From Lapdogs to Watchdogs: Selecting Monitors in Multi-layered Organizations

Silvia Vannutelli*

This Version: October 2, 2024
First Version: February 2020

Abstract

A fundamental challenge in any hierarchy is aligning incentives across layers. While monitoring can mitigate agency problems, it may itself be rendered ineffective if monitors are corruptible. In this paper, I evaluate the consequences of changes in the design of monitoring institutions for organizational performance. I exploit the staggered introduction of a reform that removed the control of municipal auditors' appointments from local politicians and introduced a random assignment mechanism. I obtain four main findings. First, random matching severs auditors-mayors connections. Second, treated municipalities significantly and persistently improve their net surpluses and debt repayments, in line with national government objectives. Third, the fiscal improvement results from a sizeable increase in tax capacity. Fourth, treatment effects are significantly larger for municipalities that were more at risk of collusion before the reform, and for those that are matched to a more distant or less connected auditor. These findings highlight the value of monitors' independence and illustrate how changes in organizational design can substantially improve governance. **JEL**: H72, D73, H83, H11, H71, H81, H77, M42.
Keywords: Monitoring, Conflict of Interest, Local Public Finance, Political connections

*Northwestern University, Department of Economics and NBER. Email: silvia.vannutelli@northwestern.edu. A preliminary version of this paper was circulated under the title "From Lapdogs to Watchdogs: Random Auditor Assignment and Municipal Fiscal Performance". I am very grateful to my main advisors in this project, Ray Fisman and Daniele Paserman. I am also grateful for the careful comments and discussions of Marianne Bertrand, Michael Best, Erika Deserranno, Namrata Kala, Emir Kamenica, Monica Martinez-Bravo, Raffaella Sadun, Bruno Strulovici, Edoardo Teso and Guo Xu. For helpful feedback I thank Edoardo Maria Acabbi, Emanuele Colonnelli, Ernesto Dal Bo, Gemma Dipoppa, Fred Finan, Leander Heldring, Ilyana Kuziemko, David Lakagos, Christian Leuz, Dilip Mookherjee, Andy Newman, Andrea Prat, Pablo Querubin, Nancy Qian, Maddalena Ronchi, Juan Carlos Suarez Serrato, Joel Slemrod, Luigi Zingales, and seminar participants at Berkeley Haas, Booth, Boston University, Carlos3, CEMFI, Columbia University, Harvard University, HEC Montreal, LMU, Northwestern University, NBER Summer Institute, PSE, Stanford University, Stanford GSB, SciencePo, SSE, USC, UCLA, University of Chicago, Yale University. I am grateful to the Manuel Abdala Fund for financial support. Costas Lambros and Martina Cuneo provided excellent research assistance. I also thank Lorenzo Casaburi, Simone Paci as well as Massimilano Baragona and Carmine La Vita (Ministry of the Interior) for data sharing. All remaining errors are my own.

1 Introduction

A classic way in which decentralized organizations try to mitigate agency problems is by introducing an intermediate layer of monitors between principals and agents. This raises the dilemma of who should select the monitors. Most commonly, organizations delegate this decision to the agents.[1] Discretion of appointment could allow agents to choose someone they trust, with whom they communicate well, who has a deep understanding of their specific circumstances and can thus provide tailored advice.[2] On the other hand, giving agents discretion might lead to conflicts of interest, favoritism, and monitors' capture. The tension reflects a classic delegation trade-off (Aghion and Tirole 1997, Dessein 2002, Alonso and Matouschek 2008), which is further complicated by the presence of a double agency problem, where both the supervisor and the agent's interests could be mis-aligned relative to the principal's ones. Given the theoretical ambiguity, understanding how to effectively select monitors requires empirical evidence which is difficult to obtain, as monitoring institutions are endogenous and it is rare to observe changes in their design that would allow to establish causality.

This paper fills this gap by providing empirical evidence on the consequences of changes in the rules for appointing monitors of local governments. This setting is of direct interest in itself, given the central role of monitoring in multi-layered public sector institutions, where agency problems are complicated by political economy considerations (Mookherjee 2015). Local officials might actively abuse discretion for their private benefit (Bandiera et al. 2009), but also because of misalignment between local constituents' interests and the national government's objectives (Martinez-Bravo 2014, Martinez-Bravo et al. 2022, Kala 2019, Bandiera et al. 2023, Deserranno et al. 2019).

I leverage a unique large-scale reform in Italy that removed discretion from local politicians and introduced a random assignment mechanism for the appointment of municipal auditors. These professional auditors act as an intermediate layer of oversight between municipalities and the central monitoring performed by the National Court of Accounts, akin to the U.S. Government Accountability Office. The audit system aims to ensure sound financial practices and compliance with national fiscal rules, so as to alleviate the moral hazard problems typical of decentralized governments.[3] Differently from third-party and anti-corruption

[1] For example, in corporate settings, shareholders frequently delegate the choice of auditors and board members to management. Similarly, regulators delegate the choice of auditors and rating agencies to the regulated institutions.

[2] Indeed, a large literature in organizational economics provides evidence of the benefits of delegating authority to managers and local agents (Garicano et al. 2005, Kala 2019, Bandiera et al. 2021).

[3] The key issue here is the so-called *soft-budget constraint* (Rodden 2002, Dovis and Kirpalani 2020), whereby local entities might undertax, overspend, and accumulate debt in the expectation of being eventually bailed out by higher levels of government. This behavior does not serve the broader national interest, yet it

auditors, who merely certify documents or verify expenditures ex-post, municipal auditors provide guidance along the entire budgetary process and can thus influence municipal financial decisions ex-ante and in a continuous manner throughout their three-year long terms. Their advising role is particularly important, given the limited financial expertise of local politicians and bureaucrats. Furthermore, auditors' local expertise allows flexibility in the enforcement of fiscal rules.[4] However, their repeated interactions with the mayor pose a risk of bias. For example, auditors might be lenient and allow mayors to keep taxes low and accumulate debt.

This setting is thus particularly relevant because the intervention maps directly onto the theoretical delegation trade-off. If discretion was previously misused to appoint lenient or corruptible auditors, the reform should lead to improved fiscal sustainability. Conversely, if mayors used discretion wisely to appoint knowledgeable auditors who could provide well-tailored advice, the reform might have little to no effect, or even a negative one. My findings show that the reform significantly improves local governments' fiscal performance and tax collection, aligning local agents' behavior to the objectives of the principal, i.e. the national government. Thus, the results suggest that the positive effects of reducing the risk of capture outweigh any potential costs of removing autonomy from local agents.

The reform under study was adopted amidst the European Sovereign Debt Crisis of 2011. Facing pressure to reduce the national debt and improve credibility vis-a-vis the financial markets, the national government became increasingly concerned about local governments' fiscal health, given that local entities in Italy, as in many countries, cannot default on their debts, and their deficits are consolidated in the national budget. Thus, in August 2011 Italy adopted a reform that removed mayors' discretion in auditors' appointments to introduce a random assignment mechanism, aimed at strengthening the oversight of municipal budgets.

The implementation of the new appointment system took place gradually, when the incumbent auditor's term expired, which occurred at different moments across municipalities between 2012 and 2015 for historical reasons unrelated to the reform and beyond mayors' control. This gave rise to exogenous, staggered variation in the timing of the introduction of the new appointment that I exploit for identification, in a similar vein to Martinez-Bravo et al. (2017). I apply a generalized difference-in-differences (DiD) methodology and compare the outcomes of municipalities treated earlier ($Treatment$) to those treated later ($Control$), before and after the arrival of a randomly assigned auditor.[5] Under the identifying assump-

benefits each individual subnational government. These issues led to bailouts in many decentralized countries in history, such as Argentina and Brazil in the 1980s and 1990s, and also in supra-national institutions - with the bailouts of Greece, Portugal and Ireland from the European Union and the ECB during the European debt crisis.

[4]Specifically, auditors might provide a nuanced view to the Court as to whether local non-compliance with fiscal rules is due to unexpected, temporary local shocks or to intentional self-serving behavior.

[5]The reform does not apply to municipalities in the 5 "Special Status" Regions of Italy (Valle d'Aosta,

tion that the treatment timing is uncorrelated with the evolution of outcomes over time, this approach allows me to identify the effect of the change in the appointment system on the financial health of municipalities. In support of this assumption, I first provide evidence that there was no differential selection into treatment timing across a wide range of municipal characteristics that might potentially affect outcomes. I then show that early and later-treated municipalities display similar pre-treatment levels and trends for all the outcome variables. To address concerns raised by the recent econometric literature on differences-in-differences designs with variation in treatment timing, I adopt a stacked-by-event design (Cengiz et al. 2019; Deshpande and Li 2019) which ensures that the treatment effects are estimated only based on comparisons of units switching into treatment to not-yet-treated units.[6]

I start by investigating how random assignment affects the selection and allocation of auditors. I document that the reform severs connections between municipal auditors and mayors along several dimensions likely to impair monitoring effectiveness. The share of auditors residing in the municipality they were hired to monitor declines from 31% before the reform to less than 1% afterward, and the percentage of those sharing the same surname as a local politician falls from 7% to 1%. The reform also induces a change in the composition of the auditors' pool: almost 14,000 new auditors join the list, representing around 60% of the pool of auditors. While these new entrants have no municipal auditing specific experience, they look otherwise very similar to the pre-reform pool of auditors along other observable dimensions. Most importantly, they are of the same age and are equally experienced accountants.

I then study the causal effects of random auditors' assignments on municipal finances. I find that, when subject to more external oversight, municipalities significantly improve their financial performance, as per national government objectives, increasing their expenditure in debt repayments by 9% and their net budget surplus by 8%.[7] This result suggests that randomly appointed auditors are more effective enforcers of fiscal rules. These findings thus underscore the importance of taking into account enforcement mechanisms for rules' effectiveness, a point that has been highlighted in recent theoretical work (Halac and Yared 2019, Dovis and Kirpalani 2020), but that hasn't received empirical support so far. Inspecting the dynamics of the treatment effects, I find that the effects materialize immediately upon the

Trentino Alto Adige, Friuli Venezia Giulia, Sardinia, and Sicily). However, these regions are granted extra autonomy along many margins, including the possibility of setting their own fiscal rules and fiscal monitoring procedures, and therefore do not represent a viable control group and are excluded from the analysis.

[6]Results are robust to the use of alternative estimators, such as those of de Chaisemartin and D'Haultfœuille (2020) and Callaway and Sant'Anna (2021), as well as to a large set of robustness checks.

[7]I also look at off-balance-sheet debt repayments, a rare but critical indicator of budget imbalances. While I do not detect any effect on average, there are significantly heterogeneous patterns based on the risk of pre-reform auditors' capture, suggesting that this is an important proxy for inappropriate practices.

arrival of the (randomly assigned) auditor and remain essentially constant over the auditor's three-year term, suggesting a sudden and persistent shift in budgetary practices. I further provide suggestive evidence that the effects persist in the long run, up to seven years after the adoption of the reform.

I then move to investigate the margins of adjustment. To improve their fiscal stance, municipalities can either increase revenue (e.g., by collecting more taxes) or cut spending (e.g., by cutting investments or current expenditures). I find that municipalities sizeably expand revenues from the main local tax, that is the property tax, by over 20%. I provide evidence that the increase in revenue comes from an expansion in the tax base rather than from a change in tax rates, consistent with an improvement in the enforcement of tax collection upon the arrival of randomly assigned auditors.[8] In line with this explanation, I find that the effects are sizably larger in municipalities with a higher share of housing units and buildings undeclared to the tax authorities.[9] I further show that treated municipalities sizably increase their investments in administrative capacity and that this response is stronger in municipalities with a higher risk of tax evasion, as proxied by the aforementioned measures. Overall, the evidence suggests that, when facing tighter budget constraints, mayors have stronger incentives to fight property tax delinquents they might have tolerated previously for political economy considerations.[10] By providing evidence on the role of monitoring and political economy concerns for tax enforcement, these findings expand our understanding on the determinants of fiscal capacity (Besley and Persson 2009; 2013, Sánchez De La Sierra 2020, Slemrod 2019, Allen et al. 2022, Dzansi et al. 2022, Aman-Rana and Minaudier 2024).

On the other hand, I do not detect any reduction in spending. I further examine whether the reform impacted public good provision and the local economy and fail to detect any significant effects. Furthermore, exploiting new data on personnel composition and task allocation within municipalities, I investigate whether the reform potentially altered the internal functioning of governments. While I detect some evidence of bureaucratic discontent, with an increase in sick leaves and strike days, the internal functioning of local governments remained largely stable, with no adverse effects on local governance. Overall, these results suggest that the reform did not negatively impact local welfare.

[8]Italian municipalities have full discretion on the procedures for local tax collection. One of the critical responsibilities of auditors is to rigorously examine these procedures and to monitor municipalities' effectiveness and efficiency in revenue collection. Auditors can propose changes aimed at improving tax collection, such as expanding the use of cross-checks across data sources to detect evasion or using nudges to increase tax compliance.

[9]I also show that the effects are not simply driven by an improvement in the so-called speed of tax collection (Gagliarducci and Nannicini 2013), meaning that municipalities are not simply improving at collecting revenues that they had already assessed, but they are effectively expanding the tax base, i.e. increasing the size of the assessed revenues.

[10] In closely related theoretical work, Snowberg and Ting (2019) show how political incentives and heterogeneous preferences can lead politicians to alter the state structure and diminish fiscal capacity.

In the last part of the paper, I perform a series of analyses that allow me to shed light on the mechanisms behind the estimated effects. Guided by the insights of models of delegation in organizations (Alonso and Matouschek 2008, Mookherjee 2012), in which the costs of local discretion increase with agents' misalignment with the organizational objectives, I investigate how the treatment effect varies with several proxies of local mis-alignment in the pre-reform period.[11] I find significantly stronger treatment effects of the reform along all these dimensions, suggesting that indeed the removal of discretion in monitors' appointment was more important in places with a higher risk of auditor's capture.

While these results are informative about the pre-reform potential abuse of discretion in appointments, it is also important to understand the factors which reduced the risk of capture in the post-reform period. The switch from discretionary appointment to random assignment changes both the selection of auditors—who gets hired and who decides to participate in the auditors' pool – as well as their allocation to municipalities. To assess the role of selection, I explore heterogeneity along municipality size, as the law imposes minimum experience requirements that allow only incumbent auditors to be assigned to municipalities above 5,000 inhabitants. Focusing on small municipalities, I compare the effect of receiving a new entrant versus an incumbent, thus isolating the role of changes in auditors' selection and separating the causal effect of the random appointment from that of the minimum experience requirement introduced by the reform, which only affect larger municipalities. I find that the treatment effects are largely driven by localities that randomly receive a new entrant as opposed to an incumbent auditor, suggesting that selection indeed plays an important role in small municipalities. That said, I still detect sizable effects when focusing only on large municipalities that are mandated to receive an auditor drafted among the pool of incumbent auditors, thus ruling out the possibility that selection is the only driver of fiscal improvement, and pointing to an important role played by the change in the allocation of auditors. I exploit the exogenous variation in commuting distance arising from random assignment and find that the treatment effects are driven by places where the randomly assigned auditor resides in a municipality at least one hour away from the audited one. This result highlights the importance of social ties and emphasizes the potential role of home bias in adversely affecting agents' behavior (Garicano et al. 2005, Xu et al. 2021, Bandiera et al. 2023). Investigating heterogeneity along these margins also allows me to shed further light on the bias-information trade-off. More distant auditors, as well as new entrants, are less likely to be captured by mayors, but also potentially less knowledgeable. My results suggest that the benefits of reduced capture risk outweigh any potential costs of reduced local knowledge.

[11] I consider (a) indicators of lax financial practices, (b) whether any auditor appointed before the reform was a local resident, (c) a measure of local corruption investigations, (d) a measure of auditor's career concerns, (e) whether the treatment occurs when the mayor is term-limited and thus more likely to misbehave.

Finally, building on the insight from standard models of corruption monitoring (Becker and Stigler 1974, Banerjee et al. 2012, Olken and Pande 2012), the strengthening in monitoring effectiveness can improve outcomes through a direct effect, whereby randomly assigned auditors improve the *detection* of inappropriate practices and induce local politicians to change behavior; and an indirect *deterrence* effect, leading to a change in behavior even without any action of auditors. I detect sizable spillover effects: being closer to a treated municipality has an impact on a yet-to- be-treated municipality that is around 60% of the size of the main effects. This confirms that the change in monitoring institutions also acts through deterrence.

Overall, these findings point to the importance of careful consideration of the monitoring institution design to mitigate agency problems in multi-layered organizations. The evidence suggests that removing appointment discretion of monitors from agents can significantly enhance fiscal performance by preventing capture and ensuring better alignment of agents' behavior with the principal's objectives.

Related Literature: This is the first paper to provide empirical evidence on the consequences of changes in the design of monitoring institutions for organizational performance. In doing so, I contribute to a largely theoretical literature in organizational economics that has investigated the optimal design of monitoring institutions, which stresses the importance of balancing the trade-off between bias and information, and the complications that arise from the presence of a three-tier hierarchy (principal-supervisor-agent), as auditors and auditees might enter into collusive agreements (Tirole 1986, Prendergast 1993, Laffont and Martimort 1997, Strausz 1997, Faure-Grimaud et al. 2003, Ortner and Chassang 2018, Rahman 2012, Strulovici 2021, Angelucci and Russo 2022, Mookherjee and Tsumagari 2023). Despite the large body of theoretical work and the numerous real-world instances of monitoring failures, very little is known empirically about what works, and what doesn't, to improve monitoring institutions. A notable exception is Duflo et al. (2013), who conducted an RCT with 236 industrial plants in India providing estimates of the combined effects of random assignment and changes in the monetary incentives to improve effectiveness of third-party auditors. In this paper, I provide evidence on the effects of an at-scale reform in the public sector that similarly aims at improving the independence of monitors but in a drastically different context. This allows me to expand their findings in several ways. First, the features of the reform I study allow me to isolate the consequences of changes in the appointment mechanism of auditors, while keeping monetary incentives constant. Second, I focus on a public sector setting, in which the oversight problem is complicated by political economy considerations. Third, I study a more complex type of audit regime, in which auditors are hired for longer terms and serve a dual monitoring and advising role, influencing budget decisions both ex-ante and ex-post. The wider scope of their mission also implies higher margins of discretion,

introducing a stronger trade-off between bias and information.

