



IJSRIS
journal
ISSN: 2820-7157

★ INTERNATIONAL JOURNAL OF ★

**SCIENTIFIC RESEARCH AND
INNOVATIVE STUDIES**

Advancing Knowledge, Inspiring Innovation

 www.ijsrisjournal.com



**INDEXED
JOURNAL**

 **PEER REVIEWED**
Rigorous & Transparent

 **DOI: CROSSREF**
Member

 **OPEN ACCESS**
For Global Visibility

 **BI-MONTHLY**
Timely Publication

 **WIDE READERSHIP**
High Impact & Visibility

 **AUTHOR FRIENDLY**
Fast | Supportive | Reliable



RECEIVED
17 August 2026



ACCEPTED
09 September 2026



PUBLISHED
1 October 2026



Pages from:



October 2026



Volume 5



Number 5

690 to 704

AI-Enabled Management Control and Treasury Management for Financial Resilience, A Financial-Agility Framework, Econometric Research Design, and Reproducible Simulation

Radi Abderrahman¹, Kaiss Redouane¹, Arjani Chaimae¹, Maimoun Ahmed¹

¹Research Laboratory in Economics, Management, and Business Administration, Faculty of Economics and Management, Hassan 1st University Settat 26000, Morocco, abderrahmanradi2@gmail.com, redouane.kaiss.doc@uhp.ac.ma, chaimaearjani01@gmail.com, maimounahmed@gmail.com

LICENSE



This article is licensed under the
Creative Commons Attribution
4.0 International License (CC BY 4.0).

 <https://creativecommons.org/licenses/by/4.0/>



PREFIX :

10.63883

IJSRIS JOURNAL

OPEN ACCESS



This is an open access article
distributed under the terms of
open access.

ABSTRACT

Artificial intelligence may improve treasury forecasts without necessarily improving an enterprise's capacity to withstand financial disruption. This paper develops an implementation-

ready research method connecting artificial intelligence capability, management control, treasury quality, financial agility, and financial resilience. A structured, multi-source literature synthesis motivates a governed architecture combining accounting-data reconciliation, probabilistic cash-flow forecasting, constrained liquidity decisions, and human

authorization. The proposed Moroccan field study links a prospective enterprise panel to transaction records and repeated multiformat surveys. A randomized phased offer is preferred; cohort-specific difference-in-differences provides a conditional observational alternative. Dynamic panel models, nonlinear adjustment, and longitudinal mechanism analysis address distinct secondary questions rather than substituting for identification. Because no enterprise dataset was supplied, the numerical evidence consists exclusively of an executed Monte Carlo experiment: nine scenarios, 500 replications each, and four estimators. In the reference scenario with 300 synthetic firms and eight periods, nonlinear fixed effects and firm-block cross-fit series estimation recover an imposed coefficient of 0.18, whereas pooled and mis specified linear fixed-effects regressions remain biased. Cross-fit interval coverage is 93.8%; an omitted time-varying confounder raises bias to 0.241 and destroys coverage. For an imposed smaller effect of 0.06, simulated power increases from 44.2% with 150 firms to 96.8% with 600 firms. These findings validate limited aspects of the statistical workflow, not the business effectiveness of artificial intelligence. The contribution is a falsifiable, auditable protocol integrating forecast evaluation, causal inference, managerial mechanisms, and operational safeguards.

Keywords: management control; treasury management; artificial intelligence; financial resilience; financial agility; panel econometrics; Morocco.

Evidence status: Research proposal with executed synthetic validation; no Moroccan firm-level field results. Author names and affiliation are reproduced as supplied. The design and numerical assumptions require author approval and prospective registration before field implementation. References use APA author–date formatting with blue clickable citations and source links.

1. Introduction

1.1. The financial decision problem

Treasury management connects short-horizon cash availability to the continuity of productive activity. A profitable enterprise can nevertheless face a liquidity shortfall when customer receipts arrive after payroll, tax, debt-service, or supplier obligations. Management control supplies planning routines, responsibility structures, variance analysis, and escalation rules through which such information becomes action. Treating those practices as an interacting package is more appropriate than attributing performance to a single dashboard or software acquisition (Malmi & Brown, 2008). Working-capital research also establishes that receivables,

inventories, and payment policies are economically consequential, although evidence from Belgian firms cannot directly establish effects in Morocco (Deloof, 2003).

Artificial intelligence (AI) introduces tools for nonlinear forecasting, anomaly detection, and prioritization. Yet the relevant organizational exposure is not the existence of a purchased application. It is a capability combining usable data, analytical skills, organizational resources, and operational routines (Mikalef & Gupta, 2021). A model can produce a more accurate forecast while managers lack an available credit line, sufficient authority, or the time needed to act. Conversely, an organization can react quickly but choose costly or harmful actions. Financial agility and financial resilience must therefore be separated: agility concerns timely feasible reconfiguration, whereas resilience concerns continuity and recovery under disruption (Duchek, 2020).

The Moroccan setting offers a substantive context for examining this distinction without assuming that results from other economies transfer unchanged. The World Bank Enterprise Survey (WBES) documents formal private establishments, management practices, financing conditions, and performance, and provides a useful contextual benchmark (World Bank Group, 2025). It does not, on the documentation inspected, establish a transaction-level panel of AI-enabled treasury deployment. This paper consequently proposes original data collection rather than presenting public business-environment indicators as if they measured the complete mechanism.

1.2. Research gap, questions, and contribution

Existing research links AI capability to firm-level outcomes and relates AI assimilation to organizational agility (Fosso Wamba, 2022; Mikalef & Gupta, 2021). However, those broad links do not themselves identify whether treasury-specific forecast improvements generate lower realized liquidity disruption after implementation costs, control constraints, and managerial responses are considered. The proposed gap is a missing integrated test: prediction, action, and resilience should be evaluated within a common, time-stamped research design. This is a design opportunity, not a claim that no related study exists.

The central question is whether a governed AI-enabled treasury package improves financial continuity relative to usual treasury practice. Three subsidiary questions concern the accuracy and calibration of forecasts, the contribution of financial agility to the mechanism, and the conditions under which management control strengthens or inhibits useful action. The paper develops a modular architecture, specifies

separately identified total-effect and mechanism models, and provides an executable synthetic test of estimator behavior. Its contribution is methodological and prospective; it does not claim a measured Moroccan treatment effect.

The analysis distinguishes four evidential layers throughout: verified literature, proposed hypotheses, executed simulation, and unexecuted field validation. This separation prevents statistical sophistication from being mistaken for empirical support. The numerical results should be read as an implementation diagnostic for a restricted data-generating process, not as estimates of gains that Moroccan firms can expect.

