The Value of Information for Regulatory Enforcement*

Elliott Ash
ETH Zurich

Maddalena Ronchi
Northwestern University

Elena Stella
Northwestern University

Silvia Vannutelli
Northwestern University

November 24, 2025

Abstract

How does the quality of information affect the allocation of regulatory enforcement? We study this question in the context of the Italian Court of Auditors, a large bureaucracy that oversees municipal finances through a hierarchical structure. Local auditors gather information on municipalities and report to judges, who decide whether to issue enforcement deliberations. We exploit a reform that randomized auditor assignments, increasing auditor independence without changing judicial rules. Combining novel administrative records with a machine-learning measure of predicted municipal default, we find that judicial enforcement increased overall, with the largest gains in high-risk municipalities. This targeting improvement arises through two channels. First, randomly assigned auditors report more financial irregularities, especially in high-risk municipalities and where pre-reform local ties were stronger. Second, experienced judges use these improved signals to focus deliberations on high-risk cases, highlighting the complementarity between information quality and decision-makers' expertise.

*Elliott Ash: ETH Zurich (email: ashe@ethz.ch); Maddalena Ronchi: Northwestern University (email: maddalena.ronchi@kellogg.northwestern.edu); Elena Stella: Northwestern University (email: elena.stella@kellogg.northwestern.edu); Silvia Vannutelli: Northwestern University (email: silvia.vannutelli@northwestern.edu). All errors are solely our responsibility.

1 Introduction

Governments rely on monitoring institutions to detect fiscal irregularities and enforce corrective action. Yet monitoring effectiveness crucially depends on the quality of information that regulators and enforcers can rely on. In many systems, the actors in charge to *gather* information (auditors) are different from those who *act* on it (judges or regulators), and the message that travels between them can be incomplete or strategically manipulated in the presence of conflicting incentives.

We study a national reform of municipal auditing in Italy that centralized and randomized the assignment of municipal auditors to local governments. The reform tightened auditors' independence without altering the National Court of Auditors' formal decision rules. We show that this change increased judicial action—the probability that the Court opens a deliberation against a municipality— improved the targeting of decisions and in the end positively affected downstream fiscal risk, and we trace these effects to an upgrade in the *judge-visible* signal produced by auditors.

We outline a simple conceptual framework that makes explicit the information pipeline from truth to action. Each municipality-year has a latent fiscal state T (irregularity or impending distress). Auditors form a private assessment $S = S(X, Z; \kappa)$ by combining characteristics of the municipality observable to both the auditor and the judge (X, e.g., balance-sheet indicators, arrears, demographics) with information that the auditor can acquire through on-site work but that remains unobservable to the judge (Z). The parameter κ captures the auditor's ability to detect and diagnose local fiscal problems, which depends on their skills and/or effort. Auditors then transmit a *reported* signal $R = g(S; \tau, \mathit{ties})$ to the Court via a standardized questionnaire; where τ captures differences in auditors' preferences/strictness, and *ties* capture their local entanglement. Judges observe (R, X) (but not Z) and decide whether or not to start an enforcement action $D = 1\{\mathbb{E}[T \mid R, X; \phi] \geq \bar{t}\}$, where ϕ is judicial capital/skill. By removing control of appointment from mayors and introducing random assignment, the reform plausibly reduced auditors' incentives to conceal information as well as the probability of having pre-existing ties with the mayors. This translates into a likely increase in κ and lowering of effective τ/ties, while holding judicial rules fixed. Our model thus predicts that judges would receive a more informative R, which will push borderline high-risk municipalities across the judicial enforcement threshold, especially where ϕ is high.

We take the framework to the data leveraging unique administrative micro-data that link (i) the full universe of municipal auditors questionnaires - which represents the way in which we measure R empirically; (ii) the timing and content of Court deliberations, corresponding

to D in the conceptual framework, and the identity of judges, which we can use to get a measure of their experience corresponding to ϕ in the model, (iii) municipal balance sheets and ledgers used to construct judge-invisible anomaly indices (T^*) and a machine-learning risk score $M(X)$ pre-dating the reform, and (iv) exogenous "adverse events"–formal entry into fiscal distress recorded by the Treasury.

Our identification strategy exploits the quasi-exogenous variation in treatment timing arising from the staggered introduction of the random auditor assignment reform across municipalities, based on pre-determined exogenous variation in expiration of previously-appointed auditors' terms (Vannutelli, 2021). This represents a unique experiment, as both the timing and the identity of the new auditors are orthogonal to current municipal shocks. The key challenge absent this design would be endogenous auditor selection and turnover— with mayors choosing different types of auditors depending on local economic conditions—which would confound the role of auditors' effort or preferences/strictness in her reporting with local municipal shocks. Furthermore, we observe both the *judge-visible* reports and the judicial actions, letting us tie any observed changes in enforcement to the role played by changes in the information available to judges when making their enforcement decisions.

We establish four results. First, random assignment of auditors increases judicial action. Event studies around the first random-draft year display flat pre-trends and a discrete, persistent rise in the probability of a Court deliberation. The average effect corresponds to 2 percentage points (about 7.1% of the pre-reform mean) in the first year after the appointment of the new auditor and stabilizes around 8.5 p.p. afterwards. Placebo leads are jointly insignificant; estimates are robust to staggered-adoption estimators and reweighting by pre-trend covariates.

Second, judicial enforcement sharpens where it matters most. Interacting treatment with municipalities' pre-reform ML risk score, we find effects concentrated in the upper half of the $\mathcal{M}(X)$ distribution: deliberations rise by 1.83 p.p. among high-risk municipalities versus -0.18 p.p. among low-risk ones, with a near-monotonic gradient across quintiles. This pattern matches the model's prediction that an improved R crosses a fixed judicial threshold precisely where latent risk $b(X)$ is high.

Third, the mechanism runs through the *reported* signal. Following the reform, auditors *report* more irregularities: the probability any irregularity is flagged in the questionnaire rises by 1.27 p.p. (a 107% increase over the baseline). These changes are larger in high-risk municipalities and in places with strong pre-reform local ties of the auditor, consistent with a reduction in concealment ($S \to R$ truer) and an increase in diagnostic effort ($T \to S$ sharper).

Fourth, the identity and characteristics of auditors and judges matter in the directions

implied by the framework. On the supply side, gains in reported irregularities are larger when auditors are younger at appointment, belong to post-reform cohorts, or are out-of-province relative to the municipality audited. On the demand side, rapporteurs with greater experience and local procedural capital exhibit steeper D-in-R responses: the same increase in R triggers more deliberations under higher ϕ.

This paper makes three key contributions to the litetaure. First, we open the black box of communication inside a hierarchical organization. While classic organizational theories in economics and finance posit that agency problems distort information transmission (Aghion and Tirole 1997, Tirole 2001, Dessein 2002, Rahman 2012, Strulovici 2021), direct measurement has been rare because internal messages are seldom observable and their informational content is hard to score. In our setting—municipal auditors reporting to the national Court of Auditors—we observe standardized audit questionnaires over time and, uniquely, a reform that randomized auditor assignment. This lets us quantify communication at scale and cleanly isolate the role of information supply from other factors.

Second, our design delivers a rare empirical laboratory for multi-layered governance. Earlier studies often face selection and unobserved heterogeneity in who communicates with whom. Our setting exploits (quasi-)random auditor assignment with staggered term expirations and a fixed reporting template, letting us isolate how the same organization performs when the messenger's independence and incentives change. This design cleanly separates auditor effects from persistent local organizational quality.

Furthermore, our unique design also allows us to study how judicial experience and local knowledge mediate the translation of information into action, contributing to the broader literature on judicial decision-making (Coviello et al., 2019; Abrams et al., 2022; Ash and MacLeod, 2024).

2 Institutional Setting

Italy is a highly decentralized democracy, with two levels of subnational governments: 20 regions and over 8000 municipalities. Each municipality has its local government composed of an elected mayor (Sindaco), an executive body (Giunta) appointed by the mayor, and an elected council (Consiglio Comunale). The mayor is directly elected for a five-year mandate with a two-term limit, holds executive power at the municipal level, and is responsible for all key policy decisions, including the annual budget.

Municipalities are granted large autonomy, they manage around 8% of total public expenditure (over 55 billion Euros) and have full control of a wide range of essential public services, such as waste management, social services, childcare and nursery schools, school-

related services, local police, road maintenance, housing, culture, recreation, and economic development. Spending is financed by municipal fiscal revenues (87%) plus transfers from the central government (13%), while borrowing is allowed only to finance investment expenditures and is subject to strict quantitative limits.

Given the high degree of decentralization in spending and revenue collection decisions, municipalities are subject to national fiscal rules and a stringent monitoring system led by the National Court of Auditors. This body is a Supreme Audit Institution, responsible for overseeing and ensuring the proper management of public finances and the legality of public administration activities. It operates as an independent body, reporting directly to Parliament, and its main functions include auditing the accounts and financial management of the state, regions, and other public entities, as well as adjudicating matters related to public accounting and financial responsibility. As such, it can initiate judicial proceedings against municipalities that fail to comply with the rules or engage in improper financial practices. The Court is articulated in 20 regional offices. Judges are selected through a process that ensures both merit and independence and are rotated between different chambers or offices periodically, to ensure a fresh perspective in auditing and judicial functions while maintaining the integrity of the court's operations.

The Court of Auditors operates across the entire national territory, with the central sections based in Rome and regional sections located in the capitals of each region, as well as in the autonomous provinces of Trento and Bolzano. The composition of each regional panel (or section) is determined by the Council of the Presidency of the Court of Auditors, the body responsible for managing the careers of judges and the internal organization of the Court. This Council ensures the balanced allocation of judges and resources to each regional section.

Each regional section is led by a President, appointed based on merit and seniority, who is responsible for assigning cases to judges. The assignment of judges to regional offices is institutionalized and regulated. By contrast, the allocation of judges to municipalities within each region follows more informal practices. In most cases, judges are assigned to the municipalities of a specific province, with this assignment reviewed annually. This internal distribution—though widespread—is not codified by formal rules. We learned about this province-based allocation through direct conversations with judges. Judges remain in the same region for an average period of five years, with a maximum limit of ten years. After this period, they may be reassigned to a different region to ensure a rotation of territorial competences, or to the central office.

Once a case is assigned, the designated judge conducts the investigation on the budget of the assigned entities. When the investigation is complete, the judge submits the case,

along with a report, to the President. The President then convenes the entire panel and the involved municipality to review the case. The process ensures compliance with the principles of adversarial proceedings and provides the possibility to appeal the decision to a central section, specifically through the special composition of the united sections. While all judges are summoned, the final deliberation is carried out by a commission of no fewer than three members from the panel. During the control year this commission deliberates on the budget and verifies the financial compliance of the municipality.

Italy's 8,000 municipalities pose a significant oversight challenge for the Court of Auditors, which counts 477 judges—only half of whom are assigned to regional control sections. Given these constraints, the monitoring system relies on a hierarchical structure. Approximately 14,000 local auditors (*Revisori dei conti*) act as the Court's operational backbone, conducting first-line financial checks and feeding critical information into the judicial process.

These auditors, embedded within local governments across the country, serve as both advisors and monitors of municipal finances. Remunerated by the municipalities themselves, with contracts renewed annually, they operate in close proximity to local political actors. They conduct over 200 regulatory and compliance checks each year, focusing on financial records, accounting practices, transactions, and internal controls. Their most critical responsibility is the annual audit of the municipal balance sheet, the findings of which are compiled in a report reviewed by the municipal council during budget approval. While not legally binding, the auditors' recommendations must be addressed or formally justified by the council.