Second, I contribute to the broad debate on the optimal design of organizations (Aghion and Tirole 1997, Dessein 2002, Bolton and Dewatripont 2013, Gibbons et al. 2013, Garicano and Rayo 2016). The allocation of decision-making power across layers and the structure of monitoring are key elements of organizational design. This paper focuses on the allocation of authority over a particularly consequential decision, that is the selection of monitors.[12] My research demonstrates that delegating this decision to agents can be counterproductive, as they might select individuals whose interests align with their own rather than with those of the principal. Two recent empirical papers are of particular relevance. First, my results complement the evidence of Bandiera et al. (2021), which experimentally document the positive effects of removing power of monitors who hold-up procurement purchases in Pakistan. Crucially, their experiment affects a different margin - reducing monitors' authority, holding their selection constant - and focuses on a different type of friction - the misalignment of monitors with both the principal and the agent's objectives. Second, my results closely relate to Deserranno et al. (2023), who show how changes in the allocation of incentives between agents and supervisors matter for performance. In this sense, the papers offer three alternative ways — incentive distribution, the allocation of decision-making power or monitors' selection— in which organizations could improve performance, and highlight the importance of carefully considering the interaction and the relative degrees of misalignment of different agents across layers of the hierarchy.[13]

More generally, by providing evidence on the substantial consequences of changes in monitoring effectiveness for state capacity, this paper also contributes to an important debate in public finance about the organization of the state and its implications for state capacity and tax enforcement. My results complement previous evidence on the role of tax design (Best et al. 2015), individual ability (Best et al. 2023), local knowledge (Balan et al. 2022), technology (Dzansi et al. 2022, Aman-Rana and Minaudier 2024), and tax collectors' incentives (Khan et al. 2016; 2019) for fiscal performance. Most closely, three studies have looked specifically at how changes in the structure of the state matter for state capacity. Mastrorocco and Teso

[12] In this sense, I also contribute to the growing literature on principal-agent problems in the public sector and their implications for state performance (Finan et al. 2017, Besley et al. 2022). While some empirical papers have documented the effects of *introducing* monitoring tools to reduce corruption or shirking and improve accountability (Olken 2007, Ferraz and Finan 2011, Bobonis et al. 2016, Dal Bó et al. 2021), research is silent about the determinants of monitoring effectiveness. Recent work by Henn et al. (2020) provides evidence on the consequences of *ineffective* monitoring. They show how a certification program for local governments' effectiveness fails to improve actual public service delivery in Mexico, as corrupt local politicians are likely to manipulate the certification process.

[13] My results also relate to recent empirical papers analyzing the consequences of shifting other types of decision-making rights within firms or governments. Kala (2019) explore changes in decision-making rights from government officials to managers of SOEs, Fenske et al. (2023) study the re-allocation of authority from senior to junior officers for granting environmental permits.

(2023) provide evidence on the importance of reducing information costs, which plausibly improved monitoring capacity, for expanding the state in the US. Chambru et al. (2021) empirically show how local administrative presence matter for fiscal performance, based on the establishment of new local capitals in 19th century France, while Sánchez De La Sierra (2020) show how differences in the ability to hide the tax base lead to the development of different ways of organizing local taxation by bandits in the DRC. I expand this work by providing evidence of a new, key mechanism behind improvements in local tax collection, that is the presence of effective monitoring institutions that allow for a realignment of local leaders' incentives with national objectives and that limit moral hazard problems that may lead to under-investment in tax enforcement.

2 Institutional Context

2.1 Fiscal Monitoring

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) 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.[14] The central government also allows municipalities to undertake new debt to refinance existing debt or to refund previously emitted bonds, provided that this allows them to achieve debt service savings and that the new funds are still used to finance investment spending (*Law 311/2004*).

Fiscal revenues come from two main sources: (1) local taxes, among which the most relevant is the property tax; (2) local fees (e.g. building permits, traffic fines, and fees for other services). One of the main responsibilities of mayors is to propose the annual provisional

[14]Municipalities can undertake new debt only if the total amount of debt service does not exceed 15% of current revenues of the two preceding years. While in theory they can borrow from private banks, the vast majority of debt (89%) is granted by the so-called "Cassa Depositi & Presiti", a state-owned body that operates to promote national and local governments' investment projects. Most of the municipal debt is therefore implicitly guaranteed directly by the national government.

budget and final budget to the municipal council that approves them with majority rule. The mayor enjoys a substantial amount of executive power and discretion over tax collection, tax rates, and budget allocations. From 1999 onwards, all Italian sub-national entities have been subject to the so-called "Domestic Stability Pact" (DSP), the national counterpart of the European Union's Stability and Growth Pact, adopted in 1997.[15] The pact prescribes a set of fiscal rules that has undergone several changes over time, but which generally requires municipalities to limit the growth of their so-called fiscal gap - defined as the deficit, net of transfers, and debt service - below a given threshold (Grembi et al. 2016). It is important to note that in Italy, as in many countries in the European Union, local entities cannot default on their debts, deficits (or surpluses) run by local governments are consolidated in the national budget, and the bailout of sub-national entities is formally regulated by law. Subnational government debt accounts for 11% of GDP and represents 7% of general government debt.[16] Therefore, especially during the economic crisis, the national government has frequently made use of the pact by strengthening its requirements to shift part of the national debt reduction required by the European Union towards local entities.[17] In this context, enforcing the respect of fiscal rules and ensuring the fiscal sustainability of local governments became critical for the national government, motivating a surge in attention to fiscal monitoring procedures.

2.2 The Role of Municipal Auditors

Since 1990, all municipal budgets need to be audited and certified by a board of auditors.[18] The board is composed of one member for municipalities with less than 15,000 inhabitants - that represent 90% of municipalities - and three members for larger municipalities. Prior to the reform, the board used to be nominated with a majority vote of the municipal council,[19] for a 3-year term, renewable for an extra term. The board cannot be dismissed for any reason unless the council can prove auditors' faulty inaction or breach of official duties. Each of the board members can have at most eight contemporaneous appointments. The Italian market for auditors is characterized by a high degree of competition. Differently from the U.S., where auditors are mostly belonging to large private firms, the auditing job in Italy is carried out

[15] *Legge Finanziaria* 23 December 1998, n. 448, modified by the subsequent *Legge Finanziaria* 23 December 2000, n. 388.

[16] In the OECD, in 2013, sub-central government debt represented on average 17% of total public debt. See https://www.oecd-ilibrary.org/economics/national-accounts-at-a-glance-2015_na_glance-2015-en.

[17] For example, between 2010 and 2017, the national government cut resources for local governments by over €12 billion: partly via a reduction of inter-governmental transfers (9 billion), and partly through a tightening of the rules of the pact (3.3 billion).

[18] *Law 142/1990, Article 57.*

[19] Given that a mayor always enjoys a majority in the council, she is the one in charge of the decision.

by self-employed Certified Public Accountants, who work for municipal governments but also offer 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. Working as a municipal auditor provides visibility and is likely to allow auditors to expand their private clients' network, and thus represents a potentially attractive career choice, beyond its monetary compensation. As common in markets for experts (Ronen 2010, White 2010), however, discretionary appointments by mayors likely limited the reliance on a restricted set of "connected" auditors based on patronage networks.

From its very first inception, the board of auditors was assigned a double role of both monitoring and advising. The original law instituting the board of auditors describes its role as follows: *The board of auditors collaborates with the municipal council in its control and governing function, supervises the accounting and financial regularity of the management of the entity, and certifies the correspondence of the balance sheet to the economic outcomes, drawing up a specific report, which accompanies the proposed board resolution of the balance sheet. The auditors shall provide advice and proposals aimed at achieving better efficiency, productivity, and cost-effectiveness in local government management.*

The tasks of auditors have progressively been expanded over time between 1990 and the early 2000s,[20] but they were not affected by the reform. Table A1 provides a summary list of these tasks, as indicated in the law. Importantly, and differently from the typical auditor in a private sector context, municipal auditors are not simply auditing and signing off the balance sheets ex-post, but they have a significant role in influencing policy-making ex-ante. According to the National Association of Municipal Auditors (ANCREL), auditors need to perform over 200 tasks every year, summing up advising and supervising ones. For example, in his advising role, the auditor is required to provide written opinions on several important acts, including municipal protocols and regulations about tax collection. The auditor is also charged with verifying the accuracy of the information reported in the budget/balance sheet in terms of both sources of revenues as well as chapters of expenditures and could ask for clarifications regarding revenue projections. Auditors are also required to advise municipalities on ways to improve efficacy, efficiency, and cost-effectiveness in the management of public funds. Last but not least, auditors are required to verify the respect of fiscal rules.

The results of the audit review, as well as the auditors' opinions and suggestions, are included in an audit report. This report is reviewed by the council at the time of final budget or balance sheet approval. While the auditors' opinions are not legally binding, the council is encouraged to either follow their suggestions or formally justify any inaction. After the approval, the auditor is required to fill in a detailed questionnaire and transmit all the relevant

[20] *art. 239, TUEL, law, and d.p.r. n. 97/2003, attachment 17*

documents, including the approved balance sheet and the audit review, to the *National Court of Auditors*. This body is a Supreme Audit Institution, it serves as the central monitor and it reviews all the material sent by the auditors. 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. For example, the first question in the Court's questionnaire to auditors pertains to whether they have identified any irregularity and, if so, whether the municipality has followed the recommendations of the auditors. On average, the Court issues around 2000 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 Courts' remarks.

2.3 The Auditors' Appointment Reform

In August 2011, at the peak of the European Sovereign Debt Crisis, the national government suddenly changed the mechanism of auditors' appointment (*Law 148/2011*). According to the new law, *"Starting from the first renewal after the adoption of the law, financial auditors of local entities will be chosen by a random draw from a list in which the following subjects can ask to be included: a) those currently included in the regional list of auditors, b) any officially authorized Certified Public Accountant"*. In February 2012, the government outlined the procedures to participate in the list, including experience requirements varying with population thresholds, so as to avoid sending inexperienced auditors to audit large municipalities. For municipalities up to 4,999 inhabitants, auditors shall have been certified public accountants for at least 2 years. For those between 5000 and 14,999 inhabitants, they shall have been certified public accountants for at least 5 years and have been appointed as municipal auditors at least once before. For those above 15,000 inhabitants, auditors shall have been certified public accountants for at least 10 years and have been appointed as municipal auditors at least twice before.

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), which are granted extra-autonomy along many margins, including the possibility of setting their own fiscal rules and fiscal monitoring procedures.[21] My analysis

[21]In 2016, Friuli Venezia Giulia, Sardinia, and Sicily adopted laws introducing similar drafting procedures that generally preserve some discretion for local governments. In Friuli, municipalities can choose among the $3 \times N$ drafted candidates. In Sicily, the drafting procedure was adopted upon the pressure of the Regional

is therefore limited to municipalities in the 15 "Ordinary" regions.

The random drafting procedure is carried out by prefectures, the local offices of the Ministry of the Interior, via a standardized, computer-based system. Municipalities are required by law to inform their local prefecture at least 2 months before the current auditor's term expires. For each draft, the number of drafted candidates is equal to $3 \times N$, where N is the number of auditors to be appointed. The prefecture immediately informs the municipality of the draft results, and the municipal council shall proceed with the appointment of the first drafted candidate(s), after having checked that she is eligible for the specific appointment.[22] The formal appointment notice also contains the wage offered to the auditor for the three-year mandate. Auditors' compensation is regulated nationally, including a minimum floor that is estimated to cover the expenses to effectively perform the job, and maximum caps that vary by population threshold, leaving mayors very limited margins of discretion in setting compensation within this small bandwidth, and thus resulting in very limited variation in the actual compensation across municipalities of similar size. This aspect of the auditors' appointment was not affected by the reform. Shall the first among the drafted auditors refuse the appointment or be deemed to be incompatible, the municipality proceeds to contact the second-drafted auditor. Thus, mayors cannot choose among the set of drafted auditors but have to proceed following the order of the draft.[23]

3 Data

To measure financial outcomes, I use detailed data on all municipal budgets provided by the Italian Ministry of the Interior. These are the so-called "final balance sheets" (*Bilanci Consuntivi*), reporting effective revenues and expenses for the previous year, which need to be presented and approved by April 30th of the following year. The data include detailed information about municipal governments' spending and revenue sources, such as local taxes, current expenditures, investments, debts, and transfers. For the sake of my analysis, I want

Council of Auditors, which repeatedly denounced political pressures toward the local auditors. In contrast to the national system, however, the drafting is carried out directly by the City Council, which exposes it to a higher risk of illicit acts.

[22]Causes of ineligibility include: a) being a member (or have been a member in the past 2 years) of the municipal council or executive body; b) being a local bureaucrat; c) having already reached the maximum number of concurrent appointments (8).

[23]The data shows that rejections are a relatively rare occurrence: around 20% of first-drafted auditors in total reject the appointment, and in most cases the second among the drafted auditors accepts the appointment. In less than 3% of municipalities, a new draw was needed as none of the candidates accepted the position. Table A2 investigates the determinants of auditors' rejections. The most significant predictor is the distance between the municipality of residence and the municipality of appointment. All other characteristics have very limited explanatory power. Importantly, rejections are uncorrelated with indicators of the financial performance of municipalities.

to inspect whether randomly-assigned auditors provide more independent monitoring and thus improve the fiscal sustainability of municipalities, as per the national government's objectives. Therefore, I focus on a set of indicators that are used by the national government and/or by the National Court of Auditors to monitor the fiscal stance of local governments. First, I look at the *NetSurplus*, which is calculated as total revenues, net of transfers from other levels of government, minus total expenditure, net of interest payments on outstanding debt. The *NetSurplus* is a clean measure of the actual fiscal sustainability of municipal finances, as it nets out factors that are not under the direct control of the current local administration, such as changes in the level of transfers from higher levels of government or the consistency of interests on debts undertaken by previous administrations. Furthermore, this variable is the main target of the national government's fiscal rules (Grembi et al. 2016).

Second, I look at *DebtRepayments*, capturing the overall amount of municipal finances devoted to reducing the stock of existing debt. This variable is also closely monitored, as municipal debt is part of the national debt. Municipalities are required by the national government to use any surplus towards the repayment of existing debt.[24] Third, I look at *OBSDebtRepayments*, representing the total amount of Off-Balance-Sheet (OBS) Debt that is recognized and re-payed by municipalities in a given year. This variable is a signal of potential profound imbalances. The National Courts of Account defines OBS Debt as any liability that is undertaken in violation of the municipal budget rules, such as any expense incurred that was not previously authorized and accounted for in the provisional budget. [25] Municipal bureaucrats that approved unexpected expenses that count as OBS liabilities but cannot be recognized in the balance sheet remain liable for those expenses. In other words, they should pay out of their pockets for those expenses. Municipalities are allowed to recognize OBS liabilities that arose in previous years and had not been previously recognized. To finance *OBSDebtRepayments*, municipalities may employ any existing current surplus and, residually, they may undertake new debt, provided they justify the impossibility of financing *OBSDebtRepayments* out of their own resources. Auditors are explicitly asked by the Court to closely monitor the presence of OBS debts and the process of recognition. They are asked to review the reasons why the OBS debts arose and evaluate the proposed repayment methods. The auditors are also required to report about *OBSDebtRepayments* in a specific section of the questionnaire they need to complete for the Court after the balance sheet has been approved. Given that they represent an exceptional violation of the rules, *OBSDebtRepayments* are a relatively rare phenomenon, and this variable is equal to 0 for over 80% of the sample.

[24] Particularly from 2010 onwards, the national government pushed municipalities to quickly repay their outstanding debts undertaken in the past, as they were subject to significantly higher interest rates.

[25] These include, for example, unexpected expenses incurred for emergency interventions or other expenses motivated by local public necessities, but also expenses arising from recapitalizing SOEs.

All variables are measured in per-capita terms (winsorized at the 1%), expressed in 2018 euros equivalent, and reflect accrual accounting. All spending variables, which are always greater than 0, are then transformed using the inverse hyperbolic sine transformation. Table 1 shows the summary statistics of all the outcome variables used in the analysis. The key treatment assignment variable for my analysis is the date when each municipality appoints an auditor with the drafting procedure for the first time. All information about currently randomly appointed auditors is available on the website of the Ministry of the Interior, separately for each municipality. I created a complete historical database for the universe of municipalities by scraping the website and obtained a dataset containing information on the draft dates and identities of all the auditors drafted from 2012 onwards. I combine this data with information from the candidates' pool, published yearly by the Ministry of the Interior. The list contains information about age, gender, municipality of birth, municipality of residence, region of residence, level of professional experience (proxied by the number of years they have been certified as public accountants), and information on the previous service as an auditor. This allows me to partially overcome the fact that I don't have information about auditors' appointments for the pre-reform period. By exploiting the information reported by auditors when they sign up for the lists, I am able to recover the identity of the last auditor appointed before the reform for over 40% percent of my sample.[26] I supplement the above information with further data to account for municipal characteristics that might affect fiscal sustainability. For a detailed description of controls as well as sample restrictions, see Section A2. My main analysis sample consists of 5603 municipalities in Ordinary regions for the years 2007-2015.

4 Empirical Strategy

My identification strategy leverages the staggered introduction of the auditors' appointment reform across municipalities. As explained in Section 2, while the reform became effective in December 2012, the treatment date varies across municipalities depending on the expiration date of the incumbent auditor's term. For example, municipalities that appointed auditors in August 2011, just before the announcement of the reform, would be "treated" with a draft-appointed auditor only from September 2014 onwards. Two key aspects allow me to exploit the introduction of the reform for my identification. First, there exists wide variation across municipalities in the auditor's appointment date before the reform due to historical reasons, which are uncorrelated with other municipal characteristics.[27] Second, mayors cannot affect

[26] I have information for those municipalities that appointed auditors who: a) decide to participate in the drafting procedure, b) want to participate in drafts for larger municipalities.

[27] Most importantly, treatment timing is not correlated with other factors that might affect local fiscal policy, such as the election cycle, which also varies across municipalities but in a way that does not overlap

the auditors' appointment date or auditors' term length in any way. Thus, treatment timing is plausibly exogenous. My design, therefore, compares the evolution over time of outcomes of municipalities that are treated earlier to municipalities that are treated later. Figures 1 and A1 display the sources of identifying variation. Figure 1 shows the total number of municipalities with randomly-assigned auditors over time. Figure A1 displays their geographic distribution, highlighting the lack of any geographic patterns by treatment timing.

Given the presence of staggered treatment assignment, I adopt a stacked-by-event design to perform my estimation (Deshpande and Li 2019, Cengiz et al. 2019, Aneja and Xu 2023). This estimator allows me to account for the recently-highlighted potential pitfalls of two-way fixed effects estimators in the presence of staggered adoption (Goodman-Bacon 2021, de Chaisemartin and D'Haultfœuille 2020, Borusyak et al. 2021a).[28] In Section 5.3, I also show the robustness of my estimates to the use of alternative estimators recently developed by Callaway and Sant'Anna (2021) and de Chaisemartin and D'Haultfœuille (2020).

The "stacked" method simply amounts to consider each reform's wave of adoption as a separate sub-experiment, creating a separate dataset for each of the three treatment waves before the last one (2012, 2013, 2014). In each of these datasets, municipalities that receive the randomly-drafted auditor in that year are considered treated, while municipalities that will experience the treatment in later years serve as a control, and event-time dummies relative to the year of treatment for the specific sub-experiment.[29] Municipalities that experience treatment in the last year, 2015, serve only as controls. My main estimating equation is as follows:

$$Y_{mt} = \alpha_m + \delta_{pt} + \beta_0 Treated_{mc} + \beta_{DD} Treated_{mc} \times Post_{mt} + \sum_{k=-7}^{k=2} \beta_k * D^k + X'_{mt}\zeta + \epsilon_{mt} \quad (1)$$

Where $Treated_{mc}$ is a dummy that takes the value of 1 if the municipality m is a treated municipality in cohort c. This variable is not collinear with the municipality fixed effect as, given the data structure, the same municipality can appear multiple times both as treated and as control;[30] $Post_{mt}$ is a dummy equal to 1 for the years in which a randomly-assigned auditor

with auditors' terms - see Figure A12.

[28]The main concern raised by the literature is that the estimate of the average treatment effect β_{DDtwfe}, estimated from a more "standard" two-way-fixed effect specification of the form $Y_{mt} = \alpha_m + \delta_t + \beta_{DDtwfe} Treated_{mt} + X'_{mt}\zeta + \epsilon_{mt}$ is a weighted average of all the possible 2x2 comparisons in my sample. Therefore, it is also estimated using comparisons among already-treated units and not-yet-treated units, where the already-treated units serve as controls. This induces a bias in the presence of heterogeneous treatment effects across groups experiencing treatment at different points in time.