2. Background and Related Work

2.1. Management control, analytics, and dynamic capabilities

Management control encompasses more than financial reporting. The package perspective emphasizes the coexistence of planning, cybernetic, reward, administrative, and cultural controls (Malmi & Brown, 2008). In treasury, these elements translate into rolling budgets, cash thresholds, delegated authority, responsibility for variance explanations, and rules governing the use of financial headroom. This perspective motivates the proposed emphasis on control quality rather than simply control intensity. A rapid escalation routine can enable action, while redundant approval steps can delay it. The sign of a control–agility relationship should therefore be tested, not imposed.

The review of business intelligence and analytics in management accounting by Rikhardsson and Yigitbasioglu (2018) provides a direct bridge between information systems and decision support. It supports examining how analytical information is integrated into accounting routines; it does not imply that predictive accuracy alone is a sufficient organizational outcome. The present design extends that bridge by making cash-flow forecasts, subsequent managerial actions, and realized payment continuity separately observable. This separation permits failures to be localized: poor information, delayed action, unavailable resources, or inappropriate policy can each interrupt the pathway.

Dynamic-capability theory concerns the renewal and coordination of competences under changing conditions (Teece et al., 1997). For the proposed method, sensing corresponds to detecting emerging cash imbalances, seizing to selecting feasible funding or working-capital responses, and reconfiguration to adjusting routines and authority. These labels organize the architecture but do not constitute

measurements. Observable actions, timestamps, and outcomes are needed before a capability explanation becomes empirically informative.

Recent work remains relevant to this distinction. Al-Dhubaibi (2026) reports a cross-sectional survey of 161 Saudi medium and large firms linking dynamic management control, strategic management accounting, knowledge practices, learning, and agility. That evidence strengthens the motivation to examine complementary organizational arrangements. Its cross-sectional setting and different national context, however, do not identify the causal effect of an AI treasury package in Morocco. The proposed longitudinal design is intended to address part of this limitation rather than to contradict the study.

2.2. AI capability and the agility–resilience distinction

Mikalef and Gupta (2021) conceptualize and calibrate AI capability as an organizational resource configuration. This supports measuring data readiness, skills, and deployment routines alongside actual system use. Counting software licenses or labeling all automation as AI would conflate capability with procurement. The field protocol therefore records active forecasting coverage, model version, data availability, staff competence, and the share of forecasts reviewed in documented treasury meetings.

Fosso Wamba (2022) explicitly places organizational and customer agility between AI assimilation and performance. The proposed financial-agility construct is narrower: it concerns the time and feasible options needed to respond to cash risks. Its indicators exclude realized resilience outcomes to avoid defining the mediator with the dependent variable. A shorter approval time can coexist with worse payment continuity if managers choose an unsuitable response.

Duchek (2020) distinguishes anticipation, coping, and adaptation as elements of resilience. Treasury forecast quality primarily supports anticipation; feasible credit access and operational action support coping; revisions to policy after a disruption support adaptation. Financial resilience is consequently not operationalized as a cash ratio alone. Continuity of essential obligations is the primary outcome, with recovery and liquidity buffers analyzed separately.

2.3. Cash holdings, working capital, and econometric credibility

Cash buffers can absorb shocks, but higher holdings are not automatically an efficiency improvement. Bates et al. (2006) document changing cash holdings in US firms, while Deloof (2003) analyzes working-capital relationships in Belgium.

Together, these sources motivate examining precautionary liquidity and resource costs jointly. They do not justify maximizing idle cash or delaying supplier payment as universal prescriptions. The proposed intervention must preserve payment obligations while reducing preventable shortfalls and avoidable financing costs.

A separate literature clarifies the identification problem. Dynamic panel methods address persistence under moment restrictions (Arellano & Bond, 1991); excessive instruments can weaken empirical diagnostics (Roodman, 2009). Double/debiased machine learning combines orthogonal scores, nuisance estimation, and sample splitting (Chernozhukov et al., 2018), but it does not remove unobserved confounding. Likewise, staggered adoption comparisons require explicit counterfactual assumptions and estimators robust to heterogeneous effects (Callaway & Sant’Anna, 2021; Sun & Abraham, 2021). These methods are complementary tools with different assumptions, not interchangeable signals of methodological quality.

Table 1: Related-method landscape and the proposed design response.

Research stream	Useful contribution	Remaining limitation	Design response
Control packages	Complementary governance and accounting routines	Configuration is not a causal adoption effect	Baseline control audit; moderation; documented approvals
AI capability and agility	Organizational resources and adaptation mechanisms	Broad outcomes; potential common-source bias	Separate informants, transaction outcomes, event timing
Cash and working capital	Liquidity trade-offs and resource costs	Ratios alone omit disruption and recovery	Shortfall days plus costs, buffers, and recovery
Causal panel methods	Counterfactual and dynamic estimands	Assumptions may fail; instruments may proliferate	Randomized offer first; cohort DiD fallback; diagnostics

2.4. Evidence searches and limits of synthesis

The evidence searches used MCP-mediated Crossref, Google Scholar, and arXiv retrieval, supplemented by official institutional pages and repository documentation, accessed on 16 October 2026. Query families combined management control, treasury and working capital, AI capability, agility and resilience, and the econometric methods required by the design. Exact-title follow-up searches reconciled authors, publication versions, page ranges, and identifiers. Twenty scholarly works and one institutional dataset catalog were selected for the bibliography. Search results and normalized citation mappings are supplied with the reproducibility package.

Selection favored foundational theory, directly relevant management and finance research, and primary methodological sources. Broad AI-finance items with weak contextual relevance were excluded from the synthesis. This was a structured purposive review, not a preregistered systematic review or an exhaustive census. No PRISMA screening counts are invented. Full-text or detailed institutional passages were inspected when accessible; some publisher restrictions limited inspection to verified metadata and abstracts. Accordingly, the paper avoids cross-study effect-size pooling and unsupported numerical claims about AI performance. Repository documentation establishes implementation availability, not independent evidence of economic effectiveness.

3. Methods and Proposed Approach

3.1. Mechanism and falsifiable hypotheses

The proposed AI-enabled Treasury Resilience and Agility framework (A-TRA) treats AI as a governed decision support capability. Figure 1 summarizes the main mechanism. AI capability should improve the timeliness, calibration, and relevance of treasury information; better information may enable faster feasible action; and successful action may preserve continuity during a disruption. Management control conditions the translation from information to authorized action. Human authorization and fallback rules remain present throughout the chain.



Figure 1: Explanatory framework for A-TRA.

Solid arrows represent proposed mechanisms; dashed arrows denote moderation by management control quality. No arrow is an estimated causal effect. The diagram is an editable vector schematic.

Hypothesis H1 states that access to the governed package improves out-of-sample treasury forecast quality relative to usual practice. H2 states that the offer improves financial continuity over a prespecified six-month window. H3 states that financial agility is a temporally ordered candidate mechanism linking forecast quality to continuity. H4 states that baseline control quality modifies the total effect, with the direction estimated rather than constrained. H5 states that effects differ by baseline financing constraints and exposure to customer payment delays. H1 and H2 are confirmatory; H3–H5 are secondary and interpreted with multiplicity adjustment and appropriate identification cautions.

