Leading like Scientists

N. Ricardo Laverde-Cubillos
Assistant Professor, Insper
rlaverde@insper.edu.br

Gabriel Caser dos Passos
Research Assistant, World Bank
gabrielcaser9@hotmail.com

Abstract

Decision making under uncertainty can improve when leaders define problems, test hypotheses, and update beliefs as evidence accumulates. We capture this logic in a simple analytical model where scientific training and experience increases the returns to experimentation by improving Bayesian updating. We test its implications during COVID-19 using a regression discontinuity design around close leader elections, comparing municipalities narrowly won by STEM-trained leaders with those narrowly won by non-STEM leaders. The estimates indicate about 36 fewer deaths per 100,000 inhabitants under STEM-educated mayors. Consistent with the model, the advantage is strongest where mayors have longer experience in STEM occupations and appears linked to more feedback-rich policy adjustment.

KEYWORDS

Decision-Making, Scientific-approach, Leadership, STEM, Regression Discontinuity Design, Scientific Intensity

INTRODUCTION

Decision-making under uncertainty poses critical challenges for organizations because established routines and past performance often offer limited guidance in volatile environments. These challenges stem from bounded information processing capacities and information overload, which constrain attention, search, and inference (Simon, 1956; March and Simon, 1993; Cyert and March, 1963; Walsh, 1995). In this setting, recent work advances a scientific approach to managerial decision-making: leaders explicitly frame problems, derive testable hypotheses, and confront predictions with disciplined evidence (Camuffo et al., 2020; Felin and Zenger, 2017; Zellweger and Zenger, 2023). This ”entrepreneurs-as-scientists” (E-a-S) perspective integrates insights on how mental models guide strategic search (Csaszar and Levinthal, 2016; Ott et al., 2017) with evidence that disciplined experimentation accelerates learning under uncertainty (Camuffo et al., 2020; Felin and Zenger, 2017). The core premise is that explicit theorization combined with disciplined testing mitigates cognitive biases and improves strategic choices, particularly when uncertainty is severe and feedback is ambiguous (Camuffo et al., 2020; Zellweger and Zenger, 2023).

Although evidence from entrepreneurial and experimental settings suggests that scientific reasoning can improve performance (Camuffo et al., 2020, 2024a; Valentine et al., 2024; Novelli and Spina, 2024; Agarwal et al., 2025), some boundary conditions remain unspecified. First, *when* does scientific reasoning generate the greatest performance gains, particularly, how do different forms and levels of environmental uncertainty condition its value? Second, *how* does effectiveness vary with leaders' prior training and their behavioral consistency in applying scientific routines? Third, *what mechanisms* link scientific capabilities to performance outcomes in high-stakes organizational contexts beyond entrepreneurial experimentation? Existing randomized studies demonstrate clear benefits of scientific reasoning but do not differentiate between *capacity* (possessing scientific or analytical training) and *deployment* (consistently enacting theorization–experimentation–updating cycles), nor do they examine whether certain educational backgrounds, such as STEM (Science, Technology, Engineering and Math) training, provide distinct advantages under high uncertainty.

We investigate how a leader's scientific training and professional background can shape crisis management outcomes in environments characterized by high uncertainty. We rely on a quasi-experimental setting, close leader elections, to examine how STEM-trained leaders perform during crises marked by high uncertainty. This context enables us to separate pre-existing differences from the effects of STEM leadership while directly measuring how intensively leaders deploy scientific routines. Our study offers three contributions. First, we identify *when* scientific capabilities matter by showing that the value of analytical and hypothesis-driven reasoning depends on the level of environmental uncertainty. Second, we distinguish *capacity* (STEM-based analytical skill) from *deployment* (scientific intensity), showing that while STEM training endows leaders with strong analytical abilities (Ackerman et al., 2013; Wai et al., 2009), systematic engagement in theorization–experimentation–updating cycles is the behavioral mechanism through which these capacities translate into superior decision-making (Camuffo et al., 2020, 2024a; Novelli and Spina, 2024). Third, we provide novel evidence linking *scientific intensity* to crisis performance, isolating evidence-based reasoning from prior beliefs. Our findings suggest that STEM-trained leaders outperform peers during crises *only when* they consistently enact scientific routines, consistent with the comparative statics of our analytical framework.

The context of our study is municipal governance in Brazil during COVID-19, marked by high uncertainty, tight timelines, and varying technical expertise. As the main decision makers, mayors had to turn shifting evidence into action, choosing and enforcing public health measures, allocating scarce resources, and keeping citizens informed. Like corporate executives, elected leaders act as chief executives of complex organizations, setting priorities, coordinating departments, and bearing responsibility for outcomes (Bertrand and Schoar, 2003; Schoar et al., 2024). This makes political leadership a useful setting to study how leader behavior shapes strategic decisions. STEM-trained mayors provide a relevant contrast: their backgrounds are plausibly aligned with theorization, experimentation, and evidence-based updating—the core routines of a scientific approach (Camuffo et al., 2020; Coali et al., 2024). STEM education is also associated with strong analytical reasoning and facility with abstract problem solving (Ackerman et al., 2013; Wai et al., 2009), which can support disciplined hypothesis-driven decision-making under uncertainty (Zellweger and Zenger, 2023).

Empirically, we implement a regression discontinuity design comparing municipalities where STEM-trained candidates narrowly won to those where they narrowly lost, estimating sizable local effects at the cutoff (Lee, 2008). In our preferred specification, STEM-led municipalities experienced *36 fewer deaths per 100,000 inhabitants*, together with more active non-pharmaceutical interventions such as mobility restrictions. The estimated effects are sizable, statistically robust, and consistent across alternative bandwidths, STEM classifications, and specification checks. However, these results should be interpreted cautiously, as they reflect local average treatment effects at the electoral cutoff. The performance differences appear to be partly explained by greater experimentation in policy responses, as STEM-led municipalities implemented a broader and timelier set of non-pharmaceutical interventions. The effectiveness of the scientific approach is also associated with higher levels of *scientific intensity*, understood as the extent to which leaders engage in theorization, experimentation, and evidence-based updating, rather than with differences in resource endowments (Novelli and Spina, 2024; Coali et al., 2024).

Our findings therefore indicate that the effectiveness of scientific decision-making hinges primarily on scientific intensity, how consistently leaders enact theorization–experimentation–updating cycles, rather than on resource endowments per se (Camuffo et al., 2024b; Valentine et al., 2024). Although discretionary resources can underwrite search and innovation (Bourgeois III, 1981; George, 2005; Vanacker et al., 2017), our evidence points to methodological discipline as the binding margin. This pattern aligns with theory-based views emphasizing disciplined inference and belief testing (Felin and Zenger, 2017; Zellweger and Zenger, 2023) and with research on strategy-making under uncertainty (Ott et al., 2017; Levinthal, 2017). It clarifies when evidence-based routines create organizational value: when leaders apply them intensively and update decisions as signals accumulate. More broadly, the results suggest that realizing the entrepreneur-as-scientist logic requires the systematic enactment of scientific routines, with implications for leadership selection and organizational design in uncertain environments (Csaszar and Levinthal, 2016; Zellweger and Zenger, 2023).

The paper is structured as follows. Section 2 presents the related literature and the contribution of this study. Section 3 describes the data and variables used in the empirical analysis. Section 4 presents the empirical results. Sections 5 provide the discussion and conclusion.

BACKGROUND AND CONTRIBUTION OF THE STUDY

This section develops a framework for understanding why some leaders navigate crises more effectively than others. We argue that a scientific approach to leadership, characterized by explicit hypothesis formulation, disciplined testing, and evidence-based updating, offers a distinctive model of rational decision-making under acute uncertainty. The argument builds on the ”entrepreneurs-as-scientists” perspective and extends it to organizational leadership in crisis contexts. We propose that STEM training cultivates cognitive routines that make the scientific approach more accessible and behaviorally salient, but that effectiveness depends critically on scientific intensity, the consistency with which leaders enact theorization-experimentation-updating cycles. Building on these insights, we develop an analytical model of scientific leadership under uncertainty and derive hypotheses about when and why leaders with STEM backgrounds should exhibit performance advantages in turbulent environments.

A Scientific Approach to Organizational Decision-Making

The scientific approach to decision-making treats strategic choices as testable propositions and organizes action into four core routines, explicit theorization, operationalization, experimentation, and belief updating (Camuffo et al., 2020, 2024a). This logic, often described as the “entrepreneurs-as-scientists” (E-a-S) perspective (Felin and Zenger, 2009, 2017), represents a shift from intuition-led search toward theory-driven learning. Zellweger and Zenger (2023) use an approach that emphasizes how entrepreneurs use theorization to produce value from uncertainty by articulating causal beliefs, designing experiments to test them, and updating beliefs as evidence accumulates. It builds on evidence that structured managerial practices improve performance (Bloom and Van Reenen, 2007; Feldman et al., 2019) and integrates action-based learning traditions (Bingham and Eisenhardt, 2011; Thomke, 1998) with cognition-based theories emphasizing explicit hypothesis formation and guided evidence collection (Csaszar and Laureiro-Martínez, 2018; Walsh, 1995). In this view, disciplined experimentation provides decision makers with more diagnostic signals, enabling faster and more coherent adaptation in uncertain environments.

Although the scientific approach can be implemented as an organizational routine, it ultimately relies on individual cognitive capabilities. Articulating causal theories requires abstraction and analytical reasoning (Schmidt and Hunter, 1998), while translating theories into testable hypotheses draws on rule formation and working memory (Ackerman et al., 2013; Baddeley, 2020). Running disciplined tests depends on metacognitive monitoring and the ability to regulate intuitive judgments (Dane and Pratt, 2007), and effective updating requires cognitive flexibility to shift frames and alternate between exploration and exploitation as signals evolve (Laureiro-Martínez and Brusoni, 2018). These abilities form the microfoundations for the analytical model developed below, where scientific reasoning shapes evidence generation, belief revision, and adaptive choice.

How Educational Background Shapes Strategic Decision-Making

Research on Top Management Teams (TMTs), originating with Hambrick and Mason (1984), shows that executives' educational and professional backgrounds shape the cognitive frames through which they interpret strategic challenges. Education conveys not only technical knowledge but also deeper cognitive orientations (Finkelstein et al., 2009). Engineers, for instance, tend to adopt analytical and systems-oriented reasoning, whereas legal or humanities training often aligns with interpretive or normative styles. Because educational and career choices reflect stable preferences, the composition of a TMT influences how executives collectively scan environments and respond to uncertainty (e.g., Bigley and Wiersema, 2002; Chin and Hambrick, 2013; Crossland et al., 2014; Hambrick, 2007).