[29]For example, in the dataset for the first cohort of treatment, event-time dummies are defined in time relative to 2012.

[30]For example, a municipality treated in 2013 would appear as a control municipality for the cohort of 2012 and as treated for the cohort of 2013 only.

is active, while the D^k is a set of relative event-time dummies, that take the value of 1 if year t is k periods after (or before, if k is negative) the treatment. The inclusion of these event-time dummies allows me to control for event-time trends that are not captured by the calendar year fixed effects. I also estimate separate calendar year effects δ_{pt} for seven population bins so that municipalities belonging to different population size classes are allowed to evolve along different trends. The matrix of time-varying controls X_{mt} includes election-cycle fixed effects, and a set of pre-determined characteristics of the mayor in power at the time of auditors' appointment, such as a dummy for whether the mayor is term limited, a dummy for whether the mayor was born in the municipality, as well as the gender and the age of the mayor at the time of election (measured in logs). Standard errors are clustered at the municipality level (Bertrand et al. 2004), accounting for the possibility of serial correlation over time and the repeated appearance of municipalities in the datasets as both treatment and control units.

To investigate pre-trends, as well as the dynamic evolution of the treatment effect, I also estimate a non-parametric event-study specification:

$$Y_{mt} = \alpha_m + \delta_{pt} + \beta_0 Treated_{mc} + \sum_{k=-7}^{k=2} \gamma_k * D^k \times Treated_{mc} + \sum_{k=-7}^{k=2} \beta_k * D^k + X'_{mt}\zeta + \epsilon_{mt} \quad (2)$$

In this specification, the coefficients of interest are the γ_k's, measuring the change in outcomes of treated municipalities k years after treatment, relative to the pre-treatment year, relative to the change in outcomes of control municipalities, who have yet to be treated. I estimate treatment effects up to three periods from treatment onset (i.e. $k = 2$) thus covering the entire term of the draft-appointed auditor.

4.1 Identifying Assumptions and Threats to Identification

The key identifying assumption of my design is the absence of differential trends between municipalities experiencing treatment at different points in time; that is, absent the change in auditor's appointment, the outcomes of municipalities belonging to different treatment cohorts would have evolved similarly over time. My empirical strategy already allows me to rule out various concerns that could otherwise impair my ability to interpret the results as causal. First, the inclusion of municipality fixed-effects rules out that the results are driven by time-invariant differences in outcomes across municipalities. Second, the inclusion of calendar time effects allows me to exclude that results are driven by factors that affect the evolution of outcomes over time in a way that is common across municipalities, such as changes in macroeconomic conditions.

Given the exogenous and historical nature of the difference in audit cycles across municipalities, the timing of expiration of the existing auditor, and thus the timing of treatment, should be uncorrelated with the evolution of outcomes over time. To provide evidence in favor of this assumption, in Table A3 I look at whether municipalities treated at different times display any significant differences in their observable characteristics (Panel A) and in the levels of outcomes (Panels B and C) in the pre-reform period. The first four columns of the table show the average levels of each variable, by cohort of treatment. I then perform a regression of each variable on cohort dummies and in column 5 I display the p-value of a joint F-test for the null hypothesis that all are jointly equal to zero. I fail to reject that all dummies are jointly equal to zero for all variables, except for population.[31]

The presence of significant coefficients should not be interpreted as evidence of differential selection of municipalities across cohorts, as mayors cannot manipulate in any way the expiration date of their current auditor, which is the only variable determining the assignment across cohorts. Results simply suggest that, for historical reasons, larger municipalities are slightly more likely to have their auditors' terms expiring in 2012 or 2013, as opposed to later years. To account for this imbalance, in all my analyses, I control non-parametrically for the presence of differential trends by population size, by including population-by-year fixed effects.[32] These controls also allow me to account for potential effects induced by changes in the composition of the pool of auditors for municipalities of different sizes due to the minimum experience requirements for municipalities above 5,000 and 15,000 inhabitants.

Panels B and C of Table A3 repeat the same exercise as Panel A but look at the levels of outcomes in 2010. Here, I add an additional column, where I report the p-value of the joint F-test of the joint significance of cohort dummies in a regression that also includes the observable characteristics of Panel A. While some outcomes display small significant unconditional differences, after conditioning on the set of covariates, no significant difference can be detected. Notably, testing for the absence of a significant difference in pre-treatment *levels* of outcomes by treatment timing is a stronger test, as the standard difference-in-differences assumptions only require no counterfactual differences in the trends of the outcomes.

Third, and most importantly, one can inspect whether outcomes exhibit parallel trends in the pre-reform period as a way to assess the plausibility of the assumption of parallel

[31]This is further confirmed in Table A4, where I directly look at predictors of treatment timing. Here, I perform two different F-tests, whose p-values are reported at the bottom of the Table. When excluding population dummies from the test, across columns, I fail to reject the null hypothesis that all coefficients are jointly equal to zero.

[32]I construct a set of dummies for the following population bins (reflecting the ones used by the Italian Statistical Office to classify municipalities along population size): 0 to 1,000 inhabitants; 1,000 to 5,000; 5,000 to 10,000; 10,000 to 15,000 inhabitants; 15,000 to 20,000; 20,000 to 60,000 and above 60,000 inhabitants. I then interact each category with year fixed-effects to estimate the population-by-year fixed effects δ_{pt}.

trends in the post-reform period.[33] In Figure 2, I estimate the flexible dynamic specification of Equation 2 and show that I do not detect any evidence of differential pre-event trends for any of the outcomes. Another source of potential concern is the presence of anticipation effects (Malani and Reif 2015), due to the fact that, at least in theory, municipalities knew exactly the timing of treatment. There are two types of anticipation effects that could be happening in this setting. On the one hand, one could think that later-treated units, in expectation of the arrival of a stricter auditor, would start responding earlier, in order to smooth the adjustment of their finances. This type of anticipatory behavior is not particularly problematic, as it would lead to an under-estimate of my treatment effect, as I would compare my treated units to later-treated units who are already starting to behave like treated ones before they become treated. On the other hand, one might think of a scenario in which the later-treated units respond in the opposite direction, meaning that they start over-spending and accumulating debt in the expectation of not being able to do so anymore in the future. This second type of anticipatory behavior is more worrisome, as it would potentially lead to an over-estimate of the treatment effects. I exploit one of the stacked-by-event design features to provide evidence of the type of anticipation effects potentially at play. As can be seen in Equation 2, the design allows me to separately control for both calendar-time fixed effects (the δ_{pt}) and event-time fixed effects (the terms $\sum_{k=-7}^{k=2} \beta_k * D^k$). The latter are estimated off of the control group.[34] By looking at the β_k coefficients, one can thus inspect if, and in which direction, the control units change their behavior before actually becoming treated.[35] Figure A3 shows the results. The important thing to notice in the figures is that, if anything, control units seem to display an anticipatory behavior of the first kind, meaning that they start behaving in a way that is similar to the one of the treatment group, before being treated. In fact, the coefficients of Figure A3 are of the same sign, though of a much smaller magnitude, as those of Figure 2, which displays the coefficients γ_k, which represent the event studies for the treated group. In this sense, my estimates could be considered conservative lower-bound estimates of the true effects in the absence of anticipatory effects. These results are further confirmed by the presence of positive spillover effects on yet-to-be-treated nearby municipalities, which I present and discuss in Section 6.3.

Finally, a source of concern might be the presence of other time-varying shocks that

[33] Kahn-Lang and Lang (2020), Roth (2020), and Rambachan and Roth (2020) have cautioned against the use of pre-event trends testing as these tests are frequently underpowered and can lead to type-2 errors. I have conducted the sensitivity tests as in Rambachan and Roth (2020), and my results are robust to allow for large degrees of non-linearity in the violation of parallel trends.

[34] In particular, they are estimated thanks to the fact that the same municipality appears multiple times in the dataset. As an alternative, I also estimate a specification that includes fully saturated cohort-specific municipality fixed effects and obtain nearly identical results, as discussed later in Section 5.3.

[35] In my stacked-by-event design, when constructing the dataset, for each event, control units are assigned a "placebo" shock at event time = 0 for that specific event, which ensures that their potentially anticipatory responses are accounted for explicitly in the estimation of the dynamic treatment effect.

occur during the same period, in a way that is correlated with the staggered introduction of the reform across municipalities. The most significant change that occurred during this period was the extension of the rules of the Domestic Stability Pact to municipalities with a population below 5000 inhabitants in 2014. This change is accounted for by the inclusion of population-size-by-year fixed effects. In Section 6, I also investigate the heterogeneity in treatment effects by municipality size and show that the effects are present also when restricting the sample only to municipalities above 5,000, which have been subject to the same rules since 2001. Relatedly, one might worry about the fact that the change in auditors' appointment was adopted as part of a large emergency reform, which might have affected local governments' finances in other ways beyond the change in auditor's appointment.[36] The most important one was the temporary inclusion of owner-occupied dwellings among houses subject to the property tax in 2012. However, it is important to remember that these changes affected contemporaneously all municipalities, while the change in auditors' appointments entered into place in a staggered way, depending on the timing of the expiration of the previous municipal auditor. Therefore, the effect of any other change related to the reform is captured by the presence of calendar time effects in my regressions.

5 Results

5.1 The Impact of Random Assignment on the Selection and Allocation of Auditors

I begin by examining how the reform affected the selection and allocation of auditors across municipalities. This analysis serves two purposes: confirming whether the random assignment was effectively implemented and identifying key changes brought by the reform with potential implications for monitoring effectiveness.

Table 2 compares the characteristics of the auditors and auditor-municipality match under discretionary appointments (pre-reform) as opposed to the ones resulting from random assignment. Columns (1) and (2) show average characteristics in the pre and post-reform periods, respectively. Column (3) shows the difference in means between the two, and column (4) the standardized differences, to make it easier to compare the magnitude of the changes across different variables.

The first notable result is the significant increase in the number of auditors following the

[36]The reform also introduces a) a tightening of fiscal rules and cuts of national government transfers for the year 2012, b) progressive tax rates for the municipal additional income tax, c) mandatory collective management of public goods and public services for municipalities up to 1000 inhabitants, d) a reduction in the number of members of the municipal council e) collection of property tax on owner-occupied units.

reform, with 22,000 auditors joining the list, 60% of whom are new entrants. The large entry response implies that auditors in the post-reform period are, on average, less experienced in municipal auditing compared to their pre-reform counterparts. On the other hand, these new auditors share similar characteristics with incumbents in other important areas such as age and professional experience in accounting. This similarity suggests that the pre-reform barriers likely restricted the entry of competent auditors. Despite their lack of municipal specific experience, the new entrants could be of higher unobserved quality along other dimensions that are relevant for monitoring efficacy, such as integrity. Consequently, the reform may have broadened the market to include a wider range of auditors, potentially improving the quality of oversight and reducing leniency or bias in municipal auditing practices. [37]

I then look at the average probability that auditors are re-appointed for a second term. In the pre-reform system, this happened in 45% of the cases, while in the post-reform system, it becomes an extremely rare occurrence (1%), as it can only happen as a result of the same auditor being drafted again for the same municipality. The large change in re-appointment rates highlights how the reform might have also changed incumbent auditors' incentives, a mechanism that I explore in Section 6.2.

Finally, looking at the characteristics of the auditor-municipality match, the introduction of random assignment makes it significantly less likely for municipalities to have a local resident as a municipal auditor. Under the pre-reform discretionary appointment system, 25% of municipalities appointed a hometown auditor. This number goes down to 1% by virtue of the random assignment. One consequence of this change is the large increase in the average distance between the auditors' municipality of residence and the one where he operates as municipal auditor from 17 to over 60 minutes. Finally, I look at the probability that the auditor and the mayor share some relevant characteristics. In the pre-reform period, 18% of the auditors were born in the same place as the mayor, and 7% shared the same surname with a local politician. These shares also go down to 2% and 1% after the reform.

Together, these findings provide evidence that the randomization was effectively implemented. To further confirm that this is the case, in Table A5 I show that, for each of the observable characteristics discussed above, there are no significant differences between the drafted auditors and the pool of eligible auditors. Furthermore, in Table A6 and Figure A4 I move beyond a simple comparison in average characteristics in the pre vs. post reform period and report the static and dynamic difference-in-differences estimates using the main

[37] A distinct and interesting question is whether the reform caused some incumbent auditors to exit by choosing not to sign up for the new list. While data limitations prevent me to inspect directly this response, I can rule out that this type of differential sorting is the sole driver of the observed effects. In fact, significant improvements in fiscal policy are still detected when focusing on municipalities with over 5,000 inhabitants, which, due to minimal experience requirements prescribed by law, could only receive randomly-assigned auditors selected from incumbent auditors. These results are discussed further in Section 6.2.

empirical specifications in Equation 5 and 2, where the characteristics of the auditors and auditor-municipality matches are on the left-hand side.[38]. The event studies displayed in Figure A4 and confirm a lack of pre-trends and a sudden and persistent change in auditors' characteristics.

Overall, the results suggest that random assignment significantly changed the selection and allocation of auditors across municipalities in a way that likely severed previously existing ties between auditors and mayors. Previous research has shown that these types of connections are likely to negatively affect monitoring performance (Chu et al. 2020). Thus, it would be plausible to expect that the reform improved monitoring effectiveness, which could translate into improvements in municipal fiscal performance. On the other hand, if connected auditors appointed under the status-quo discretionary system had higher municipality-specific knowledge and were thus able to provide more tailored advice on local fiscal practices, we could observe a null or even a worsening of fiscal performance. In the following section, I shed light on this question by investigating the causal effect of the reform on municipal fiscal performance.

5.2 The Impact of Random Assignment on Municipal Fiscal Performance

Table 3 displays the main regression results for the effect of the arrival of a randomly-assigned auditors on indicators of municipal fiscal sustainability. Odd columns present results from Equation 1 without controls, while even columns include controls. Results indeed provide evidence that the random assignment of auditors contributes significantly to improving the fiscal stance of municipal finances. In columns (1) and (2), I investigate the effect on $NetSurplus$. Treated municipalities increase their net surplus by €16 per capita, an 8% rise relative to the pre-treatment mean, upon the arrival of a randomly-assigned auditor. In columns (3) and (4), I look at $DebtRepayments$. I find a significant and positive effect of comparable magnitude, with treated municipalities increasing their debt repayments by 8% after the arrival of draft-appointed auditors. In the last two columns, I show results for $Off-BalanceSheetDebtRepayments$. Here, I do not find any significant effect. As I will explain in later sections, this is likely due to the presence of strongly heterogeneous dynamics across different types of municipalities. [39]

[38] All the outcome variables are standardized, to make magnitudes more comparable across columns/figures.

[39] To better understand the patterns of adjustment, in Figure A2, I investigate changes in the distribution of outcomes, by comparing histograms of the variables by treatment status. Given that municipalities belonging to different cohorts of treatment display similar pre-treatment levels of outcomes, these graphs are not confounded by potential level differences across early vs. later treated municipalities. The top panel shows a clear rightward shift in the $NetSurplus$ distribution for values below 0 under the random assignment system, indicating improved fiscal stance for deficit-running municipalities. The distributions overlap for values above

Figure 2 displays coefficients from the event-study specification in Equation 2, allowing inspection of differential pre-trends and treatment effect dynamics over the auditor's three-year term. The graphs show no significant differential pre-trends across all three outcomes. For both $NetSurplus$ and $DebtRepayments$, treatment effects appear in the first year of the auditor's arrival and remain consistent across the three years. This suggests a sudden and persistent shift in municipal budgetary practices in response to the treatment. The figure also confirms no significant effect on $OBSDebtRepayments$.

Overall, these results underscore the effectiveness of randomly assigning auditors in strengthening oversight and improving municipal fiscal health, thus emphasizing the importance of organizational design changes in enhancing monitoring in multi-layered organizations.

5.3 Robustness and Persistence of the Effects

5.3.1 Alternative Estimators and Specifications

In this section, I discuss a series of empirical tests that I perform to corroborate the robustness of my analyses. First, I apply three alternative estimators, the standard TWFE estimator, as well as two alternative robust estimators recently proposed by de Chaisemartin and D'Haultfœuille (2020) and Callaway and Sant'Anna (2021) to solve issues with treatment effect heterogeneity in TWFE estimators. Results are displayed in Table A7 and A8 and Figure A9. Reassuringly, estimates are all very similar to the ones in Table 3. Further details on these alternative estimators are provided in Appendix A3. I then perform a series of alternative specifications to further corroborate the robustness of my results. In Table A9, I re-estimate the main model but include non-parametric trends in the log of population, municipal income and all the other municipal characteristics included in Table A4. This ensures that results are not driven by the differential effect over time of different pre-treatment municipalities' characteristics.

In Table A10, I estimate a version of Equation 1 fully saturated with cohort-specific municipality fixed effects and obtain nearly identical estimates. In Table A11, I include a set of 135 region-times-year fixed effects, which ensure that the results are not driven by differences in region-specific trends over time. While lower in magnitude, results remain highly significant and qualitatively similar. In Table A12 I show that results are robust to account for the presence of spatial correlation in the data, by re-estimating the model using Conley (1999; 2010) standard errors.

0, suggesting auditors do not cause overshoots in municipalities already running surpluses. The second panel shows a rightward shift in $DebtRepayments$, consistent with fiscal rules requiring timely debt repayment. No clear changes are observed in the $OBSDebtRepayments$ distribution, aligning with the lack of treatment effects.

Furthermore, I verify whether effects are potentially driven by specific groups of municipalities. First, I show that I am able to detect significant effects for both small and large municipalities (see Table A13). This specification compares treated and control municipalities of similar sizes (say, all those below 5,000 inhabitants), ensuring that results are not driven by potentially differential selection in treatment timing by population size. It also allows me to rule out the possibility that the effects are entirely driven by the minimum experience requirements introduced by the reform, which only apply to municipalities above 5,000 inhabitants.Futhermore, I perform a heterogeneity split by geography and show that significant effects are present across the whole country (see Table A14).

5.3.2 Alternative Design: Fiscal Performance with Discretionary Appointments

To further corroborate my results, I investigate whether, in the pre-reform period, granting discretion in auditors' appointments indeed induced adverse effects in terms of fiscal sustainability. To do so, I exploit the fact that, even before the reform, mayors did not always control auditor appointments due to the differing lengths of their mandates—mayors serve 5-year terms, while auditors serve 3-year terms, renewable once. This creates exogenous variation in mayoral control over appointments (see the bottom panel of Figure A12 and A11 for an example), which allows us to analyze the use or abuse of discretion. Intuitively, if mayors were appointing friendly and lenient auditors who would allow them leeway vis-a-vis fiscal rules and spending, we should observe a worsening of fiscal sustainability during years in which mayors had control of the appointment. On the other hand, if mayors were using discretion effectively to appoint efficient and experienced auditors, then we should not observe any significant difference. I test this hypothesis performing the following analysis:

$$Y_{mt} = \alpha_m + \delta_t + \beta Control_{mt} + X'_{mt}\zeta + \epsilon_{mt} \tag{3}$$

The probability of controlling appointments increases throughout the electoral term, which might confound the results with political budget cycle effects. Therefore, I include a dummy variable for the two pre-election years among the controls. This ensures β is estimated by comparing mayors at similar points in the election cycle who either did or did not appoint the current auditor due to exogenous reasons. Results are shown in Table A15. Odd columns include all second-term mayors and a control for second-term mayors interacted with the pre-election dummy. Even columns focus on first-term mayors. The findings indicate that when mayors control appointments, municipalities run higher deficits and decrease debt repayments. This suggests that, indeed, mayors who have discretionary appointment power act in a less fiscally responsible way. It is interesting to compare the magnitude of the effect

of the arrival of an independent, draft-appointed auditor in Table 3 with the effects obtained in Table A15, with the caveat of the difference in the estimating samples and equations. The positive effect on the *NetSurplus* of the independent auditor's appointment is twice as large as the negative effect found for the discretionary appointment effect, while the effect on *DebtRepayments* is over 4 times as large. This suggests that the draft appointment mechanism is a more effective disciplining tool than the simple term misalignment between mayor and auditor in the pre-reform period.