The framework deliberately allows null and adverse effects. More accurate forecasts may have no continuity benefit if financing alternatives are unavailable. Highly centralized controls may create delays despite useful information. Excessive sensitivity to adverse forecasts may produce costly precautionary borrowing. Empirical support requires jointly considering accuracy, action, cost, and service continuity, rather than retaining only favorable paths.

3.2. Design rationale and expected, not established, benefits

The architecture favors auditable tabular models and probabilistic outputs over a complex model chosen for novelty alone. Treasury users need to distinguish uncertainty about receipt dates from uncertainty about the amount of a payment, inspect reconciliation failures, and reproduce the information available at a decision time. Model performance should therefore be evaluated at several forecast horizons and compared with ordinary statistical and expert baselines.

The expected benefits are earlier detection of potential shortfalls, clearer allocation of responsibility, and better use of permitted liquidity options. These are proposed advantages

that remain contingent on data quality, calibration, management response, and implementation cost. Neither the literature synthesis nor the synthetic experiment demonstrates that these benefits occur in the target population.

3.3. Module architecture and implementation contract

A-TRA has five modules with explicit inputs, outputs, and failure responses (Table 2). The ingestion layer accepts bank balances, ledger entries, receivable and payable schedules, credit limits, and calendar information. Each record retains economic date, ingestion timestamp, source identifier, currency, and revision status. The feature store exposes only information available at the historical forecast origin. Late reconciliation and backdated corrections must never enter earlier model snapshots.

Table 2: Implementation-ready module boundaries and operational controls.

Module	Inputs and output	Default implementation	Failure response
M1: Reconciliation	Bank/ledger/schedule snapshots to validated cash state	Daily balance identity, duplicate checks, currency validation	Quarantine mismatch; manual reconciliation
M2: Forecasting	As-of features to 1-, 4-, 13-week cash distributions	Quantile regression and boosted-tree quantiles	Seasonal baseline if calibration or data gate fails
M3: Decision layer	Forecast scenarios and limits to feasible actions	Rolling constrained optimization with tail-risk penalty	Reject infeasible plan; escalate
M4: Control interface	Proposed actions to signed decisions and explanations	Role-based approval; override reasons; immutable log	No autonomous payment release
M5: Monitoring	Forecasts, realized receipts, actions, outcomes	Coverage, error, drift, outages, incident register	Recalibrate, roll back, or suspend

Invoice-level features include contractual due dates, customer-specific payment histories computed strictly before the origin, invoice age, dispute flags, and observed collection events. Enterprise-level features include trailing receipt and payment aggregates, known payroll dates, tax calendars, holidays, and approved credit headroom. Identifiers should be pseudonymized, and counterparty history should be computed

only within authorized data boundaries. All monetary modeling is performed in Moroccan dirhams after documented currency conversion when required.

3.4. Probabilistic forecasts and feasible cash decisions

For horizon h , the forecasting module estimates quantiles $b_{qit,h}(u)$ for $u \in \{0.05, 0.50, 0.95\}$ of cumulative net operating cash flow. Quantile regression provides a transparent baseline (Koenker & Bassett, 1978); boosted tree quantile models provide a nonlinear candidate. The inspected LightGBM repository documents an implementable tree-based framework. Version and commit identifiers must be frozen at deployment; documentation inspection is not a treasury benchmark.

The decision layer propagates cash scenarios through $B_{h+1} = B_h + R_h - P_h + d_h - r_h$, where B is cash, R receipts, P due payments, d approved borrowing, and r repayment. It minimizes financing cost, idle-cash cost, and a penalty on conditional value-at-risk of liquidity shortfalls. Credit ceilings, legal payment dates, minimum service levels, and approval rules are hard constraints. Thus, predicted cash improvement cannot be generated by silently postponing due obligations. Optimization weights, funding rates, and tail-risk tolerances are policy inputs requiring local calibration, not values estimated by the present simulation.

4. Field Methodology and Data Strategy

4.1. Population, recruitment, and prospective design

The proposed population is formal Moroccan non-financial enterprises with at least five employees and sufficiently traceable cash records. A planned confirmatory cohort comprises 600 enterprises, with an initial recruitment target of 750 to allow an assumed 20% loss before complete follow-up. These are planning assumptions, not observed sample counts. A feasibility pilot of 30 firms will assess extraction burden, record availability, consent, and the distinction between firm and establishment identifiers. The design should be scaled down or re-powered if recruitment fails; a smaller convenience sample must not be relabeled nationally representative.

Recruitment will stratify by sector, enterprise size, and region, while recording ownership, export activity, and baseline credit access. The study requires 12 months of historical records plus 24 months of prospective observation. Daily treasury records support forecasting; monthly outcomes support causal analysis; quarterly surveys measure organizational constructs. The finance director, treasury operator, and management controller provide different information where staffing

permits. For small firms with overlapping roles, the degree of informant overlap is recorded and sensitivity analyses exclude single-informant cases.

The preferred allocation randomizes the phased offer at enterprise level across five equal groups: access at months 0, 6, 12, and 18, and a waitlist receiving access after month 24. Randomization occurs within prespecified strata and is disclosed before implementation. All enterprises retain usual treasury services. The primary outcome window is months 0–5 after assigned offer, which is observable for every treated cohort. Later-horizon effects are reported only where the required comparison support exists.

Table 3: Data sources, feasibility, and limits of the proposed empirical validation.

Source	Intended use	Access requirement	Limitation
Prospective bank/ERP records	Cash trajectories, invoice timing, obligations, credit limits	Consent; enterprise agreement; secure extraction	Not yet obtained; heterogeneous systems
Repeated multi-informant survey	AI capability, controls, financial agility	Translation, cognitive pilot, repeated consent	Measurement error and response burden
WBES Morocco 2023 catalog	Population context; financing and management benchmarks	Published documentation; microdata subject to access terms	Not a verified AI treasury adoption panel
Official payment-delay reports	Institutional context and stress scenarios	Public institutional reports	Aggregates cannot replace firm causal outcomes
Synthetic panel in this paper	Estimator behavior and design sensitivity	Fully supplied generator and outputs	No observed Moroccan firms; not calibration data

4.2. External benchmarks and contextual evidence

The WBES catalog describes stratified sampling, formal-sector eligibility, establishment identifiers, and survey weighting (World Bank Group, 2025). A WBES establishment is not automatically identical to a multiestablishment enterprise in the proposed study. Any

benchmarking requires explicit unit harmonization and appropriate survey weights; no record linkage is assumed. Public payment-delay reporting by Morocco's Ministry of Economy and Finance offers institutional context, but this paper extracts no unverified numerical claim from it. No public source is presented as a substitute for the necessary treasury and rollout records.