Following the review, auditors submit the approved balance sheet, audit report, and a detailed questionnaire to the Court. These materials serve as screening tools for judges, who use them to evaluate fiscal soundness and, when necessary, trigger more in-depth investigations. In cases where issues arise, the Court can initiate a judicial proceeding, culminating in a formal deliberation.

On average, the Court issues around 3000 judicial acts of ex-post verifications per year (Corte dei Conti, 2022). The majority of these relate to initial warnings designed to alert the municipal governments and encourage them to take corrective actions within 60 days. By contrast, the number of subsequent judicial prosecutions is much smaller, averaging around 300 cases per year, indicating a high degree of compliance with the Court's remarks.

2.1 Monitoring System and Assignment of Auditors

Italy is a decentralized democracy, with two levels of subnational governments: 20 regions and over 8000 municipalities. Municipalities are autonomous – they manage around 8%

of total public expenditure (over €55 billion) and maintain control of a range of public services, such as waste management, social services, childcare and nursery schools, school-related services, local police, road maintenance, housing, culture, recreation, and economic development.

Since 1990, all municipal budgets must be audited and certified by a board of auditors. This board of auditors is overseen by a national public organization called the Court of Auditors (Corti dei conti). Each municipality's board of auditors is the primary liaison between the Court and each local entity. Every year, the board issues an audit report, which is reviewed by both the municipal council and the national Court of Auditors. The National Court of Auditors is the central and ultimate monitor of Italian municipal finances: it can initiate judicial proceedings against municipalities that fail to comply with the rules or engage in improper financial practices. Non-compliance with auditors' recommendations can lead to increased scrutiny from the Court.

The board of auditors is composed of one member for municipalities with less than 15,000 inhabitants (about 90% of municipalities fall under this threshold) and three members for larger municipalities. The Italian market for auditors is characterized by a high degree of competition. Unlike in the U.S., where many municipalities do not face audits and most auditors are employed by large private firms, most municipal auditors in Italy are self-employed Certified Public Accountants. These auditors can work for municipal governments and provide freelance professional accounting and auditing services to individuals and small businesses. There are over 150,000 registered CPAs that are authorized to work as auditors in the public and the private sector.

Prior to the reform, the board of auditors and its municipality were closely involved: each board was selected by the municipality, the board's contracts were renewed every three years, and the boards' remuneration was directly sourced from the municipality's budget and left to the municipality's discretion.

2.2 The Policy Reform

After the reform, the board of auditors for each municipality was selected according to a random draft, and the boards' renumeration was set according national regulations.

In August 2011, at the peak of the European Sovereign Debt Crisis, the national government changed how the board of auditors is selected. New laws mandated that auditors of local entities be chosen by random draw from a list formed by those from the following groups who request to be included: a) those currently included in the regional list of auditors, b) any socially authorized Certified Public Accountant. In February 2012, the government

outlined the procedures to participate in the list, including experience requirements varying with population thresholds (to avoid sending inexperienced auditors to audit large municipalities). Due to administrative constraints and the time needed to form the public lists, the new drafting procedure entered into effect on December 10, 2012. The reform does not apply to the 5 Special Status Regions of Italy – Valle d'Aosta, Trentino Alto Adige, Friuli Venezia Giulia, Sardinia, and Sicily. These Special Status Regions maintain extra autonomy along many margins, which includes setting their own fiscal rules and monitoring procedures

The random auditor draft is organized by local offices of the Ministry of the Interior, via a standardized computerized system. For each draft, the number of drafted candidates is equal to $3N$, where N is the number of auditors to be appointed. The Ministry of the Interior informs the municipality of the results of the draft, and the municipality appoints the drafted auditors in order, provided they have the required experience. The order of the drafted auditors is strict: if the first-drafted auditor refuses the appointment, the municipality can only contact the second-drafted auditor next. Auditors' compensation became regulated nationally, including a minimum floor that is estimated to cover the expenses to effectively perform the job, and maximum caps that vary by municipality population. Mayors have limited discretion in setting compensation within these constraints; ultimately, there is little variation in the actual compensation across municipalities of similar size.

3 Conceptual Framework

We model the Court judicial enforcement as a testing problem. Specifically, we study a two-tier enforcement technology in which *auditors* produce information and *judges* act on what they observe. Let each municipality–year (i, t) have a latent fiscal state

$$T_{it} \in \{0, 1\} \qquad \text{(true irregularity or impending distress)}.$$

Let X_{it} denote observables available to the econometrician and to judges ex ante: balance-sheet indicators, arrears, demographics, etc. Let Z_{it} denote auditor-acquirable but judge-invisible details revealed by on-site work and document review. An assigned auditor with diagnostic parameter κ_a maps (X_{it}, Z_{it}) into a private assessment

$$S_{it} \equiv S(X_{it}, Z_{it}; \kappa_a),$$

which we interpret as the auditor's *best read of the truth*. The auditor then transmits a standardized *reported signal* R_{it} (questionnaire ticks, severity flags, narrative red flags) to

the Court according to

$$R_{it} \equiv g(S_{it}; \tau_a, \text{ties}_{im}) \,,$$

where τ_a is a reporting threshold (lower τ_a means that auditors have a lower threshold for reporting) and ties_{im} captures potential local entanglement. Judges observe (R_{it}, X_{it}) but not Z_{it} and choose whether to open a deliberation:

$$D_{it} = \mathbb{1}\Big\{\mathbb{E}[T_{it} \mid R_{it}, X_{it}; \phi_j] \geq \bar{t}\Big\} \,,$$

where ϕ_j summarizes *judicial capital* (rapporteur experience and local procedural knowledge) that improves the translation of R into action.

Two frictions separate truth from action:

1. **Information production (diagnostic skill/effort):** the mapping $T \to S$ depends on κ_a, i.e. how well the auditor uncovers the underlying state through competent procedures. Variation in production creates cross-auditor dispersion in the precision of S. Higher κ_a increases the likelihood that S correctly reflects T.

2. **Information transmission threshold (strictness/concealment):** even with informative S, auditors can differ in their strictness, i.e. in what they regard as being worth reporting to the court. They can also strategically downplay or omit problems because of the fact that they have strong local ties_{im}. Thus the mapping $S \to R$ can be attenuated when τ_a is high or ties_{im} are strong (strategic downplaying/omission).

The random-draft reform plausibly shifts both margins:

1. $\kappa_a \uparrow$ (standardized procedures; fresher cohorts)

2. $\tau_a \downarrow$ and impact of $\text{ties}_{im} \downarrow$ (weaker local capture).

Because judges act on R, any improvement in production or transmission raises the likelihood ratio $\mathcal{L}(R \mid T)$, pushing borderline high-risk cases over the judicial threshold.

$$\underbrace{\boldsymbol{T}}_{\text{truth}} \xrightarrow[\text{skill/effort } \kappa_a]{\text{production}} \underbrace{\boldsymbol{S}(X,Z)}_{\text{auditor's private read}} \xrightarrow[\text{pliability } \tau_a]{\text{transmission}} \underbrace{\boldsymbol{R}}_{\text{judge-visible}} \xrightarrow[\text{judicial capital } \phi_j]{\text{decision}} \underbrace{\boldsymbol{D}}_{\text{action}} \,.$$

3.1 Testable implications

Let $\mathcal{M}(X)$ be our pre-reform ML risk score. The framework yields the following predictions, each mapped to an empirical test:

1. **Enforcement intensity** The probability of deliberation D_{it} rises following the random audit reform, as the judges receive a more credible R

2. **Targeting.** The increase in D_{it} is larger at higher $\mathcal{M}(X)$ (upper quantiles of predicted risk), consistent with a more informative R crossing a fixed judicial threshold.

3. **Transmission/pliability.** Holding case mix fixed, the *reported* probability and intensity of accounting irregularities (components of R) increase post-reform, especially (i) in high-risk municipalities and (ii) where pre-reform ties were strongest (e.g., same-province match).

4. **Production/skill.** Gains in irregularity detection concentrate among auditors with proxies for higher diagnostic effort/skill.

5. **Judicial capital.** The $R \to D$ link steepens with rapporteur experience/local procedural capital ϕ_j: for a given improvement in R, deliberations rise more under higher ϕ_j.

6. **Downstream outcomes.** Composite fiscal-risk indicators improve with a lag (budget cycle), especially in high-risk cells and when a deliberation actually occurs, consistent with the chain $T \Rightarrow S \Rightarrow R \Rightarrow D$.

4 Data

4.1 Data Sources

Auditor Questionnaires 2010-2015 Each year, Italian municipalities are assigned a board of auditors. This board prepares an audit report that is submitted to the national Court of Auditors. One part of this audit report is a standardized questionnaire that remains largely consistent year to year. The questionnaire is a structured text document that solicits information about financial distress, structural deficit indicators, and accounting irregularities.

Between 2010 and 2015, the format and content of these questionnaires were highly standardized, with only minor revisions introduced to reflect changes in financial regulation or oversight priorities. The dataset we assembled includes approximately one million observations, corresponding to about forty-four thousand unique questionnaires. Each entry is associated with a specific municipality, fiscal year, and question, and derives from the standardized forms submitted by auditors to the Court of Auditors. We digitized and systematized these forms to create a consistent panel suitable for empirical analysis.

Each questionnaire provides a comprehensive account of the financial condition of the municipality. It begins with an identification header reporting basic information about the local entity, population, and the responsible auditor, followed by references to relevant financial documents and deliberations approving the budget or the final account. Subsequent sections focus on the main dimensions of municipal finance, including budget and cash-flow results, debt evolution and guarantees, compliance with internal stability constraints, and the management of participations in publicly owned entities. The questionnaires also contain sections on personnel expenditure, assessing compliance with statutory limits, and on patrimonial and economic accounts, which verify the consistency of balance-sheet items and the valuation of assets and liabilities. Each form concludes with a certification statement, signed by the auditors, attesting that all reported data correspond to the official municipal accounts.

The questionnaires combine both structured and narrative elements. Closed-ended questions (typically Yes/No/Not Applicable) are complemented by numerical tables and qualitative explanations that allow auditors to clarify specific irregularities or exceptional circumstances. This hybrid format enables systematic comparison across municipalities and over time, while retaining valuable qualitative insights into local financial management.

Examples of the standardized questionnaires used during the period 2010–2015 are reproduced in Appendix A4.

Court Deliberations, 2009–2021 The Italian Court of Auditors (Corte dei conti) produces deliberations that follow a (fairly) structured format. These deliberations are formal judicial documents assessing the financial management of local governments and other public entities. Each deliberation corresponds to an official judgment concerning the compliance of a municipality's financial statements with accounting and fiscal regulations. The structure and language of these documents follow a relatively standardized institutional format, allowing for systematic textual analysis and extraction of key information.

The dataset comprises approximately thirty-six thousand observations, each corresponding to a specific deliberation–municipality pair. For each document, we identify the regional section issuing the judgment, the names and positions of the participating magistrates (including the president, councilors, and rapporteur judge), and the fiscal years of the municipal accounts under review. The deliberations were collected directly from the regional websites of the Court of Auditors and processed to extract relevant information through automated text recognition and structured parsing.

Each deliberation begins with a formal header reporting the unique identification number, the regional control section (e.g., "Regional Control Section for Tuscany"), and the

composition of the judging panel. The main body of the document is divided into sections that follow a quasi-legal structure. The introductory paragraphs ("*VISTO*" or "*VISTA*") list the relevant legal references and precedents. The subsequent section ("*FATTO*") presents the factual elements emerging from the audit, including irregularities or imbalances detected in municipal accounts. The following section ("*DIRITTO*") provides the legal reasoning that connects the factual findings to the applicable laws and accounting standards. The deliberation concludes with a decision section ("*DELIBERA*" or "*DECIDE*"), often followed by a final operative part ("*P.Q.M.*" or "*ORDINA*"), which specifies the measures to be adopted by the municipality or other competent authorities.