Early career experiences also leave lasting cognitive and behavioral imprints (e.g., Jung and Shin, 2019; Marquis and Qiao, 2020; Marquis and Tilcsik, 2013). These formative periods shape the interpretive lenses that leaders apply to novel problems, reinforcing the behavioral foundations of bounded rationality (Cyert and March, 1963). Extensive empirical evidence links executive characteristics, including educational specialization, functional background, and career variety, to systematic differences in decision-making and firm outcomes (e.g., Chatterjee and Hambrick, 2007; Crossland et al., 2014). This work provides the conceptual grounding for examining how STEM backgrounds translate into distinct leadership behaviors in crises.

STEM Training and Scientific Reasoning Capabilities

STEM education provides foundational training in the core routines of scientific reasoning, hypothesis formulation, quantitative analysis, and systematic problem solving, which promote comfort with uncertainty and disciplined belief revision (Lubinski et al., 2014; Ackerman et al., 2013; Stanovich and West, 2000). Mathematical and scientific curricula strengthen logical reasoning and the ability to navigate complex trade-offs (Wang et al., 2017). Engineering and technology programs emphasize problem decomposition, rapid prototyping, and iterative testing (Dym et al., 2005), practices that mirror the theorization–experimentation–updating cycle central to the scientific approach (Novelli and Spina, 2024). This training shapes durable cognitive schemas that leaders can deploy in high-stakes, ambiguous environments.

More generally, leadership capabilities are shaped not only by innate traits but also by training, practice, and structured developmental experiences (Day and Antonakis, 2012). Consistent with this view, early exposure to leadership roles predicts subsequent leadership attainment (Kuhn and Weinberger, 2005), and economic models conceptualize leadership as an accumulable skill with measurable returns (Lazear, 2012). Complementing these arguments, quasi-experimental evidence shows that reductions in early-life opportunities for structured training can lower the probability of assuming leadership roles later in life (Gagliardi and Mariani, 2022).

The Value of Scientific Reasoning Under High Uncertainty

Crisis contexts heighten the value of matching leaders' capabilities to the demands of the task. When uncertainty is extreme and feedback is noisy, the environment moves outside the range where routine expertise or intuition can reliably guide action. Under such conditions, STEM capabilities become binding constraints: leaders must convert fragmented information into structured problem representations, formulate testable conjectures, and interpret signals that arrive irregularly or with low diagnostic quality. Uncertainty therefore moderates the effectiveness of technical training by raising the premium on practices that transform ambiguity into informative evidence through controlled probes, rapid experimentation, and iterative learning (Eisenhardt, 1989; Kahneman and Klein, 2009). Scientific intensity captures this practical deployment of training—the extent to which leaders rely on hypothesis-driven reasoning, disciplined measurement, and recursive updating to navigate volatile conditions.

These dynamics also clarify why STEM leadership can yield superior crisis outcomes through several interlocking mechanisms. First, leaders with scientific reasoning skills tend to articulate sharper hypotheses, which improves initial sensemaking and problem decomposition when events evolve in unexpected ways (Weick, 1988; Danziger and Levav, 2017). Second, they design more diagnostic experiments and information-gathering strategies, accelerating learning in environments where the costs of delay are high and feedback is often delayed or distorted (Morgeson et al., 2015; Morgeson, 2005; Argote, 2012). Third, they engage in more disciplined belief updating, integrating new evidence even when it challenges established assumptions, thereby mitigating biases that commonly impair decision performance in crises (Kapoor and Wilde, 2023).

When STEM Capabilities Matter: Scientific Intensity and Decision Accuracy Under Uncertainty

The preceding discussion suggests that STEM-trained leaders are more likely to apply the routines of scientific reasoning, formulating hypotheses, running disciplined tests, and updating beliefs, when facing uncertainty. Yet capabilities alone do not ensure deployment. What remains unclear is how consistently leaders enact these routines and whether this consistency, scientific intensity, conditions the performance returns to STEM training under uncertainty.

We use scientific intensity to capture the frequency and consistency with which leaders engage in theorization–experimentation–updating cycles (Valentine et al., 2024). This construct reflects behavioral enactment rather than latent ability. Recent work shows that the benefits of scientific decision-making hinge on how intensively such routines are practiced, with systematic hypothesis articulation, structured tests, and disciplined updating each contributing to performance gains (Coali et al., 2024; Camuffo et al., 2024b).

Distinguishing between capacity (STEM training) and deployment (scientific intensity) is essential: analytical skill may exist but only improve decisions when it is exercised through sustained, task-relevant engagement and practice, while less formally trained leaders can partially compensate through disciplined repetition (Dane and Pratt, 2007). This logic implies complementarity—the returns to STEM training should be amplified when scientific intensity is high and muted when intensity is low—consistent with strategic cognition research showing that adaptive decision-making is built through repeated, feedback-rich use of analytical routines (Laureiro-Martínez and Brusoni, 2018; Park, 2024). In our framework, STEM training supports sharper hypothesis formation, scientific intensity increases the diagnostic value of testing, and higher-quality evidence enables more effective belief updating, so leaders who consistently enact these routines—testing, interpreting evidence, and updating priors as signals arrive—should therefore be better positioned to navigate uncertainty and scale timely interventions in crises (Kapoor and Wilde, 2023).

A formal model allows us to articulate these mechanisms precisely: How do scientific routines translate into differences in decision accuracy? How does environmental uncertainty moderate the value of these capabilities? Under what conditions do the costs of experimentation outweigh its informational benefits? The analytical framework below formalizes these trade-offs and derives comparative statics linking STEM background, scientific intensity, and environmental uncertainty to performance outcomes.

Analytical model

The organization faces an unknown state θ that can be high or low, and a leader chooses between two policies, A or B. Each policy is correct in one state and wrong in the other, so performance depends on how accurately the leader infers θ. Leaders start from a prior belief about the state and then run experiments that generate binary signals about whether θ is high or low. Following the Bayesian learning frameworks of Jovanovic and Nyarko (1995) and Cohen and Levinthal (1994), each signal is correct with probability $\alpha > 0.5$, and after observing a sequence of signals the leader updates beliefs using Bayes' rule. Decision accuracy therefore increases in two key features of the learning process: the number of experiments n that the leader runs and the diagnostic quality of each experiment, captured by α.

Leader attributes shape these two levers. Following research on ability-task matching (Braguinsky et al., 2012), the model distinguishes leaders with a STEM background ($D_S = 1$) from those without ($D_S = 0$), and introduces a continuous measure of scientific intensity I that captures the extent of prior scientific or technical experience. These attributes determine how many experiments leaders run and how informative they are. Formally, the number of experiments is written as

$$n(D_S, I) = n_0 + \kappa_1 D_S + \kappa_2 I + \kappa_3 D_S \times I,$$

and the diagnostic quality as

$$\alpha(D_S, I) = \tfrac{1}{2} + (\bar{\alpha} - \tfrac{1}{2})\left[\lambda_1 D_S + \lambda_2 I + \lambda_3 D_S \times I\right],$$

with $\bar{\alpha} \in (0.5, 1]$.[1]

A STEM background and greater scientific intensity thus increase both the cadence and the quality of testing, and the interaction terms κ_3 and λ_3 capture complementarity: having both a STEM background and high intensity yields more *and* better experiments than either alone.

Decision Accuracy Under Environmental Uncertainty: Decision accuracy q is a function of leader attributes and environmental uncertainty U. The role of U reflects empirical findings from crisis contexts, where high uncertainty and ambiguity make structured decision-making particularly valuable (Eisenhardt, 1989). When uncertainty is low (U close to 0), feedback is clear and all leaders can distinguish good from bad policies, so q is close to 1 regardless of background. When uncertainty is high (U close to 1), accuracy depends heavily on experimental capacity.

[1] This functional form makes signal quality endogenous to leader attributes, extending the information design framework of Kamenica and Gentzkow (2011). It formalizes how STEM training improves the capacity to design diagnostic tests, consistent with research on cognitive abilities (Stanovich and West, 2000).

Formally, a leader observes the information content per signal, $\Lambda(D_S, I) = \log[\alpha(D_S, I)/(1 - \alpha(D_S, I))]$, which is strictly increasing in both D_S and I. Environmental uncertainty dampens the effective informativeness of each signal by a factor $1/(1+\varphi U)$, where $\varphi > 0$ quantifies the strength of uncertainty's signal-degrading effect. The posterior log-odds shift after n experiments with net evidence Δk (correct minus incorrect signals) is thus

$$\text{log-odds shift} = \frac{n\Delta k \cdot \Lambda(D_S, I)}{1+\varphi U}.$$

Decision accuracy follows from Bayesian belief updating applied to this log-odds mechanism. The proposed (nonlinear) form is

$$q_{\text{exact}}(D_S, I, U) = \frac{1}{2} + \frac{1}{2}\tanh\left[\frac{\Delta k \cdot \Lambda(D_S, I)}{2(1+\varphi U)}\right].$$

Linear Approximation. For the comparative-static analysis, we adopt a linear approximation to the nonlinear accuracy function. This approach is appropriate under conditions that typically characterize early-to-mid crisis periods, when uncertainty is elevated but not overwhelming. The first assumption concerns the magnitude of posterior evidence: we require that the log-odds argument remains in a region where the sigmoid function behaves approximately linearly, allowing the substitution $\tanh(x) \approx x$. This reflects environments in which uncertainty affects judgment but does not push beliefs toward extreme values. A second assumption concerns the effect of uncertainty on signal quality. We assume that the influence of uncertainty on signals is sufficiently mild to justify a first-order expansion of the dampening term, replacing $1/(1+\varphi U)$ with its linear approximation $1 - \varphi U$. This is a standard simplification when uncertainty degrades signals without dominating them entirely.

The third assumption restricts the variation in signal quality so that the logit transformation $\Lambda(D_S, I) = \log[\alpha/(1-\alpha)]$ behaves approximately linearly in leader attributes. This excludes situations in which signals approach either perfection or randomness, conditions under which feedback structure breaks down and the approximation becomes unreliable. Collectively, these assumptions describe crisis phases in which leaders receive imperfect but still informative feedback—precisely the conditions under which linearization permits tractable comparative-static analysis.