4.3. Construct operationalization and primary outcomes

The primary resilience outcome is financial continuity: $Y_{it} = -100\text{Sit}/\text{Oit}$, where Sit is the number of observed days in month t on which unrestricted cash plus currently drawable committed credit is insufficient for essential obligations due that day, and Oit is the number of eligible observed days. Higher values are better; zero indicates no shortfall day. Essential obligations and drawable-credit conditions are frozen before allocation. A model-generated prediction is never substituted for an observed obligation or bank balance.

Table 4: Measurement specification. Proposed survey items are adaptations, not already validated Moroccan scales.

Construct	Observable indicators	Source and timing	Measurement safeguard
AI capability (A)	Data availability, analytical skills, deployment routines; active-use coverage	Quarterly staff survey plus usage logs	Distinguish capability, offer, and actual use
Control quality (C)	Rolling reviews, delegated limits, reconciliation, exception escalation	Baseline controller survey and audit	Freeze baseline C0 for total-effect moderation
Treasury quality (Q)	Scaled forecast error, quantile loss, interval coverage, reporting latency	Out-of-sample weekly forecasts	Equal origins and information across models
Financial agility (G)	Time to approved feasible response; available action set; execution latency	Time-stamped decisions plus quarterly survey	No resilience outcome embedded in items

Financial continuity (Y)	Negative percentage of shortfall days	Monthly reconciled bank and obligations data	No ratio calculated from missing days
Secondary resilience	Recovery time, covenant events, liquidity coverage, essential on-time payment share	Records; event adjudication	Prespecified shock onset and recovery rule

A minimum observation rule requires at least 80% eligible daily coverage within a month; otherwise, the outcome is missing rather than zero. The threshold is a proposed quality rule and will be sensitivity-tested at 90%. Secondary outcomes include the share of essential payments made on time, financing cost relative to sales, and days to recovery after a documented liquidity incident. Recovery is defined as the first 30 consecutive days without a shortfall after the incident, with right censoring at follow-up end. These outcomes help detect gaming through cash accumulation or shifted payment dates.

4.4. Measurement quality and statistical preparation

Survey preparation includes French and Arabic translation, back-translation, cognitive interviews, and pilot item review. Confirmatory factor analysis evaluates reflective subscales using ordinal estimators when appropriate. Reliability, factor loadings, and discriminant validity are reported; capability domains that form a resource bundle are analyzed as formative components rather than forced into a reflective factor. Longitudinal and language-group invariance are examined before score comparisons. Failure of invariance triggers domain-specific analyses, not automatic pooled means.

Common-source bias is reduced through informant separation, timing, and objective outcome records, following the methodological concerns reviewed by Podsakoff et al. (2003). Missingness is described by source and allocation group. Multiple imputation for covariates, inverse-probability sensitivity for attrition, and pattern mixture analyses are planned under explicitly stated assumptions. Outcome fabrication or mean replacement is prohibited. Winsorization, if used for extreme financial ratios, is fitted on training data only and reported alongside untrimmed robust analyses.

5. Econometric Model and Identification

5.1. Estimands and dynamic structural specification

The primary policy estimand is the six-month intention-to-treat effect of an assigned offer of the governed package, not the effect of a hypothetical one-unit increase in an AI questionnaire score. Let g denote an assigned offer month and $Y_{it}(g)$ potential continuity under an offer at g , versus $Y_{it}(\infty)$ under no offer through t . The primary contrast averages $Y_{i,g+e}(g) - Y_{i,g+e}(\infty)$ equally over event months $e = 0, 5$ and the four assigned offer cohorts. Training and governance support are part of the package. Non-use after an offer is noncompliance and remains in the assigned group.

For explanatory analysis, a dynamic panel specification separates baseline controls, realized treasury quality, and financial agility:

$$Q_{it} = a_0 + I_t + a_1 A_{i,t-1} + a_2 C_{i0} + a_3 A_{i,t-1} C_{i0} + a_4' X_{i,t-1} + u_{it}, \quad (1)$$

$$G_{it} = b_0 + m_t + b_1 Q_{i,t-1} + b_2 A_{i,t-1} + b_3 Q_{i,t-1} C_{i0} + b_4' X_{i,t-1} + v_{it}, \quad (2)$$

$$Y_{it} = c_0 + n_t + \rho Y_{i,t-1} + \beta A_{i,t-1} + \gamma Q_{i,t-1} + \delta G_{i,t-1} + \kappa A_{i,t-1} C_{i0} + \eta' X_{i,t-1} + \varepsilon_{it}. \quad (3)$$

Here i indexes enterprises, t months, and X includes prespecified lagged size, leverage, sales change, seasonality exposure, and financing conditions. Enterprise effects absorb time-invariant differences; calendar effects capture common shocks. Sector-by-time effects are a sensitivity analysis. A time-invariant main effect such as C_{i0} is absorbed by enterprise fixed effects and is not separately identified in the within estimator. Its interactions remain estimable when exposure varies.

Equations (1)–(3) describe mechanism-oriented associations unless additional identification assumptions are justified. Conditioning on Q and G , which can be affected by treatment, does not identify the total rollout effect. The primary randomized-offer analysis therefore excludes post-offer mediators from its adjustment set. This distinction prevents overcontrol from obscuring the very pathway that the framework proposes.

5.2. Moderation, mediation, and persistence

In a linear static mechanism without interactions, the serial indirect path is represented by $a_1 b_1 \delta$. With baseline control moderation it becomes $(a_1 + a_3 c)(b_1 + b_3 c) \delta$, conditional on $C_0 = c$, provided the posited ordering and linearity hold. A stable dynamic outcome can yield a long-run multiplier divided by $1 - \rho$, but only under time-invariant coefficients, $|\rho|$

< 1 , and valid dynamic identification. Neither formula should be interpreted as evidence that the assumptions hold.

Mediation requires stronger assumptions than random assignment of the offer, including absence of unmeasured mediator–outcome confounding and careful handling of treatment-induced confounders (Imai et al., 2010). The protocol measures proposed mediators at successive quarterly windows, reports cluster-bootstrap uncertainty, and performs sensitivity analysis. If those assumptions are not credible, the pathways are presented as descriptive mechanism evidence rather than causal indirect effects. Resilience in the lower tail is additionally examined using prespecified quantile analyses (Koenker & Bassett, 1978); they are not equated with heterogeneous individual treatment effects.

5.3. Primary causal analysis and staggered observational fallback

For randomized timing, the analysis respects assigned offer dates and strata, compares supported cohorts with not-yet-offered or waitlisted enterprises, and reports a prespecified weighted six-month contrast. Inference uses the actual allocation unit: enterprise, or the treasury-center group if enterprises share allocation. Folds and resampling retain whole allocation groups; power must be recalculated using their independent count. Randomization inference is a robustness check. Realized use is a secondary exposure. An instrumental-variable interpretation of actual AI use is not automatic: the offer includes training and governance, which may affect resilience without passing through a narrow software-use measure. Exclusion would need independent justification.