Beyond their formal structure, the deliberations frequently discuss recurring issues in local public finance, such as the accumulation of arrears, the misclassification of expenditures, or breaches of fiscal discipline under the internal stability pact. They also reference the work of municipal auditors, linking the oversight activity of the *revisori dei conti* to the subsequent judicial evaluation of compliance. This makes the deliberations a valuable complement to the auditor questionnaires, providing an ex post institutional assessment of local financial behavior.

An example of deliberation issued by the Court of Auditors is reproduced in Appendix A5.

Municipal Balance Sheets, 2000-2022 Each year, each Italian municipality must also prepare a municipal balance sheet. This balance sheet is submitted to the board of auditors and also submitted to the national Court of Auditors as a part of the audit report. The municipal balance sheet represents the official accounting document summarizing the financial position, operating performance, and patrimonial situation of the local government. It provides the quantitative foundation for the assessments carried out by the auditors and for the subsequent judicial review by the Court of Auditors.

Our dataset covers the period 2000–2022 and contains approximately two hundred thousand observations at the municipality–year level. For each municipality, the data include a wide range of accounting variables that together describe the structure of revenues, expenditures, assets, and liabilities. These variables allow for the construction of financial indicators that capture the fiscal sustainability, liquidity, and performance of local governments. The balance sheets were collected from official administrative sources and harmonized over time to ensure consistency across changes in accounting frameworks.

Each municipal balance sheet is composed of three main sections: the budget account, the economic account, and the patrimonial account. The budget account reports realized revenues and expenditures, distinguishing between current and capital operations, and provides measures of the annual budget balance and the overall management result. The economic

account records flows of costs and revenues on an accrual basis, reflecting the economic performance of the municipality. The patrimonial account provides a stock measure of assets and liabilities, including the composition of debt, receivables, and patrimonial resources. Together, these components allow for a comprehensive assessment of both the financial and economic equilibrium of local administrations.

Using these data, we compute a set of key indicators that summarize the financial condition of municipalities. These include measures of indebtedness and debt service costs, compliance with legal borrowing limits, off-balance-sheet obligations, liquidity ratios, budget balances, and indicators of revenue collection and payment execution efficiency. We also derive measures of fiscal independence, investment activity, and per capita debt burden, which provide additional insights into local fiscal behavior and sustainability.

Random Audits, 2012-2019 These datasets report the names of all municipalities that receive randomly assigned auditors in each year. The key treatment assignment variables for our analysis are the dates when municipalities appoint auditors through the random draw procedure for the first time. All information on currently appointed auditors is available on the website of the Ministry of the Interior, separately for each municipality. We constructed a comprehensive historical database covering the entire universe of municipalities by scraping the Ministry's website, obtaining datasets that contain the draw dates and identities of all auditors selected through the random assignment system since 2012. We then combined these datasets with information from the yearly candidate pools published by the Ministry of the Interior. These lists include details on candidates' age, gender, municipality and region of birth and residence, level of professional experience (measured by the number of years they have been certified as public accountants), and previous service as auditors.

Municipal Financial Distress Procedures, 1990-2021 This dataset reports all municipalities that activated either the bankruptcy (*dissesto*) or financial rebalancing (*riequilibrio*) procedures between 1990 and 2021 (Figure A1). *Dissesto* represents the more severe financial distress procedure, where municipalities formally declare bankruptcy and cannot meet their financial obligations. *Riequilibrio*, introduced as a preventive alternative, allows municipalities to propose recovery plans while maintaining greater autonomy, enabling them to address financial difficulties before formal bankruptcy. After the highest bankruptcy incidence in the 1990s, these instruments stayed mostly unused until the 2011 reform, when riequilibrio became the more frequently used mechanism as municipalities sought to avoid

the drastic consequences of formal bankruptcy.[1]

4.2 Sample Statistics

To measure financial outcomes, we rely on detailed data on municipal budgets collected by the Italian Ministry of the Interior. These are the so-called "final balance sheets" (*Bilanci Consuntivi*), which report realized revenues and expenditures for the previous fiscal year and must be submitted and approved by April 30 of the following year. The dataset provides comprehensive information on municipal finances, including revenues from local taxes, current and capital expenditures, investments, debt positions, and intergovernmental transfers.

In our analysis, we aim to examine whether the random assignment of auditors results in more independent oversight and, consequently, greater fiscal sustainability of municipalities, consistent with the objectives of the national government. For this purpose, we focus on a set of fiscal indicators that the central government and the National Court of Auditors routinely use to monitor the fiscal stance of local administrations.

First, we consider *Off-balance Sheet Debts*, defined as municipal liabilities not covered by any revenues. These include the absolute value of off-balance-sheet debts that arise outside the approved budget framework and are not absorbed into current or capital expenditures. Their presence signals violations of budgetary discipline and typically reflects accounting irregularities or unauthorized commitments that must eventually be recognized and repaid.

Second, we analyze the *Debt Limit Index*, which measures the ratio of expenditures on debt repayments to total available revenues. This indicator reflects the extent to which a municipality's resources are absorbed by servicing outstanding obligations and provides a forward-looking measure of fiscal pressure associated with compliance with nationally mandated debt limits.

Third, we examine *Cash Advances*, which represent short-term cash inflows used to cover temporary liquidity shortages. Municipalities rely on these advances when the timing of incoming revenues is insufficient to meet ongoing expenditure needs. Although these advances must be repaid, recurrent reliance on them highlights underlying liquidity constraints and weaknesses in cash-flow management.

Finally, we study the *Budget Surplus* (or deficit), calculated as the ending cash balance plus active residuals, minus passive residuals and restricted funds. This indicator summarizes a municipality's year-end fiscal position by measuring the resources available after accounting for all recognized commitments. A higher surplus reflects stronger fiscal management and greater alignment with the balanced-budget targets emphasized by national fiscal rules.

[1] *Riequilibrio* is similar to Chapter 11 reorganization in the US, while *dissesto* resembles Chapter 7 liquidation.

All variables are expressed in per-capita terms and winsorized at the 1% level to limit the influence of extreme values. Table A2 presents summary statistics for these fiscal indicators in the pre-reform period.

5 Methods

5.1 Machine Learning

We use machine learning (ML) methods to predict municipal fiscal distress based on financial statement data. The goal is to build a reliable early-warning system that identifies municipalities most at risk of entering formal financial distress, enabling prioritized oversight by the Court of Auditors. This predictive exercise faces two main challenges. First, fiscal distress is a multi-dimensional phenomenon that reflects both structural and cyclical imbalances in municipal finances. Second, distress events are rare—occurring in only about two percent of municipality-year observations—so the problem involves severe class imbalance. ML algorithms are well-suited to address these issues because they can flexibly model non-linearities and interactions among financial indicators while employing regularization to prevent over-fitting and handle sparse positive outcomes (e.g., Hastie et al., 2009; Mullainathan and Spiess, 2017; Athey, 2018; Ash et al., 2025).

Data Preparation. Our dataset combines annual municipal balance sheets and audit questionnaires submitted to the Court of Auditors over 2005–2022. We focus on 24 financial indicators that are routinely monitored by the Court and capture the main dimensions of fiscal capacity, debt exposure, liquidity, and revenue performance (see Table 1). These include variables such as current and capital account balances, off-balance-sheet liabilities, tax revenues, and debt service ratios.

To account for temporal dynamics in fiscal deterioration, we augment the dataset with one- to five-year lags of each financial indicator. Missing values are imputed using within-municipality means across available years—a conservative approach that preserves cross-sectional comparability while avoiding spurious temporal patterns. We also include demographic controls (population, income per capita, and dependency ratios) and geographic fixed effects to capture persistent structural differences across regions and municipality size classes.

The resulting panel provides a rich set of fiscal and demographic predictors for each municipality-year observation. In contrast to traditional fiscal risk assessments that rely on single thresholds (e.g., debt-to-revenue ratios), this dataset allows the model to learn

complex, non-linear combinations of indicators that jointly predict distress. The final matrix X contains roughly 24 base features, 120 lagged features, and several demographic controls, for a total of about 150 predictors and over 100,000 municipality-year observations. The binary outcome variable $Y_{it} \in 0, 1$ equals one if municipality i enters formal financial distress in year t, and zero otherwise.

Machine Learning Approach. We aim to learn a conditional expectation function $Y(X) = \mathbb{E}\{Y|X\}$ that provides a predicted probability, based on observed financial indicators X, that a municipality-year will enter fiscal distress. The feature space is moderately high-dimensional, with correlated fiscal and demographic predictors relative to the limited number of distress events. Classical linear models often over-fit and fail to capture important non-linearities or interactions, limiting their ability to generalize to unseen data (Hastie et al., 2009). Given our goal of predicting distress in municipalities that have not yet exhibited financial problems, we must move beyond standard statistical models used in applied economics (Mullainathan and Spiess, 2017; Athey, 2018).

Machine learning (ML) algorithms provide two key advantages for this task. First, they introduce *regularization*—algorithmic penalties that constrain model complexity and reduce over-fitting. For example, in a regularized logistic regression, parameters are estimated by minimizing a penalized loss function,

$$\min_{\theta} \sum(-L(\theta) + R(\theta)),$$

where $L(\theta)$ is the log-likelihood and $R(\theta)$ is a penalty on the magnitude of the coefficients, such as the L1 (Lasso) or L2 (Ridge) norm (Tibshirani, 1996; Zou and Hastie, 2005). Second, ML models can efficiently capture complex, non-linear relationships among fiscal indicators without requiring the researcher to manually specify interactions or polynomial terms (Hastie et al., 2009). This flexibility motivates the use of ensemble methods such as random forests and gradient boosting.

The first approach is *penalized logistic regression*, a linear classifier that estimates a logit link between fiscal variables and the distress indicator. Regularization shrinks coefficients toward zero, improving out-of-sample stability and interpretability. Despite its simplicity, this model provides a transparent benchmark for comparison with more flexible algorithms.

The second approach is the *random forest* (Breiman, 2001), an ensemble of decision trees trained on bootstrap samples of the data. Each tree recursively partitions the predictor space to reduce classification error, and the ensemble aggregates their predictions by majority vote. Random forests are robust to multicollinearity and capture non-linear and interaction effects

Table 1: List of Variables and Their Descriptions

#	**Variable**	**Description**
1	Off-balance sheet debts	Liabilities not covered by any revenues. Absolute value of off-balance-sheet debts over sum of current and capital expenditures.
2	Total debt service	Sum of principal and interest payments due in the year.
3	Debt limit index	Expenditures on debt repayments to total available revenues.
4	Cash advances	Short-term cash inflows used to cover liquidity shortages.
5	Budget surplus (deficit)	Ending cash balance plus active residuals, minus passive residuals and restricted funds.
6	Management balance	Total receipts minus total payments.
7	Current account balance	Ongoing revenues minus routine expenditures.
8	Capital account balance	Capital revenues minus long-term investments.
9	Passive carryovers	Outstanding payment obligations from previous periods.
10	Active carryovers	Uncollected revenues from previous periods.
11	Active carryovers to current revenues	Ratio of carryovers to current year revenues.
12	Active carryovers to total revenues	Ratio of carryovers to total revenues.
13	Total expenditures	Sum of all municipal spending.
14	Total capital expenditures	Investment spending on infrastructure and assets.
15	Total current expenditures	Operating expenses for municipal services.
16	Revenues from taxes	Income from local taxation.
17	Other revenues	Non-tax income sources.
18	Cash	Liquid assets available.
19	Financial autonomy index	Own-source revenues to total revenues.
20	Debt per capita	Total outstanding debt divided by population.
21	Investments	Capital expenditures relative to current expenditures.
22	Collecting capacity	Tax collections relative to total tax commitments.
23	Ability to pay	Total payments to total commitments.
24	Resident population	Number of residents in municipality.

through hierarchical splits, though they can be less precise in ranking continuous risk scores.