Under these assumptions, the accuracy function simplifies to a clean linear form:

$$q(D_S, I, U) \;=\; q_0 \;+\; U\left[\beta_1 D_S + \beta_2 I + \beta_3 D_S \times I\right],$$

where β_i are composite parameters capturing the effect magnitudes. The critical feature is that *all effects on accuracy are multiplied by U*: when $U = 0$ (clear feedback), scientific background and intensity provide no advantage; when $U = 1$ (extreme uncertainty), the advantage is maximal. This multiplicative structure embodies the ability-task matching principle: scientific capacity becomes a binding constraint precisely when task demands (uncertainty) are highest. The linear form enables tractable comparative-static analysis while preserving the core economic mechanism linking scientific training to crisis performance.

Detailed derivations of these assumptions, their boundary conditions, and the corresponding nonlinear results are provided in Online Appendix B. The linear approximation is most accurate in the early stages of a crisis, when evidence is still accumulating, uncertainty is meaningful but not overwhelming, and experimentation generates signals that are informative without being extreme. It becomes unreliable once strong evidence pushes beliefs into the nonlinear region of the sigmoid, in environments where uncertainty so severely degrades feedback that signals lose informational value, or when signals become either nearly perfect or nearly random, causing the feedback structure itself to collapse.

Belief Convergence Mechanism: This accuracy function is rooted in Bayesian belief updating. Leaders update posterior beliefs through log-odds: each experiment with quality α shifts the log-odds by $\log[\alpha/(1-\alpha)]$. Critically, this shift is larger when α is higher (closer to 1). Consequently, STEM-trained leaders who design better tests (higher α) shift their log-odds faster per experiment, converging to accurate posterior beliefs more quickly. The parameter φ quantifies how environmental uncertainty corrupts signals, making high signal quality increasingly valuable as U rises.

Thus, $q(D_S,I,U)$ implicitly captures how leader attributes determine the rate of posterior belief convergence under uncertainty.

Expected net performance Π combines the accuracy benefits of correct decisions with the costs of experimentation. A leader earns $+1$ for a correct choice and -1 for an incorrect one; thus, if the probability of being correct is q, the expected payoff is $2q-1$.[2] Incorporating experimentation costs yields an overall performance measure that balances improved diagnostic accuracy against the resource and political costs of running tests:

$$\Pi(D_S,I,U) \;=\; 2q(D_S,I,U)-1 \;-\; c_n n(D_S,I),$$

where $c_n > 0$ is the cost per experiment and $n(\cdot)$ denotes the number of tests induced by scientific orientation. The first term captures how greater ability to diagnose raises expected payoff, while the second deducts the resource and political costs associated with experimentation. This formulation makes explicit the central trade-off: experimentation increases decision accuracy but is costly, so the optimal level depends on uncertainty and the leader's scientific orientation.

Differentiating this performance function with respect to leader attributes and uncertainty yields comparative-static results that underpin our hypotheses.

First, the marginal effect of a STEM background on performance is

$$\frac{\partial \Pi}{\partial D_S} \;=\; 2U\Delta q(\beta_1+\beta_3 I)-c_n(\kappa_1+\kappa_3 I).$$

The first term represents accuracy gains from having a STEM background under uncertainty; the second term represents higher experimentation costs. When uncertainty U is high and costs c_n are not too large, the accuracy gains dominate, especially when scientific intensity I is substantial. Mechanistically, STEM-trained leaders design better tests (higher α), which through the log-odds mechanism enables faster convergence to accurate posterior beliefs.

This leads to our first hypothesis:

[2] The expected payoff is $q(+1)+(1-q)(-1)=q-(1-q)=2q-1$.

H1. *In high-uncertainty crises, organizations led by top managers with a STEM background exhibit superior performance relative to organizations led by non-STEM top managers.*

Second, the model clarifies the mechanism through which STEM leadership improves performance. Define *effective experimentation* as

$$E(D_S, I) \;=\; n(D_S, I) \times \alpha(D_S, I),$$

the total volume of usable evidence entering the Bayesian update. A STEM background and higher scientific intensity increase n and α, which raises E, which in turn improves accuracy q and thus performance Π. The mediation operates through posterior belief convergence: the cumulative log-odds shift is $n\log[\alpha/(1-\alpha)]$ (which improves with E), enabling faster convergence to accurate posterior beliefs and improving decision accuracy. In other words, the causal chain runs from leader attributes to experimentation, then to posterior convergence, then to belief accuracy, and finally to outcomes:

$$(D_S, I) \;\to\; (n, \alpha) \;\to\; E \;\to\; \text{log-odds convergence} \;\to\; q \;\to\; \Pi.$$

This causal chain, from attributes through experimentation to posterior convergence to accuracy to performance, reflects decades of work on action-based learning (Bingham and Eisenhardt, 2011) and the scientific approach to management (Camuffo et al., 2024a). This logic motivates our second hypothesis:

H2. *In high-uncertainty crises, the performance advantage of organizations led by STEM-trained top managers is mediated by their more intensive use of disciplined scientific experimentation.*

Third, the cross-partial derivative of performance with respect to STEM background and scientific intensity is

$$\frac{\partial^2 \Pi}{\partial D_S \, \partial I} = 2U \Delta q \, \beta_3 - c_n \kappa_3.$$

When the complementarity parameters β_3 and κ_3 are such that the accuracy gains from combining STEM training and high intensity outweigh the extra experimentation costs, this cross-partial is positive. In that case, the performance benefit of a STEM background is larger when scientific intensity is high than when it is low. This captures the idea of *scientific intensity* as the degree to which leaders consistently apply the scientific method in practice: a STEM degree without sustained scientific experience may not suffice, whereas a STEM-trained leader with high intensity can translate training into frequent, high-quality experiments. Under high uncertainty, leaders must generate large log-odds shifts to reach confident posterior beliefs about the correct policy. Scientific intensity amplifies this convergence: higher n provides more learning opportunities, while higher α makes each opportunity more informative (larger log-odds shift per experiment). As uncertainty U increases, the positive cross-partial term becomes more important, magnifying this complementarity. This triple interaction, whereby the joint benefit of STEM training and intensity grows with environmental turbulence, is grounded in research showing that structured reasoning becomes increasingly valuable when feedback is noisiest (Eisenhardt, 1989).

H3. *Scientific intensity positively moderates the effect of STEM training in high-uncertainty crises: organizations led by STEM-trained top managers will exhibit a larger performance advantage when scientific intensity is high and a smaller (or negligible) advantage when scientific intensity is low.*

These three hypotheses translate the analytical model's comparative statics into empirically testable claims about when, how, and through which mechanisms STEM-trained leaders using a scientific approach outperform others in managing crises.

EMPIRICAL ANALYSIS

Context: Political leaders as CEOs

This section frames politically elected leaders, such as governors and mayors, as chief executives, akin to corporate CEOs in their decision-making authority and discretion. Grounded in Upper Echelons Theory (UET), which argues that leaders' backgrounds shape their strategic decisions, this study examines how these factors influenced governance during the COVID-19 pandemic. Despite operating in different domains, political and corporate leaders occupy top positions with significant power and discretion (Hambrick and Mason, 1984; Finkelstein et al., 2009). Governors, like CEOs, balance external responsibilities, representing their jurisdictions to higher levels of government, the media, and stakeholders, and internal duties, including managing agencies, enforcing laws, and overseeing budgets (Bigley and Wiersema, 2002).

This environment of competing demands and asymmetric information provides fertile ground for applying UET, which emphasizes leaders' "latitude of action" in shaping organizational responses. Although formal authority is institutionally defined, evidence shows that executives exercise considerable discretion in policymaking, and that their backgrounds and experience systematically influence strategic choices (Jones and Olken, 2005). In a crisis such as COVID-19, attributes like scientific training or political orientation may therefore help explain why leaders diverge in responsiveness and policy choices.

Brazil's municipal elections, held across 5,570 municipalities in the world's third-largest democracy, offer an ideal empirical setting to examine these dynamics. We now turn to the data used to evaluate these relationships.

Data

We assemble a comprehensive dataset that integrates granular electoral records, municipal demographic and socioeconomic indicators, employer–employee registers from RAIS to identify STEM occupational backgrounds, detailed epidemiological time series from SIVEP-Gripe, and municipal logs of non-pharmaceutical interventions (NPIs) that capture both policy timing and scope. STEM backgrounds are defined using the CBO-based taxonomy of Machado et al. (2021), coding candidates as STEM if they accumulated at least six months of verified employment in STEM occupations, with one-month spells incorporated to correct for zero-tenure reporting errors. Deaths due to COVID-19 are reported to the national SRAG surveillance system (Síndrome Respiratória Aguda Grave), which records all severe respiratory infections, including those resulting in hospitalization or death.[3] Demographic, socioeconomic, and political covariates are drawn from the 2010 Census, IEPS health indicators, and Power and Rodrigues-Silveira (2019). To examine mechanisms, we incorporate municipal NPI data from the National Confederation of Municipalities, which document local containment measures reported by 72.3% of mayors between May and July 2020.[4]

"INSERT FIGURE 1 HERE"

Before turning to the sample description, it is informative to examine raw differences in crisis trajectories across municipalities. Figure 1 plots the evolution of cumulative COVID-19 deaths over 2020 for municipalities led by STEM and non-STEM leaders. Both groups begin the year with similarly low fatality levels, but their trajectories diverge sharply from March onward as the pandemic intensifies. Municipalities governed by non-STEM leaders exhibit a much steeper and sustained increase in cumulative deaths, whereas those led by STEM-trained leaders show a consistently slower rise in fatalities. By December, this divergence results in a substantial gap: non-STEM municipalities approach nearly 100 cumulative deaths per 100,000 inhabitants, whereas STEM municipalities remain closer to 60. These descriptive patterns provide preliminary evidence that STEM-led municipalities experienced markedly better health outcomes during the first year of the pandemic.

"INSERT FIGURE 2 HERE"

[3] SRAG is the most comprehensive and standardized national source for monitoring severe respiratory outcomes in Brazil.