Overall, these results go in the same direction as my previous findings, confirming the importance played by changes in the control of auditors' appointments for monitoring effectiveness and municipal fiscal performance. From a policy perspective, these results also provide evidence that the introduction of a gap between the political and the audit cycle could be an alternative system to strengthen oversight and improve auditor independence.

5.3.3 Long-Run Persistence of Effects

A natural question to ask is whether the observed effects would persist in the longer run, or if instead they just represent a temporary improvement. On the one hand, we would expect that, by changing the selection and allocation of auditors, the removal of discretion of appointment should have a persistent impact on public finances. On the other hand, mayors might learn over time how to influence, and potentially capture, even randomly-assigned auditors. Policy events suggest the reform's long-lasting significance, as the randomized assignment system remained a political concern for mayors seven years after its implementation. In December 2019, during fiscal negotiations with the national government, mayors of large municipalities secured a partial reversal of the reform by reintroducing patronage appointments for the head of the auditors' committee. However, it is essential to empirically confirm the reform's persistent impact on budgetary outcomes.

While the features of my design do not naturally lend themselves to analyzing long-run outcomes, I provide suggestive evidence of the persistence of effects for 7 years after treatment, that is until the second round of randomly-assigned auditors is completed. To do so, I extend the sample until 2019 and estimate the following dynamic specification:

$$Y_{mt} = \alpha_m + \delta_t + \beta \sum_{k=-7}^{k=6} \gamma_k * D^k + X'_{mt}\zeta + \epsilon_{mt} \tag{4}$$

Given that all units are treated by 2015, the estimation of longer-run effects exploits already-treated units as controls, and rest on the following two additional assumptions, as discussed in Borusyak et al. (2021b): a) lack of negative weights, and b) homogeneity of short run effects across cohorts. As discussed in the previous section, I provide evidence in

support of the first assumption in Table A7. As for the second assumption, in Table A16 I re-estimate the same model but allow for cohort-specific treatment effects.[40] I perform an F-test and fail to reject the hypotheses that the treatment effects are the same across cohorts for all three outcomes of interest. Given these premises, the estimates are unlikely to be biased and can be interpreted as the long run effects of the reform. Overall, results reported in Figure A7 suggest that the treatment effects of the reform are remarkably stable and persistent over the long run, consistent with a long-lasting shift in budgetary practices induced by the improvement of monitoring effectiveness.[41]

5.4 Margins of Adjustment: Revenues vs. Spending

The results above show that, upon the arrival of randomly-assigned auditors, treated municipalities run higher net surpluses and reduce their outstanding debts, in accordance with the national government's objectives. How did municipalities manage to improve their fiscal stance? Intuitively, they could either increase their revenues (e.g. by raising taxes) or cut expenditures (e.g. cut investments and/or current expenditures), or a combination of both.

In Table 4, I investigate the margins of adjustments of municipal governments by looking at their spending and revenue choices. In columns (1) and (2) I look separately at total $CurrentExpenditures$ and $CapitalExpenditures$. The results reveal that fiscal sustainability is not achieved through expenditure cuts. While current expenditures remain essentially unaffected, capital expenditures increase by over 4%, although this effect is only marginally statistically significant. The absence of an effect on current expenditures is plausibly explained by the fact that the vast majority of current expenditures are due to personnel costs and are therefore rigid, as local governments cannot fire local bureaucrats, nor can they adjust their wages. On the other hand, the observed increase in capital investment is a potentially desirable outcome. A frequently debated side-effect of fiscal rules' constraints is that municipalities responded by cutting investments (Chiades and Mengotto 2015). A potential explanation for the observed increase is that it results from political bargaining between the auditor and the mayor. the observed increase in capital expenditures may result from political bargaining between the auditor and the mayor. In exchange for significant

[40] In this exercise, I and restrict attention to the short run effect in $t = 0$, so as to maximize comparability across cohorts, as the cohort treated in 2014 is only observed as treated for one year, before municipalities in the control group, that is those treated in 2015, receive a rndomly-assigned auditor.

[41] An alternative way of estimating long run effects could be to use municipalities in special regions, which are exempted from the initial application of the reform. In a recent working paper, Barone et al. (2020) similarly use a DiD approach to evaluate the reform's effects on a different set of outcomes, using Special Regions as controls. However, these regions are also granted special autonomy in setting their own fiscal rules. Furthermore, they were allowed to choose whether or not to opt in into the reform, which lead to endogenous sorting into treatment which could be correlated with their trends of outcomes over time. For these reasons, they do not represent a viable control group and are excluded from the analysis.

improvements in fiscal sustainability, primarily through revenue enhancements, mayors may be permitted to increase investment spending. This spending can help justify the revenue consolidation efforts to local constituents. Notably, significant increases are observed in investment categories related to general administration, education, culture, and social welfare - areas directly observable and beneficial to local residents and thus likely to be politically relevant (see Table A17).

In the subsequent columns, I investigate adjustments on the revenue side. Municipalities' main sources of current revenues are the property tax and the local income tax surcharge. Column (3) shows a significant increase in $PropertyTax$ revenues, rising by over 20%, equivalent to an average increase of over 30 euros per capita. This analysis focuses on non-owner-occupied properties, as most owner-occupied dwellings were exempt from the property tax during most of the sample period.[42] On the other hand, I fail to detect any significant change in the local income tax (column (4)). Finally, in columns (5) and (6) I look at $CapitalRevenues$ and $NewDebt$. $CapitalRevenues$ increase by 4%, while $NewDebt$ increases by over 22 %.[43] Given that the observed increase in total revenues to be used for investment financing ($NewDebt$ and $CapitalRevenues$) is higher than the increase in capital expenditures, it is likely that municipalities are using part of the $NewDebt$ to re-finance pre-existing debt, thus partially explaining the observed increase in $DebtRepayments$. As explained in Section 2, according to Italian law, municipalities can incur new debt to finance new investments or refinance existing debt if it results in substantial debt service savings. New Debt is more advantageous than old debt, as most pre-existing debt was contracted in the 1990s and 2000s when interest rates were higher. Post-reform (2012-2015), interest rates were near zero due to ECB policies.

Figure 3 displays results from event-study specifications for the outcomes in Table 4. While property tax revenues increase immediately with the auditor's arrival, the effect on new debt only appears initially and then dissipates in the last period. On the other hand, both capital revenues and expenditures show similar delayed response patterns from period two onwards, indicating expanded investment activities.[44]

Overall, results show that, upon the arrival of a randomly-assigned auditor, municipalities improve their fiscal stance through revenue-based adjustment, specifically by increasing

[42]Due to data availability, the analysis for this variable is restricted to 2008-2014.

[43]This result should be taken with some caution as $NewDebt$ is the only variable for which I detect a significant negative pre-trend for the relative event coefficients -5 and -6, leading to a slight overestimate in the magnitude of the effect.

[44]As a robustness, in FigureA8, I also investigate the response of transfers from higher levels of government. Previous research on Brazil (Brollo and Troiano 2016, Brollo 2008) shows that audited municipalities who are found liable for misbehavior receive fewer transfers from the federal government. I find suggestive evidence of a very short run mild decrease of transfers from the state, that disappear from the first period onwards, while transfers from other levels of government (Regions or the EU) are unaffected.

revenues from the local property tax. This result is in line with previous evidence from Italy, showing that municipalities mainly respond to cuts in tax transfers from the central government through revenue-based adjustments (Grembi et al. 2016, Marattin et al. 2019).

How did municipalities manage to expand tax revenues? The observed expansion could be driven by either an increase in the property tax rate or an expansion of the tax base, due to an improvement in the enforcement of tax collection and reduction in tax evasion. I provide evidence of the drivers behind the expansion in tax revenue in Table 5. Column (1) replicates the results from Table 4, where I look at the total amount of property taxes assessed in a given year. In column (4), I look at the total amount of property tax revenues that is effectively collected by the end of the year. Here, the results look almost identical to column (1), suggesting that the effect is coming from an expansion of the reported tax base, rather than an improvement in collection ability. In Column (7), I test explicitly for whether there is any change in the speed of tax collection, that is the ratio between collected and assessed property tax revenue (Gagliarducci and Nannicini 2013), and I fail to detect any significant average improvement. Finally, column (10) shows no significant change in the property tax rate, further suggesting the revenue increase is due to a broader tax base. In line with this explanation, in columns (2) and (3), and similarly in columns (4) and (6), I show that the effects on both the assessed and collected revenues are significantly stronger in municipalities that had a higher share of non-declared units[45], and where there was a higher share of so-called "ghost buildings", as documented in Casaburi and Troiano (2016).[46] Overall, this pattern supports the hypothesis that mayors are expanding tax revenues by combating past tax evaders. In line with this theory, in columns (8) and (9), I find a negative heterogeneous effect on the speed of tax collection, suggesting that, despite the effort to fight evasion by taxing previously un-assessed properties, it is slightly more difficult to effectively collect taxes from former tax evaders. Overall, the evidence presented so far clearly indicates that the observed expansion in tax revenue is primarily due to significant improvements in tax enforcement and collection from potential evaders.

To put this finding in context, it is important to note that mayors have complete discretion over tax collection procedures, and auditors are tasked with reviewing the efficacy and efficiency of these procedures. For instance, the responsibility of verifying if a property qualifies for the owner-occupied property tax exemption lies with the municipal police.

[45] This metric is based on the matching between tax declarations and cadastral data done by the National Tax Agency.

[46] In 2007, the national government started a large anti-evasion program to identify buildings that were previously not included in the land registry and thus hidden from tax authorities. Thanks to the use of aerial photography, the program detected over 2 million parcels with at least one ghost building. As documented in Casaburi and Troiano (2016), "the program led to a substantial wave of building registration and a sizable increase in total tax revenues (almost half a billion €), with a large share of these extra revenues coming from local property taxes.[...] The intensity of the additional tax enforcement varied significantly across towns".

According to the Italian Revenue Agency, the false declaration of residence in a different property to benefit from the owner-occupied exemption is one of the most common forms of property tax evasion and is estimated to affect almost 150,000 properties. Consistent with an increased effort of mayors in increasing tax collection capacity, in the last three columns of the Table, I show that treated municipalities increase by 7% their investment spending for administrative functions, which include investments targeted towards expanding tax collection functions. Again, the treatment effects are sizably larger in municipalities with a higher share of non-declared units. Overall, the evidence suggests that, when facing more independent auditors that enforced tighter budget constraints, mayors have stronger incentives to fight property tax delinquents — a behavior they might have previously overlooked due to political or economic considerations.

5.5 The Broader Consequences of the Reform: Public Good Provision, the Real Economy and the Internal Organization of Municipalities

The results presented so far provide robust empirical evidence that the arrival of draft-assigned auditors improved monitoring effectiveness and spurred more fiscal stability, with municipalities running higher surpluses and achieving the adjustment through a sizable improvement in tax revenue and collection capacity. This raises the question of whether such a reform also affects the provision of public goods and the broader economy. Given the previous evidence showing no reduction in spending, it's reasonable to hypothesize that public goods provision should not suffer adverse effects. However, the notable increase in tax collection could potentially have negative repercussions on the broader economy. To shed light on the potential broader consequences of the reform, in Table 6 I look at a wide range of outcomes related to public good provision and economic activity.

In columns 1 to 5, I look at a series of measures coming from the local public goods' performance assessment survey that municipalities have to fill in alongside their balance sheet and report to the Ministry of the Interior. In columns 1 and 2, I start by looking at two metrics of basic local public good provisions, the percentage of housing units with access to waste collection services and the extent of public lighting coverage across local streets. I then move to investigate the impact on one of the most important public services provided by local governments, public nurseries for kids between ages 0 and 3. Here, I look at three different metrics. First, I investigate whether there are changes in the extensive margin, i.e. in the presence of nurseries in the municipality. Then, I look at the acceptance rate and the teacher-pupil ratio.[47] In column 6, I look at the number of sports facilities per 10,000

[47]These two outcomes have a lower number of observations as they are, respectively, conditional on having

inhabitants. In the last three columns, I turn to investigate the consequences on the real economy. Here, I exploit data coming from the Ministry of the Interior on the universe of tax declarations. In column 7, I look at total income per capita, a close proxy of GDP per capita. In column 8, I look at the total number of taxpayers. Finally, in column 9, I look at the log of the total population, which would capture any migration responses. Across all these metrics, the findings indicate negligible and statistically insignificant impacts.

Another natural question that arises is how the arrival of a randomly-assigned auditor affected the internal functioning of the local bureaucracy. Removing discretion of appointment could come at the cost of poorer coordination, lower information transmission, and in turn disrupt the functioning of the bureaucracy. While these are hard things to measure, I exploit the availability of rich data on the personnel composition and task allocation coming from the annual census of public personnel from the Ministry of Economics and Finance. To the best of my knowledge, this is the first paper to use this data for economic research. Results are reported in Table 7. In the first column of Panel A, I look at the total number of bureaucrats per 10,000 inhabitants, while in columns 2 and 3, I look at the hiring and separation rates, to assess whether the reform induces changes in the size of the local buraucracy. Across all the three outcomes, I find small and insignificant effects. In column 4 and 5, I look at the total number of hours worked per capita, and the share of workers with a short-term contract. These two variables allow me to capture whether there were changes, beyond the extensive margin, on the overall amount of worktime of local bureaucrats. Again, I fail to detect any significant changes. Finally, in columns 6 to 8, I move to investigate proxies of bureaucratic discontent: the number of absences, sick days, and strike days per capita. While I do not detect any significant change in absenteeism overall, I find a significant 10 % increase in the number of sick days and an even larger increase in the number of strike days. These results seem suggestive of a negative reaction of local bureaucrats to the arrival of external auditors and the strengthening in monitoring. Several reasons could explain this result. First, it could be that external auditors are worse at communicating and understanding the functioning of the local bureaucracy. Second, it could be that they require greater effort on the job and disrupt the regular oganization and allocation of tasks. Third, external auditors might hold local bureaucrats responsible for pre-existing poor fiscal practices.

To better understand the potential changes in the internal bureaucratic organization, in Panel B of Table 7 I look at the allocation of hours across different tasks. I examine the allocation of time across eight categories, that reflect the main functions performed by local bureaucrats: public records and clerk's office (column 1), human resources (column

at least one application and conditional on having a nursery in town. Not all municipalities have a nursery within their boundaries, as given the average small size of municipalities, this is frequently a service that is shared across towns.

2), political support (column 3), housing and local public works (column 4), public good and service provision (column 5), economic and financial management and procurement (column 6), and taxation (column 7). Interestingly, I find a reallocation of time away from political support activities and towards local housing and public works. The evidence in Panel A of increased sick and strike days and the reduction in time allocation towards political support activities in Panel B might potentially adversely affect the functioning of the municipal governing bodies. To investigate this possibility, in column 8, I look at the per-capita number of laws and regulations adopted by the municipal council, which is a proxy of overall local government productivity. Reassuringly, I do not find any evidence of a negative effect. If anything, results point towards a positive, albeit insignificant, effect of the reform on municipal council productivity.

Overall, these findings suggest that while the reform improved fiscal stability and revenue collection, it did not negatively impact public good provision or the broader economy. Additionally, although there was some bureaucratic discontent, the internal functioning of local governments remained largely stable, with no adverse effects on government effectiveness.

6 Mechanisms

The findings presented up to this point show a significant improvement in municipal financial performance after the removal of local discretion in auditors' appointments and the introduction of random assignment. In this section, I perform a series of additional analyses and heterogeneity tests that allow me to inspect the mechanisms through which the reform most plausibly operates.[48]

6.1 Local Misalignment and Risk of Auditors' Capture

From an organizational point of view, leaving discretion of appointment locally can be considered a classic example of a delegation problem, which brings up a canonical capture-versus-information trade-off, where mayors might use discretion for good—to select better auditors—or abuse of it—to hire more corruptible ones. A useful insight from optimal delegation models (Alonso and Matouschek 2008, Mookherjee 2012) is that the benefits of discretion increase with the agents' alignment with the organizational objectives. Thus, we might expect random assignment to be more valuable in places where mayors are less aligned

[48] All heterogeneity analyses are performed using the same specification of the main results, the one of Equation 1, augmented with a triple interaction term for the relevant heterogeneity margin, as well as interactions between the *Treated* and *Post* indicators with the relevant heterogeneity covariate. This ensures that heterogeneous effects are not driven by differential trends along the specific heterogeneity margin. As a robustness check, I also repeat all the analyses using a standard two-way fixed effects DID estimator in Table A18 and obtain very similar results.

with the interests of the central government and thus more likely to abuse discretion in the pre-reform period. I adopt several proxies to identify places where lax or corrupt monitoring was more likely to occur before the reform.

I start by investigating how the treatment effects vary with pre-reform financial health. Intuitively, lax or corrupt monitoring might have allowed for more irresponsible financial behavior. In Panel A of Table 8, I repeat the analysis of Table 3 but include an interaction term with an indicator that is equal to one for municipalities for which the dependent variable was below the median value in 2011.[49] The results indeed show that the treatment effects are significantly larger in municipalities that were less fiscally sustainable before the reform. The treatment effect for the *NetSurplus* is twice as large for municipalities that had a *NetSurplus* below the median value in 2011. An even more striking difference emerges when looking at the effects on *DebtRepayments*, where the treatment effect is four times as large. Finally, the last column shows that municipalities without prior off-balance sheet debt replayments show a substantial and significant positive treatment effect, contrasting with a significant negative effect in those with previous debt repayments. The sign divergence explains the overall zero effect in Table 3 and suggests the presence of two different underlying phenomena. *OBSDebtRepayments* represent a self-declaration of an illicit spending act. Thus, auditors can act on two different margins. On the one hand, they can pressure municipalities with existing undeclared debts to report them, thus increasing reported *OBSDebtRepayments*. On the other hand, they can discourage municipalities from committing illicit spending, thus reducing the occurrence of *OBSDebtRepayments*. While I cannot effectively separate the two mechanisms without having information about the unobserved amounts of illicit spending, the divergent signs of the effects suggest the presence of both phenomena across different contexts. In "honest" municipalities—those that correctly report illicit spending in the pre-reform period—the arrival of a randomly assigned auditor reduces the occurrence of the illicit spending phenomena, thus leading to a decrease in *OBSDebtRepayments*. Whereas in "dishonest" municipalities, the reform induces the reporting of illicit spending, thus leading to an increase in *OBSDebtRepayments*.