The third and preferred approach is the *gradient boosted tree model*, implemented as XGBoost (Friedman, 2001; Chen and Guestrin, 2016). Gradient boosting sequentially fits trees to the residuals of prior trees, gradually improving predictive accuracy while controlling over-fitting through shrinkage and subsampling. XGBoost adds additional regularization and computational enhancements that yield state-of-the-art performance on structured tabular data (Feurer et al., 2018; Grinsztajn et al., 2022). We therefore take XGBoost as our main specification for predicting fiscal distress.

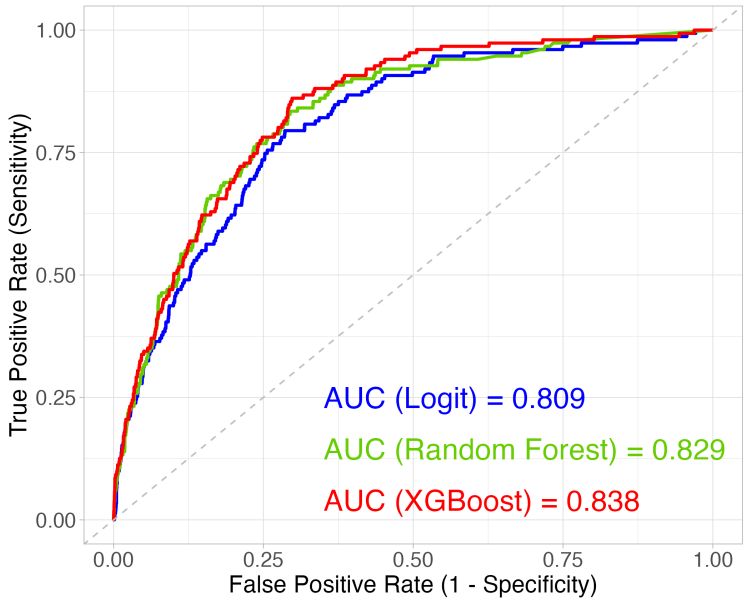
Model Training and Selection. Model training is restricted to post-2011 data, when financial reporting standards became consistent following national accounting reforms. We partition the data into training and test sets using five-fold cross-validation at the municipality level, ensuring that all observations for a given municipality fall within the same fold. This design prevents information leakage from lagged predictors and provides unbiased estimates of out-of-sample predictive performance. The rare-event nature of fiscal distress—roughly two percent of observations—creates substantial class imbalance, which we address by weighting the loss function inversely to class frequencies (King and Zeng, 2001; He and Garcia, 2009). For tree-based models, we also implement `scale_pos_weight` parameters equal to the imbalance ratio.

For each model, hyperparameters are tuned by grid search within the cross-validation folds (Varma and Simon, 2006; Cawley and Talbot, 2010; Feurer et al., 2018). For logistic regression, we select the optimal regularization strength λ and the elastic-net mixing parameter using the validation loss. For the random forest, we tune the number and depth of trees, the number of features sampled per split, and the minimum node size. For XGBoost, we tune the number of boosting rounds, learning rate, maximum tree depth, and regularization parameters (α and λ). Early stopping is applied when the validation loss fails to improve over a fixed number of iterations.

Across models, performance is evaluated on held-out test data using accuracy, precision, recall, F1 score, and the area under the receiver operating characteristic curve (AUC-ROC). As in other economic prediction tasks with tabular data, we find that gradient-boosted trees (XGBoost) consistently outperform linear and bagging-based baselines in both accuracy and calibration (Friedman, 2001; Chen and Guestrin, 2016; Feurer et al., 2018; Grinsztajn et al., 2022). Consequently, XGBoost serves as our preferred predictive model for fiscal distress.

Model Evaluation. We evaluate model performance using held-out test data to assess predictive accuracy and robustness. Figure 1 presents Receiver Operating Characteristic

Figure 1: ROC Curves for Financial Distress Prediction Models



Notes. Receiver Operating Characteristic curves comparing the discriminative performance of XGBoost, Logistic Regression, Random Forest, and Neural Network models on the test set. Area Under the Curve (AUC) values indicate each model's ability to distinguish between municipalities that will and will not enter financial distress.

(ROC) curves comparing the discriminative ability of the three classifiers, while Table 2 reports confusion matrices at a 0.5 classification threshold. These diagnostics summarize the trade-off between sensitivity (the ability to detect true distress cases) and specificity (the ability to avoid false alarms). The area under the ROC curve (AUC-ROC) provides a threshold-independent measure of performance, interpretable as the probability that a randomly selected distressed municipality is assigned a higher predicted risk than a non-distressed one.

To better understand classification behavior, Table A3 reports how precision, recall, and F1 scores vary across different decision thresholds for the XGBoost model. This analysis highlights how policy priorities—such as minimizing false positives versus maximizing early detection—map onto alternative decision boundaries. In our baseline results, we report outcomes at a balanced operating point that optimizes the F1 score, while appendix figures present full threshold-sensitivity curves.

Table 2: Confusion Matrices for Financial Distress Prediction Models

	Prediction	
	Not Distress	Distress
Panel A. Logistic Regression		
Not Distress	3358	11
Distress	148	3
Panel B. Random Forest		
Not Distress	3368	1
Distress	151	0
Panel C. XGBoost		
Not Distress	3201	168
Distress	100	51

Notes. Confusion matrices from model predictions using a 0.5 classification threshold. Rows represent actual financial distress status, columns represent predicted status. Not Distress = municipality did not enter financial distress; Distress = municipality entered financial distress procedure.

5.2 Empirical Design

Our identification strategy leverages the staggered introduction of the random auditor assignment reform across municipalities. The reform became effective in December 2012, but treatment timing varies across municipalities depending on when their incumbent auditor's term expired. This variation creates a natural experiment where municipalities are randomly assigned to different treatment cohorts based on historical appointment patterns unrelated to municipal characteristics. Figure A2 displays municipalities geographic distribution, highlighting the lack of any geographic patterns by treatment timing.

To estimate the causal effects of random audits, we employ an event study specification:

$$Y_{it} = \sum_{k \neq -1} \beta_k \cdot \mathbf{1}[t - t_i^* = k] + \alpha_i + \gamma_{qrt} + \varepsilon_{it} \tag{1}$$

where Y_{it} represents our outcome of interest for municipality i in year t, t_i^* denotes the year municipality i first receives a random auditor, and k represents event time relative to treatment. The coefficients β_k capture the dynamic treatment effects. Municipality fixed effects α_i control for time-invariant characteristics, while γ_{qrt} represents region-risk-year fixed effects allowing municipalities in different risk quartiles and regions to follow distinct trends.

The main identifying assumption underlying our empirical design is that municipalities exposed to the reform at different points in time would not have followed distinct trends in

the absence of treatment; in other words, without the change in the auditor's appointment, the trajectories of outcomes across treatment cohorts would have evolved in a comparable manner over time.

Our empirical specification already addresses several potential issues that might otherwise challenge the causal interpretation of our findings. First, by including municipality fixed effects, we remove the influence of any time-invariant municipal characteristics that could explain differences in outcomes. Second, the inclusion of year fixed effects ensures that we are not capturing shocks that affect all municipalities simultaneously, such as macroeconomic fluctuations or nationwide policy changes.

Given the exogenous and historically determined variation in audit cycles across municipalities, the timing of auditor expiration—and hence the treatment onset—should be uncorrelated with the evolution of municipal outcomes. To support this assumption, Table A1 examines predictors of treatment timing. The table reports two F-tests at the bottom. When population dummies are excluded, the null hypothesis that all coefficients are jointly zero cannot be rejected across specifications.

Third, and perhaps most crucially, we assess whether outcomes evolve in parallel before the reform, thereby testing the plausibility of post-reform parallel trends.[2] Figure A3 presents the event-study estimates from Equation 1, showing no evidence of differential pre-treatment dynamics across cohorts.

A further potential issue concerns anticipation effects (Malani and Reif, 2015), given that municipalities may have known the timing of their auditor changes. Two opposite anticipation patterns could theoretically emerge. On one hand, later-treated municipalities might start adjusting early in anticipation of a stricter audit, gradually adapting their fiscal behavior; such a response would bias our estimates downward, as the "control" group would already begin acting like the treated one. On the other hand, they could instead increase spending or debt accumulation before treatment, expecting future restrictions—an effect that would bias the estimates upward.

Finally, we address the possibility that other contemporaneous, time-varying shocks coincide with the staggered reform rollout. The main event of concern is the 2014 extension of the Domestic Stability Pact to municipalities under 5,000 inhabitants, which we capture through the inclusion of population-size-by-year fixed effects.

Another possible worry is that the change in auditor appointments occurred alongside a broader emergency reform package that may have influenced local finances through other

[2] Kahn-Lang and Lang (2020), Roth (2020), and Rambachan and Roth (2020) warn that pre-trend tests often have low power and may produce type-II errors. We follow Rambachan and Roth (2020) and conduct the robustness checks using the approach implemented in the *HonestDiD* package, finding that our conclusions remain robust even under substantial deviations from linearity.

channels.[3] However, these concurrent measures were implemented simultaneously across all municipalities, whereas the change in auditor appointments occurred gradually, depending on the expiry date of the incumbent auditor. Consequently, any effects of the broader reform are absorbed by the inclusion of year fixed effects in our specifications.

6 Results

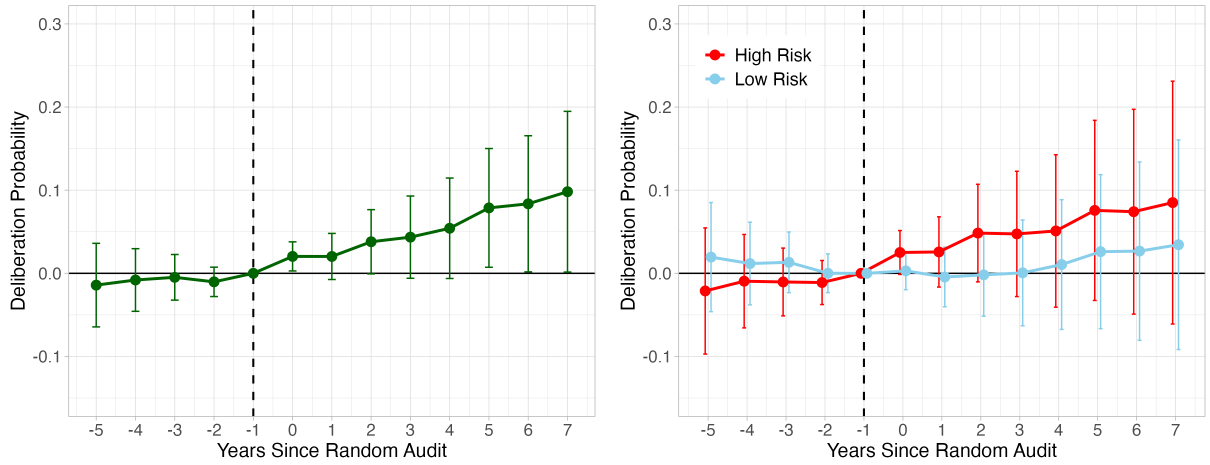
6.1 Effect of Random Audits on Court Deliberation

We begin by examining our primary outcome of interest, namely whether random auditor assignment affects the probability that municipalities are targeted by Court's enforcement actions. Figure 2(a) presents our main results. The event-time coefficients display flat pre-trends and a discrete, persistent post-treatment rise in the probability that the Court opens a deliberation. To gauge a sense of the magnitudes, in Appendix Table A4, we perform a static DID specification to obtain an estimate of the ATE in the post period. In our preferred specification, the probability of a deliberation increases by 7.1% relative to the pre-reform mean.