[4] These data capture the timing and scope of locally adopted NPIs, including mask mandates, activity restrictions, circulation limits, and transport controls. It is also used by from de Souza Santos et al. (2021)

Building on this suggestive evidence, we define our sample of empirical strategy on Brazilian municipalities where a STEM candidate finished first or second or third in mayoral elections with single-round contests overseen by the Tribunal Superior Eleitoral (Superior Electoral Court) [5]. Figure 2 shows the geographical distribution of municipalities with a STEM-trained candidate in the 2016 mayoral elections in Brazil. The map reveals that STEM candidates are scattered across the country, with no clear regional concentration or clustering of leadership with heterogeneous backgrounds. Notably, the lack of any distinct spatial pattern suggests that leaders with STEM backgrounds are dispersed across diverse regions, rather than being concentrated in specific urban or rural areas.

"INSERT TABLE 1 HERE"

The descriptive evidence in Table 1 shows that municipalities in the RD sample exhibit the variation required for our empirical tests. The sample is nearly evenly divided between jurisdictions that elected a STEM-trained leader and those that elected a non-STEM leader. 140 municipalities in the analytic sample selected a STEM mayor. Mortality outcomes also vary substantially, with cumulative deaths per 100,000 inhabitants displaying wide dispersion even within the restricted RD sample, an essential condition for identifying heterogeneous crisis performance.

A central feature of the data for evaluating Hypothesis 3 is the heterogeneity in scientific intensity[6] among STEM mayors. Experience ranges from zero to more than a decade, with a standard deviation of 2.91 years, even considering that Non-STEM mayors have all 0 years, providing the within-group variation needed to assess whether STEM leaders' effectiveness depends on the depth of prior scientific work. These patterns, dispersed mortality outcomes, and variation in STEM experience and a large number of municipalities motivate our close-election design, in which we compare municipalities narrowly won by STEM candidates to those narrowly won by non-STEM competitors.

[5]To capture potential effects during the leadership transition, we include outcomes through February 28th, 2021, ensuring full coverage of mortality patterns associated with the outgoing incumbents.

[6]Measured as the number of years a leader spent working in STEM occupations.

Methodology

Identifying the causal impact of a leader's professional background on organizational performance under uncertainty presents significant challenges. Simply comparing municipalities led by STEM-trained and non-STEM leaders risks biased inference, as performance outcomes may be correlated with unobserved factors that also shape the selection of leaders with scientific training. For instance, socioeconomic conditions, institutional capacity, or political environments may jointly influence both the likelihood of electing a STEM-trained mayor and the effectiveness of crisis management, confounding the relationship between leadership background and performance.

Although the regression discontinuity (RD) design helps account for municipality-specific unobserved variables, it does not automatically control for individual characteristics of the mayors themselves. Attributes such as educational attainment, political ideology, and previous experience may influence the policy decisions and outcomes we are studying (Bruce et al., 2022). Therefore, a key challenge in our research is ensuring that we control for these variables to better understand the mechanisms through which STEM backgrounds may influence epidemiological outcomes. To address this, we incorporate controls for mayors' personal characteristics that show discontinuities near the vote margin cutoff, following the methodological approach proposed by Calonico et al. (2019).

To estimate the effect of STEM leadership on epidemiological outcomes, we employ a sharp regression discontinuity (RD) strategy. This approach leverages the vote margin between the most-voted STEM and non-STEM candidates, allowing us to compare municipalities where a STEM candidate narrowly won or lost, thereby approximating a quasi-random assignment of treatment. The strategy enables a more credible estimation of the causal effects of STEM backgrounds on health outcomes, given the assumption that near the cutoff, municipalities are comparable in all respects except for the professional background of the elected mayor. Our empirical specification for this RD approach is detailed below.

$$y_{ms} = \alpha + \beta \cdot STEM_{ms} + f(Margin_{ms}) + \lambda_s + Z_{ms} + \varepsilon_{ms} \tag{1}$$

In the previous equation, m denotes a municipality and s denotes a state. The parameter $Margin_{ms}$ is the vote margin between the STEM mayor candidate in an election and a non-STEM candidate, where these were among the top 3 voted. This variable assumes a positive value if the winner was a STEM candidate and the second or third place was a non-STEM candidate. The Margin is a running variable in our estimation strategy, representing the vote difference between the first and second/third most-voted candidates. We constructed this variable using data from the Superior Electoral Court (TSE), assigning a positive value if a STEM candidate won and a negative value if a non-STEM candidate won.

Likewise, it assumes a negative value if the opposite takes place. The independent variable $STEM_{ms}$ receives the value 1 if the running variable $Margin_{ms} \geq 0$ and zero otherwise. We assume that $f(\cdot)$ is a flexible polynomial on both sides of the threshold. We estimate an optimal bandwidth using the non-parametric procedure from Calonico et al. (2014b). Our coefficient of interest β estimates the effect of electing a STEM candidate on outcome y_{ms}. We denote λ_s as a state election fixed-effect term. Finally, Z is a vector of control variables that include leader' personal characteristics, such as age, education, ideology and gender.

In all our specifications we consider state-fixed effects, as shown in Equation (1). We chose a specification with state-fixed effects because of a few reasons. First, it allows us to control for unobserved heterogeneity across states. Besides that, adding fixed effects turn our estimates more efficient without biasing them (Calonico et al., 2019). Moreover, we compare the performance of cities subject to similar state regulations since governors have autonomy in enforcing NPIs. Finally, Brazil is a large, continental country with many regions. Because of this, when adding state fixed effects we decrease the chances of comparing areas where COVID-19 waves behaved differently at the time we selected the outcomes. We applied different (7%, 8% and 9%) fixed winning margins to test the robustness of the optimal bandwidth and also to see the effect of the covariates inclusion on the same sample.

Recent methodological work cautions that close-election RD designs may not always cleanly isolate the effects of politicians' personal attributes. Marshall (2024) shows that when candidate characteristics are correlated with the vote margin, even in a narrow window around the cutoff, the as-if random assignment assumption may be weakened, potentially biasing estimates of individual trait effects. This insight highlights the need to verify the smoothness of covariates and assess balance in observable leader characteristics near the threshold. In line with these recommendations, our design examines whether attributes such as education, gende, age, and political alignment vary discontinuously at the cutoff to evaluate the plausibility of the identifying assumptions. These diagnostics help ensure that differences associated with STEM backgrounds are not driven by systematic sorting but are instead consistent with local quasi-experimental variation around the electoral margin.

Randomization Tests

The key assumption behind a sharp RD design is that the probability of treatment assignment changes discontinuously, from 0 to 1, at the cutoff. Therefore, it is possible to identify causal effects for the individuals whose scores are near this threshold (Cunningham, 2021). In our case, the key assumption is that the probability of $STEM_{ms} = 1$ changes discontinuously at $Margin_{ms}$ near 0. However, to interpret our coefficient of interest β as causal we need to satisfy the two validity conditions of an RD design: (i) the treatment does not affect baseline covariates; and (ii) there is no manipulation of the running variable near the threshold. The (i) condition means that our sample must be balanced between treated and untreated observations in their baseline characteristics. In Table 2 we show that all our baseline characteristics are balanced considering a 95% statistical margin.

"INSERT TABLE 2 HERE"

MAIN RESULTS

To test the first hypothesis, we examine whether selecting a top manager with a STEM background improves organizational performance during crises characterized by uncertainty. This hypothesis builds on the premise that leaders with STEM expertise are more likely to adopt systematic, data-driven approaches to problem-solving, enabling them to navigate complex and volatile environments more effectively. By leveraging analytical skills and evidence-based decision-making, STEM-trained leaders may enhance organizational resilience and adaptability, leading to improved outcomes during periods of heightened uncertainty.

“INSERT TABLE 3 HERE”

Table 3 reports the estimated impact of electing a STEM-trained mayor on COVID-19 mortality using local linear regression discontinuity (RD) designs. Across all five specifications, the point estimates are consistently negative, indicating that municipalities narrowly electing a STEM mayor experienced lower mortality than those narrowly electing a non-STEM mayor.

Column (1) presents the estimate obtained using the optimal bandwidth selected according to the Calonico–Cattaneo–Titiunik (CCT) procedure (Calonico et al., 2014a). This specification reflects the standard RD bias–variance trade-off: the CCT bandwidth minimizes mean squared error by correcting for boundary bias in local polynomial estimation. The coefficient of approximately -36 deaths per 100,000 inhabitants is statistically significant at the 5% level, suggesting a substantial mortality reduction near the cutoff. In relative terms, this magnitude corresponds to nearly half of the standard deviation of the death count in Table 1. Interpreted literally, this implies a substantial life-saving effect, and therefore calls for additional rigorous robustness checks and alternative specifications.

Columns (2)–(4) present estimates based on fixed bandwidths between 7 and 9 percentage points. The estimated effects are similar in size, ranging from about -28 to -39 deaths per 100,000 inhabitants, with p-values spanning 0.02 to 0.09. While the level of precision differs somewhat across specifications, the estimated effects are uniformly negative and economically sizable. These results suggest that the conclusions are robust to alternative bandwidth choices and specifications. Column (5), which uses the full set of effective observations, produces a smaller and statistically insignificant estimate, although its sign remains negative. The loss of precision aligns with the anticipated increase in variance when the bandwidth constraint is relaxed.

To complement, Figure 3 offers a visual complement to the RD estimates by depicting the discontinuity at the electoral threshold. The fitted regression lines, adjusted by mayor level covariates, reveal a clear downward shift in COVID-19 deaths precisely at the cutoff where the STEM candidate narrowly wins. Municipalities just above this threshold, those that barely elected a STEM mayor, exhibit lower mortality rates compared to observationally similar municipalities just below it that narrowly elected a non-STEM mayor. Although the underlying scatter is naturally noisy, the direction and magnitude of the jump align closely with the negative RD estimates in Table 3, reinforcing the conclusion that STEM-led municipalities experienced improved epidemiological performance even in the immediate neighborhood of the electoral margin.

"INSERT FIGURE 3 HERE"

Mechanism

Hypothesis 2 proposes that leaders with scientific training should differ from other leaders not only in their overall crisis performance but in the processes through which they make decisions. The scientific-approach perspective argues that scientific training equips leaders with habits of mind oriented toward disciplined experimentation: defining problems clearly, formulating testable predictions, collecting relevant evidence, and updating beliefs in light of new information. Under this view, leaders with scientific backgrounds should be more inclined to adopt policy responses that generate informative feedback and help reduce uncertainty over time. During a rapidly evolving public-health crisis, such as the COVID-19 pandemic, this decision style should manifest in greater deployment of Non-Pharmaceutical Interventions (NPIs), including mask mandates, activity restrictions, circulation limits, and transport controls, because these measures provide low-cost opportunities to test interventions, monitor their effects, and iteratively adjust policy. To evaluate this mechanism, we examine whether STEM-trained leaders adopted more extensive and stringent NPIs than their non-STEM counterparts.