In Panels B and C, I look at two other proxies of local misalignment and risk of auditors' capture in the pre-reform period. I first use an indicator for whether the last auditor appointed before the reform was born or resided in the same municipality. Intuitively, a "local" auditor is more likely to have connections with the mayors and/or be more sensitive to local interests and thus be willing to collude to favor local spending vis-a-vis fiscal sustainability. Furthermore, using restricted-access data from the *Sistema D'Indagine Interforze* (SDI)[50]

[49]For off-balance-sheet debt repayments (*OBSDebtRepayments*), the median value in 2011, and in all pre-treatment years, is always 0. Therefore, for this variable, I construct an indicator equal to 1 if the municipality has never had a positive amount of *OBSDebtRepayments* in the pre-period and 0 otherwise.

[50]The SDI is a centralized investigation archive that contains reports of all individuals investigated by any

I construct an indicator variable flagging whether, in any given municipality, there was at least one investigation for corruption-related crimes (bribery, corruption, malfeasance, and embezzlement) before treatment. Figure A10 shows a map of municipalities that have the *Corrupt* dummy equal to 1. While the inspection and detection of corrupt practices is not the auditors' main task, corruption cases usually involve illegal transactions between private parties and members of the local bureaucracy and are thus a proxy of an environment where illegal practices occur, that might or might not involve directly local politicians. Furthermore, corruption-related crimes are frequently linked to local public spending, which are factors that should be subject to auditors' monitoring. Thus, a context in which corruption-related crimes occur is more likely to be one in which the mayor might have the incentive to hire a lenient local auditor who would not uncover illicit practices. Table 8, Panels B anc C display the results. Treatment effects are generally significantly stronger for municipalities that had a local auditor or corruption investigations. For example, for *DebtRepayments*, the treatment effect is four times larger in corrupt municipalities (0.06 versus 0.24). Furthermore, I again detect diverging patterns when looking at *OBSDebtRepayments* similar to the one detected in panel A, suggesting a reduction in the actual size of off-balance sheet debts in non-corrupt places, and the emergence of a large amount of previously hidden OBS debts in corrupt places that are brought to the table by the arrival of randomly assigned auditors.

In Panel D, I compare the effect of the reform for municipalities in which the last appointed auditor was in his first versus second term. Under the pre-reform system, auditors could be reappointed for a second term by mayors. The possibility of re-appointment likely gave strong incentives to the auditors to please the mayor who appointed them during their first term, and it also provided the mayor with a clear retaliation mechanism.[51] Intuitively, auditors in their second term should have had fewer incentives to be lenient toward the mayor as, for them, career incentives were already muted. This would mean that we should expect to observe stronger effects for municipalities with first-term auditors. Indeed, I find suggestive evidence consistent with this hypothesis, as the effects are stronger in municipalities with a first-term auditor as opposed to a second-term one.

Finally, I investigate heterogeneity by the strength of mayors' political incentives. Elec-

of the Italian police forces, which include state police (*Polizia di Stato*), finance police (*Guardia di Finanza*), military police (*Carabinieri*), and environmental police (*Guardia Forestale*). The SDI data have been previously used in research by Decarolis et al. (2019). The data contains information, for each municipality, on the total number of investigations for all types of crimes in the years 2004-2013. See Decarolis et al. (2019) for further details on the data.

[51] In addition, auditors had a more general incentive to build a reputation for "leniency" under the old patronage system since they could be appointed by other mayors. Thus, under the old system, auditors experienced a clear tension: career incentives pushed them to be lax to please the mayor while being at risk of getting sanctioned by the National Court of Auditors, shall they be found liable for lax or corrupt monitoring. In the post-reform system, auditors have incentives to monitor effectively in order to please the Court as well as to preserve their reputation in the private market.

toral incentives may have an ambiguous effect on local behavior. On the one hand, we might expect that mayors subject to strong re-election pressures would have stronger incentives to perform well, restraining from inappropriate or wasteful financial practices. On the other hand, local electoral incentives may push mayors to be less financially responsible, particularly if improving their financial performance comes at the cost of cutting local spending or raising taxes for local residents. Furthermore, previous evidence has shown that mayors facing lower re-election pressures tend to behave worse (Ferraz and Finan 2011). To shed light on the role played by mayors' political incentives, I exploit the existence of a two-term limit for Italian mayors to see whether mayors who can be re-elected respond differently than mayors who are serving their second and last term. Results are displayed in Panel E of Table 8. Here I find that, for both $NetSurplus$ and $DebtRepayments$, treatment effects are larger for mayors that could not run for re-election at the moment of treatment, suggesting that audits played a disciplining effect that substitutes for the lack of electoral accountability incentives.

6.2 Selection and Allocation of Auditors

In this section, I aim at investigating how much the characteristics of the new auditor-municipality match explain the observed effects. The switch from discretionary appointment to random assignment is potentially changing both the selection of auditors—who gets hired and who decides to participate in the auditors' pool – as well as their allocation to municipalities.

Selection: By removing barriers to entry based on patronage networks, the new appointment system changes the pool of potential auditors, thus affecting the selection margin. As is common in markets for experts (Ronen 2010, White 2010), despite the high degree of potential competition in the market, the pre-reform discretionary appointment system was characterized by a relatively limited number of auditors who performed audit tasks for multiple municipalities.Removing discretionary appointments allowed new auditors to enter the market. If these new entrants are of higher quality, less corruptible, or more likely to enforce national government objectives, then the selection margin might drive the observed treatment effects. Indeed, as discussed in Section 5.1, the reform induced a change in the composition of the auditor pool: almost 14,000 new auditors joined the list, representing around 60% of the pool of potential auditors. While these new entrants have no municipal-specific experience, they look otherwise very similar to the existing pool of incumbents on other observable dimensions (see Table 2). Most importantly, they are equally experienced in terms of general private-sector accounting, so the change in the selection method does not seem to come at the cost of compromising quality. However, it could well be that new

entrants are different from incumbent auditors along other unobservable dimensions, for example, honesty.

To assess the role of selection in explaining my results, I gauge the exogeneity of random assignment in the post-reform period to compare the outcomes of municipalities that, by chance, are assigned a new entrant to those that are assigned an already experienced one. In this comparison, I am limiting attention to municipalities with less than 5,000 inhabitants as these are the only ones that could be assigned new entrants. Table 9, Panel A displays the results and shows that, at least for the subsample of smaller municipalities, the effects are essentially driven by the assignment of new entrants. On the other hand, results presented in Table A13 show the presence of sizable treatment effects also in municipalities above 5,000 inhabitants that, due to the presence of minimum experience requirements, could only be assigned incumbent auditors, thus rejecting the possibility that selection is the only driver of the effects. Overall, the results suggest that selection indeed plays an important role but cannot explain the entirety of the effects.

Allocation: By removing control of appointment from the mayor, the reform affects the allocation of auditors to municipalities, moving from endogenous matches to random ones. This change has two potential effects. On the one hand, random matching reduces the risk of auditors' capture. On the other hand, it might come at the cost of sacrificing the potential value of discretionary appointments and inducing misallocation. As shown in Table 2, one of the most critical observable changes induced by the reform is a substantial increase in travel distance between auditors' residence and the municipality of assignment, from around 15 minutes to over 60. The role of distance for monitoring effectiveness is theoretically ambiguous. On the one hand, a local auditor is more likely to have connections with the mayors and to be more sensitive to local interests, and thus be more willing to favor local spending vis-a-vis fiscal sustainability. On the other hand, closer auditors might have better local knowledge and face lower costs in acquiring information, which might make them more effective monitors. I perform two sets of analyses to shed light on the role of changes in auditors' allocation. First, the results presented in Table A13 already suggest that, indeed, changes in auditors' allocation matter. The presence of significant treatment effects in municipalities above 5,000 inhabitants, that can only receive an already-experienced auditor, suggests that changing the allocation of incumbents across municipalities significantly improves outcomes. This suggests that, indeed, breaking potential local ties matters. I then exploit the exogenous nature of the post-reform municipality-auditor match arising as a result of random assignment to more directly explore the role of distance. The results in Panel B of Table 9 show that the effect of the reform is substantially larger in municipalities that, by chance, receive an auditor who lives at least one hour away, the median distance in the post-reform period. This result further sheds light on the importance of local ties and on how home bias

might adversely affect bureaucratic performance even in the presence of random assignment. Combined, the results suggest that both changes in the selection and allocation of auditors are important drivers of the overall effects, both pointing toward the importance of ensuring the independence of monitors.

6.3 Deterrence and the Role of Spillovers

Finally, in the lenses of the standard models of crime (Becker and Stigler 1974), the reform can improve outcomes directly through an improvement in the detection of lax financial practices, but also act indirectly through deterrence, inducing mayors to change their behavior in expectation of potentially facing a stricter auditor in the future. By design, my baseline estimates capture the effects of the reform, as all municipalities became aware of the future change in the appointment system of auditors at the moment of the reform adoption. Thus, a general deterrence effect cannot immediately be isolated and would likely lead to a downward bias in my baseline estimates of Table 3. However, we can hypothesize that the consequences of the change in auditors' appointment are not fully clear or salient to mayors before a nearby municipality is subject to the rule change, but they become more salient afterwards. Under this assumption, I can exploit the staggered introduction across municipalities and the geographic variation in the exact timing of the audits to isolate the role played by deterrence effects. Following Colonnelli and Prem (2021), I proceed in two steps. First, I identify the spillover effects of the reform on municipalities that are yet-to-be-treated but neighboring a treated municipality. I consider a municipality to be exposed to the spillover effects of the treatment whenever a neighboring municipality has been treated. The control group consists of all municipalities, aside from the neighboring ones, that are yet-to-be treated. Results are reported in Table A19. As evident from Panel A, I detect sizable spillover effects, that are over 60% the size of the main effects, a magnitude that is similar to the one found by Colonnelli and Prem (2021) for Brazilian anti-corruption audits. Given the quantitative importance of spillover effects and deterrence, I then re-estimate my main equation in order to include both the direct and indirect effects of the audits. To do so, in Panel B of Table A19, I exclude from the control group all not-yet-treated municipalities that neighbour the one experiencing treatment in a given year. This ensures that the estimates are based on comparisons of treated municipalities relative to municipalities that were neither directly nor indirectly affected. The estimated effects of the reform in this context are almost twice as large as the average effects of the reform in Table 3, as one would expect given the size of the spillover effects detected in Panel A, thus highlighting that the results estimated in Table 3 represent a lower bound of the overall effects of the reform.

Taken together, these results suggest that the effects estimated in Section 5.2 represent

a lower bound for the overall effects of the reform, and highlight the importance of taking into account both direct and indirect effects when assessing the effectiveness of monitoring.

7 Discussion and Conclusion

Most organizations are multi-layered and crucial decisions are frequently delegated to local agents, giving rise to agency problems. While considered a crucial tool to align incentives across layers, effective monitoring is however frequently impaired by conflicts of interest, as the audited party itself is usually charged with appointing the auditor. Although this discretionary system may reduce information frictions and promote cooperation, it inherently introduces potential bias into the monitoring process. This paper highlights that changes in the design of monitoring institutions can significantly improve governance outcomes in multi-layered organizations. Taking advantage of a large-scale reform that changed the appointment system of auditors for municipal government budgets in Italy, I investigate the causal effects of removing appointment control from mayors and introducing a random assignment of auditors. Four findings arise. First, random matching severs pre-existing auditors-mayors connections that likely impaired monitoring effectiveness. Second, treated municipalities improve their financial health, in accordance with national government objectives. Third, the improvement is achieved through revenue-based consolidation rather than by cutting expenditures, as treated municipalities show substantial improvements in their tax capacity, a key measure of organizational performance in governments. Fourth, the treatment effects are significantly larger for municipalities where the misalignment between national and local incentives lead to higher risk of auditors' capture before the reform and for those matched to a more distant or less connected auditor in the post-reform period.

These findings can inform the global debate about policies to curb mis-beheavior and solve conflicts of interest in monitoring, and they provide two important policy takeaways. One pertains to the reform's cost-effectiveness. Rather than introducing a new auditing system (e.g., federal audits in Brazil) or an additional supervisory board (e.g., the PCAOB for U.S. audit firms), the reform I study operates by introducing frictions that make potential collusive agreements harder while leaving features of institutional design unchanged. In this sense, the benefits come with very low implementation costs. The second takeaway relates to the auditors' identity. As distinct from other types of reforms or experimental studies, the auditors here are not central government bureaucrats—as in the well-known case of Brazil (Ferraz and Finan 2008) or Puerto Rico (Bobonis et al. 2016) —nor are they village members—as in the case of Indonesia (Olken 2007). They are Certified Public Accountants hired directly by municipalities to carry out auditing duties, and auditing governments is only a limited part of their work activities. This suggests a potentially important role for

external professionals to carry out audit roles in other settings as well. For example, a similar system is in place in the United States, where local governments are required to hire external auditors to conduct so-called "single audits", an important audit process aimed at reviewing and certifying the proper spending of federal awards (Tassin et al. 2019, Cuneo et al. 2023).

More generally, the results of this paper have implications for a long-standing debate in accounting and finance around the challenges posed by the conflicts of interest in the context of private-sector auditing (Moore et al. 2006; Lennox and Wu 2018) and bank supervision (Agarwal et al. 2014, Fernández-Villaverde et al. 2013). While most research has focused on diagnosing and quantifying the problems of regulatory and auditors' capture, little is known about potential solutions, as it is rare to observe changes in the design of regulatory institutions. While focusing on the public sector, this paper provides evidence on how mandatory rotation policies, if combined with random assignment and a large pool of potential supervisors, can be an effective policy solution to limit conflicts of interest and improve monitoring effectiveness.

References

Agarwal, Sumit, David Lucca, Amit Seru, and Francesco Trebbi, "Inconsistent regulators: Evidence from banking," *The Quarterly Journal of Economics*, 2014, *129* (2), 889–938. 37

Aghion, Philippe and Jean Tirole, "Formal and Real Authority in Organizations," *Journal of Political Economy*, 1997, *105* (1), 1–29. 1, 7

Allen, Robert C., Mattia C. Bertazzini, and Leander Heldring, "The Economic Origins of Government," Working Paper 2022. 4

Alonso, Ricardo and Niko Matouschek, "Optimal delegation," *The Review of Economic Studies*, 2008, *75* (1), 259–293. 1, 5, 30

Aman-Rana, Shan and Clement Minaudier, "Spillovers in State Capacity Building: Evidence from the Digitization of Land Records in Pakistan," 2024. 4, 7

Aneja, Abhay and Guo Xu, "Strengthening State Capacity: Civil Service Reform and Public Sector Performance during the Gilded Age," 2023. 15

Angelucci, Charles and Antonio Russo, "Petty corruption and citizen reports," *International Economic Review*, 2022, *63* (2), 831–848. 6

Balan, Pablo, Augustin Bergeron, Gabriel Tourek, and Jonathan L. Weigel, "Local Elites as State Capacity: How City Chiefs Use Local Information to Increase Tax Compliance in the Democratic Republic of the Congo," *American Economic Review*, March 2022, *112* (3), 762–97. 7

Bandiera, Oriana, Andrea Prat, and Tommaso Valletti, "Active and Passive Waste in Government Spending: Evidence from a Policy Experiment," *The American Economic Review*, 2009, *99*, 1278–1308. 1

_ , **Michael Carlos Best, Adnan Qadir Khan, and Andrea Prat**, "The allocation of authority in organizations: A field experiment with bureaucrats," *The Quarterly Journal of Economics*, 2021, *136* (4), 2195–2242. 1, 7

_ , **Robin Burgess, Erika Deserranno, Ricardo Morel, Munshi Sulaiman, and Imran Rasul**, "Social incentives, delivery agents, and the effectiveness of development interventions," *Journal of Political Economy Microeconomics*, 2023, *1* (1), 162–224. 1, 5

Banerjee, Abhijit, Sendhil Mullainathan, and Rema Hanna, "Corruption," Technical Report, National Bureau of economic research 2012. 6

Barone, Guglielmo, Laura Conti, Gaia Narciso, and Marco Tonello, "Auditors' conflict of interest: does random selection work?," Trinity Economics Papers tep0820, Trinity College Dublin, Department of Economics April 2020. 25

Becker, Gary S. and George J. Stigler, "Law Enforcement, Malfeasance, and Compensation of Enforcers," *The Journal of Legal Studies*, 1974, *3* (1), 1–18. 6, 35

Bertrand, Marianne, Esther Duflo, and Sendhil Mullainathan, "How much should we trust differences-in-differences estimates?," *The Quarterly journal of economics*, 2004, *119* (1), 249–275. 16

Besley, Timothy and Torsten Persson, "The origins of state capacity: Property rights, taxation, and politics," *American economic review*, 2009, *99* (4), 1218–1244. 4

_ **and** _ , "Chapter 2 - Taxation and Development," in Alan J. Auerbach, Raj Chetty, Martin Feldstein, and Emmanuel Saez, eds., *handbook of public economics, vol. 5*, Vol. 5 of *Handbook of Public Economics*, Elsevier, 2013, pp. 51–110. 4

_ , **Robin Burgess, Adnan Khan, and Guo Xu**, "Bureaucracy and development," *Annual Review of Economics*, 2022, *14*, 397–424. 7

Best, Michael Carlos, Anne Brockmeyer, Henrik Jacobsen Kleven, Johannes Spinnewijn, and Mazhar Waseem, "Production versus revenue efficiency with limited tax capacity: theory and evidence from Pakistan," *Journal of political Economy*, 2015, *123* (6), 1311–1355. 7

_ , **Jonas Hjort, and David Szakonyi**, "Individuals and organizations as sources of state effectiveness," 2023. 7

Bó, Ernesto Dal, Frederico Finan, Nicholas Y Li, and Laura Schechter, "Information technology and government decentralization: Experimental evidence from paraguay," *Econometrica*, 2021, *89* (2), 677–701. 7

Bobonis, Gustavo J., Luis R.Cámara Fuertes, and Rainer Schwabe, "Monitoring corruptible politicians," *American Economic Review*, 2016, *106* (8), 2371–2405. 7, 36

Bolton, Patrick and Mathias Dewatripont, *9. Authority in Organizations: A Survey*, Princeton: Princeton University Press, 7

Borusyak, Kirill, Xavier Jaravel, and Jann Spiess, "Revisiting event study designs: Robust and efficient estimation," *arXiv preprint arXiv:2108.12419*, 2021. 15

_ , _ , **and** _ , "Revisiting event study designs: Robust and efficient estimation," *arXiv preprint arXiv:2108.12419*, 2021. 24

Brollo, Fernanda, "Who is punishing corrupt politicians: Voters or the central government? Evidence from the Brazilian anti-corruption program," Working Paper, IGIER 2008. 26

_ **and Ugo Troiano**, "What happens when a woman wins an election? Evidence from close races in Brazil," *Journal of Development Economics*, 2016, *122*, 28–45. 26

Callaway, Brantly and Pedro H.C. Sant'Anna, "Difference-in-Differences with multiple time periods," *Journal of Econometrics*, 2021, *225* (2), 200–230. Themed Issue: Treatment Effect 1. 3, 15, 22, 7, 26, 27

Casaburi, Lorenzo and Ugo Troiano, "Ghost-house busters: The electoral response to a large anti–tax evasion program," *The Quarterly Journal of Economics*, 2016, *131* (1), 273–314. 27, 50

Cengiz, Doruk, Arindrajit Dube, Attila Lindner, and Ben Zipperer, "The Effect of Minimum Wages on Low-Wage Jobs*," *The Quarterly Journal of Economics*, 05 2019, *134* (3), 1405–1454. 3, 15

Chambru, Cédric, Emeric Henry, and Benjamin Marx, "The dynamic consequences of state-building: evidence from the French Revolution," 2021. 8