We then move to investigate whether the reform affects not only the degree of enforcement, but its targeting. To do so, we exploit our ML-produced risk score measures and investigate whether our treatment effects vary significantly by municipal risk level. Figure 2(b) demonstrates that high-risk municipalities experience substantially larger increases in deliberation probability compared to low-risk municipalities, suggesting that random audits indeed improve the targeting of judicial oversight.

We then move to investigate whether the reform affects not only the degree of enforcement, but its targeting. To do so, we exploit our ML-produced risk score measures and investigate whether our treatment effects vary significantly by municipal risk level. Figure 2(b) demonstrates that high-risk municipalities experience substantially larger increases in deliberation probability compared to low-risk municipalities, suggesting that random audits indeed improve the targeting of judicial oversight.

[3]The 2012 reform also introduced (a) stricter fiscal rules and cuts in central transfers, (b) progressive municipal income tax rates, (c) compulsory inter-municipal management of public services for towns below 1,000 residents, (d) a reduction in the size of municipal councils, and (e) temporary taxation of owner-occupied dwellings.



(a) **Overall Effect** (b) **By Municipal Risk Level**

Figure 2: Random Audit Effects on Court Deliberation Probability
Panel (a) shows event study estimates of the causal effect of random audits on the probability that a municipality receives a Court deliberation. Panel (b) compares high-risk municipalities (above median predicted financial distress, shown in red) with low-risk municipalities (below median, shown in light blue). The reference period is $t = -1$ (one year before the first random audit). Standard errors are clustered at the municipality level. 95% confidence intervals are shown.

6.2 Mechanisms: Improvements of Auditors' Signals

To understand the mechanisms driving the targeted increase in court deliberations, we examine whether the reform improves the informational content of audit reports, following our conceptual framework's predictions about the information pipeline from truth (T) to judicial action (D).

Improved transmission ($S \to R$) Figure 3(a) shows that random audits significantly increase the probability that auditors report accounting irregularities, consistent with our framework's prediction that reduced local ties and capture ($\downarrow \text{ties}_{im}$ and $\downarrow \tau_a$) improve the transmission of auditors' private assessments to judge-visible reports.

Better targeting at high risk This improvement in reporting is concentrated among high-risk municipalities, as shown in Figure 3(b), indicating that the reform leads to better reporting that more accurately reflects the underlying fiscal risk of municipalities as measured by our ML-based distress index. This pattern aligns with our model's prediction that a more informative reported signal R pushes borderline high-risk cases over the judicial threshold, concentrating enforcement gains where $\mathcal{M}(X)$ is highest.

Breaking local capture The largest increases in irregularities reporting occur in municipalities that were employing local auditors (residents of the municipality) in the year before the reform (Figure 4). This concentration of effects among municipalities with local auditors provides direct evidence supporting our framework's transmission channel: random assignment helps break local capture (reducing ties_{im}), enabling auditors to report problems they might otherwise have concealed due to personal or professional ties with municipal officials.

Improved production capacity ($T \to S$) Figure 5 reveals that the increase in irregularities reporting is concentrated among municipalities that previously employed auditors with limited professional and territorial experience. To establish this result, we construct four measures of auditor experience, each capturing different aspects of professional competence. We classify each municipality based on characteristics of their last auditor before the reform (measured at $t = -1$) and estimate separate event studies for municipalities with above- and below-median auditor experience along each dimension.

The experience measures capture distinct aspects of auditor capability. Experience as a court auditor (median: 13 years) measures specialized knowledge of municipal oversight procedures and legal requirements. Experience across different provinces (median: 1 province) measure territorial breadth, reflecting exposure to diverse institutional contexts and administrative practices.[4] The results consistently show that the largest increases in irregularities reporting occur in municipalities whose pre-reform auditors had below-median experience across all four dimensions. This pattern supports our framework's production channel, indicating that random assignment systematically upgrades audit quality by improving diagnostic skill and effort ($\uparrow \kappa_a$), with the largest benefits occurring where auditor expertise was previously weakest.

Judicial capital and enforcement decisions ($R \to D$) We test our framework's prediction that judicial capital (ϕ_j) mediates how judges translate improved information into enforcement action. We decompose deliberations according to the rapporteur's pre-reform characteristics. We classify judges based on their baseline experience (2009-2011) across two dimensions that capture different aspects of judicial capital. Total deliberations as rapporteur reflects experience as the primary judge on each case. Rapporteurs handle the investigation and draft the final decision, making this the most demanding judicial role. Prior experience in the specific province measures local institutional knowledge and familiarity with regional administrative practices.

[4]Figure A7 displays an additional measure of experience: general accounting experience (median: 17 years), which captures broader professional competence in financial analysis and auditing standards.

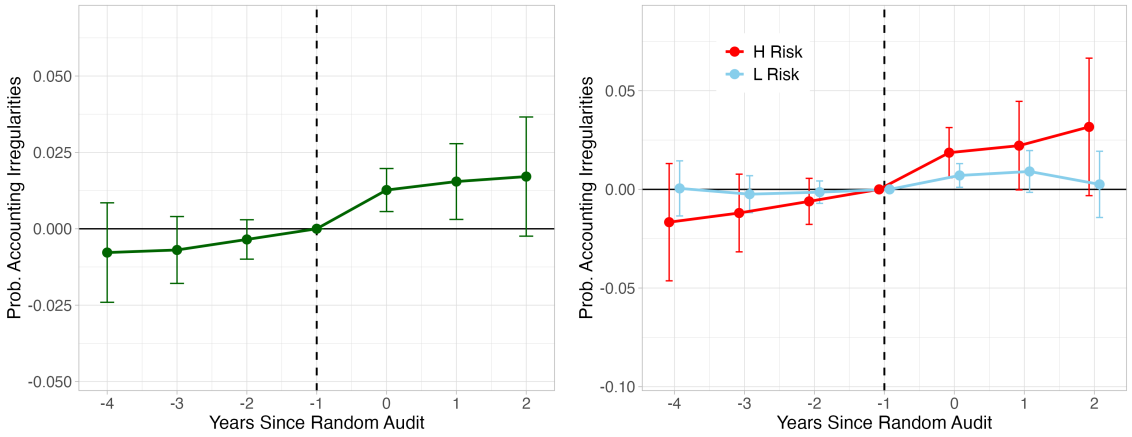
Figure 6 shows the results for the full sample of municipalities. Experienced judges consistently drive larger increases in deliberations following random audits. This pattern holds across both measures of judicial experience, suggesting that different dimensions of judicial capital enhance enforcement effectiveness. The results provide initial evidence that higher judicial capital ($\uparrow \phi_j$) enhances judges' ability to act on improved information signals.

We next examine how this relationship varies with underlying municipal risk. We estimate separate event studies for high-risk and low-risk municipalities, splitting deliberations by judge experience within each risk category. This approach allows us to test whether judicial capital interacts with case characteristics in determining enforcement responses.

The results reveal an important interaction between judicial experience and municipal risk level. Among high-risk municipalities (Figures 6(b),6(e)), deliberations increase sharply when conducted by experienced judges but show only modest increases under inexperienced judges. We interpret this as evidence that higher judicial capital ($\uparrow \phi_j$) enables judges to effectively recognize and act upon credible signals from improved auditor reports. This effect is particularly valuable in complex cases where interpretive skill matters most.

Differential filtering capacity Among low-risk municipalities (Figures 6(c),6(f)), we observe the opposite pattern. Experienced judges show little response to improved information. Inexperienced judges exhibit notable increases in deliberation activity. We interpret this as evidence of differential filtering capacity. Experienced judges efficiently screen out cases unlikely to warrant action. They recognize that even improved signals (R) from low-risk municipalities rarely justify enforcement. Inexperienced judges increase enforcement activity even in marginal cases when information quality improves. This suggests potentially inefficient over-enforcement by judges with lower ϕ_j.

We leverage our theoretical framework to interpret these findings. The differential response provides direct evidence of our framework's $R \xrightarrow[\phi_j]{\text{decision}} D$ channel. Judicial capital amplifies information improvements in a risk-dependent manner, enhancing both efficiency and targeting of judicial oversight.



(a) **Overall Effect** (b) **By Municipal Risk Level**

Figure 3: Random Audit Effects on Accounting Irregularities Reporting
Panel (a) shows event study estimates of the causal effect of random audits on the probability that a municipal auditor reports accounting irregularities in the yearly questionnaires for the Court of Auditors. Panel (b) compares high-risk municipalities (H Risk, above median predicted financial distress, shown in red) with low-risk municipalities (L Risk, below median, shown in light blue). The reference period is $t = -1$ (one year before the first random audit). Standard errors are clustered at the municipality level. 95% confidence intervals are shown.

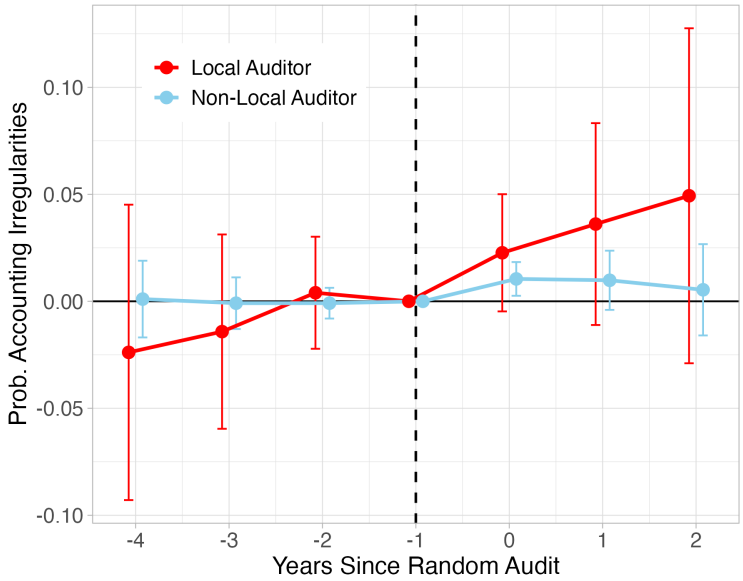
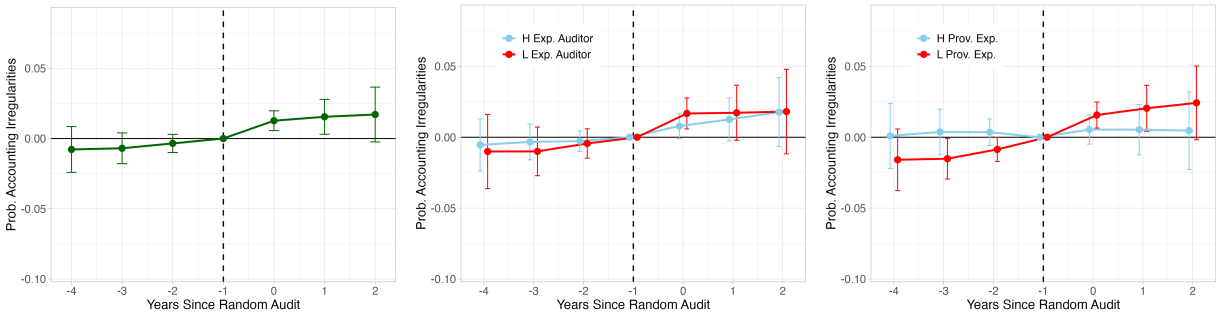