If random allocation is infeasible, the observational fallback defines G_i as the first verified active-deployment month and targets cohort–time effects:

$$ATT(g, t) = E[Y_{it}(g) - Y_{it}(\infty) | G_i = g], \quad t \geq g. \quad (4)$$

Conditional parallel trends, no anticipation, adequate overlap, consistent exposure, and no material interference are required (Callaway & Sant'Anna, 2021). Comparison enterprises must be not yet treated or never treated at each date; previously treated firms are not used as untreated controls. The primary aggregation weights the first six event months equally and cohorts by their prespecified baseline target-population shares; realized weights are reported. Event-time plots include simultaneous intervals and counts of contributing firms. Conventional two-way fixed-effects lead–lag coefficients are not treated as definitive causal estimates when timing and effects are heterogeneous (Sun & Abraham, 2021).

Pretrend checks, placebo offer dates, baseline balance, and adoption-propensity overlap diagnose aspects of credibility but cannot prove identification. Anticipatory training is logged, and sensitivity windows shift exposure to the earliest material intervention. Enterprises sharing treasury centers are assigned together or excluded from independent-unit inference. Customer–supplier links are recorded to assess spillovers; substantial spillovers change the estimand from a direct isolated effect to a network-dependent policy effect.

5.4. Nonlinear adjustment and dynamic robustness

The secondary panel score requires strict exogeneity conditional on full observed covariate and exposure histories and firm effects; lagging alone is insufficient. Define $eY = Y - \bar{Y}_i$, $eD = D - \bar{D}_i$, and fW as the within-transformed nonlinear basis and time indicators. With $\ell(fW) = E(eY | fW)$ and $m(fW) = E(eD | fW)$, estimate X :

$$\sum_{i,t} \{ \bar{D}_{it} - \bar{m}^{(-k(t))}(\bar{W}_{it}) \} [\bar{Y}_{it} - \bar{\ell}^{(-k(t))}(\bar{W}_{it}) - \theta \{ \bar{D}_{it} - \bar{m}^{(-k(t))}(\bar{W}_{it}) \}] = 0. \quad (5)$$

Whole independent allocation groups determine folds. The basis is transformed after constructing its nonlinear terms. Outcome-to-future-exposure feedback violates the within-score exogeneity condition; then this analysis is descriptive, while randomized-offer inference remains primary. For prospective prediction, rolling temporal separation is also required. Orthogonality protects against certain nuisance-estimation errors, not omitted variables or invalid timing (Chernozhukov et al., 2018). The DoubleML implementation offers reusable score and inference machinery (Bach et al., 2022); panel transformations and clustered resampling must still be specified and tested explicitly.

Difference GMM is reserved for the secondary lagged-outcome model under defensible sequential moment conditions (Arellano & Bond, 1991). The plan collapses instruments, restricts lag windows, reports instrument count relative to firms, and examines second-order serial correlation and overidentification diagnostics. A favorable Hansen statistic cannot certify validity under instrument proliferation (Roodman, 2009). Bias corrected fixed-effects and alternative lag structures provide robustness checks. None of these procedures replaces randomized timing or the counterfactual assumptions behind Equation (4).

6. Evaluation and Experimental Design

6.1. Prediction benchmarks, metrics, and temporal separation

Forecast evaluation uses identical eligible enterprises, forecast origins, and information sets. Models train on the earliest available history, tune through rolling-origin validation, and are evaluated on a locked final prospective period. A 13-week gap is imposed where necessary to prevent overlapping target windows from leaking into validation. Tests cover 1-, 4-, and 13-week cumulative net cash flow. The confirmatory H1 endpoint is enterprise-averaged four-week pinball loss, averaged over the three prespecified quantiles; cash-weighted and other-horizon metrics are secondary so large firms do not silently dominate. Models use a maximum of 50 tuning configurations per rolling validation split; the pilot fixes the final candidate before confirmatory testing.

Table 5: Prespecified experiments. All field experiments in this table are proposed, not executed results.

Experiment	Comparator or manipulation	Main metric	Intended inference
Forecast benchmark	Seasonal naïve, expert rolling budget, AR/ETS, regularized linear model, quantile model, boosted trees	MAE, scaled error, pinball loss, interval score	Out-of-sample predictive usefulness
Causal package effect	Assigned phased offer versus usual practice	Shortfall-day percentage; essential on-time payment	Six-month ITT effect and cost
Control complementarity	Treatment interaction with baseline control quality	Effect contrast and uncertainty	Prespecified moderator, not post-treatment subgroup
Architecture ablation	Remove calibration, customer-history features, action optimizer, or escalation interface	Accuracy, coverage, simulated decision cost	Contribution of a component, subject to feasible design
Stress and transport	Payment delays, credit tightening, data outage, sector holdout	Coverage, constraint violations, continuity	Robustness and boundary conditions

Qualitative mechanism	Divergent forecast/action/outcome cases	Coded decision sequences and explanations	Mechanism plausibility, not effect estimation
-----------------------	---	---	---

For quantile u , pinball loss is $Lu(e) = e\{u - \mathbb{I}(e < 0)\}$, where $e = y - \text{bq}(u)$. Reported uncertainty metrics include empirical 90% interval coverage, interval width, and interval score. Coverage alone rewards excessively wide intervals, so it is never the sole criterion. Mean absolute percentage error is avoided for near-zero or sign-changing cash flows. Loss-difference uncertainty uses resampling by enterprise and calendar blocks; equal-predictive-accuracy logic follows Diebold and Mariano (1994). Adaptive calibration is motivated by Gibbs and Candès (2021), but arbitrary regime-specific coverage is not promised under distribution shift.

6.2. Ablations, qualitative analysis, and decision criteria

Offline ablations retain the same train/test origins and hyperparameter budget. Removing a governance component from a live system may be unsafe; such ablations are first evaluated in shadow mode, without executing altered payment decisions. A realized-action replay assumes counterfactual receipts and credit availability are correctly modeled. Its output is a simulation, not a substitute for the randomized field outcome.

Approximately 20–30 purposively selected interviews are proposed across concordant and discordant cases: accurate forecasts with poor continuity, inaccurate forecasts with successful intervention, and unexpectedly large or adverse effects. Two analysts code decision timelines, financing restrictions, overrides, and perceived usefulness, maintain an audit trail, and examine disconfirming cases. Deployment is recommended only if forecast quality and continuity improve without unacceptable cost, supplier harm, or compliance failures. The precise economic loss threshold must be agreed with participating firms before outcome inspection.

7. Executed Statistical Validation: Simulation Methods

7.1. Data-generating process and reproducibility

The executed experiment tests estimator behavior under synthetic confounding and serial dependence. It does not simulate the full cash optimizer, latent-variable measurement model, randomized rollout, causal mediation, or staggered-adoption estimator. Those analyses remain field-validation tasks. The generator uses seed 20261016, nine scenarios, 500 independent replications per scenario, eight periods per firm, and four estimators, producing 18,000 fitted estimates. The

simulation is intentionally transparent and low-dimensional, permitting exact inspection of what the statistical evidence establishes.