Chiades, Paolo and Vanni Mengotto, "Il calo degli investimenti nei Comuni tra Patto di stabilità interno e carenza di risorse," *Economia Pubblica*, 2015, (2). 25

Chu, Jian, Raymond Fisman, Songtao Tan, and Yongxiang Wang, "Hometown favoritism and the quality of government monitoring: Evidence from rotation of Chinese auditor," Boston University - Department of Economics - The Institute for Economic Development Working Papers Series dp-343, Boston University - Department of Economics February 2020. 21

Colonnelli, Emanuele and Mounu Prem, "Corruption and Firms," *The Review of Economic Studies*, 07 2021, *89* (2), 695–732. 35

Conley, Timothy G, "GMM estimation with cross sectional dependence," *Journal of econometrics*, 1999, *92* (1), 1–45. 22, 9

_ , "Spatial econometrics," in "Microeconometrics," Springer, 2010, pp. 303–313. 22, 9

Corte dei Conti, "Banca Dati della Corte dei Conti della Giurisdizione e del Controllo," 2022. 11

Cuneo, Martina, Jetson Leder-Luis, and Silvia Vannutelli, "Government Audits," Working Paper 30975, National Bureau of Economic Research February 2023. 37

de Chaisemartin, Clément and Xavier D'Haultfœuille, "Two-Way Fixed Effects Estimators with Heterogeneous Treatment Effects," *American Economic Review*, September 2020, *110* (9), 2964–96. 3, 15, 22, 7, 21, 25, 26

Decarolis, Francesco, Raymond Fisman, Paolo Pinotti, and Silvia Vannutelli, "Rules, Discretion, and Corruption in Procurement: Evidence from Italian Government Contracting," *SSRN Electronic Journal*, 01 2019. 32

Deserranno, Erika, Miri Stryjan, and Munshi Sulaiman, "Leader selection and service delivery in community groups: Experimental evidence from uganda," *American Economic Journal: Applied Economics*, 2019, *11* (4), 240–267. 1

_ , Stefano Caria, Gianmarco Leon-Ciliotta, and Philipp Kastrau, "The Allocation of Incentives in Multi-Layered Organizations: Evidence from a Community Health Program in Sierra Leone," Technical Report, CEPR Discussion Papers 2023. 7

Deshpande, Manasi and Yue Li, "Who Is Screened Out? Application Costs and the Targeting of Disability Programs," *American Economic Journal: Economic Policy*, November 2019, *11* (4), 213–48. 3, 15

Dessein, Wouter, "Authority and communication in organizations," *The Review of Economic Studies*, 2002, *69* (4), 811–838. 1, 7

Dovis, Alessandro and Rishabh Kirpalani, "Fiscal rules, bailouts, and reputation in federal governments," *American Economic Review*, 2020, *110* (3), 860–88. 1, 3

Duflo, Esther, Michael Greenstone, Rohini Pande, and Nicholas Ryan, "Truth-telling by Third-party Auditors and the Response of Polluting Firms: Experimental Evidence from India," *Quarterly Journal of Economics*, 2013, pp. 1499–1545. 6

Dzansi, James, Anders Jensen, David Lagakos, and Henry Telli, "Technology and tax capacity: evidence from local governments in Ghana," Technical Report, National Bureau of Economic Research 2022. 4, 7

Faure-Grimaud, Antoine, Jean-Jacques Laffont, and David Martimort, "Collusion, delegation and supervision with soft information," *The Review of Economic Studies*, 2003, *70* (2), 253–279. 6

Fenizia, Alessandra, Raffaele Saggio et al., "Can the Mafia's Tentacles Be Severed? The Economic Effects of Removing Corrupt City Councils," Technical Report 2020. 25

Fenske, James, Muhammad Haseeb, and Namrata Kala, "How Rules and Compliance Impact Organizational Outcomes: Evidence from Delegation in Environmental Regulation," Technical Report, National Bureau of Economic Research 2023. 7

Fernández-Villaverde, Jesús, Luis Garicano, and Tano Santos, "Political credit cycles: the case of the Eurozone," *Journal of Economic perspectives*, 2013, *27* (3), 145–166. 37

Ferraz, Claudio and Frederico Finan, "Exposing corrupt politicians: the effects of Brazil's publicly released audits on electoral outcomes," *The Quarterly journal of economics*, 2008, *123* (2), 703–745. 36

_ and _ , "Electoral accountability and corruption: Evidence from the audits of local governments," *American Economic Review*, 2011, *101* (4), 1274–1311. 7, 33

Finan, F., B.A. Olken, and R. Pande, *The Personnel Economics of the Developing State*, Vol. 2, Elsevier Ltd, 2017. 7

Gagliarducci, Stefano and Tommaso Nannicini, "Do better paid politicians perform better? Disentangling incentives from selection," *Journal of the European Economic Association*, 2013, *11* (2), 369–398. 4, 27

Garicano, Luis and Luis Rayo, "Why organizations fail: Models and cases," *Journal of Economic Literature*, 2016, *54* (1), 137–192. 7

_ , **Ignacio Palacios-Huerta, and Canice Prendergast**, "Favoritism under social pressure," *Review of Economics and Statistics*, 2005, *87* (2), 208–216. 1, 5

Gibbons, Robert, Niko Matouschek, and John Roberts, "Decisions in organizations," *The handbook of organizational economics*, 2013, pp. 373–431. 7

Goodman-Bacon, Andrew, "Difference-in-differences with variation in treatment timing," *Journal of Econometrics*, 2021, *225* (2), 254–277. Themed Issue: Treatment Effect 1. 15

Grembi, Veronica, Tommaso Nannicini, and Ugo Troiano, "Do Fiscal Rules Matter?," *American Economic Journal: Applied Economics*, July 2016, *8* (3), 1–30. 9, 13, 27

Halac, Marina and Pierre Yared, "Fiscal rules and discretion under limited enforcement," Technical Report, National Bureau of Economic Research 2019. 3

Henn, Soeren, Horacio Larreguy, and John Marshall, "You get what you pay for: When do Certification Programs improve Public Service Delivery?," 2020. 7

Kahn-Lang, Ariella and Kevin Lang, "The Promise and Pitfalls of Differences-in-Differences: Reflections on 16 and Pregnant and Other Applications," *Journal of Business & Economic Statistics*, 2020, *38* (3), 613–620. 18

Kala, Namrata, "The impacts of managerial autonomy on firm outcomes," Technical Report, National Bureau of Economic Research 2019. 1, 7

Khan, Adnan Q, Asim I Khwaja, and Benjamin A Olken, "Tax farming redux: Experimental evidence on performance pay for tax collectors," *The Quarterly Journal of Economics*, 2016, *131* (1), 219–271. 7

_ , **Asim Ijaz Khwaja, and Benjamin A Olken**, "Making moves matter: Experimental evidence on incentivizing bureaucrats through performance-based postings," *American Economic Review*, 2019, *109* (1), 237–270. 7

Laffont, Jean-Jacques and David Martimort, "Collusion under asymmetric information," *Econometrica: Journal of the Econometric Society*, 1997, pp. 875–911. 6

Lennox, Clive S. and Xi Wu, "A Review of the Archival Literature on Audit Partners," *Accounting Horizons*, 2018, *32* (2), 1–35. 37

Malani, Anup and Julian Reif, "Interpreting pre-trends as anticipation: Impact on estimated treatment effects from tort reform," *Journal of Public Economics*, 2015, *124*, 1–17. 18

Marattin, Luigi, Tommaso Nannicini, and Francesco Porcelli, "Revenue vs Expenditure Based Fiscal Consolidation: The Pass-Trough from Federal Cuts to Local Taxes," Technical Report 2019. 27

Martinez-Bravo, Monica, "The role of local officials in new democracies: Evidence from Indonesia," *American Economic Review*, 2014, *104* (4), 1244–1287. 1

_ , **Gerard Padró i Miquel, Nancy Qian, and Yang Yao**, "The rise and fall of local elections in China," *American Economic Review*, 2022, *112* (9), 2921–58. 1

_ , **Priya Mukherjee, and Andreas Stegmann**, "The non-democratic roots of elite capture: Evidence from Soeharto mayors in Indonesia," *Econometrica*, 2017, *85* (6), 1991–2010. 2

Mastrorocco, Nicola and Edoardo Teso, "State Capacity as an Organizational Problem. Evidence from the Growth of the US State Over 100 Years," Technical Report, National Bureau of Economic Research 2023. 7

Miani, Massimo, Remigio Sequi, and Davide Di Russo, *Principi di Vigilanza e Controllo dell'Organo di Revisione degli Enti Locali*, 1 ed., Fondazione Nazionale di Ricerca dei Commercialisti, 2019. 1

Mookherjee, Dilip, "Incentives in Hierarchies," in Robert Gibbons and John Roberts, eds., *The Handbook of Organizational Economics*, Princeton University Press, 2012. 5, 30

_ , "Political Decentralization," *Annual Review of Economics*, 2015, *7* (1), 231–249. 1

_ **and Masatoshi Tsumagari**, "Regulatory mechanism design with extortionary collusion," *Journal of Economic Theory*, 2023, *208*, 105614. 6

Moore, Don A., Philip E. Tetlock, Lloyd Tanlu, and Max H. Bazerman, "Conflicts Of Interest And The Case Of Auditor Independence: Moral Seduction And Strategic Issue Cycling," *Academy of Management Review*, 2006, *31* (1), 10–29. 37

ODCEC, *L'Organo di Revisione economico-finaziaria negli Enti locali* 2009. 1

Olken, Benjamin A., "Monitoring corruption: Evidence from a field experiment in Indonesia," *Journal of Political Economy*, 2007, *115* (2), 200–249. 7, 36

Olken, Benjamin A and Rohini Pande, "Corruption in developing countries," *Annu. Rev. Econ.*, 2012, *4* (1), 479–509. 6

Ortner, Juan and Sylvain Chassang, "Making corruption harder: Asymmetric information, collusion, and crime," *Journal of Political Economy*, 2018, *126* (5), 2108–2133. 6

Prendergast, Canice, "A theory of" yes men"," *The American Economic Review*, 1993, pp. 757–770. 6

Rahman, David, "But who will monitor the monitor?," *American Economic Review*, 2012, *102* (6), 2767–97. 6

Rambachan, Ashesh and Jonathan Roth, "An Honest Approach to Parallel Trends," 2020. 18

Rodden, Jonathan, "The dilemma of fiscal federalism: Grants and fiscal performance around the world," *American Journal of Political Science*, 2002, pp. 670–687. 1

Ronen, Joshua, "Corporate audits and how to fix them," *Journal of Economic Perspectives*, 2010, *24* (2), 189–210. 10, 33

Roth, Jonathan, "Pre-test with Caution: Event-study Estimates After Testing for Parallel Trends," 2020. 18

Sierra, Raúl Sánchez De La, "On the origins of the state: Stationary bandits and taxation in eastern congo," *Journal of Political Economy*, 2020, *128* (1), 000–000. 4, 8

Slemrod, Joel, "Tax compliance and enforcement," *Journal of economic literature*, 2019, *57* (4), 904–954. 4

Snowberg, Erik and Michael Ting, "An organizational theory of state capacity," *NBER Working Paper*, 2019. 4

Strausz, Roland, "Delegation of monitoring in a principal-agent relationship," *The Review of Economic Studies*, 1997, *64* (3), 337–357. 6

Strulovici, Bruno, "Learning and corruption on monitoring chains," in "AEA Papers and Proceedings," Vol. 111 American Economic Association 2014 Broadway, Suite 305, Nashville, TN 37203 2021, pp. 544–548. 6

Sun, Liyang and Sarah Abraham, "Estimating dynamic treatment effects in event studies with heterogeneous treatment effects," *Journal of Econometrics*, 2021, *225* (2), 175–199. 27

Tassin, Kerri L., Tammy R. Waymire, and Christopher S. Hines, "A Historical Evaluation of the Single Audit: Thirty Years from Initial Legislation to Uniform Guidance," *Journal of governmental & nonprofit accounting*, 2019, *8* (1), 21–35. 37

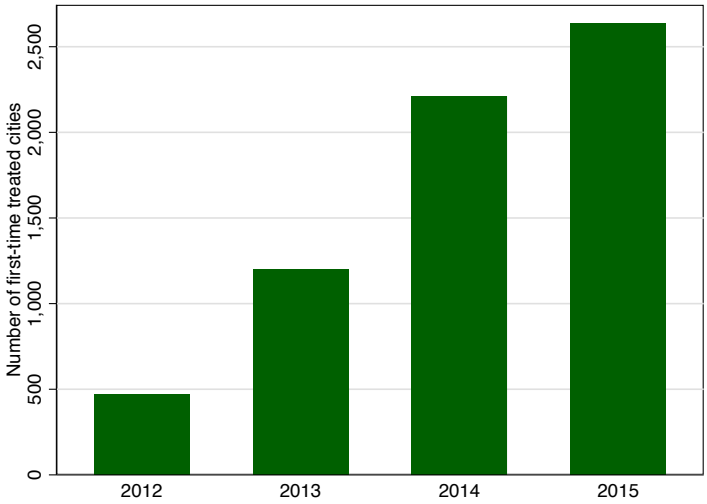
Tirole, Jean, "Hierarchies and Bureaucracies: On the Role of Collusion in Organizations," *Journal of Law, Economics, and Organization*, 1986, *2* (2), 181–214. 6

White, Lawrence J, "Markets: The credit rating agencies," *Journal of Economic Perspectives*, 2010, *24* (2), 211–26. 10, 33

Xu, Guo, Marianne Bertrand, and Robin Burgess, "Organization of the State: Home Assignment and Bureaucrat Performance," *The Journal of Law, Economics, and Organization*, 10 2021. ewab022. 5

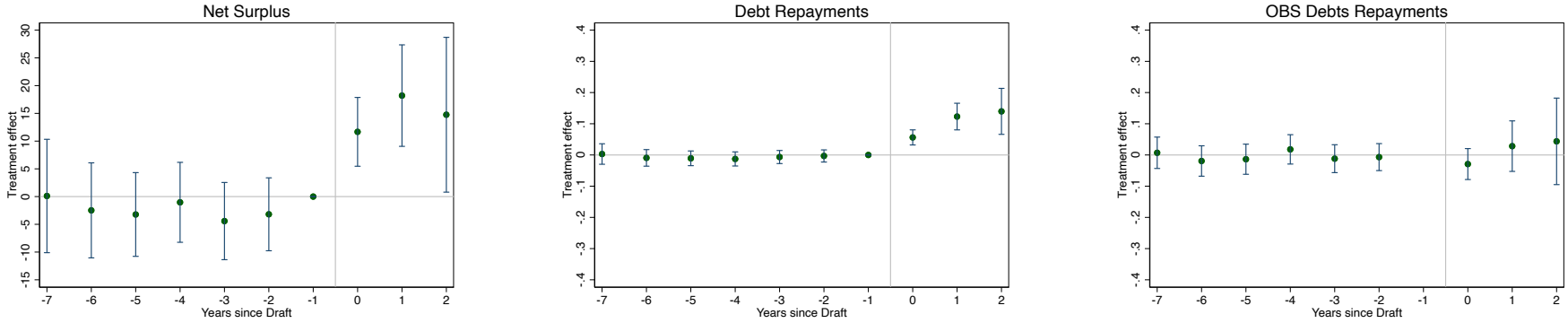
Tables and Figures

Figure 1
Number of Municipalities Receiving a Draft-appointed Auditor for the First Time in a Given Year



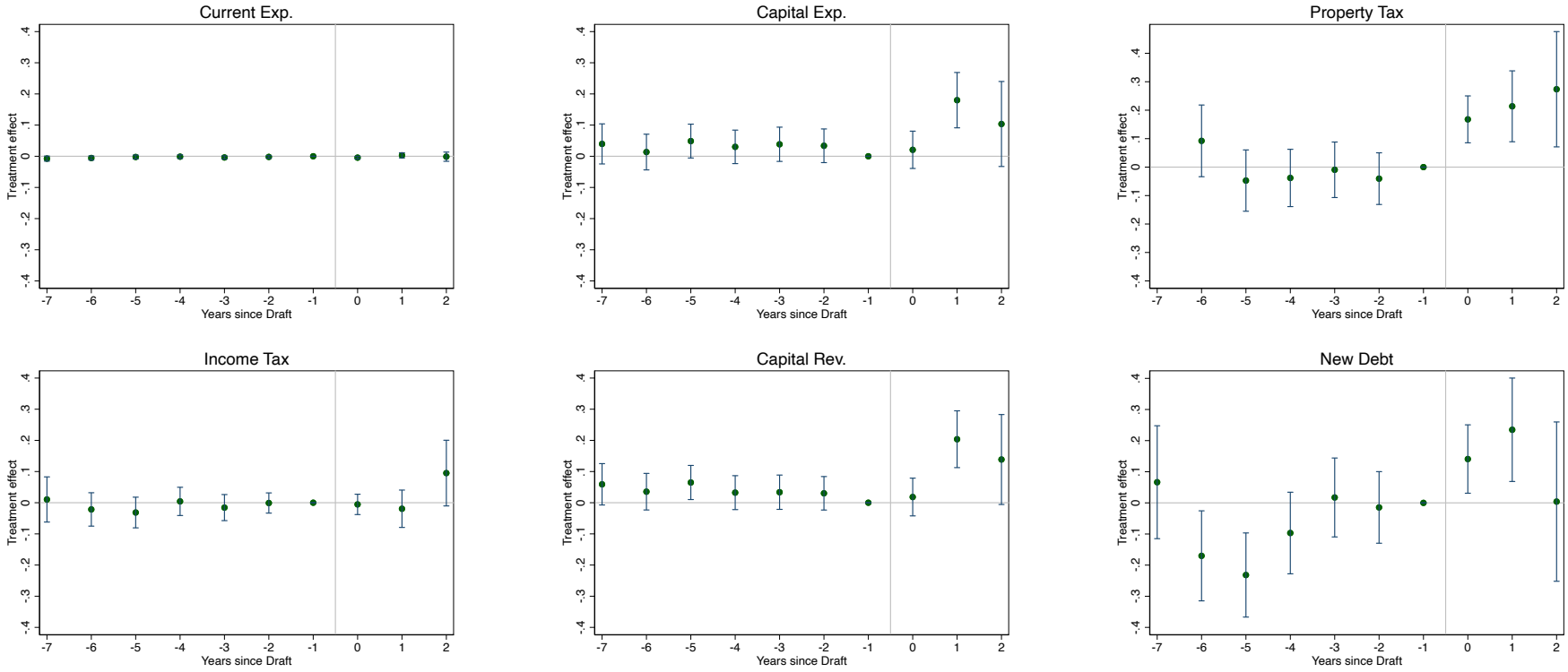
Notes: The figure shows the main margin of variation used in the empirical analysis. The bar graph shows the number of municipalities (y-axis) that had a draft-appointed auditor operative for at least 3 months at the moment of the balance-sheet adoption, in a given year (x-axis).

Figure 2
The Dynamic Effect of Auditor's Independence on Fiscal Sustainability



Notes: The graphs report coefficients and 95% confidence intervals estimated according to specification 2. 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.

Figure 3
The Dynamic Effect of Auditor's Independence on Aggregate Spending and Revenue Choices



Notes: The graphs report coefficients and confidence intervals estimated according to specification 2. Standard errors are clustered at the municipality level. All dependent variables are in per capita terms, and 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.