Figure 4: Random Audit Effects on Accounting Irregularities Reporting: Heterogeneity by Auditor Municipality of Origin

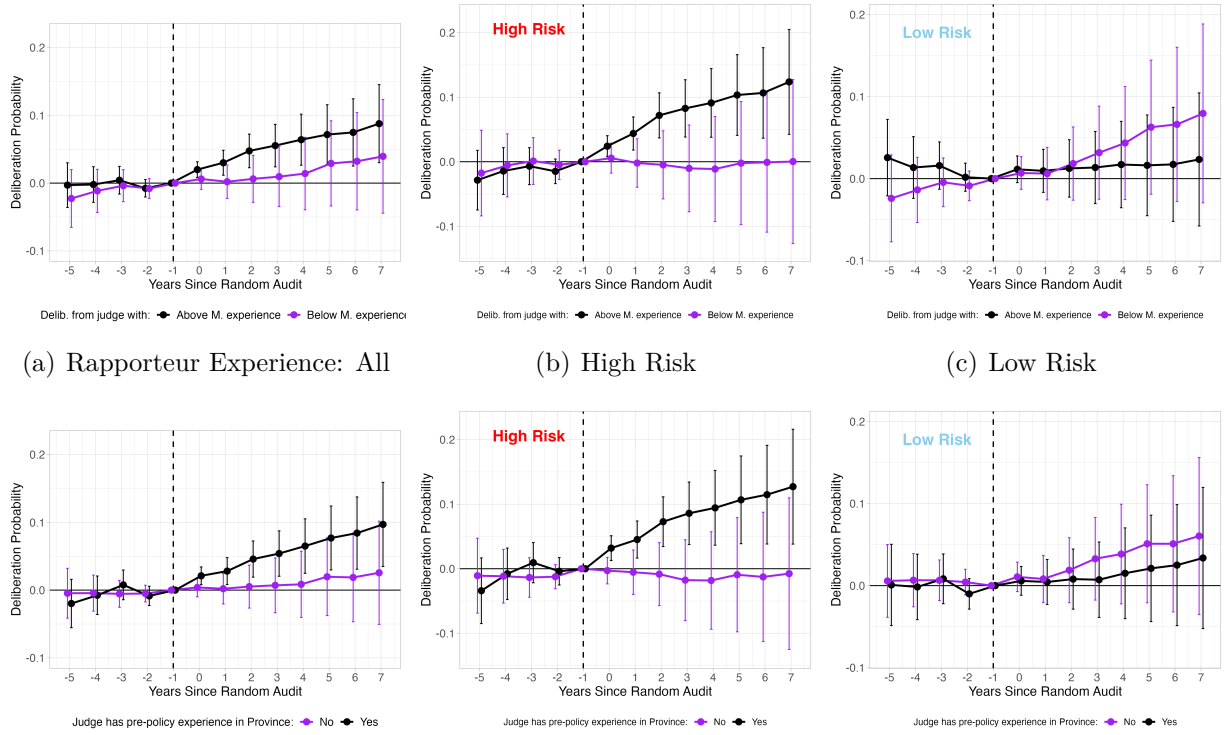
This panel compares municipalities that had local auditors in the year before reform (Local Auditor, shown in red) with those that had non-local auditors (Non-Local Auditor, shown in light blue). An auditor is considered local if he or she was hired in the same municipality in which they have their residence. The reference period is $t = -1$ (one year before the first random audit). Standard errors are clustered at the municipality level. 95% confidence intervals are shown.



(a) **Main Effect** (b) **Experience as an Auditor** (c) **Experience in Different Provinces**

Figure 5: Random Audit Effects on Accounting Irregularities Reporting: Heterogeneity by Auditor Experience

Panel (a) shows the main event study estimate of random audit effects on accounting irregularities reporting probability. Panels (b) and (c) examine heterogeneity based on characteristics of the *last auditor before the reform* (measured at $t = -1$). Panel (b) splits municipalities by their last auditor's experience as a court auditor (median: 13 years), while panel (c) splits by number of different provinces previously served (median: 1 province). In panels (b) and (c), red lines show municipalities whose last auditor had below-median experience/diversity (L), while light blue lines show those with above-median experience/diversity (H). The reference period is $t = -1$. Standard errors are clustered at the municipality level. 95% confidence intervals are shown.



(a) Rapporteur Experience: All (b) High Risk (c) Low Risk

(d) Local Experience: All (e) High Risk (f) Low Risk

Figure 6: Random Audit Effects by Judge Experience

Event study estimates of random audit effects on deliberation probability by judge experience type. Panels (a)-(c) examine rapporteur experience: experienced judges are those above the pre-period median of 44 deliberations as rapporteur (2009-2011). Panels (d)-(f) examine local experience: experienced judges had prior experience in the municipality's province during 2009-2011 (4.2% of deliberations). Black lines show effects for experienced judges, purple lines for inexperienced judges. Each panel reports coefficients from equation 1 where the dependent variable equals 1 if the municipality received deliberations from the respective judge type. Reference period is $t = -1$. Standard errors clustered at municipality level. 95% confidence intervals shown.

References

Abrams, D., R. Galbiati, E. Henry, and A. Philippe (2022). When in rome... on local norms and sentencing decisions. *Journal of the European Economic Association 20*(2), 700–738.

Ash, E., S. Galletta, and T. Giommoni (2025). A machine learning approach to analyze and support anticorruption policy. *American Economic Journal: Economic Policy 17*(2), 162–193.

Ash, E. and W. B. MacLeod (2024). Mandatory retirement for judges improved the performance of us state supreme courts. *American Economic Journal: Economic Policy 16*(1), 518–548.

Athey, S. (2018). The impact of machine learning on economics. In *The economics of artificial intelligence: An agenda*. University of Chicago Press.

Breiman, L. (2001). Random forests. *Machine learning 45*(1), 5–32.

Cawley, G. C. and N. L. C. Talbot (2010). On over-fitting in model selection and subsequent selection bias in performance evaluation. *Journal of Machine Learning Research 11*, 2079–2107.

Chen, T. and C. Guestrin (2016). Xgboost: A scalable tree boosting system. In *Proceedings of the 22nd acm sigkdd international conference on knowledge discovery and data mining*, pp. 785–794.

Corte dei Conti (2022). Banca dati della corte dei conti della giurisdizione e del controllo.

Coviello, D., A. Ichino, and N. Persico (2019). Measuring the gains from labor specialization. *The Journal of Law and Economics 62*(3), 403–426.

Feurer, M., K. Eggensperger, S. Falkner, M. Lindauer, and F. Hutter (2018). Practical automated machine learning for the automl challenge 2018. In *International Workshop on Automatic Machine Learning at ICML*, pp. 1189–1232.

Friedman, J. H. (2001). Greedy function approximation: a gradient boosting machine. *Annals of statistics*, 1189–1232.

Grinsztajn, L., E. Oyallon, and G. Varoquaux (2022). Why do tree-based models still outperform deep learning on tabular data? *arXiv preprint arXiv:2207.08815*.

Hastie, T., R. Tibshirani, and J. Friedman (2009). *The elements of statistical learning: data mining, inference, and prediction*. Springer Science & Business Media.

He, H. and E. A. Garcia (2009). Learning from imbalanced data. *IEEE Transactions on Knowledge and Data Engineering 21*(9), 1263–1284.

Kahn-Lang, A. and K. Lang (2020). The promise and pitfalls of differences-in-differences: Reflections on 16 and pregnant and other applications. *Journal of Business & Economic Statistics 38*(3), 613–620.

King, G. and L. Zeng (2001). Logistic regression in rare events data. *Political Analysis 9*(2), 137–163.

Malani, A. and J. Reif (2015). Interpreting pre-trends as anticipation: Impact on estimated treatment effects from tort reform. *Journal of Public Economics 124*, 1–17.

Mullainathan, S. and J. Spiess (2017). Machine learning: an applied econometric approach. *Journal of Economic Perspectives 31*(2), 87–106.

Rambachan, A. and J. Roth (2020). An honest approach to parallel trends.

Roth, J. (2020). Pre-test with caution: Event-study estimates after testing for parallel trends.

Tibshirani, R. (1996). Regression shrinkage and selection via the lasso. *Journal of the Royal Statistical Society Series B: Statistical Methodology 58*(1), 267–288.

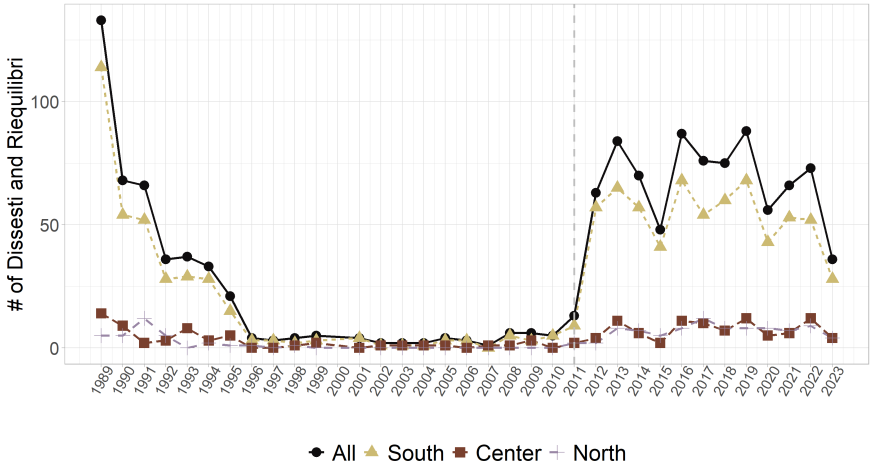
Vannutelli, S. (2021). From lapdogs to watchdogs: Random auditor assignment and municipal fiscal performance in italy. *Job Market Paper: Yale University*.

Varma, S. and R. Simon (2006). Bias in error estimation when using cross-validation for model selection. *BMC Bioinformatics 7*(1), 91.

Zou, H. and T. Hastie (2005). Regularization and variable selection via the elastic net. *Journal of the Royal Statistical Society Series B: Statistical Methodology 67*(2), 301–320.

A Appendix

Figure A1: Municipal Financial Distress Procedures Over Time



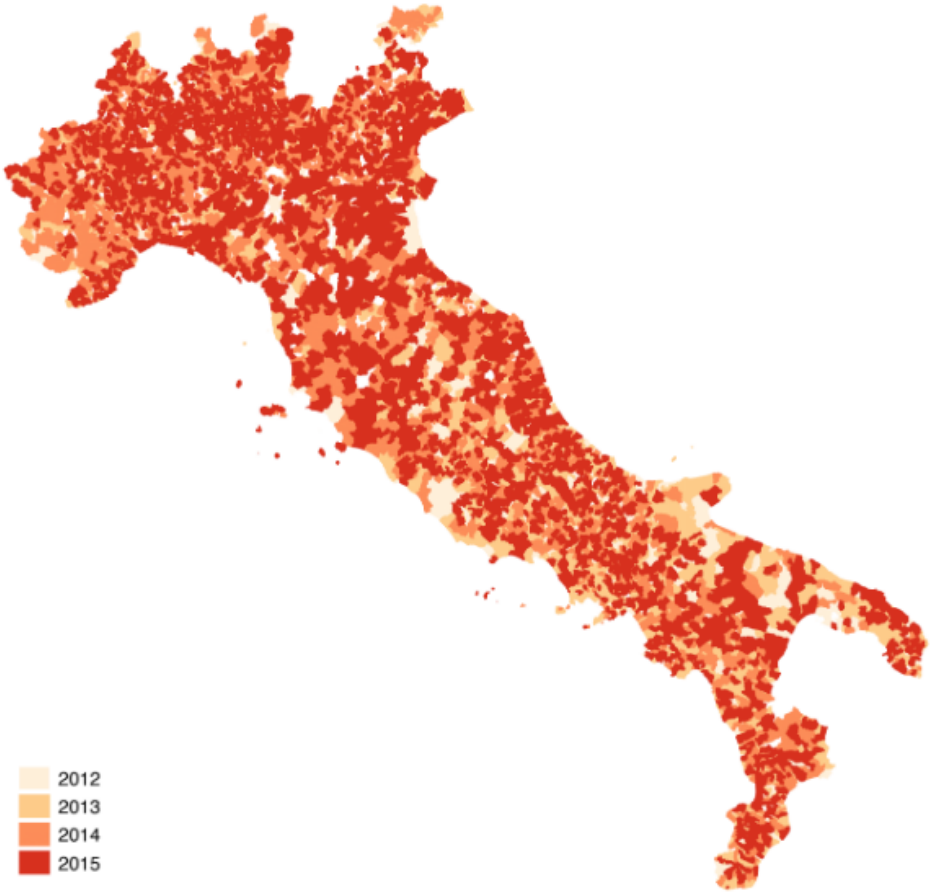
Notes. Time series of municipal financial distress procedures from 1989 to 2023, showing the number of dissesti (bankruptcy) and riequilibri (financial rebalancing) by geographic region. The black line with circles shows national totals, while triangular markers represent the South, square markers the Center, and cross markers the North.