For firm i and period t , independent standard normal draws generate a_i , X_{it} , V_{it} , U_{it} , v_{it} , and innovations ζ_{it} . The process is:

$$D_{it} = 0.6\alpha_i + 0.5X_{it} + 0.45(X_{it}^2 - 1) + 0.3\sin V_{it} + 0.5HU_{it} + v_{it}, \quad (6)$$

$$Y_{it} = \theta D_{it} + 0.4X_{it} + 0.35(X_{it}^2 - 1) + 0.25\sin V_{it} + \alpha_i + \lambda_t + e_{it} + 0.6HU_{it}, \quad (7)$$

$$e_{it} = \rho e_{i,t-1} + \zeta_{it}, \quad e_{i1} \sim N(0, (1 - \rho^2)^{-1}). \quad (8)$$

The deterministic λ_t sequence is evenly spaced between -0.2 and 0.2 . The hidden-confounding switch H is zero except in the omitted-confounder scenario. U is never provided to the estimators. D is a dimensionless continuous synthetic exposure, not actual adoption, and Y is an unbounded synthetic score, not the bounded field continuity measure. Thus, coefficient values are illustrative units rather than forecasted percentage-point policy effects.

7.2. Estimators, scenarios, and finite-simulation uncertainty

Pooled linear ordinary least squares (OLS) controls for X , V , and time effects but omits firm heterogeneity and nonlinear terms. Within-linear estimation removes firm means but still omits X^2 and $\sin V$. Within-nonlinear estimation adds the correct transformed basis and time effects. Firm-block cross-fit series estimation first demeans the outcome, exposure, nonlinear basis, and time indicators within firms, then estimates outcome and exposure nuisance regressions on four fifths of firms, predicts held-out firms, and solves Equation (5). This is an orthogonal series estimator with known relevant transformations, not a benchmark of a fitted neural network or boosted-tree deployment.

All intervals use a normal critical value of 1.96 and enterprise-clustered score covariance; OLS models use the finite-cluster correction specified in the supplied code. The experiment reports mean estimate, bias, root mean squared error (RMSE), 95% coverage, and rejection frequency. Reference scenarios vary firms across $N \in \{150, 300, 600\}$ with $\theta = 0.18$ and $\rho = 0.5$. Additional scenarios set $\theta = 0$, increase ρ to 0.8, introduce $H = 1$, or reduce θ to 0.06 at each sample size.

500 for mean estimates and p ;

Monte Carlo standard errors (MCSEs) are $s(b\theta)/\sqrt{p}$;

$p(1 - p)/500$ for coverage or rejection proportions. Zero observed coverage does not prove population coverage is mathematically zero; with 500 replications, it signals a severe observed failure. The released replication file preserves every estimate and standard error. Regression tests verify equivalence to explicit firm/time dummies, noiseless coefficient recovery, and the scalar cluster sandwich; all three pass.

8. Results and Evidence

8.1. Reference-scenario coefficient recovery

Table 6 reports executed results for 300 synthetic firms and eight periods with imposed coefficient $\theta = 0.18$, serial correlation $\rho = 0.5$, and no hidden time-varying confounder. Pooled linear OLS produces a mean coefficient of 0.700, and within-linear estimation produces 0.408. Removing time-invariant heterogeneity alone therefore does not recover the target when the remaining confounding relationship is nonlinear.

Table 6: Executed synthetic estimates: $N = 300$, $T = 8$, $\theta = 0.18$, 500 replications.

Estimator	Mean	Bias	RMSE	Coverage	Rejection
Pooled linear OLS	0.700	0.520	0.521	0.000	1.000
Firm FE, linear controls	0.408	0.228	0.229	0.000	1.000
Firm FE, nonlinear controls	0.182	0.002	0.023	0.946	1.000
Firm-block cross-fit series	0.182	0.002	0.023	0.938	1.000

Coverage is the proportion of nominal 95% intervals containing the imposed coefficient. Rejection is the two-sided rejection frequency for $H_0: \theta = 0$. These are Monte Carlo summaries, not enterprise survey estimates.

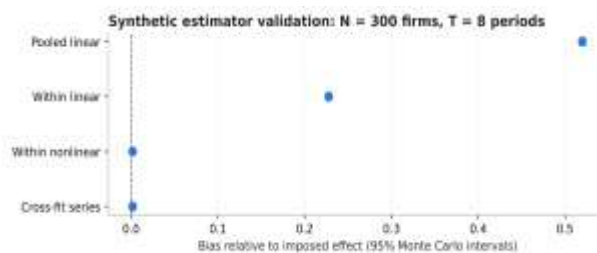


Figure 2: Executed synthetic bias relative to the imposed coefficient.

Bars show 95% Monte Carlo intervals for mean bias, not the confidence interval from a single regression. The nonlinear and cross-fit estimates overlap near zero. Numeric values are provided in Table 6.

8.2. Interpretation of uncertainty and method comparison

Within-nonlinear estimation gives a mean of 0.182, bias 0.002 after rounding, and coverage 94.6%. Cross-fit series estimation gives a mean of 0.182, bias 0.002, RMSE 0.023, and coverage 93.8%. Its bias MCSE is approximately 0.0010 and its coverage MCSE approximately 0.0108. Thus, the observed 1.2 percentage-point difference from nominal coverage is small relative to finite-simulation uncertainty. Figure 2 highlights the much larger errors from misspecification.

The cross-fit estimator does not demonstrate superiority over correctly specified nonlinear fixed effects in this design: their performance is nearly identical. That is an informative result, because the correct low-dimensional basis is deliberately available to both methods. The experiment supports specification discipline and clustered inference, not a general claim that more complex AI is better. All four estimators reject a zero effect in every reference replication; such rejection is not evidence of validity because the two misspecified estimators also reject while being badly biased.

8.3. Power, persistence, and hidden-confounding sensitivity

For the smaller imposed effect of 0.06, cross-fit power is 44.2%, 75.4%, and 96.8% at 150, 300, and 600 firms, respectively (Table 7). The corresponding power MCSEs are approximately 2.2, 1.9, and 0.8 percentage points. Under the null scenario, rejection is 3.8% with an MCSE of 0.9 percentage points. This does not establish exact size control beyond the tested process. Higher error persistence produces coverage of 95.2% when firmclustered uncertainty is used.

Table 7: Executed sensitivity results for firm-block cross-fit series estimation.

Scenario	Firms	True θ	Bias	Coverage	Rejection
Small150	150	0.06	-0.001	0.926	0.442
Small300	300	0.06	-0.001	0.944	0.754
Small600	600	0.06	-0.001	0.950	0.968
N300	300	0.18	0.002	0.938	1.000
Null	300	0.00	0.000	0.962	0.038
Persistent	300	0.18	0.000	0.952	1.000
Hidden	300	0.18	0.241	0.000	1.000

Small scenarios use $r = 0.5$, $H = 0$, and $\theta = 0.06$. Persistent uses $r = 0.8$. Hidden adds omitted U to exposure and outcome; all other displayed scenarios use $N = 300$, $r = 0.5$ unless indicated.