Table 1
Summary Statistics for the Main Outcomes in the Pre-Reform Period

A. Fiscal Sustainability						
	Mean	S.D.	Median	p10	p90	p99
Net Surplus	-306.97	204.81	-258.73	-554.19	-125.73	31.30
Debt Repayments	59.96	48.95	47.72	21.63	107.07	279.40
OBS Debts Repayments	3.61	13.01	0.00	0.00	7.74	89.05

B. Revenue and Spending Choices						
	Mean	S.D.	Median	p10	p90	p99
Current Exp.	886.77	401.74	770.17	566.10	1,312.69	2,804.00
Capital Exp.	556.88	783.00	295.62	88.94	1,230.57	5,012.36
Property Tax	140.82	140.24	123.58	0.00	268.57	779.23
Income Tax	38.32	28.84	37.61	0.00	76.26	113.14
Capital Rev.	487.38	744.31	232.78	69.08	1,112.87	4,704.39
New Debt	132.34	255.13	25.94	0.00	379.03	1,446.98

Note: The table shows summary statistics of outcome variables for years up to 2010. All variables are in per-capita terms, expressed in 2018 euros, and winsorized at the 1% to remove outliers.

Table 2
Reform-induced Changes in Auditors' Characteristics

	Pre Draft	Post Draft	Difference	Standardized Difference
	(1)	(2)	(2)-(1)	(3)/S.D(1)
Female	0.26	0.27	0.01 *	0.023
Age	53.49	53.87	0.38 ***	0.046
Experience as auditor	5.95	2.21	-3.75 ***	-0.513
Experience as accountant	15.08	15.64	0.56 ***	0.063
Re-appointed	0.45	0.04	-0.41 ***	-0.824
Local resident	0.24	0.01	-0.23 ***	-0.535
Distance from Residence (in minutes)	16.12	63.41	47.29 ***	2.711
Same gender of mayor	0.66	0.67	0.01	0.021
Same birthplace of mayor	0.15	0.02	-0.12 ***	-0.339
Same surname of local politician	0.06	0.01	-0.05 ***	-0.217
Observations	12,541	10,553		

Notes: The table shows the average characteristics of appointed auditors, before and after the draft appointment was introduced. Column (3) (resp. (4)) shows the difference (standardized difference) in means between columns (1) and (2), as well as the significance level (p-values are calculated using heteroskedasticity-robust standard errors). Significance: * significant at 10%; ** significant at 5%; *** significant at 1%,. *Experience as auditor* refers to the number of years the individual has previously served as municipal auditor, *Experience as accountant* refers to the number of years as CPA, *Re-appointed* is an indicator equal to 1 if the auditor serves for 2 consecutive terms in the same municipality, *Local resident* is an indicator equal to 1 if the auditor is resident in the same municipality where she serves as an auditor, *Distance from residence* represents the distance in minutes between the municipality of appointment and the municipality of residence.

Table 3
The Effect of Random Auditor Assignment on Fiscal Sustainability

	Net Surplus		Debt Repayments		OBS Debt Repayments	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	16.45***	16.02***	0.0885***	0.0869***	-0.00218	-0.00222
× Post	[3.240]	[3.236]	[0.0140]	[0.0139]	[0.0253]	[0.0253]
Controls	No	Yes	No	Yes	No	Yes
Dep. Var Mean	-194.8	-194.8	65.29	65.29	3.255	3.255
Observations	114028	114028	114028	114028	114028	114028
Adj. R-sq	0.699	0.700	0.757	0.757	0.412	0.413

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. All dependent variables are in per capita terms. *DebtRepayments* and *OBSDebtRepayments* are transformed using the inverse hyperbolic sine transformation, but their mean is reported as non-transformed. All regressions include municipality, event time fixed effects, election cycle fixed effects, population bins-times-year fixed effects, and the following controls: mayor initial age (in logs), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Table 4
The Effect of Random Auditor Assignment on Aggregate Spending and Revenue Choices

	(1)	(2)	(3)	(4)	(5)	(6)
	Current Exp.	Capital Exp.	Property Tax	Income Tax	Capital Rev.	New Debt
Treated	0.000184	0.0423*	0.205***	0.00758	0.0445*	0.225***
× Post	[0.00309]	[0.0254]	[0.0434]	[0.0254]	[0.0266]	[0.0560]
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Dep. Var Mean	875.0	465.3	158.7	43.73	398.8	123.4
Observations	114028	114028	99040	114028	114028	114028
R-sq	0.944	0.514	0.458	0.844	0.506	0.445

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. All dependent variables are in per capita terms and transformed using the inverse hyperbolic sine transformation, but their non-transformed mean is reported at the bottom of the table. All regressions include municipality, relative event-time fixed effects, election cycle fixed effects, population bins-times-year fixed effects, and the following controls: mayor's initial age (in log), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Table 5
The Effect of Random Auditor Assignment on Tax Capacity

	Property Tax Revenue, Assessed			Property Tax Revenue, Collected			Speed of Property Tax Collection			Property Tax Rate			Investments for Administrative Capacity		
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)	(11)	(12)	(13)	(14)	(15)
Treated	0.201***	0.186***	0.183***	0.183***	0.171***	0.170***	0.406	0.451	0.423	-0.00320	-0.00243	0.00177	0.0669*	0.0629*	0.0517
× Post	[0.0436]	[0.0432]	[0.0430]	[0.0423]	[0.0420]	[0.0417]	[0.413]	[0.411]	[0.411]	[0.0222]	[0.0222]	[0.0221]	[0.0373]	[0.0376]	[0.0370]
TreatXPostXGBI		0.126***			0.111***			-0.645*			-0.00680			0.0350	
		[0.0351]			[0.0339]			[0.336]			[0.0161]			[0.0255]	
TreatXPostXUndeclared			0.164***			0.151***			-0.679*			-0.0452***			0.142***
			[0.0337]			[0.0329]			[0.406]			[0.0173]			[0.0292]
Dep. Var Mean	158.9			158.6			183.3			7.218			3.238		
Het. Var Mean		0.0268	0.0182		0.0267	0.0181		0.0262	0.0174		0.0268	0.0182		0.0268	0.0182
Het. Var SD		0.0209	0.0137		0.0209	0.0137		0.0205	0.0132		0.0209	0.0137		0.0209	0.0137
Observations	98401	98401	98401	97531	97531	97531	84267	84267	84267	97677	97677	97677	98401	98401	98401
R-sq	0.458	0.458	0.458	0.544	0.544	0.544	0.657	0.657	0.657	0.875	0.875	0.875	0.496	0.496	0.496

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. All regressions include municipality, relative event-time fixed effects, election cycle fixed effects, population bins-times-year fixed effects, and the following controls: mayor's initial age (in log), gender, term in office and a dummy equal to one if the mayor was born in the municipality. *Property Tax Revenue, Assessed* is the total amount of property taxes assessed in a given year, *Property Tax Revenue, Collected* is the total amount of property taxes collected in a given year, *Speed of Property Tax Collection* is the ratio between collected and assessed property tax revenues within the year; *Investments for Administrative Capacity* is the total amount of investment spending devoted to adminstrative affairs, which include investments for expanding the local tax office. All outcomes variables, except for *Speed of Property Tax Collection*, are transformed using the inverse hyperbolic sign transformation. *GBI* is the standardized share of "Ghost Buildings Intensity" at the municipality level, as calculated by Casaburi and Troiano (2016); *Undeclared* is the standardized share of houses that the National Tax Agency found to be undeclared by cross-checking tax declarations with cadastral data.

Table 6
The Effect of Random Auditor Assignment on Public Good Provision and the Real Economy

	Housing Units w. Waste Collection (%)	Public Lighting Coverage	Any Public Nursery	Nursery Acceptance Rate	Teacher-Pupil Ratio	Sports Facilities per 10,000 Inhabitants	Total Income per Capita	Total Number of Taxpayers	Total Population
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)
Treated	0.000952	-0.000832	-0.00552	-0.000947	0.00390	-0.0646	0.00118	0.000379	-0.000506
× Post	[0.00413]	[0.00315]	[0.00448]	[0.00657]	[0.00292]	[0.128]	[0.000901]	[0.000465]	[0.000906]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-T. DV Mean	0.899	0.580	0.277	0.860	0.199	5.477	9.429	0.708	7.805
Pre-T. DV SD	0.207	0.308	0.448	0.202	0.110	9.712	0.267	0.0681	1.254
Observations	103914	109570	114028	26452	19262	114028	114028	114028	114028
Adj. R-sq	0.636	0.897	0.833	0.737	0.838	0.786	0.987	0.953	1.000

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. All regressions include municipality, relative event-time fixed effects, election cycle fixed effects, population bins-times-year fixed effects, and the following controls: mayor's initial age (in log), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Table 7
The Effect of Random Auditor Assignment on the Internal Organization and Performance of Municipal Bureaucracy

PANEL A: Personnel Composition, Worktime and Beaurocratic Discontent								
	Bureaucrats per 10,000 Inhabitants	Hiring Rate	Separation Rate	Hours Worked per Capita	% of Workers w. Short-Term Contracts	Strike Days per Capita	Absences per Capita	Sick Days per capita
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated	-0.315	-0.000140	0.000895	2.170	0.00118	0.0169***	0.348	0.317***
× Post	[0.231]	[0.00145]	[0.00193]	[8.747]	[0.00186]	[0.00275]	[0.231]	[0.118]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-T. DV Mean	60.48	0.0403	0.0537	1568.4	0.0397	0.0687	37.31	3.456
Pre-T. DV SD	31.48	0.0757	0.0896	444.7	0.0969	0.167	11.51	5.185
Observations	105645	105645	105645	105645	105645	105645	105645	105645
Adj. R-sq	0.947	0.166	0.135	0.546	0.464	0.585	0.528	0.508

PANEL B: Allocation of Hours across Tasks and Overall Productivity								
	Public Records & Clerks' Office	Human Resources	Political Works	Housing & Public Works	Public Goods & Service Provision	Financial Management & Procurement	Taxation	Legislative Productivity
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
Treated	0.000485	0.00443**	0.00219	-0.000591	-0.00552***	-0.00130	0.000507	20.64
× Post	[0.00108]	[0.00196]	[0.00229]	[0.000684]	[0.00187]	[0.00128]	[0.000604]	[17.38]
Controls	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Pre-T. DV Mean	0.109	0.208	0.357	0.0331	0.204	0.102	0.0482	5.620
Pre-T. DV SD	0.0706	0.125	0.139	0.0324	0.114	0.0687	0.0296	258.1
Observations	106169	106169	106169	106169	106169	106169	106169	112668
Adj. R-sq	0.712	0.722	0.673	0.382	0.628	0.590	0.752	0.0980

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. All regressions include municipality, relative event-time fixed effects, election cycle fixed effects, population bins-times-year fixed effects, and the following controls: mayor's initial age (in log), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Table 8
The Effect of Random Auditor Assignment on Fiscal Sustainability, by Pre-reform Risk of Capture

	Net Surplus	Debt Repayments	OBS Debt Repayments
		PANEL A: Fiscal Stance	
	(1)	(2)	(3)
Treated × Post	10.90***	0.0350**	-0.176***
	[3.559]	[0.0150]	[0.0466]
Treated × Post × Below Median	12.42**	0.106***	0.314***
	[5.616]	[0.0237]	[0.0461]
Dep. Var Mean	-194.791	65.292	3.255
Sum of Coefs.	23.315***	.141***	.137***
Observations	114028	114028	114028
R-sq	0.709	0.769	0.442
		PANEL B: Local Auditor	
	(1)	(2)	(3)
Treated × Post	10.35**	0.0876***	-0.0432
	[4.984]	[0.0226]	[0.0416]
Treated × Post × Local Auditor	0.461	0.0755*	0.110
	[7.483]	[0.0428]	[0.0791]
Dep. Var Mean	-180.278	61.281	4.072
Sum of Coefs.	10.811*	.163***	.067
Observations	62394	62394	62394
R-sq	0.727	0.734	0.453
		PANEL C: Corruption Investigations	
	(1)	(2)	(3)
Treated × Post	15.88***	0.0637***	-0.0330
	[3.483]	[0.0137]	[0.0249]
Treated × Post × Corrupt	-3.545	0.181***	0.282***
	[8.432]	[0.0551]	[0.102]
Dep. Var Mean	-194.791	65.292	3.255
Sum of Coefs.	12.332	.245***	.249**
Observations	114028	114028	114028
R-sq	0.715	0.770	0.442
		PANEL D: Auditor's Career Concerns	
	(1)	(2)	(3)
Treated × Post	10.280**	0.133***	-0.0544
	[4.827]	[0.0268]	[0.0462]
Treated × Post × Second	-1.844	-0.0757**	0.0690
	[6.407]	[0.0295]	[0.0545]
Observations	91846	91846	91846
Sum of Coefs.	8.436***	.057***	.014
R-sq	0.693	0.757	0.427
		PANEL E: Mayor's Career Concerns	
	(1)	(2)	(3)
Treated × Post	15.70***	0.0778***	-0.0136
	[3.629]	[0.0155]	[0.0283]
Treated × Post × Term-limited	4.116	0.0532*	0.0470
	[6.210]	[0.0308]	[0.0550]
Dep. Var Mean	-194.8	65.29	3.255
Observations	114028	114028	114028
R-sq	0.715	0.769	0.442

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are in parentheses. $BelowMedian$ is a dummy variable equal to 1 if the value of the outcome variable in a given municipality was below the cohort-specific median value in 2011. $LocalAuditor$ is equal to 1 if any of the municipality's appointed auditors before the reform were either born or resident in the municipality. $Corrupt$ is an indicator flagging whether there was at least one investigation for corruption-related crimes from 2004 to 2013 in the municipality, $Second$ is an indicator equal to 1 if the last auditor before the introduction of the random draft was in her second term (as opposed to the first). $Term-Limited$ is a dummy equal to 1 if the mayor in power at the moment of treatment could not run for re-election due to binding term limits. All regressions include municipality, relative event-time fixed effects, election cycle fixed effects, population bins-times-year fixed effects, and the following controls: mayor's initial age (in log), gender, term in office and a dummy equal to one if the mayor was born in the municipality, as well as interactions between the $Treated$ and $Post$ indicators with the relevant heterogeneity covariate.

Table 9
The Effect of Random Auditor Assignment on Fiscal Sustainability, the Role of Selection and Allocation of Auditors

	PANEL A - Selection:	New Entrants vs. Incumbent Auditors (Post-Reform)	
	Net Surplus	Debt Repayments	OBS Debt Repayments
	(1)	(2)	(3)
Treated ×	2.311	-0.00252	-0.0174
Post	[6.719]	[0.0208]	[0.0382]
Treated ×	21.29***	0.0439**	0.00218
Post × New Entrant	[6.974]	[0.0217]	[0.0386]
Dep. Var Mean	-224.226	70.303	2.540
Sum of Coefs.	23.606***	.041***	-.015
Observations	82041	82041	82041
R-sq	0.701	0.800	0.357
	PANEL B - Allocation:	Local vs. Distant Auditors (Post-Reform)	
	(1)	(2)	(3)
Treated ×	4.831	0.0427**	-0.0355
Post	[4.230]	[0.0184]	[0.0300]
Treat × Post	22.97***	0.0528**	0.0226
× At least 1 hr	[5.226]	[0.0210]	[0.0360]
Dep. Var Mean	-198.950	65.926	2.923
Sum of Coefs.	27.803***	.095***	-.013
Observations	105477	105477	105477
R-sq	0.712	0.773	0.401

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level in parentheses. All dependent variables are in per capita terms. *DebtRepayments* and *OBSDebtRepayments* are transformed using the inverse hyperbolic sine transformation. All regressions include municipality, year, election cycle fixed effects, and the following controls: mayor's initial age (in logs), gender, term in office and a dummy equal to one if the mayor was born in the municipality. *NewEntrant* is an indicator equal to 1 if the municipality is assigned as an auditor an individual who has never been an auditor before. In Panel B, the sample is restricted to municipalities below 5,000 inhabitants as only those can receive non-experienced auditors. *AtLeast1hr* flags whether the driving distance between the auditor's residence and the municipality is at least 1 hr. Here, sample is restricted to municipalities below 15,000 inhabitants as they are the one receiving only one (instead of 3) auditors.

Appendix For Online Publication

A1 Additional Tables and Figures

Table A1
Auditor's Main Tasks *ex lege 239 TUEL & dlgs 118/2011*

Advising	Monitoring
Mandatory written opinions on:	**Audit and Certification of:**
mayoral three-years programmatic document	preliminary budget
budgetary planning tools	balance sheet
preliminary budget draft	prospectus of mayor's political campaign and entertainment expenses
management of public services	end-of-mandate mayoral report
creation or participation in city-owned-enterprises	prospectus on the respect of fiscal rules
assignment of public services management to private firms or city-owned-enterprises	prospectus on the respect of other limits to public spending
proposals of new debt undertakings	prospectus on the respect of limits to personnel costs
proposals of underwritings of derivatives and other exceptional financial instruments	prospectus on the respect of limits to advertising expenses
proposals of inclusion of off-balance-sheet debt in the financial statements	prospectus on the respect of limits to employees' training expenses
changes to the local tax code	prospectus on the respect of limits to expenses for public vehicles
changes to the local tax management regulations	prospectus on the respect of limits to new real estate acquisitions
re-assessments of un-collected credits	prospectus on the government participation in municipal-owned enterprises
financial recovery plans	three-year local procurement and public works plans
debt repayment plans	three-year hiring plan
extraordinary maintenance plans of local public buildings, roads, and utilities	
	Verification of:
budget adjustments	respect of mandatory transparency regulations and transmission of data
anti-evasion tax collection plans	reliability of tax revenue projections
plans to achieve more efficient management of public services	sustainability of debt repayment plans
plans to reduce personnel costs	existence of off-balance-sheet debts
three-year hiring plans	current account budget balance
three-year public works plans	capital account budget balance
three-year procurement plans	reliability of anti-evasion tax collection plans
	tax collection procedures
	public spending procedures and respect of payment times
	proper use of earmarked revenue sources

Notes: The table represents a summary of the main tasks of auditors, as regulated by Italian law, *ex lege 239 TUEL & dlgs 118/2011*. Additional details can be found in Miani et al. (2019) and ODCEC (2009).