Table A1: Municipal Characteristics that Predict Treatment Timing.

	2012 Cohort	2013 Cohort	2014 Cohort	2015 Cohort
	(1)	(2)	(3)	(4)
Election Cycle	0.00628**	0.000807	-0.00827	0.00118
	[0.00311]	[0.00448]	[0.00522]	[0.00556]
1-5k pop.	-0.00791	-0.0149	0.0137	0.00909
	[0.00779]	[0.0127]	[0.0162]	[0.0163]
5-10k pop.	0.0174	-0.0160	-0.125***	0.124***
	[0.0111]	[0.0156]	[0.0187]	[0.0206]
10-15k pop.	0.0440***	0.00723	-0.161***	0.109***
	[0.0154]	[0.0198]	[0.0222]	[0.0255]
15-20k pop.	0.115***	0.0791***	-0.195***	0.000487
	[0.0249]	[0.0287]	[0.0261]	[0.0322]
20-60k pop.	0.183***	0.114*	-0.187***	-0.110*
	[0.0589]	[0.0621]	[0.0516]	[0.0620]
Above 250k pop.	0.366*	-0.0267	-0.379***	0.0395
	[0.188]	[0.115]	[0.0287]	[0.163]
Mayor Age (log)	0.00537	0.0317	0.0196	-0.0567*
	[0.0158]	[0.0257]	[0.0312]	[0.0331]
Male Mayor	-0.00562	0.0105	-0.0170	0.0122
	[0.0107]	[0.0153]	[0.0207]	[0.0210]
Local Mayor	0.00316	0.00739	-0.00289	-0.00766
	[0.00839]	[0.0122]	[0.0146]	[0.0154]
Mayor Term-limited	-0.00966	-0.00564	0.0286**	-0.0133
	[0.00686]	[0.0105]	[0.0130]	[0.0135]
Observations	5603	5603	5603	5603
R-sq	0.0292	0.0354	0.0568	0.0327
P-value Joint F-test, w/o pop.	0.25	0.58	0.15	0.43
P-value Joint F-test, w. pop.	0.00	0.04	0.00	0.00

Note: The table displays results from 4 separate OLS regressions where the dependent variables are indicators for independent auditor appointment starting in 2012, 2013, 2014, and 2015. The explanatory variables are measured in 2010. The specification also includes region fixed-effects. Robust standard errors are reported in square brackets.

Figure A2: Staggered Treatment, Geographic Variation



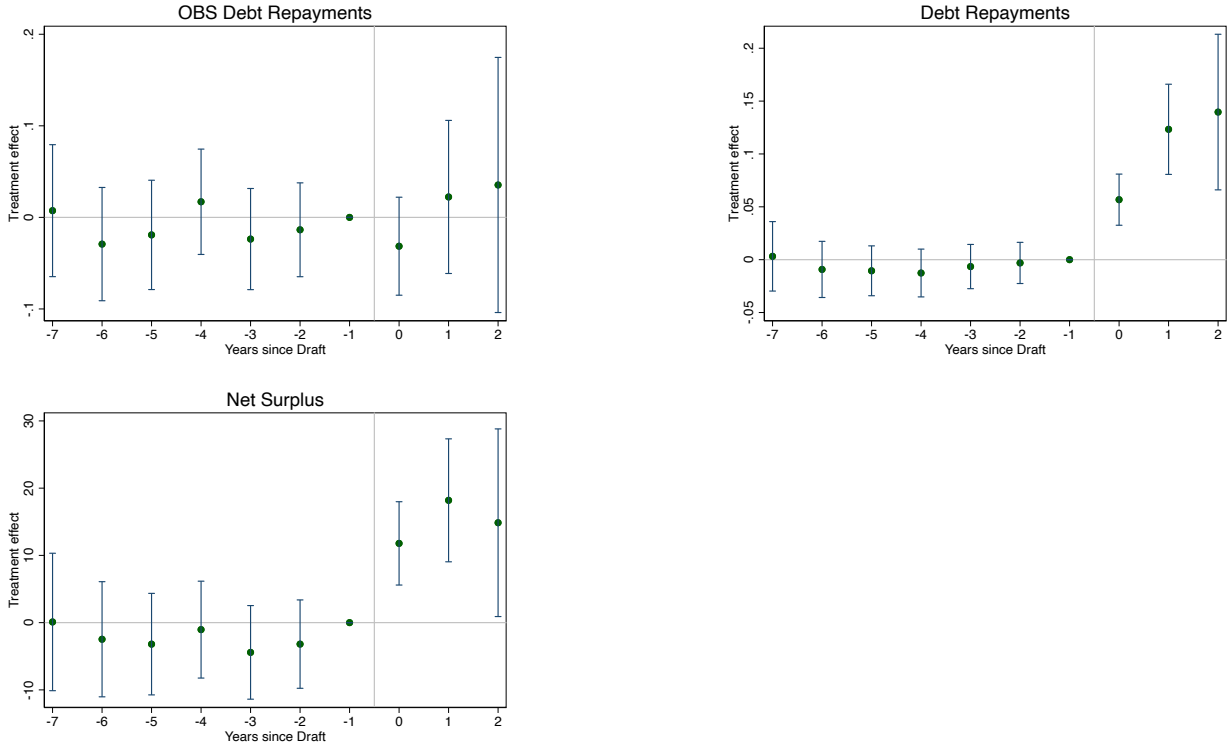
Notes. The figure shows the geographic variation in treatment timing. Darker gradation reflects later treatment timing.

Table A2: Summary Statistics for Municipal Financial Indicators

	Mean	S.D.	Median	p10	p90	p99
A. Fiscal Sustainability						
Off-balance sheet debts	0.003	0.011	0.000	0.000	0.008	0.074
Debt limit index	0.045	0.026	0.041	0.015	0.082	0.113
Cash advances	49.044	174.434	0.000	0.000	103.250	958.582
Budget surplus	171.977	311.397	77.650	5.693	404.464	1591.521
B. Revenue and Spending						
Total expenditures	7.743	1.250	7.795	7.308	8.610	9.559
Total capital expenditures	6.226	1.340	6.259	5.086	7.670	8.923
Total current expenditures	7.251	1.138	7.323	6.941	7.941	8.733
Revenues from taxes	6.331	1.092	6.459	5.666	7.140	7.920
Other revenues	5.535	1.135	5.594	4.598	6.672	7.819
Cash	412.115	655.029	216.366	4.188	921.745	3769.820
Management balance	-9.429	240.078	-0.019	-201.826	165.037	934.981
Current account balance	-53.190	216.057	-30.300	-234.473	109.857	592.055
Capital account balance	-469.742	587.142	-262.335	-1079.818	-81.160	0.000
C. Performance Indicators						
Financial autonomy	0.643	0.213	0.656	0.330	0.931	0.972
Passive carryovers	1807.409	1954.051	1114.682	441.439	3893.866	10683.087
Active carryovers	1562.628	1751.893	953.314	300.785	3480.008	9475.112
Active carryovers to current revenues	0.299	0.143	0.280	0.130	0.500	0.650
Active carryovers to total revenues	0.862	0.591	0.690	0.320	1.650	2.952
Investments	0.482	0.422	0.351	0.128	1.001	2.332
Collecting capacity	0.739	0.162	0.760	0.580	0.906	0.993
Ability to pay	0.689	0.156	0.723	0.522	0.833	0.908
Default	0.001	0.031	0.000	0.000	0.000	0.000

Notes. Summary statistics for Italian municipalities covering 2009-2011. All monetary variables are expressed in per-capita terms and winsorized at the 1% level. Expenditure variables are transformed using the inverse hyperbolic sine transformation. Sample restricted to municipalities subject to random auditor assignment.

Figure A3: The Dynamic Effect of Auditor's Independence on Fiscal Sustainability



Notes: The graphs report coefficients and 95% confidence intervals estimated according to specification 1. Standard errors are clustered at the municipality level. All dependent variables are in per capita terms, $DebtRepayments$ and $OBSDebtRepayments$ are transformed using the inverse hyperbolic sine transformation. All regressions include municipality, population-bins-by-year fixed effects, relative time fixed effects and election cycle fixed effects, as well as the following controls: mayor's age at the beginning of the term (in logs), mayor's gender, term in office and a dummy equal to one if the mayor was born in the municipality.

QUESTIONNAIRE FOR 2015 FINANCIAL STATEMENT

Report to the Regional Control Section of the Court of Auditors (art. 1, paragraphs 166 and following, Law 266/2005) by the auditing body of the Province/Metropolitan City/Municipality of...

Population as of 01/01/2015: ...

Data of the reference person/responsible for compiling the report (President of the collegiate body or single auditor): Name _____ Surname _____ Contact information: Address _____ Phone _____ Fax _____ Email address _____

Details of the report on the 2015 financial statement (to be attached): Report no. _____ dated _____

Details of the resolution approving the 2015 financial statement: Resolution no. _____ dated _____

Entities in experimentation pursuant to Legislative Decree 118/2011 (art. 78) as of 31/12/2014? Yes ☐ NO ☐

All experimental entities are required to complete this Questionnaire except for the excluded parts (questions and tables), with explicit annotation in the margin, in addition to the APPENDIX "EXPERIMENTAL ENTITIES".

We cleaned and systematized the preliminary questions

PRELIMINARY QUESTIONS:

1. Is the Entity in financial default? Yes ☐ from date_____ NO ☐
2. Has the auditing body, during 2015, detected serious accounting irregularities or serious management anomalies, and/or suggested corrective measures not adopted by the Entity? Yes ☐ NO ☐ [If yes, briefly summarize the type of irregularities detected, corrective measures suggested by the financial auditing body and reasons given by the council body to justify any failure to adopt these measures, approximately quantifying the negative impact of the irregularities.] Notes on serious irregularities. Example: "the failure to cancel accounts payable that we had indicated for an amount of approximately 160,000.00 euros, which would have modified the budget result, and the failure to verify revenue for the year 2012 as of December 31, which is being verified during the course of 2013."
3. Does the Entity, from the application of the parameters set forth in Ministerial Decree 18.02.2013, appear to be in a situation of structural financial distress? Yes ☐ NO ☐
4. Has the Entity adopted forms of account consolidation with its own companies, corporations, or other participated organizations? Yes ☐ NO ☐ NOT APPLICABLE ☐
5. Have the regulatory constraints relating to compensation and the number of directors of companies directly or indirectly participated by the Entity been respected? Yes ☐ NO ☐ NOT APPLICABLE ☐
6. Has the auditing body verified that in the 2013 financial statement the sums recorded under title V of revenue, categories 03 and 04 for the taking out of loans, opening of credit lines, securitizations, credit assignments and bond loans were exclusively allocated to financing investment expenditures?
7. Indicate whether the Entity, in the 2015 financial year, has utilized the financial instruments listed below. Do these operations constitute forms of public-private partnership?
   a) Real estate leasing: Yes ☐ NO ☐ Public-private partnership: Yes ☐ NO ☐
   b) In-progress real estate leasing: Yes ☐ NO ☐ Public-private partnership: Yes ☐ NO ☐
   c) Lease-back: Yes ☐ NO ☐
8. Has the Entity implemented project financing operations in the 2015 financial year? Yes ☐ NO ☐

We have the notes on the serious irregularities, but haven't cleaned them yet.

