparison of three methods for selecting values of input variables in the analysis of output from a computer code, *Technometrics*, vol. 21, no. 2, pp. 239–245. 10.1080/00401706.1979.10489755
61. Sargent RG (2013) Verification and validation of simulation models. *J Simul* 7(1):12–24. 10.1057/jos.2012.20

Figures

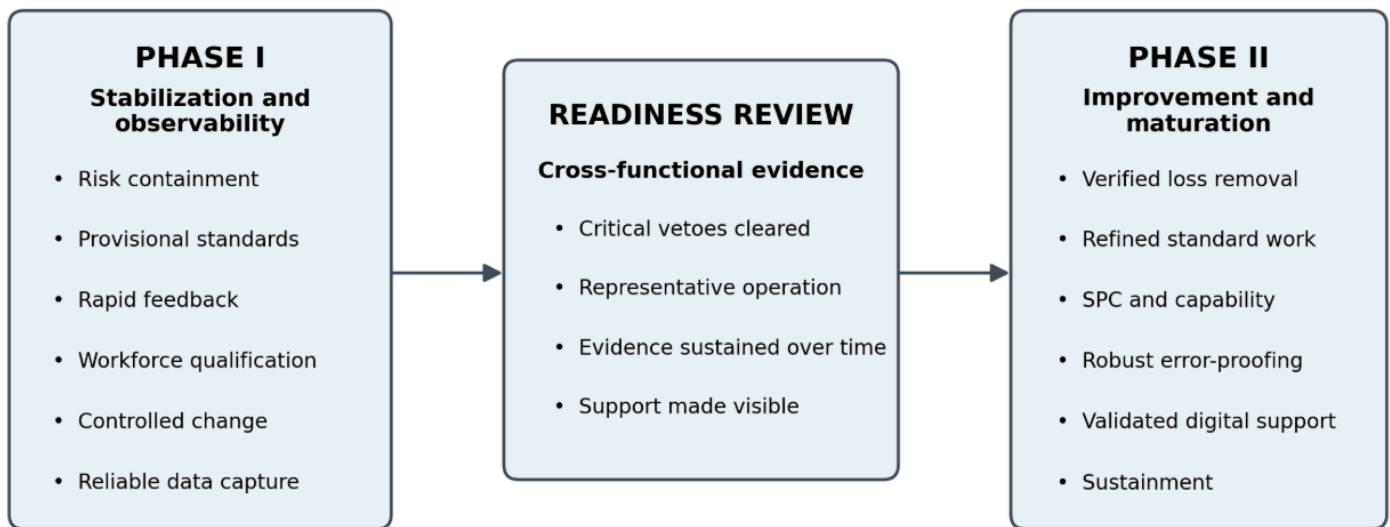


Figure 1

Readiness-gated transition from stabilization and observability to improvement and maturation. The cross-functional review authorizes, limits, defers or reverses optimization authority after critical vetoes and representative operating evidence are assessed.