Table 4 reports the RD estimates for the effect of narrowly electing a STEM candidate on the adoption of NPIs in 2020. The results indicate that municipalities led by STEM-trained leaders implemented a significantly higher number of NPIs overall, with the coefficient on Total NPI being positive and statistically significant at the 1% level. Among the specific policy categories, the most robust effects emerge for mask mandates and transport restrictions, both showing positive and highly significant coefficients. These patterns suggest that STEM leaders were more proactive in deploying measures that were widely recommended by scientific and epidemiological guidelines early in the pandemic.

"INSERT TABLE 4 HERE"

Other categories of NPIs, such as activity restrictions, circulation limits, and sanitary cordons, yield smaller and less precise estimates, though their signs remain consistent with the broader pattern of greater intervention intensity under STEM leadership. This heterogeneity is consistent with the notion that leaders with scientific training prioritize policies supported by clearer causal theories and faster feedback loops. The pattern suggests a behavioral orientation toward disciplined experimentation: rather than adhering to fixed protocols, STEM-trained mayors appear to treat policy implementation as a sequence of tests and incremental adjustments informed by observed outcomes. Figure 4 illustrates this robustness graphically by showing confidence intervals across alternative bandwidth choices.

In sum, the evidence supports Hypothesis 2, indicating that the mortality advantage observed in STEM-led municipalities may reflect a distinct decision-making style rooted in hypothesis formation, evidence evaluation, and gradual belief updating. This interpretation aligns with arguments that scientific training fosters cognitive routines conducive to experimentation, responsiveness, and adaptive management under uncertainty.

"INSERT FIGURE 4 HERE"

Moderator: Scientific Intensity

The third hypothesis (H3) examines whether the effect of a STEM background on crisis performance depends on *scientific intensity*. We proxy it with years employed in STEM occupations

(2003–2015), capturing sustained exposure beyond formal education. This moderator allows us to distinguish between mayors who are STEM-educated only and those who also have substantial STEM occupational experience. H3 predicts that, in high-uncertainty contexts, the mortality benefits of STEM leadership are amplified when scientific intensity is higher. In this OLS specification, β_3 captures whether scientific intensity strengthens or weakens the local treatment effect of electing a leader with a STEM background.

$$y_{msc} = \alpha + \beta_1\,\mathrm{STEM}_{msc} + \beta_2\,\mathrm{SciInt}_{msc} + \beta_3\,(\mathrm{STEM}_{msc} \times \mathrm{SciInt}_{msc}) + \lambda_s + \gamma Z_{msc} + \varepsilon_{msc}. \tag{2}$$

Table 5 provides evidence that supports the moderating mechanism in Hypothesis 3. The "Top 3," "Top 2," and "Top 1" samples progressively restrict the comparison to races with more presence of STEM candidates, and across the Top 3 and Top 2 models the main effect of a STEM background is positive and statistically significant, indicating that STEM education alone does not reduce mortality for leaders without STEM-related work experience. By contrast, the interaction between STEM background and scientific intensity is negative and statistically significant in all these specifications, implying that additional years of STEM occupational experience progressively offset and eventually reverse this baseline disadvantage. This pattern is consistent with evidence that task-relevant experience is often required for formal expertise to translate into better decisions under complexity (Dane and Pratt, 2007) and with the notion of "scientific intensity" as sustained exposure to scientific work that improves belief updating under uncertainty, and it provides direct empirical support for Hypothesis 3.

"INSERT TABLE 5 HERE"

In the Top 1 sample, which focuses exclusively on municipalities governed by leaders with STEM backgrounds, the coefficient on scientific intensity is large (between −4.8 and −5.5) and statistically significant at the 1 percent level. These estimates imply that each additional year of scientific experience is associated with approximately five fewer deaths per 100,000 inhabitants. Evaluated at one standard deviation of scientific intensity in the sample (2.91 years), the moderating effect reduces approximately 15 deaths per 100k inhabitants.Relative RD estimates in Table 3, which range from 27 to 36 fewer deaths, the magnitude of this moderating effect represents a sizable share of the total estimated impact corresponding to almost a half of the effect. Although these patterns should be interpreted with appropriate caution given limitations, the alignment between the moderation results and the RD effects provides suggestive evidence that the benefits of STEM leadership are not uniform across leaders. Instead, they are amplified for those with more sustained, career-based exposure to scientific work, consistent with Hypothesis 3.

Robustness Checks

To ensure that our core theoretical claims are empirically well founded, we assessed the robustness of the main results underlying our theory driven hypotheses. Because the argumentation in this study rests on identifying how scientific training shapes crisis decision making (H1–H4), it is essential to verify that the observed reductions in COVID-19 related deaths attributed to leaders with STEM backgrounds are not artifacts of measurement choices, sample composition, or model specification. Accordingly, we conducted a series of additional tests. The full set of robustness checks, including alternative classifications, model specifications, and covariate adjustments, is presented in Online Appendix A1.

In summary, the results demonstrate that electing a leader educated in STEM fields, regardless of the depth of their formal professional STEM experience, yields meaningful improvements in crisis management, as reflected in greater adoption of NPIs and reductions in COVID-19 deaths. These empirical insights set the stage for the discussion that follows, where we examine the broader theoretical and managerial implications of scientific leadership in complex, high stakes environments.

DISCUSSION

Our results suggest that leaders with scientific backgrounds exhibit superior performance in highly uncertain contexts, a pattern that remains consistent across multiple robustness checks.. Consistent with the logic of our analytical framework, municipalities narrowly electing STEM-trained mayors experienced substantially better crisis-management outcomes than those led by non-STEM peers, with estimated reductions in mortality on the order of 30–40 deaths per 100,000 residents. These effect sizes are both statistically detectable and substantively meaningful, and they are comparable to the leadership effects documented in other studies [7]. Supporting the proposed mechanism, STEM-trained leaders implemented non-pharmaceutical interventions more experimental and adjusted them more responsively as new information emerged, reflecting a more disciplined, evidence-aligned pattern of policy adaptation. The moderating effect of scientific intensity—captured by prior exposure to scientific occupations—provides further support: leaders with deeper scientific experience displayed stronger performance, underscoring the idea that scientific reasoning shapes how decision makers process signals and respond under uncertainty.

At the micro level, our results are consistent with the view that STEM-trained leaders tend to approach ill-defined problems with a more explicit, hypothesis-driven schema: they articulate testable conjectures, seek targeted evidence, and revise beliefs in light of new signals rather than relying primarily on intuition or pattern recognition (Ackerman et al., 2013; Laureiro-Martínez and Brusoni, 2018; Camuffo et al., 2020). Leaders without comparable scientific training may still update, but may do so more gradually or implicitly, especially when evidence is noisy or ambiguous. Viewed through this lens, the crisis-performance differences we document are suggestive of a deeper mechanism in which STEM expertise is associated with a stronger disposition toward interrogating assumptions, demanding evidence, and engaging in belief revision. These routines are plausibly reinforced through longer exposure to scientific norms and institutions that valorize empirical testing and falsification (Lubinski et al., 2014). This perspective also helps clarify why context matters: hypothesis-testing routines are most likely to translate into superior outcomes when decision makers have latitude to revise actions over time and when feedback from tests is interpretable enough to support learning (Camuffo et al., 2020; Novelli and Spina, 2024). In our pandemic setting—characterized by rapidly evolving information and scope for policy recalibration—these potential advantages are most likely to be expressed in observable performance

[7] See Bruce et al. (2022) and Cabral et al. (2021)

differences.

The findings also broaden the theoretical meaning of the *scientific approach to decision-making*. While prior work has examined this logic mainly in firms and entrepreneurial settings, our results suggest that the same routines—structured theorization, disciplined testing, and evidence-based updating—can enhance leadership more generally. This view is consistent with organizational-theory research showing that leaders' cognitive representations shape how they search, learn, and respond in ambiguous environments (Levinthal, 2017). At the organizational level, the results support the idea that scientific expertise equips leaders to interpret noisy signals, adjust policies as conditions evolve, and employ more systematic, evidence-informed interventions. This interpretation aligns with the scientific-approach and "entrepreneurs-as-scientists" perspectives, which emphasize hypothesis-driven experimentation and disciplined belief updating under uncertainty (Camuffo et al., 2020; Felin and Zenger, 2017).

Conceptually, the scientific approach clarifies how micro-level cognitive routines scale into organizational adaptability. Identifying scientific reasoning as a distinct cognitive schema that guides hypothesis generation and belief revision contributes to a micro-founded account of adaptive leadership under the theory-based approach (Felin and Zenger, 2017). At the individual level, it aligns with research on cognitive flexibility (Laureiro-Martínez and Brusoni, 2018; Eggers and Kaplan, 2013), emphasizing leaders' capacity to articulate hypotheses, interrogate assumptions, and update beliefs as signals evolve. At the organizational level, it illustrates how structured theorization and empirical updating generate coherent responses to uncertainty. These insights motivate the analytical model, which formalizes how leaders form priors, gather evidence through costly experimentation, and update beliefs via Bayesian inference (Jovanovic and Nyarko, 1995). By specifying belief formation, experimentation costs, and the role of signal precision, the model conceptualizes leadership as an iterative learning process that fosters adaptation and resilience across diverse crisis settings.

This mechanism aligns with upper echelons theory, which posits that leaders' observable backgrounds proxy for deeper cognitive bases and values that shape attention, interpretation, and action under uncertainty (Hambrick and Mason, 1984; Hambrick, 2007). Building on this view, we introduce scientific intensity to capture not merely whether a leader has a STEM credential, but the depth of sustained engagement in scientific or technical occupations—contexts in which work routinely involves problem decomposition, measurement, and iterative error correction—thereby strengthening the link between STEM background. Because such routines accumulate through practice, higher intensity should reflect thicker domain-specific schemas and procedural knowledge that shape what decision makers notice, how they categorize signals, and how readily they revise priors when evidence shifts (Dane and Pratt, 2007; Laureiro-Martínez and Brusoni, 2018). These experience-based advantages should matter most under high uncertainty, when effective action relies on leaders' heuristics and cognitive structures under time pressure (Eisenhardt, 1989).