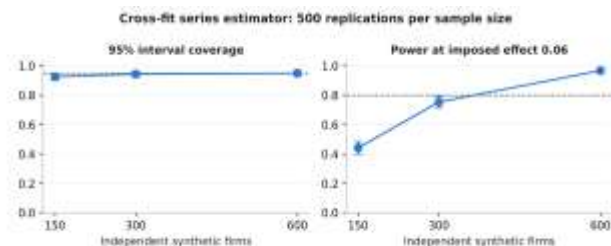


Figure 3: Executed small-effect sensitivity ($\theta = 0.06$).

Error bars are normal-approximation 95% Monte Carlo intervals for coverage and power. Dashed reference lines indicate nominal 95% coverage and a planning power target of 80%, respectively. Values derive from independent simulation replications, not observed firms.

The hidden-confounder scenario is the most consequential failure: cross-fit mean estimation rises to 0.421 when the imposed coefficient is 0.18, producing bias 0.241 and no covered intervals among 500 replications. Correctly specified observable nonlinear terms do not resolve that omission. This supports the protocol's preference for randomized access and

its caution about observational AI-adoption effects; it does not prove that an analogous hidden confounder exists in the field.

8.4. What these results do not establish

The results provide no measured forecast improvement, cost saving, mediation effect, or Moroccan resilience gain. The planning preference for 600 retained firms is informed by a transparent sensitivity example, not a field-calibrated power calculation. Pilot estimates of variance, compliance, missingness, rollout timing, cluster dependence, and a substantively meaningful continuity effect must replace the synthetic assumptions before final sample-size commitments.

9. Discussion

9.1. Theoretical and managerial implications

The framework connects control-package theory, AI capability, agility, and resilience through an observable financial decision sequence. Its main theoretical contribution is to distinguish information quality from response capability and realized continuity. This avoids a circular construct in which firms are called agile because they are resilient and resilient because they are agile. The emphasis on complementary routines is consistent with Malmi and Brown (2008), while the separation of anticipation, coping, and adaptation draws on Duchek (2020). The model remains open to contexts in which controls create delay or forecasting accuracy has little value because feasible financing alternatives are absent.

The executed evidence reinforces a narrower statistical implication. Firm effects and sophisticated nuisance adjustment are not universal solutions to selection into AI use. Nonlinear adjustment succeeds only when the relevant observed relationships are available, and hidden time-varying confounding defeats all adjusted estimators tested here. Consequently, researchers should invest in rollout design, timestamp quality, and outcome verification before expanding the model list. The recent control-agility literature (Al-Dhubaibi, 2026) motivates mechanism measurement but does not eliminate this identification requirement.

For managers, the appropriate objective is not maximum automation or minimum cash holdings. A useful system must preserve due payments, maintain approved liquidity headroom, and expose the uncertainty behind recommendations. A-TRA therefore limits automation to information processing and decision support until governance, authorization, and continuity criteria are satisfied. Model and infrastructure costs must be included in the economic

assessment. A statistically detectable improvement that costs more than its operational value is not a successful deployment.

9.2. Failure modes, risks, and mitigation

Data fragmentation is the first operational threat: incomplete bank feeds, duplicate invoices, disputed due dates, and currency inconsistencies can make a precise forecast misleading. Reconciliation gates, source timestamps, missingness reporting, and baseline fallback address this threat. Historical customer behavior can also be unstable during disruption. Monitoring error and interval coverage by horizon and sector supports recalibration, but no distribution-shift method guarantees useful interval widths under every crisis (Gibbs & Candès, 2021).

Behavioral and governance failures include automation bias, repeated unjustified overrides, excessive approval delay, and strategic manipulation of recorded obligations. Dual authorization, immutable due-date histories, override reviews, and independent outcome adjudication reduce these risks. Supplier harm is explicitly monitored through on-time payment and dispute outcomes so that treasury improvement is not achieved by transferring distress to counterparties. Cybersecurity or vendor outages require a tested manual continuity plan and exportable records, not merely a service-level promise.

Research risks include selective adoption, anticipation, attrition, common-source bias, and spillovers. Randomized timing is preferred; the observational fallback reports its assumptions and limitations. Separate informants, transaction outcomes, attrition sensitivity, and interference checks strengthen credibility. The proposed sample can still exclude informal enterprises and businesses without traceable records. External validity must therefore be argued from observed recruitment and implementation conditions rather than from national labels.

9.3. Limitations and next research decisions

This paper has no prospective field observations, and its synthesis is purposive rather than exhaustive. The simulation uses known nonlinear terms, only eight periods, no rare default events, and a simple additive hidden confounding process. It tests neither the cash optimizer nor the full mediation model. Accordingly, the next decisions are empirical: establish access and ethics arrangements, calibrate minimum meaningful effects, validate measurement, and register the allocation and analysis plan. More elaborate econometrics should follow those decisions rather than conceal their absence.

10. Conclusion

10.1. Contribution and evidential boundary

AI-enabled treasury management should be evaluated as a governed organizational intervention, not merely as a forecasting contest. A-TRA supplies an architecture connecting validated accounting data, calibrated forecasts, feasible liquidity actions, management control, and financial continuity. The proposed econometric strategy distinguishes randomized-offer effects, observational cohort effects, dynamic associations, and mechanism analyses. It also states where unmeasured confounding, post-treatment adjustment, or unsupported mediation assumptions can invalidate an interpretation.

The executed Monte Carlo experiment demonstrates recovery of an imposed coefficient under correctly specified observable adjustment and severe failure under an omitted time-varying confounder. It does not establish an empirical relationship in Moroccan firms. The article's actionable output is therefore a reproducible method and validation plan: prospective data collection and credible rollout assignment remain necessary before substantive claims about financial resilience can be made.

10.2. Deployment sequence and reporting commitments

Implementation proceeds through a feasibility pilot, a shadow forecasting phase, phased offer allocation, and locked outcome evaluation. The pilot resolves data ownership, extraction reliability, multilingual measures, and operational burden. Shadow operation tests reconciliation and uncertainty without initiating payments. Confirmatory deployment begins only after documented authorization and fallback tests. Stopping criteria include unresolved cash identities, persistent interval miscalibration, critical security incidents, and breaches of payment constraints. A suspended system retains its original assignment in intention-to-treat analysis.

The final field report must publish enterprise flow, exclusions and reasons, allocation adherence, missingness by group, baseline balance, supported event-time windows, uncertainty intervals, implementation costs, and adverse events. Confirmatory tests for forecast quality and continuity receive Holm adjustment within their declared family. Secondary mechanisms and heterogeneity analyses are labeled accordingly, with false-discovery rate control where an exploratory family is large. Null or unfavorable outcomes are reported rather than explained away through post-hoc model selection.