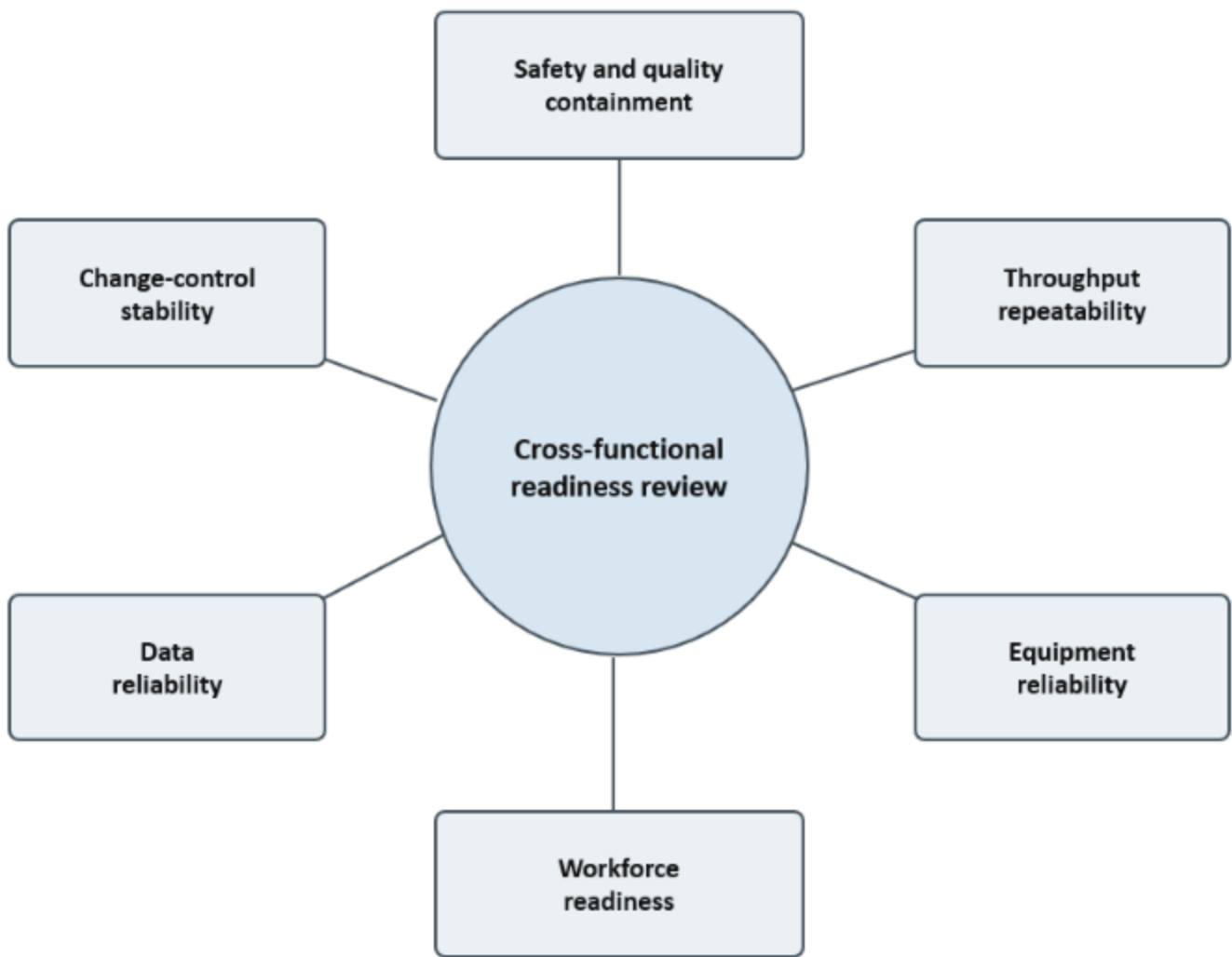


Figure 2

Six formative readiness dimensions reviewed by the cross-functional team. Safety, regulatory, critical-quality and material-data-integrity conditions are treated as non-compensatory vetoes.

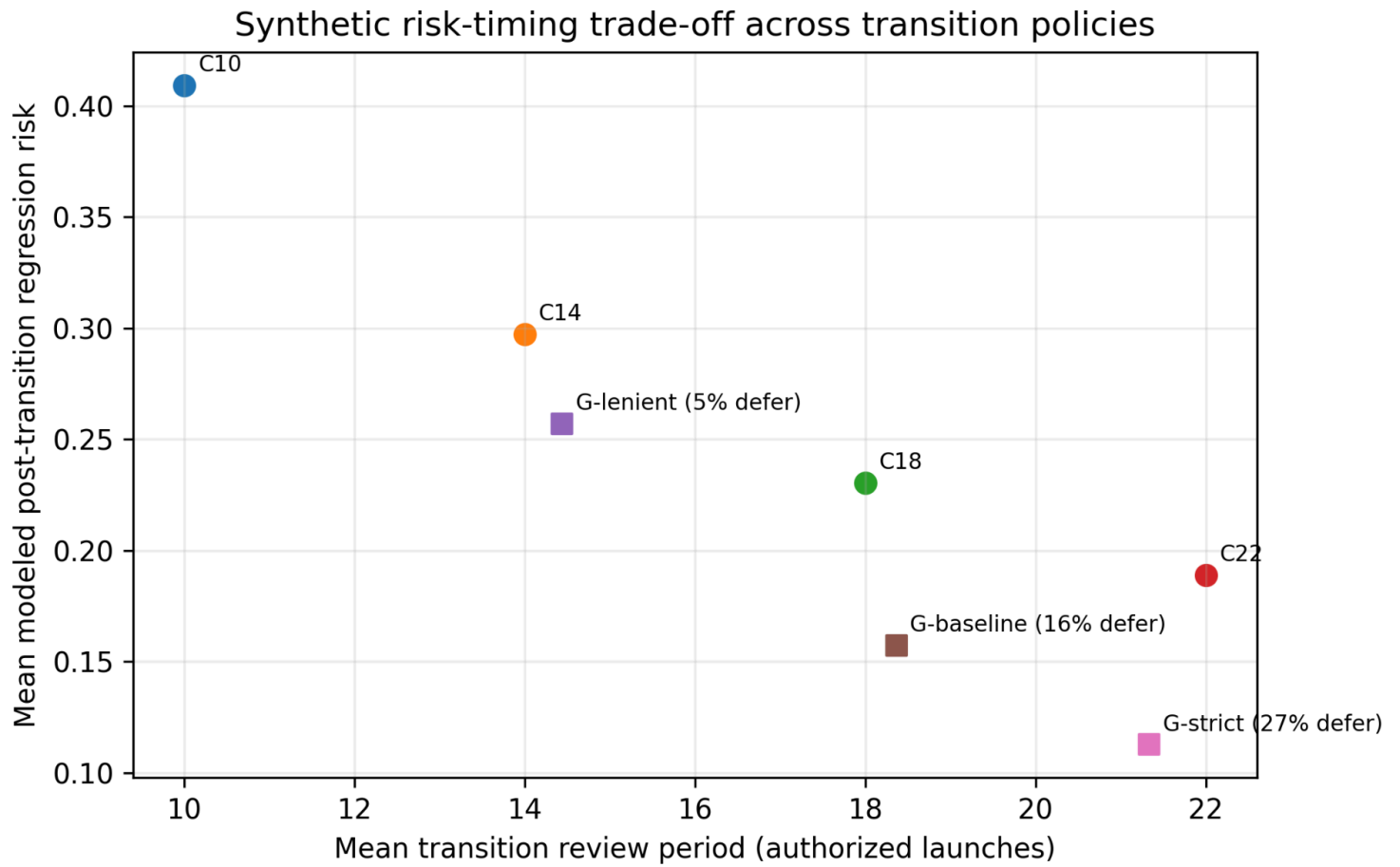


Figure 3

Synthetic risk-timing trade-off across fixed calendar and readiness-gated transition policies. Gate points show mean transition timing among authorized launches; labels also report the proportion deferred beyond the 36-period horizon.