Our study also carries several broader types of limitations that open avenues for future research. First, our close-election regression discontinuity design yields a local comparison at the vote-margin cutoff, so the estimates speak most directly to elections in which STEM and non-STEM candidates were nearly tied (Lee and Lemieux, 2010). Second, because the treatment is a leader attribute, the estimated discontinuity may conflate STEM training with other correlated characteristics that also shape both electoral competitiveness and crisis performance (Marshall, 2024). Third, our data cannot isolate which elements of STEM formation—analytic training, technical expertise, or exposure to scientific norms—drive the observed differences, nor whether selection into STEM fields reflects underlying dispositions that also influence leadership choices (Coenen et al., 2021). Fourth, we observe outcomes and policy responses but have no direct evidence on leaders' internal theorizing so any account of cognitive mechanisms should be read as suggestive rather than directly measured. More generally, these limitations point to a wider agenda for examining how different leader backgrounds perform across environments: not only when scientific training provides an advantage, but also when it may impose constraints in contexts that reward political skills, negotiation, or symbolic action (Ferris et al., 2007).

From a policy perspective, these findings connect to broader evidence that managerial types exert systematic effects on organizational outcomes (Schoar et al., 2024; Acemoglu et al., 2022;

Bertrand and Schoar, 2003). Our results suggest that scientific expertise is one such attribute: leaders with STEM training appear better equipped to process information, update beliefs, and adapt policies when conditions are volatile. This logic is relevant for the wider strategic management literature, as crises impose similar demands on leaders, requiring rapid interpretation of ambiguous signals, swift coordination, and adaptive responses under uncertainty (Eisenhardt, 1989). It also resonates with evidence that the use of research depends not only on the availability of information but on institutional capacities and incentives to generate, interpret, and act upon evidence (Hjort et al., 2021). Strengthening analytic capabilities—through training, expert collaboration, and mechanisms that embed evidence use into routine decision processes—may therefore complement the advantages associated with scientific expertise and improve organizational responsiveness in dynamic environments.

CONCLUSION

This study shows that leaders with scientific training can perform more effectively in high-uncertainty environments, highlighting how scientific intensity shapes the way decision makers generate hypotheses, design tests, and update beliefs. By combining quasi-experimental evidence with an analytical model, we identify the micro-level mechanisms—structured experimentation and disciplined Bayesian updating—through which scientific reasoning enhances adaptive performance (Camuffo et al., 2020, 2024a). The findings also align with the view that managerial cognition is a systematic source of heterogeneity in organizational outcomes, consistent with arguments that leaders' cognitive endowments influence how they search, learn, and respond to uncertainty (Levinthal, 2017). In addition, the results are consistent with recent evidence that the effectiveness of scientific decision routines depends on the surrounding organizational and informational context. Their benefits arise when leaders operate in environments that support disciplined experimentation—where information can be structured, diagnostic tests can be executed, and feedback can be incorporated into subsequent choices without major frictions (Novelli and Spina, 2024). These insights contribute to a broader understanding of leadership as an iterative learning process grounded in disciplined inquiry rather than intuition or positional authority, and they invite further research on how scientific human capital shapes adaptation in complex and evolving settings.

References

Acemoglu, D., He, A., and Le Maire, D. (2022). Eclipse of rent-sharing: The effects of managers'business education on wages and the labor share in the us and denmark.

Ackerman, P. L., Kanfer, R., and Beier, M. E. (2013). Trait complex, cognitive ability, and domain knowledge predictors of baccalaureate success, STEM persistence, and gender differences. *Journal of Educational Psychology*, 105(3):911.

Agarwal, R., Bacco, F., Camuffo, A., Coali, A., Gambardella, A., Msangi, H., Sonka, S., Temu, A., Waized, B., and Wormald, A. (2025). Does a theory-of-value add value? Evidence from a randomized control trial with Tanzanian entrepreneurs. *Organization Science*, 36(2):601–625. Publisher: INFORMS.

Argote, L. (2012). *Organizational Learning: Creating, Retaining, and Transferring Knowledge*. Springer.

Baddeley, A. (2020). Working memory. *Memory*, pages 71–111. Publisher: Routledge.

Bertrand, M. and Schoar, A. (2003). Managing with Style: The Effect of Managers on Firm Policies. *Quarterly Journal of Economics*, 118(4):1169–1208.

Bigley, G. A. and Wiersema, M. F. (2002). CEO Selection and Firm Performance: A Study of the Top Management Teams. *Strategic Management Journal*, 23(6):637–654.

Bingham, C. B. and Eisenhardt, K. M. (2011). Rational heuristics: the 'simple rules' that strategists learn from process experience. *Strategic management journal*, 32(13):1437–1464. Publisher: Wiley Online Library.

Bloom, N. and Van Reenen, J. (2007). Measuring and explaining management practices across firms and countries. *Quarterly Journal of Economics*, 122(4):1351–1408.

Bourgeois III, L. J. (1981). On the measurement of organizational slack. *Academy of Management review*, 6(1):29–39. Publisher: Academy of Management Briarcliff Manor, NY 10510.

Braguinsky, S., Klepper, S., and Ohyama, A. (2012). High-Tech Entrepreneurship. *The Journal of Law and Economics*, 55(4):869–898.

Bruce, R., Cavgias, A., Meloni, L., and Remígio, M. (2022). Under pressure: Women's leadership during the COVID-19 crisis. *Journal of development economics*, 154:102761. Publisher: Elsevier.

Cabral, S., Ito, N., and Pongeluppe, L. (2021). The disastrous effects of leaders in denial: evidence from the COVID-19 crisis in Brazil. *Available at SSRN 3836147*.

Calonico, S., Cattaneo, M. D., Farrell, M. H., and Titiunik, R. (2019). Regression Discontinuity Designs Using Covariates. *The Review of Economics and Statistics*, 101(3):442–451.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014a). Robust data-driven inference in the regression-discontinuity design. *The Stata Journal*, 14(4):909–946.

Calonico, S., Cattaneo, M. D., and Titiunik, R. (2014b). Robust nonparametric confidence intervals for regression-discontinuity designs. *Econometrica*, 82(6):2295–2326. Publisher: Wiley Online Library.

Camuffo, A., Cordova, A., Gambardella, A., and Spina, C. (2020). A scientific approach to entrepreneurial decision making: Evidence from a randomized control trial. *Management Science*, 66(2):564–586. Publisher: INFORMS.

Camuffo, A., Gambardella, A., Messinese, D., Novelli, E., Paolucci, E., and Spina, C. (2024a). A scientific approach to entrepreneurial decision-making: Large-scale replication and extension. *Strategic Management Journal*, 45(6):1209–1237. Publisher: Wiley Online Library.

Camuffo, A., Gambardella, A., and Pignataro, A. (2024b). Theory-driven strategic management decisions. *Strategy science*, 9(4):382–396. Publisher: INFORMS.

Chatterjee, A. and Hambrick, D. C. (2007). CEO characteristics and firm outcomes: The influence of personal values and experiences. *Strategic Management Journal*, 28(12):1199–1221.

Chin, M. K. and Hambrick, D. C. (2013). Executive Narcissism, Corporate Social Responsibility, and Firm Performance. *Leadership Quarterly*, 24(1):5–28.

Coali, A., Gambardella, A., and Novelli, E. (2024). Scientific decision-making, project selection and longer-term outcomes. *Research Policy*, 53(6):105022. Publisher: Elsevier.

Coenen, J., Borghans, L., and Diris, R. (2021). Personality traits, preferences and educational choices: A focus on STEM. *Journal of Economic Psychology*, 84:102361. Publisher: Elsevier.

Cohen, W. M. and Levinthal, D. A. (1994). Fortune Favors the Prepared Firm. *Management Science*, 40(2):227–251.

Crossland, C., Zyung, J., Hiller, N. J., and Hambrick, D. C. (2014). Ceo career variety: Effects on firm-level strategic and social novelty. *Academy of Management Journal*, 57(3):652–674.

Csaszar, F. A. and Laureiro-Martínez, D. (2018). Individual and organizational antecedents of strategic foresight: A representational approach. *Strategy Science*, 3(3):513–532. Publisher: INFORMS.

Csaszar, F. A. and Levinthal, D. A. (2016). Mental representation and the discovery of new strategies. *Strategic Management Journal*, 37(10):2031–2049. Publisher: Wiley Online Library.

Cunningham, S. (2021). Causal inference. In *Causal Inference*. Yale University Press.

Cyert, R. M. and March, J. G. (1963). *A Behavioral Theory of the Firm*. Prentice Hall.

Dane, E. and Pratt, M. G. (2007). Exploring intuition and its role in managerial decision making. *Academy of Management Review*, 32(1):33–54.

Danziger, S. and Levav, J. (2017). When does analysis improve decisions? *Annual Review of Psychology*, 68:263–286.

Day, D. V. and Antonakis, J. (2012). Leadership: past, present, and future. *The nature of leadership*, pages 3–25. Publisher: Thousand Oaks: Sage Publications.

de Souza Santos, A. A., Candido, D. d. S., de Souza, W. M., Buss, L., Li, S. L., Pereira, R. H., Wu, C.-H., Sabino, E. C., and Faria, N. R. (2021). Dataset on SARS-CoV-2 non-pharmaceutical interventions in Brazilian municipalities. *Scientific data*, 8(1):1–6. Publisher: Nature Publishing Group.

Dym, C. L., Agogino, A. M., Eris, O., Frey, D. D., and Leifer, L. J. (2005). Engineering design thinking, teaching, and learning. *Journal of Engineering Education*, 94(1):103–120.

Eggers, J. P. and Kaplan, S. (2013). Cognition and capabilities: A multi-level perspective. *Academy of Management Annals*, 7(1):295–340.

Eisenhardt, K. M. (1989). Making fast strategic decisions in high-velocity environments. *Academy of Management Journal*, 32(3):543–576.

Feldman, M. S., Pentland, B. T., D'Adderio, L., and Lazaric, N. (2019). Beyond routines: Systematic approaches to organizational performance. *Journal of Organizational Behavior*, 40(2):195–211.

Felin, T. and Zenger, T. R. (2009). Entrepreneurs as theorists: On the origins of collective beliefs and novel strategies. *Strategic Entrepreneurship Journal*, 3(2):127–146. Publisher: Wiley Online Library.

Felin, T. and Zenger, T. R. (2017). The theory-based view: Economic actors as theorists. *Strategy Science*, 2(4):258–271.