A. Reproducibility and Preregistration Appendix

A.1. Executable materials and scope

The supplied Python generator `data/run_simulation.py` reproduces all 18,000 synthetic estimates, summary CSV files, and both statistical figures using NumPy, pandas, and Matplotlib. It implements transparent series least-squares nuisance fits rather than claiming to have executed the DoubleML software or trained a treasury forecasting system. The manifest records the seed, versions, scenario count, and execution metadata. All tables of synthetic results are generated from those CSV outputs, and the vector framework source is provided. No private data or credentials are included.

For the future field study, a record schema will include enterprise identifier, as-of timestamp, economic date, source, currency, invoice amount, due date, receipt or payment date, opening cash, committed drawable credit, model version, forecast quantiles, action identifier, authorization time, and override reason. Training, validation, calibration, and test manifests will record exact firms and dates. Cryptographic hashes, an environment lock file, and repository commit identifiers are frozen before the confirmatory evaluation. Prospective forecasting tests require temporal separation in addition to enterprise-level separation.

A.2. Ethics, access, and author verification

No human-participant data were collected for this design-and-simulation paper. Before fieldwork, the investigators must obtain applicable institutional ethics authorization, informed consent, data-protection review, enterprise agreements, and a documented retention policy. Publication of the present draft does not certify that those approvals exist. Funding, conflicts of interest, author contributions, and the final spelling and ordering of author names require verification by the named authors before submission; none is invented here. The supplied names and affiliation identify the requested manuscript draft, not a claim that the authors have already conducted the proposed study.

References

- Al-Dhubaibi, A. A. S. (2026). Strategic pathways to agility: How dynamic management control systems, strategic management accounting and knowledge management foster organizational learning and adaptability [Advance online publication]. *Journal of Accounting & Organizational Change*, 1–33. <https://doi.org/10.1108/jaoc-07-2025-0281>
- Arellano, M., & Bond, S. (1991). Some tests of specification for panel data: Monte Carlo evidence and an application to employment equations. *The Review of Economic Studies*, 58(2), 277–297. <https://doi.org/10.2307/2297968>
- Bach, P., Chernozhukov, V., Kurz, M. S., & Spindler, M. (2022). DoubleML: An object-oriented implementation of double machine learning in Python. *Journal of Machine Learning Research*, 23(53), 1–6. <https://www.jmlr.org/papers/v23/21-0862.html>
- Bates, T. W., Kahle, K. M., & Stulz, R. M. (2006). Why do U.S. firms hold so much more cash than they used to? (No. 12534). National Bureau of Economic Research. <https://doi.org/10.3386/w12534>
- Callaway, B., & Sant’Anna, P. H. C. (2021). Difference-in-differences with multiple time periods. *Journal of Econometrics*, 225(2), 200–230. <https://doi.org/10.1016/j.jeconom.2020.12.001>
- Chernozhukov, V., Chetverikov, D., Demirer, M., Duflo, E., Hansen, C., Newey, W., & Robins, J. (2018). Double/debiased machine learning for treatment and structural parameters. *The Econometrics Journal*, 21(1), C1–C68. <https://doi.org/10.1111/ectj.12097>
- Deloof, M. (2003). Does working capital management affect profitability of Belgian firms? *Journal of Business Finance & Accounting*, 30(3-4), 573–588. <https://doi.org/10.1111/1468-5957.00008>
- Diebold, F. X., & Mariano, R. S. (1994). Comparing predictive accuracy (No. 169). National Bureau of Economic Research. <https://doi.org/10.3386/t0169>
- Duchek, S. (2020). Organizational resilience: A capability-based conceptualization. *Business Research*, 13(1), 215–246. <https://doi.org/10.1007/s40685-019-0085-7>
- Fosso Wamba, S. (2022). Impact of artificial intelligence assimilation on firm performance: The mediating effects of organizational agility and customer agility. *International Journal of Information Management*, 67, Article 102544. <https://doi.org/10.1016/j.ijinfomgt.2022.102544>
- Gibbs, I., & Candès, E. (2021). Adaptive conformal inference under distribution shift. *Advances in*

- Neural Information Processing Systems, 34, 1660–1672.
https://proceedings.neurips.cc/paper_files/paper/2021/hash/0d441de75945e5acbc865406fc9a2559-Abstract.html
- Imai, K., Keele, L., & Tingley, D. (2010). A general approach to causal mediation analysis. *Psychological Methods*, 15(4), 309–334.
<https://doi.org/10.1037/a0020761>
 - Koenker, R., & Bassett, G., Jr. (1978). Regression quantiles. *Econometrica*, 46(1), 33–50.
<https://doi.org/10.2307/1913643>
 - Malmi, T., & Brown, D. A. (2008). Management control systems as a package—opportunities, challenges and research directions. *Management Accounting Research*, 19(4), 287–300.
<https://doi.org/10.1016/j.mar.2008.09.003>
 - Mikalef, P., & Gupta, M. (2021). Artificial intelligence capability: Conceptualization, measurement calibration, and empirical study on its impact on organizational creativity and firm performance. *Information & Management*, 58(3), Article 103434.
<https://doi.org/10.1016/j.im.2021.103434>
 - Podsakoff, P. M., MacKenzie, S. B., Lee, J.-Y., & Podsakoff, N. P. (2003). Common method biases in behavioral research: A critical review of the literature and recommended remedies. *Journal of Applied Psychology*, 88(5), 879–903.
<https://doi.org/10.1037/0021-9010.88.5.879>
 - Rikhardsson, P., & Yigitbasioglu, O. (2018). Business intelligence & analytics in management accounting research: Status and future focus. *International Journal of Accounting Information Systems*, 29, 37–58.
<https://doi.org/10.1016/j.accinf.2018.03.001>
 - Roodman, D. (2009). A note on the theme of too many instruments. *Oxford Bulletin of Economics and Statistics*, 71(1), 135–158.
<https://doi.org/10.1111/j.1468-0084.2008.00542.x>
 - Sun, L., & Abraham, S. (2021). Estimating dynamic treatment effects in event studies with heterogeneous treatment effects. *Journal of Econometrics*, 225(2), 175–199. <https://doi.org/10.1016/j.jeconom.2020.09.006>
 - Teece, D. J., Pisano, G., & Shuen, A. (1997). Dynamic capabilities and strategic management. *Strategic Management Journal*, 18(7), 509–533.
[https://doi.org/10.1002/\(SICI\)1097-0266\(199708\)18:7<509::AID-SMJ882>3.0.CO;2-Z](https://doi.org/10.1002/(SICI)1097-0266(199708)18:7<509::AID-SMJ882>3.0.CO;2-Z)
 - World Bank Group. (2025). Morocco: World bank enterprise survey 2023 [Catalog metadata, reference MAR_2023_WBES_v01_M; microdata not analyzed]. Retrieved October 16, 2026, from <https://microdata.worldbank.org/index.php/catalog/6455>