FIRST SECTION:

The rest of the questionnaire asks extremely detailed questions about the municipality balance sheet. It asks the auditor to report all the relevant values indicating specifically the balance sheet items they refer to.

1. Financial Management Results

1.1 Results of Budgetary Management

- Revenue commitments
- Expenditure commitments

1.1.1 Current and Capital Budgetary Management

Current Budget Balance A) Restricted multi-year fund - current portion B) Total Revenue titles I, II, III C) Expenditures title I C1) Commitments allocated to restricted multi-year fund D) Loan repayments part of Title III (categories 2,3,4) E=Current budget difference=A+B-C-C1-D

F) Administrative surplus applied to current expenditure (+)/Deficit coverage (-) G) Capital revenue allocated to current expenditure based on specific legal provisions or accounting principles, including:

- Building permit fees
- Other revenue H) Current revenue allocated to investment expenditure, including:
- Traffic violation fines

Figure A4: Example pages from the Auditor Questionnaire (2010 and 2015 editions)

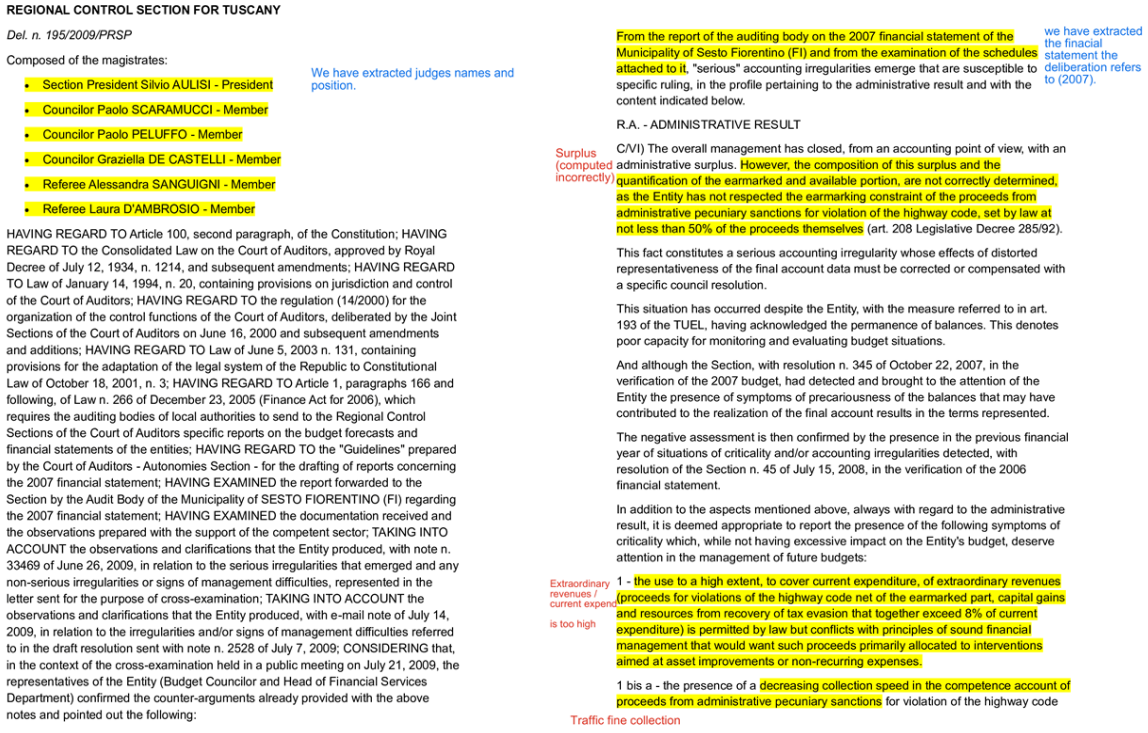


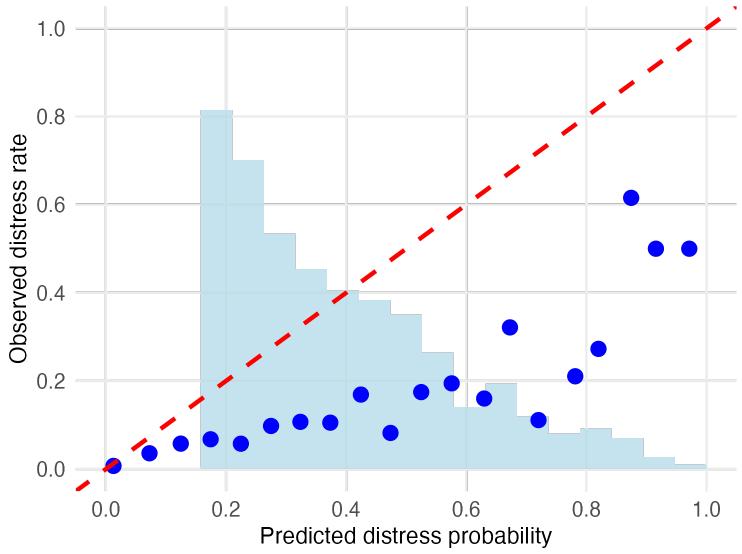
Figure A5: Excerpt from a Court of Auditors deliberation addressed to the Municipality of Sesto Fiorentino (2009)

Table A3: XGBoost Performance Across Different Classification Thresholds

Threshold	Accuracy	Precision	Recall	Specificity	F1-Score
0.10	0.716	0.113	0.821	0.711	0.199
0.15	0.768	0.126	0.742	0.769	0.215
0.20	0.804	0.137	0.675	0.810	0.228
0.25	0.839	0.156	0.623	0.849	0.249
0.30	0.860	0.166	0.563	0.874	0.257
0.35	0.879	0.178	0.503	0.896	0.263
0.40	0.896	0.194	0.450	0.916	0.271
0.45	0.909	0.200	0.371	0.934	0.260
0.50	0.924	0.233	0.338	0.950	0.276
0.55	0.936	0.256	0.265	0.966	0.261
0.60	0.942	0.275	0.219	0.974	0.244
0.65	0.947	0.305	0.192	0.980	0.236
0.70	0.949	0.299	0.132	0.986	0.183
0.75	0.953	0.367	0.119	0.991	0.180
0.80	0.957	0.467	0.093	0.995	0.155
0.85	0.958	0.579	0.073	0.998	0.129
0.90	0.957	0.500	0.020	0.999	0.038

Notes. Performance metrics calculated on test set. Accuracy = (TP+TN)/(TP+TN+FP+FN); Precision = TP/(TP+FP); Recall = TP/(TP+FN); Specificity = TN/(TN+FP); F1-Score = 2×(Precision×Recall)/(Precision+Recall). TP=True Positive, TN=True Negative, FP=False Positive, FN=False Negative.

Figure A6: Model Calibration Plot for XGBoost Financial Distress Predictions

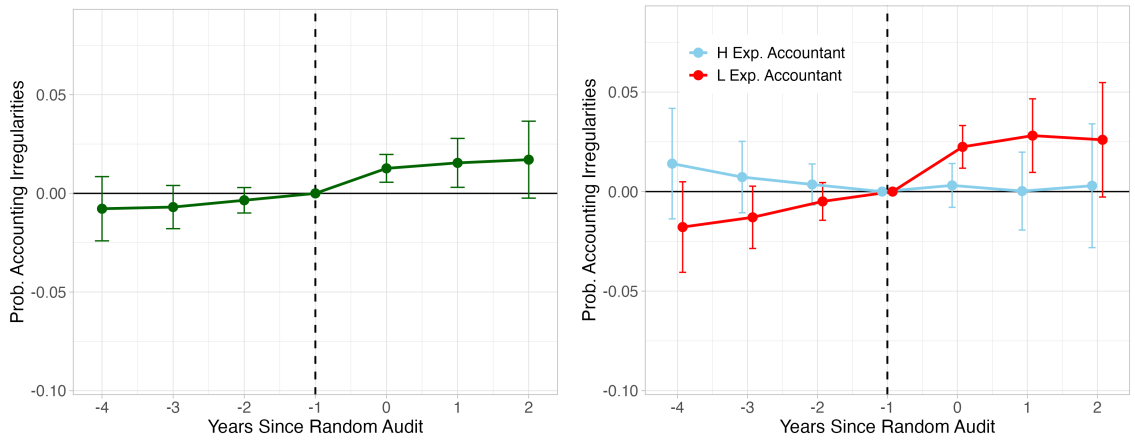


Notes. Calibration plot evaluating the reliability of predicted probabilities from the XGBoost model. The plot combines a histogram (light blue bars) showing the distribution of predicted distress probabilities across municipalities in the test set, with calibration points (blue dots) representing the observed distress rates within binned probability ranges. Each bin contains municipalities with similar predicted probabilities, and the corresponding blue dot shows the actual proportion that experienced distress in that bin. The red dashed line represents perfect calibration, where predicted probabilities would exactly match observed rates. The plot construction uses 20 probability bins of equal width (5% each) to assess model reliability across the full probability spectrum.

Table A4: Random Audit Effects on Court Deliberation Probability

	=1 if received Court deliberation		
	All	High Risk	Low Risk
	(1)	(2)	(3)
Indep. Auditor = 1	0.0113*	0.0183*	-0.0018
	(0.0068)	(0.0101)	(0.0089)
R^2	0.45443	0.44828	0.46684
Observations	84,058	42,029	42,029
Municipality fixed effects	✓	✓	✓
Risk-Region-Year fixed effects	✓	✓	✓
Dep. Var Mean	0.2893	0.3506	0.2293

Notes: Static difference-in-differences estimates of random audit effects on probability of receiving Court deliberation. Treatment = 1 for municipalities subject to random auditor assignment under the reform. Post = 1 for treatment year and after ($t \geq 0$). Column (1) shows overall effect; columns (2)-(3) split by municipal risk (above/below median predicted financial distress in the pre-period). Standard errors clustered at municipality level. Fixed effects: municipality and risk-region-year. Sample: treated municipalities, 2009-2021. Dependent variable mean calculated for pre-treatment period ($t < 0$). *** p<0.01, ** p<0.05, * p<0.1.



(a) **Main Effect** (b) **Experience as an Accountant**

Figure A7: Random Audit Effects: Accounting Experience Heterogeneity
Panel (a) shows the main event study estimate of random audit effects on accounting irregularities reporting probability. Panel (b) examines heterogeneity based on the general accounting experience of the *last auditor before the reform* (measured at $t = -1$). The analysis splits municipalities by their last auditor's general accounting experience (median: 17 years). Red lines show municipalities whose last auditor had below-median accounting experience (L), while light blue lines show those with above-median experience (H). The reference period is $t = -1$. Standard errors are clustered at the municipality level. 95% confidence intervals are shown.