Ferris, G. R., Treadway, D. C., Perrewé, P. L., Brouer, R. L., Douglas, C., and Lux, S. (2007). Political skill in organizations. *Journal of management*, 33(3):290–320.

Finkelstein, S., Hambrick, D. C., and Cannella, A. A. (2009). *Strategic Leadership: Theory and Research on Executives, Top Management Teams, and Boards*. Oxford University Press, New York, NY.

Gagliardi, L. and Mariani, M. (2022). Trained to lead: Evidence from industrial research. *Strategic Management Journal*, 43(4):847–871. Publisher: Wiley Online Library.

George, G. (2005). Slack resources and the performance of privately held firms. *Academy of management Journal*, 48(4):661–676. Publisher: Academy of Management Briarcliff Manor, NY 10510.

Hambrick, D. C. (2007). *The influence of CEOs on strategic decisions and organizational performance*. Blackwell Publishing.

Hambrick, D. C. and Mason, P. A. (1984). Upper echelons: The organization as a reflection of its top managers. *Academy of Management Review*, 9(2):193–206.

Hjort, J., Moreira, D., Rao, G., and Santini, J. F. (2021). How research affects policy: Experimental evidence from 2,150 brazilian municipalities. *American Economic Review*, 111(5):1442–80.

Jones, B. F. and Olken, B. A. (2005). Do leaders matter? National leadership and growth since World War II. *The Quarterly Journal of Economics*, 120(3):835–864. Publisher: MIT Press.

Jovanovic, B. and Nyarko, Y. (1995). A Bayesian learning model fitted to a variety of empirical learning curves. *Brookings Papers on Economic Activity: Microeconomics*, pages 247–305.

Jung, J. and Shin, Y. (2019). Recent advancements in the domain of early career experiences and their impact on cognitive patterns. *Journal of Business Psychology*, 34(2):123–135.

Kahneman, D. and Klein, G. (2009). Conditions for intuitive expertise: A failure to disagree. *American Psychologist*, 64(6):515–526.

Kamenica, E. and Gentzkow, M. (2011). Bayesian persuasion. *American Economic Review*, 101(6):2590–2615.

Kapoor, R. and Wilde, D. (2023). Peering into a crystal ball: Forecasting behavior and industry foresight. *Strategic Management Journal*, 44(3):704–736.

Kuhn, P. and Weinberger, C. (2005). Leadership skills and wages. *Journal of Labor Economics*, 23(3):395–436. Publisher: The University of Chicago Press.

Laureiro-Martínez, D. and Brusoni, S. (2018). Cognitive flexibility and adaptive decision-making: Evidence from a laboratory study of expert decision makers. *Strategic Management Journal*, 39(4):1031–1058. Publisher: Wiley Online Library.

Lazear, E. P. (2012). Leadership: A personnel economics approach. *Labour Economics*, 19(1):92–101. Publisher: Elsevier.

Lee, D. S. (2008). Randomized experiments from non-random selection in US House elections. *Journal of Econometrics*, 142(2):675–697.

Lee, D. S. and Lemieux, T. (2010). Regression discontinuity designs in economics. *Journal of economic literature*, 48(2):281–355.

Levinthal, D. A. (2017). Mendel in the C-suite: Design and the evolution of strategies. *Strategy Science*, 2(4):282–287. Publisher: INFORMS.

Lubinski, D., Benbow, C. P., and Kell, H. J. (2014). Life paths and accomplishments of mathematically precocious males and females four decades later. *Psychological Science*, 35(12).

Machado, C., Rachter, L., Schanaider, F., and Stussi, M. (2021). STEM classification in the formal labor market in Brazil. *IDRC*.

March, J. G. and Simon, H. A. (1993). *Organizations*. John wiley & sons.

Marquis, C. and Qiao, X. (2020). The role of early career experiences in shaping cognitive patterns in leadership. *Leadership Quarterly*, 31(4):556–570.

Marquis, C. and Tilcsik, A. (2013). The effect of early career experiences on CEO decision-making and organizational outcomes. *Academy of Management Journal*, 56(6):1779–1799.

Marshall, J. (2024). Can close election regression discontinuity designs identify effects of winning politician characteristics? *American Journal of Political Science*, 68(2):494–510.

Morgeson, F. P. (2005). The external leadership of self-managing teams: Intervening in the context of novel and disruptive events. *Journal of Applied Psychology*, 90(3):497–508.

Morgeson, F. P., Mitchell, T. R., and Liu, D. (2015). Event system theory: An event-oriented approach to the organizational sciences. *Academy of Management Review*, 40(4):515–537.

Novelli, E. and Spina, C. (2024). Making business model decisions like scientists: Strategic commitment, uncertainty, and economic performance. *Strategic Management Journal*, 45(13):2642–2695. Publisher: Wiley Online Library.

Ott, T. E., Eisenhardt, K. M., and Bingham, C. B. (2017). Strategy formation in entrepreneurial settings: Past insights and future directions. *Strategic Entrepreneurship Journal*, 11(3):306–325. Publisher: Wiley Online Library.

Park, C. H. (2024). Finding a road less traveled: Combining analysis and intuition to develop novel problem formulations. *Strategic Management Journal*, 45(11):2368–2392.

Power, T. J. and Rodrigues-Silveira, R. (2019). Mapping ideological preferences in Brazilian elections, 1994-2018: a municipal-level study. *Brazilian Political Science Review*, 13. Publisher: SciELO Brasil.

Schmidt, F. L. and Hunter, J. E. (1998). The validity and utility of selection methods in personnel psychology: Practical and theoretical implications of 85 years of research findings. *Psychological bulletin*, 124(2):262. Publisher: American Psychological Association.

Schoar, A., Yeung, K., and Zuo, L. (2024). The effect of managers on systematic risk. *Management Science*, 70(2):815–833.

Simon, H. A. (1956). Rational choice and the structure of the environment. *Psychological review*, 63(2):129. Publisher: American Psychological Association.

Stanovich, K. E. and West, R. F. (2000). Advancing the rationality debate. *Behavioral and brain sciences*, 23(5):701–717. Publisher: Cambridge University Press.

Thomke, S. H. (1998). Simulation, learning and R&D performance: Evidence from automotive development. *Research Policy*, 27(1):55–74. Publisher: Elsevier.

Valentine, J., Novelli, E., and Agarwal, R. (2024). The theory-based view and strategic pivots: The effects of theorization and experimentation on the type and nature of pivots. *Strategy Science*, 9(4):433–460. Publisher: INFORMS.

Vanacker, T., Collewaert, V., and Zahra, S. A. (2017). Slack resources, firm performance, and the institutional context: Evidence from privately held E uropean firms. *Strategic management journal*, 38(6):1305–1326. Publisher: Wiley Online Library.

Wai, J., Lubinski, D., and Benbow, C. P. (2009). Spatial ability for STEM domains: Aligning over 50 years of cumulative psychological knowledge solidifies its importance. *Journal of Educational Psychology*, 101(4):817–835.

Walsh, J. P. (1995). Managerial and organizational cognition: Notes from a trip down memory lane. *Organization science*, 6(3):280–321. Publisher: INFORMS.

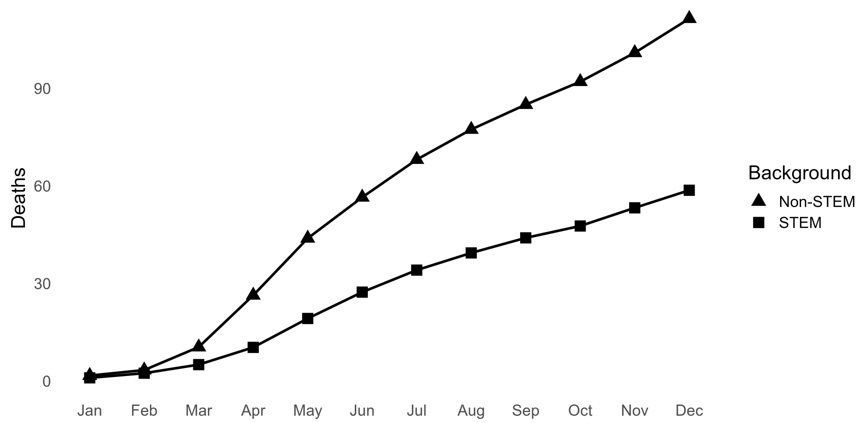
Wang, M.-T., Ye, F., and Degol, J. L. (2017). Who chooses STEM careers? Using a relative cognitive strength and interest model to predict careers in science, technology, engineering, and mathematics. *Journal of Youth and Adolescence*, 46(5):1805–1820.

Weick, K. E. (1988). Enacted sensemaking in crisis situations. *Journal of Management Studies*, 25(4):305–317.

Zellweger, T. and Zenger, T. (2023). Entrepreneurs as scientists: A pragmatist approach to producing value out of uncertainty. *Academy of Management Review*, 48(3):379–408. Publisher: Academy of Management Briarcliff Manor, NY.

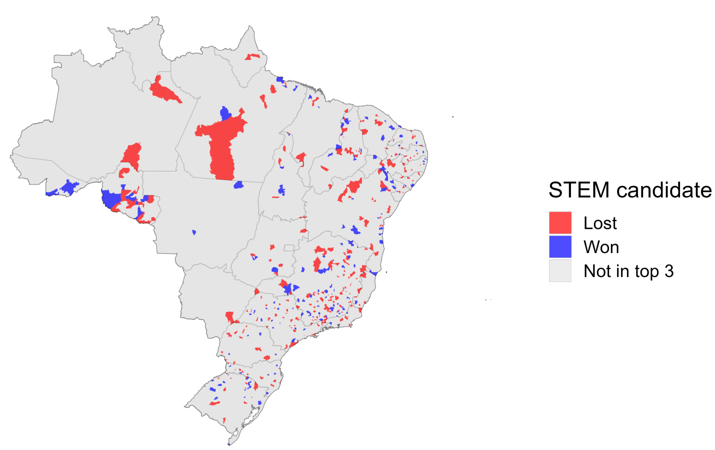
FIGURES AND TABLES

Figure 1: Cumulative Deaths by Covid-19 by Leadership Background



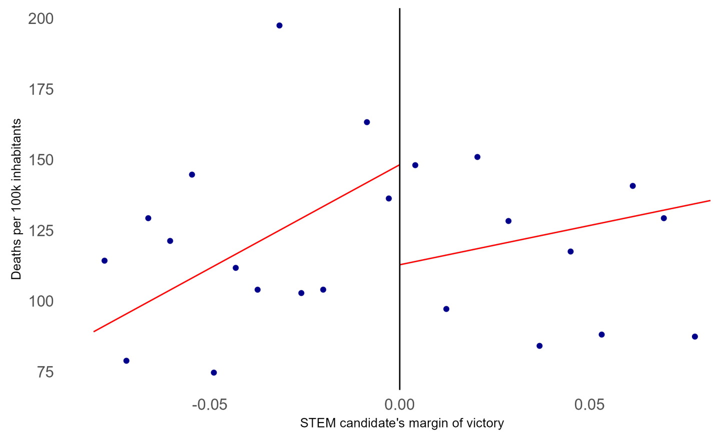
Note: This figure shows the evolution of deaths from COVID-19 (SRAG or suspected SRAG) throughout 2020 by mayoral background. The data come from Sistema de Informação de Vigilância Epidemiológica da Gripe (SIVEP-Gripe) and includes all municipalities with at least one death by Covid.

Figure 2: Municipalities with a STEM Candidate (2016)



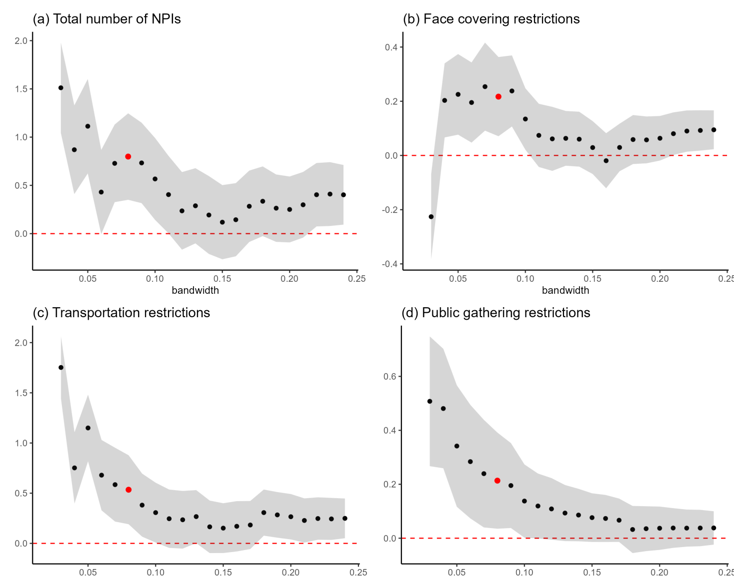
Note: In this figure, we colored all municipalities in our 2016 sample, i.e., where a STEM candidate was among the top three most voted. In red are the municipalities where the STEM candidate lost, in blue are the municipalities where the STEM candidate won, and in gray are all municipalities with no STEM candidate among the top three most voted.

Figure 3: Impact of STEM Leadership on Performance (Deaths per 100k inhabitants)



Note: This figure reports the RD estimated impact of mayors with a scientific background on deaths by COVID-19 per hundred thousand inhabitants. Municipalities chosen were those that held ordinary elections in selected years (2016) where the mayor was elected in the first round and among the top three most voted, one was a STEM candidate and the others a Non-STEM candidate.

Figure 4: Impact of STEM Leadership on Non-Pharmaceutical Interventions (NPIs) Using Different Bandwidths



Note: This figure reports the RD estimated impact of mayors with a scientific background on non-pharmaceutical interventions (NPIs). The red dots represent the optimal MSERD bandwidth, as calculated by Calonico et al. (2014), confidence intervals are 95%. The sample includes municipalities with ordinary elections in 2016 where a STEM and a non-STEM candidate were among the top three and the winner was elected in the first round. The horizontal axis varies the winning-margin bandwidth between the winner and the second/third candidate. Estimates come from our baseline RD specification: local linear, state fixed effects, uniform kernel, and mayor controls.

Table 1: Summary Statistics

Variable	N	Min	Mean	Max	SD
	(1)	(2)	(3)	(4)	(5)
Outcome					
Deaths per 100k inhabitants	413	0.00	126.15	370.77	72.12
Leader Characteristics					
STEM Background	413	0.00	0.40	1	0.49
Scientific Intensity (years in STEM)	413	0.00	1.41	12.01	2.91
Female	413	0.00	0.09	1.00	0.29
Age	413	26.00	49.98	79.00	10.54
Education (level)	413	1.00	6.05	7.00	1.50
Incumbent when elected	413	0.00	0.16	1.00	0.37
Party ideology	413	−0.69	0.26	0.76	0.38
Policy Interventions					
Cordon sanitaire	300	0.00	0.55	1.00	0.50
Face covering required	296	0.00	0.95	1.00	0.21
Closure of non-essential activities	297	0.00	0.77	1.00	0.42
Gathering prohibition	297	0.00	0.98	1.00	0.15
Public transport restriction	295	0.00	0.47	1.00	0.50
Number of Non-Pharma Interventions	294	1.00	3.72	5.00	0.90

Note: This table aggregates the summary statistics of all the observations used in the study (413). Municipalities chosen were those that held ordinary elections in selected years (2016) where the mayor was elected in the first round and among the top three most voted, one was a STEM candidate and the others a Non-STEM candidate. NPI data has null values since not all the mayors responded to the survey.

Table 2: Baseline Characteristics – RD Estimates on Deaths per 100k inhabitants

	(1) PC Income	(2) Log Pop.	(3) HDI	(4) Density	(5) % Male Pop
STEM Backg.	4.33	0.23	0.02	55.82	-0.35
SE	[2.64]	[0.33]	[0.01]	[38.93]	[0.46]
p-value	0.10	0.48	0.13	0.15	0.45
	(6) % Health Spend.	(7) Doctors	(8) CHA Prog.	(9) Hosp. Beds	(10) Mun. Ideology
RD Estimator	2.28	0.08	1.77	1.09	-0.02
SE	[1.31]	[0.13]	[6.63]	[50.07]	[0.03]
p-value	0.08*	0.57	0.79	0.98	0.49
Eff. N obs.	109	109	109	109	109
Bandwidth	8.2%	8.2%	8.2%	8.2%	8.2%

Note: This table reports the RD estimated impact of mayors with scientific background on demographic baseline characteristics. Municipalities chosen were those that held ordinary elections in selected years (2016) whose mayor was elected in the first round and among the top three most voted was a STEM candidate and a Non-STEM one. All specifications use state fixed-effects, uniform kernel and optimal bandwidths calculated following Calonico et al. (2014). We report robust p-values, estimates and standard errors.

Table 3: Impact of STEM Leadership on Performance - Deaths per 100k Inhabitants (RDD)

	(1)	(2)	(3)	(4)	(5)
STEM Backg.	-35.61	-38.75	-27.83	-32.80	-14.66
SE	[16.58]	[16.94]	[16.58]	[15.94]	[9.23]
p-value	0.03**	0.02**	0.09*	0.04**	0.11
Eff. N obs.	109	96	107	121	413
Bandwidth	8.2%	7%	8%	9%	100%
Type of Bandwidth	Optimal	Fixed	Fixed	Fixed	Fixed

Note: This table reports the RD estimated impact of mayors with a scientific background on deaths by COVID-19 per hundred thousand inhabitants. The selected municipalities are those that held ordinary elections in specific years (2016), where the elected mayor won in the first round and among the top three most voted candidates, one had a STEM background and the others did not. All specifications include state fixed effects, mayor characteristics controls, linear polynomial and use the uniform kernel. Estimation (1) uses optimal bandwidths calculated following Calonico et al. (2014), while the remaining use fixed vote margin difference between the STEM and Non-STEM candidate among top three voted. We report robust bias-corrected p-values, coefficients, and standard errors.

Table 4: STEM Leadership and the use of evidence-based policy (Non-Pharmaceutical Interventions)

	Total NPI	Masks	Activities Res.	Gathering Res.	Transportation Res.	San. barriers
	(1)	(2)	(3)	(4)	(5)	(6)
STEM Backg.	0.87	0.27	0.16	0.20	0.52	−0.26
	[0.27]	[0.09]	[0.16]	[0.10]	[0.20]	[0.17]
	<0.01***	<0.01***	0.33	0.05*	<0.01***	0.12
Eff.N.obs.	76	78	78	78	76	79
Bandwidth	8.2%	8.2%	8.2%	8.2%	8.2%	8.2%

Note: This table reports the RD estimated impact of mayors with a scientific background on the adoption of non-pharmaceutical interventions (NPIs) in 2020. The selected municipalities are those that held ordinary elections in 2016, where the mayor was elected in the first round, and among the top three most voted, one was a STEM candidate and the others a Non-STEM candidate. All specifications use state fixed effects, uniform kernel, and optimal bandwidths calculated following Calonico et al. (2014). The interventions listed, in order, are: total number of NPIs, face-covering restrictions, non-essential activities restrictions, public gathering restrictions, public transport restrictions, and cordon sanitaire restrictions (control of people entering and leaving the city). All specifications control for mayors' personal characteristics. We report robust-bias corrected p-values, estimates, and standard errors.

Table 5: Moderating effects of scientific intensity on performance (Deaths per 100k Inhabitants)

	Top 3	Top 3	Top 2	Top 2	Top 1	Top 1
	(1)	(2)	(3)	(4)	(5)	(6)
STEM Background	21.556***	18.838**	23.203**	19.982**		
	(7.383)	(8.754)	(9.217)	(9.955)		
STEM × Scientific Intensity	−3.861**	−3.882**	−3.810**	−3.871**		
	(1.693)	(1.776)	(1.647)	(1.757)		
Scientific Intensity					−4.776***	−5.515***
					(1.345)	(1.527)
Num. Obs.	413	413	338	338	164	164
Leader Controls	No	Yes	No	Yes	No	Yes
State FE	Yes	Yes	Yes	Yes	Yes	Yes

Note: This table analyzes how scientific intensity, measured as years of experience in STEM occupations, moderates the effect of mayors' STEM backgrounds on deaths per 100,000 inhabitants, using OLS regressions on municipalities with ordinary elections in specific years (2016). The models using Leader Controls control for mayors' gender, ideology, reelection status, and education. It defines three nested election samples: "Top 3" (at least one of the three most-voted candidates has a STEM background), "Top 2" (a STEM candidate is among the two most-voted), and "Top 1" (the winner has a STEM background). These progressively restrictive samples increase comparability between STEM and non-STEM candidates, sharpening the estimates of STEM background and scientific intensity on mortality outcomes.