Table A2
Predictors of Drafted Auditors' Rejecting Appointment

	Auditor Rejects Appointment				
	(1)	(2)	(3)	(4)	(5)
Long Distance	0.0664*** [0.0125]	0.0653*** [0.0124]	0.0600*** [0.0122]	0.0594*** [0.0120]	0.0600*** [0.0119]
Female		0.0278** [0.0126]	0.0278** [0.0127]	0.0280** [0.0127]	0.0279** [0.0127]
Age		-0.00371 [0.00832]	-0.00296 [0.00847]	-0.00312 [0.00844]	-0.00290 [0.00842]
Age^2		0.0000548 [0.0000717]	0.0000478 [0.0000729]	0.0000491 [0.0000725]	0.0000472 [0.0000723]
New Entrant		-0.0164 [0.0139]	-0.0173 [0.0141]	-0.0170 [0.0141]	-0.0170 [0.0141]
Experience		-0.00138 [0.00775]	-0.00199 [0.00773]	-0.00186 [0.00775]	-0.00191 [0.00774]
$Experience^2$		0.0000879 [0.000255]	0.000106 [0.000251]	0.000103 [0.000251]	0.000103 [0.000251]
Mountain			0.0234 [0.0169]	0.0212 [0.0179]	0.0211 [0.0180]
Sea			0.0375 [0.0265]	0.0347 [0.0275]	0.0358 [0.0265]
Province Capital			-0.0536 [0.0329]	-0.0553 [0.0332]	-0.0501 [0.0321]
Net Surplus Below Median				-0.0116 [0.0126]	-0.0122 [0.0126]
Debt Repayments Below Median				-0.0154 [0.0111]	-0.0153 [0.0111]
OBS Debt Repayments Below Median				0.0107 [0.00989]	0.0107 [0.00975]
Corruption					0.0110 [0.0167]
Mafia Infiltrations					-0.0570** [0.0264]
Constant	0.168*** [0.00636]	0.208 [0.182]	0.183 [0.187]	0.196 [0.182]	0.190 [0.181]
Dep. Var Mean	0.202	0.202	0.202	0.202	0.202
Observations	6705	6705	6705	6705	6705
Adj. R-sq	0.115	0.118	0.119	0.119	0.119

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. All regressions include $RegionXYearXPopulationSize$ fixed effects. These fixed effects account for the level at which the drafting is performed. The dependent variable, $AuditorRejectsAppointment$ is a dummy equal to 1 if the first drafted auditor rejects the appointment and 0 otherwise on a set of explanatory variables. $LongDistance$ is a dummy equal to 1 if the driving distance between the auditor's municipality of residence and the municipality of appointment is larger than 60 minutes, $NewEntrant$ is a dummy equal to 1 if the auditor has never been a municipal auditor before the reform, $Experience$ measures the years of general auditing experience, $Mountain$ is a dummy equal to 1 if the municipality is located in the mountains, Sea is a dummy equal to 1 if the municipality is located on the coasts, $NetSurplusBelowMedian$ is a dummy equal to 1 if the municipality had a value of fiscal surplus below the median in 2010, $Corruption$ is a dummy equal to 1 if there was at least one investigation for corruption in the municipality, $MafiaInfiltrations$ is a dummy equal to 1 if there had ever been investigations for mafia infiltrations in the municipality.

Table A3
Pre-treatment Municipal Characteristics and Levels of Outcomes, by Treatment Timing

	A. Municipal Characteristics					
	2012 Cohort	2013 Cohort	2014 Cohort	2015 Cohort	Uncond F-test	
Election Cycle	1.85	1.83	1.85	1.84	0.994	
Population	23700.84	8605.04	4040.70	6359.44	0.000	
Mayor Age (log)	3.87	3.87	3.86	3.85	0.064	
Male Mayor	0.90	0.91	0.88	0.88	0.849	
Local Mayor	0.41	0.40	0.34	0.34	0.425	
Mayor Term-limited	0.36	0.39	0.42	0.40	0.552	
	B. Fiscal Sustainability					
	2012 Cohort	2013 Cohort	2014 Cohort	2015 Cohort	Uncond. F-test	Cond. F-test
Net Surplus	-332.48	-336.70	-327.95	-310.64	0.001	0.196
Debt Repayments	60.60	61.92	64.54	62.91	0.413	0.962
OBS Debts Repayments	5.85	5.16	2.70	3.36	0.000	0.775
	C. Revenue and Spending Choices					
	2012 Cohort	2013 Cohort	2014 Cohort	2015 Cohort	Uncond. F-test	Cond. F-test
Current Expenditures	895.74	885.98	907.66	871.69	0.044	0.345
Capital Expenditures	505.93	544.56	543.26	507.68	0.367	0.780
Property Tax	126.03	119.58	144.18	144.77	0.000	0.417
Income Tax	40.05	36.82	40.61	41.11	0.001	0.207
Capital Rev.	458.40	471.34	460.69	427.97	0.295	0.780
New Debt	122.87	150.55	131.16	126.59	0.115	0.737

Note: The table shows summary statistics of outcomes in municipalities for which independent auditor appointment starts, respectively, in 2012, 2013, 2014, and 2015 (first four columns). All variables are measured in 2010. The last 2 columns display, respectively, the p-value of unconditional and conditional F-Tests for equality of means across all four groups. For each variable, the unconditional F-test is obtained by running an OLS regression of the outcome variable on region fixed effects and a set of indicators for the different cohorts and then testing the equality of the indicators' coefficients. The conditional F-test is obtained by running an analogous OLS regression but also including the pre-treatment characteristics of Panel A as controls.

Table A4
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.

Table A5
Comparison of Characteristics of Drafted Auditors vs. Pool of Eligible Auditors

	Drafted	Pool	β	Standardized Difference
	(1)	(2)	(3)	(4)
Female	0.26	0.28	-.01	-0.028
Resident in Municipality	0.008	0.004	0	0.053
Born in Municipality	0.003	0.001	0	0.049
Experience as Accountant	16.89	15.43	-.09	0.123
Age	48.56	47.17	-.07	0.105

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Column (3) displays the coefficient of a separate regression of each of the characteristics on the left-hand-side on a dummy equal to 1 if the individual has been drafted and $RegionXYearXPopulationSize$ fixed effects, which represent the randomization strata. These fixed effects account for the level at which the drafting is performed. Column (4) displays the standardized difference between Column (1) and Column (2), which is equal to the difference in means divided by the square-root of the sum of the variances.

Table A6
The Effect of Random Auditor Assignment on the Selection and Allocation of Auditors

	(1) Female	(2) Age	(3) Experience as Auditor	(4) Experience as Accountant	(5) Local Resident	(6) Distance from Residence (in minutes)	(7) Same Birthplace	(8) Same Surname	(9) Re-appointed
Treated	0.0436	0.205***	-0.480***	0.257***	-0.789***	1.217***	-0.542***	-0.341***	-0.604***
× Post	[0.0438]	[0.0418]	[0.0381]	[0.0410]	[0.0429]	[0.0345]	[0.0483]	[0.0542]	[0.0368]
Dep. Var Mean	0.266	53.25	6.364	14.64	0.217	22.52	0.146	0.0508	0.499
Dep. Var SD	0.442	7.945	8.097	8.789	0.406	18.10	0.353	0.220	0.500
Observations	34521	34521	34521	34521	34521	34521	34521	34521	34521
R-sq	0.793	0.769	0.837	0.774	0.807	0.814	0.717	0.662	0.839

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. The sample is restricted to the subset of municipalities for which I have information about auditors' characteristics for at least 1 term in the pre-period. All dependent variables are standardized. All regressions include municipality, event time fixed effects, election cycle fixed effects, population bins-times-year fixed effects, and the following controls: mayor's initial age (in logs), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Table A7
The Effect of Random Auditor Assignment on Fiscal Sustainability, TWFE model, Testing Robustness to Heterogeneous Treatment Effects as in de Chaisemartin and D'Haultfœuille (2020)

	Net Surplus		Debt Repayments		OBS Debt Repayments	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	18.04***	15.07***	0.0891***	0.0801***	-0.00817	-0.00791
	[3.385]	[3.174]	[0.0135]	[0.0135]	[0.0243]	[0.0245]
Dep. Var Mean	-201.4	-201.4	64.92	64.92	3.291	3.291
Observations	50427	50427	50427	50427	50427	50427
Adj. R-sq	0.679	0.700	0.751	0.757	0.412	0.413
% ATTs with negative weights	29.5	29.5	29.5	29.5	29.5	29.5
Sum of negative weights	-0.106	-0.106	-0.106	-0.106	-0.106	-0.106
$\underline{\sigma_{fe}}$		12.62		0.06		0.009
$\underline{\underline{\sigma_{fe}}}$		50.35		0.26		0.04

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. The table reports coefficient estimates based on the following equation: . All dependent variables are in per capita terms, $DebtRepayments$ and $OBSDebtRepayments$ are transformed using the inverse hyperbolic sine transformation. All regressions include municipality, election cycle fixed effects, population bins-times-year fixed effects, and the following controls: mayor's initial age (in logs), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Table A8
The Effect of Random Auditor Assignment on Fiscal Sustainability, Using the Estimator Developed by Callaway and Sant'Anna (2021)

	Net Surplus		Debt Repayments		OBS Debt Repayments	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	9.070***	7.494**	0.0695***	0.0587***	-0.0111	-0.00548
	[3.369]	[3.400]	[0.0134]	[0.0135]	[0.0300]	[0.0305]
Controls	No	Yes	No	Yes	No	Yes
Dep. Var Mean	-205.8	-205.8	64.47	64.47	3.435	3.435
Observations	44824	44824	44824	44824	44824	44824

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. The table reports coefficient estimates based on the estimator developed by Callaway and Sant'Anna (2021). All dependent variables are in per capita terms, $DebtRepayments$ and $OBSDebtRepayments$ are transformed using the inverse hyperbolic sine transformation. All regressions include municipality, election cycle fixed effects, population bins-times-year fixed effects, and the following controls: mayor's initial age (in logs), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Table A9
The Effect of Random Auditor Assignment on Fiscal Sustainability, additional Pre-treatment Characteristics as Controls

	Net Surplus		Debt Repayments		OBS Debt Repayments	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	16.45***	11.15***	0.0885***	0.0609***	-0.00218	0.00173
× Post	[3.240]	[3.300]	[0.0140]	[0.0135]	[0.0253]	[0.0263]
Controls	No	Yes	No	Yes	No	Yes
Dep. Var Mean	-194.8	-194.8	65.29	65.29	3.255	3.255
Observations	114028	114028	114028	114028	114028	114028
Adj. R-sq	0.699	0.704	0.757	0.764	0.412	0.413

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. 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, and the following controls: mayor's initial age (in logs), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Table A10
The Effect of Random Auditor Assignment on Fiscal Sustainability, including Cohort-by-Municipality Fixed Effects

	Net Surplus		Debt Repayments		OBS Debt Repayments	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	16.06***	15.50***	0.0860***	0.0848***	-0.00281	-0.00356
× Post	[3.228]	[3.224]	[0.0138]	[0.0138]	[0.0253]	[0.0253]
Controls	No	Yes	No	Yes	No	Yes
Dep. Var Mean	-194.8	-194.8	65.29	65.29	3.255	3.255
Observations	114028	114028	114028	114028	114028	114028
Adj. R-sq	0.672	0.673	0.736	0.737	0.359	0.359

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. 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, and the following controls: mayor's initial age (in logs), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Table A11
The Effect of Random Auditor Assignment on Fiscal Sustainability, allowing for Region-Specific Non-parametric Trends

	Net Surplus		Debt Repayments		OBS Debt Repayments	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	8.783***	8.720***	0.0395***	0.0392***	0.0108	0.0108
× Post	[3.118]	[3.117]	[0.0127]	[0.0127]	[0.0251]	[0.0251]
Controls	No	Yes	No	Yes	No	Yes
Dep. Var Mean	-194.8	-194.8	65.29	65.29	3.255	3.255
Observations	114028	114028	114028	114028	114028	114028
Adj. R-sq	0.714	0.714	0.777	0.777	0.416	0.416

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. All dependent variables are in per capita terms. *DebtRepayments* and *OBSDebtRepayments* are transformed using the inverse hyperbolic sine transformation. Regressions include municipality, population-bins-by-year fixed effects, region-times-year fixed effects, relative time fixed effects and election cycle fixed effects, and the following controls: mayor's initial age (in logs), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Table A12
The Effect of Random Auditor Assignment on Fiscal Sustainability, accounting for Spatial Correlation in the Data

	Net Surplus		Debt Repayments		OBS Debt Repayments	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	15.88***	15.88***	0.0872***	0.0872***	-0.00293	-0.00293
× Post	[3.236]	[3.566]	[0.0139]	[0.0149]	[0.0253]	[0.0192]
Conley S.E.	No	Yes	No	Yes	No	Yes
Controls	Yes	Yes	Yes	Yes	Yes	Yes
Observations	114028	114028	114028	114028	114028	114028
Adj. R-sq	0.700		0.757		0.413	

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. In odd columns, standard errors clustered at the municipality level are reported in parentheses. In even columns, I reported standard errors adjusted for spatial correlation in the data for up to 50 km using Conley (1999; 2010) methodology. 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, and the following controls: mayor's initial age (in logs), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Table A13
The Effect of Random Auditor Assignment on Fiscal Sustainability, by Population Size

	Net Surplus	Debt Repayments	OBS Debt Repayments
	(1)	(2)	(3)
Treated × Post	16.62***	0.0331**	-0.0131
	[4.383]	[0.0140]	[0.0269]
Treated × Post × Pop. 5-15 k.	-2.501	0.158***	-0.0409
	[5.919]	[0.0400]	[0.0681]
Treated × Post × Above 15k	-2.545	0.205***	0.169*
	[7.997]	[0.0558]	[0.101]
Dep. Var Mean	-194.8	65.29	3.255
Observations	114028	114028	114028
R-sq	0.715	0.770	0.442

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. All dependent variables are in per capita terms. *DebtRepayments* and *OBSDebtRepayments* are transformed using the inverse hyperbolic sine transformation. All regressions include municipality, year, election cycle fixed effects, population bins-times-year fixed effects, and the following controls: mayor's initial age (in logs), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Table A14
The Effect of Random Auditor Assignment on Fiscal Sustainability, by Geography

	Net Surplus	Debt Repayments	OBS Debt Repayments
	(1)	(2)	(3)
Treated × Post	14.67***	0.0167	-0.0329
	[3.876]	[0.0124]	[0.0220]
Treated × Post × South	-14.21**	0.123***	0.0740
	[6.549]	[0.0299]	[0.0575]
Treated × Post × Center	-2.078	0.0390	0.0797
	[7.850]	[0.0319]	[0.0665]
Dep. Var Mean	-194.8	65.29	3.255
Observations	114028	114028	114028
R-sq	0.719	0.776	0.442

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. All dependent variables are in per capita terms. *DebtRepayments* and *OBSDebtRepayments* are transformed using the inverse hyperbolic sine transformation. All regressions include municipality, year, election cycle fixed effects, population bins-times-year fixed effects and the following controls: mayor initial age (in logs), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Table A15
Discretionary Appointments and Fiscal Sustainability in Pre-Reform Period

	Net Surplus		Debt Repayments		OBS Debt Repayments	
	(1)	(2)	(3)	(4)	(5)	(6)
Control of Appointment	-7.417***	-7.360**	-0.0152*	-0.0179*	0.0199	0.0244
	[2.785]	[3.198]	[0.00885]	[0.0100]	[0.0247]	[0.0291]
Second-Term Mayors	Yes	No	Yes	No	Yes	No
Dep. Var Mean	-205.8	-208.2	64.50	64.45	3.427	3.678
Observations	39329	24442	39329	24442	39329	24442
R-sq	0.727	0.760	0.789	0.813	0.457	0.515

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. All dependent variables are in per capita terms. *DebtRepayments* and *OBSDebtRepayments* are transformed using the inverse hyperbolic sine transformation. *Control of Appointment* is a dummy equal to 1 for years in which the mayor had appointed the auditor, and is constructed using exogenous variation arising from the difference between audit (3 years) and election (5 years) cycle. See Figure A12 and Appendix Figure A11 for further details. The sample is restricted to pre-treatment years. All regressions include municipality, election cycle fixed effects, population bins-times-year fixed effects and the following controls: a dummy for dissolution of the city council, a dummy for mayor's resignation, mayor initial age (in log), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Table A16
The Effect of Random Auditor Assignment on Fiscal Sustainability, Cohort-Specific Treatment Effects

	Net Surplus		Debt Repayments		OBS Debt Repayments	
	(1)	(2)	(3)	(4)	(5)	(6)
Treated in 2012 × Post	13.19*	13.08*	0.0215	0.0201	0.0968	0.0954
	[7.920]	[7.907]	[0.0185]	[0.0185]	[0.0665]	[0.0666]
Treated in 2013× Post	11.21**	10.84*	0.0601***	0.0602***	0.00628	0.00490
	[5.659]	[5.672]	[0.0230]	[0.0230]	[0.0396]	[0.0396]
Treated in 2014× Post	12.37***	12.42***	0.0409**	0.0404**	-0.0467	-0.0470
	[4.297]	[4.298]	[0.0165]	[0.0165]	[0.0322]	[0.0322]
Controls	No	Yes	No	Yes	No	Yes
Dep. Var Mean	-210.1	-210.1	63.59	63.59	3.258	3.258
F-Stat (T.E. Heterogeneity)	0.0237	0.0346	0.881	0.950	2.050	2.011
P-Value	0.977	0.966	0.414	0.387	0.129	0.134
Observations	103498	103498	103498	103498	103498	103498
Adj. R-sq	0.701	0.701	0.786	0.786	0.423	0.423

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. All dependent variables are in per capita terms. *DebtRepayments* and *OBSDebtRepayments* are transformed using the inverse hyperbolic sine transformation. Regressions include municipality, population-bins-by-year fixed effects, region-by-year fixed effects, relative time fixed effects and election cycle fixed effects, as well as the following controls: mayor's initial age (in logs), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Table A17
The Effect of Random Auditor Assignment on Investment, by Investment Category

	Administrative Capacity	Local Police	Education	Culture	Tourism	Public Transport	Local Public Goods	Social Welfare
Treated × Post	0.0745**	0.00110	0.0860**	0.0716**	-0.0196	0.0129	0.0375	0.0718*
	[0.0370]	[0.0234]	[0.0437]	[0.0316]	[0.0263]	[0.0448]	[0.0431]	[0.0390]
Dep. Var Mean	3.309	0.920	2.068	0.762	0.399	4.039	3.818	1.932
Observations	114028	99040	114028	114028	114028	114028	114028	114028
R-sq	0.483	0.361	0.342	0.365	0.431	0.385	0.452	0.317

Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. All dependent variables are in per capita terms and transformed using the inverse hyperbolic sign transformation. All regressions include municipality, year, election cycle fixed effects, population bins-times-year fixed effects, and the following controls: mayor's initial age (in logs), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Table A18
The Effect of Random Auditor Assignment on Fiscal Sustainability, Heterogeneity Analysis using TWFE Model

	Net Surplus	Debt Repayments	OBS Debt Repayments
Treated	8.127*	0.0627***	-0.0317
	[4.601]	[0.0191]	[0.0375]
Treated	15.12**	0.135***	0.0419
× Local Auditor =1	[6.366]	[0.0240]	[0.0492]
Treated	14.05***	0.0632***	-0.0183
	[3.167]	[0.0124]	[0.0230]
Treated	12.73*	0.144***	0.0464
× Corruption=1	[6.828]	[0.0311]	[0.0656]
Treated	10.29*	0.00338	-0.0364
	[5.987]	[0.0169]	[0.0310]
Treated	8.101	0.0392**	0.0215
× New Entrant=1	[6.293]	[0.0155]	[0.0274]
Treated	12.17***	0.0608***	-0.0253
	[3.729]	[0.0145]	[0.0254]
Treated	6.679	0.0377***	0.0246
× At least 1 hr =1	[4.359]	[0.0135]	[0.0250]
Indep. Auditor=1	22.04***	0.0818***	-0.00647
	[4.854]	[0.0188]	[0.0339]
Indep. Auditor=1	-8.933*	-0.00318	-0.00894
× Re-Electable	[5.079]	[0.0179]	[0.0323]
Treated	-7.314**	0.0849***	-0.403***
	[3.562]	[0.0130]	[0.0276]
Treated	47.45***	-0.00850	0.169***
× Below Median=1	[4.361]	[0.0136]	[0.0221]
Dep. Var Mean	-205.802	64.467	3.435
Observations	50427	50427	50427
R-sq	0.661	0.751	0.438

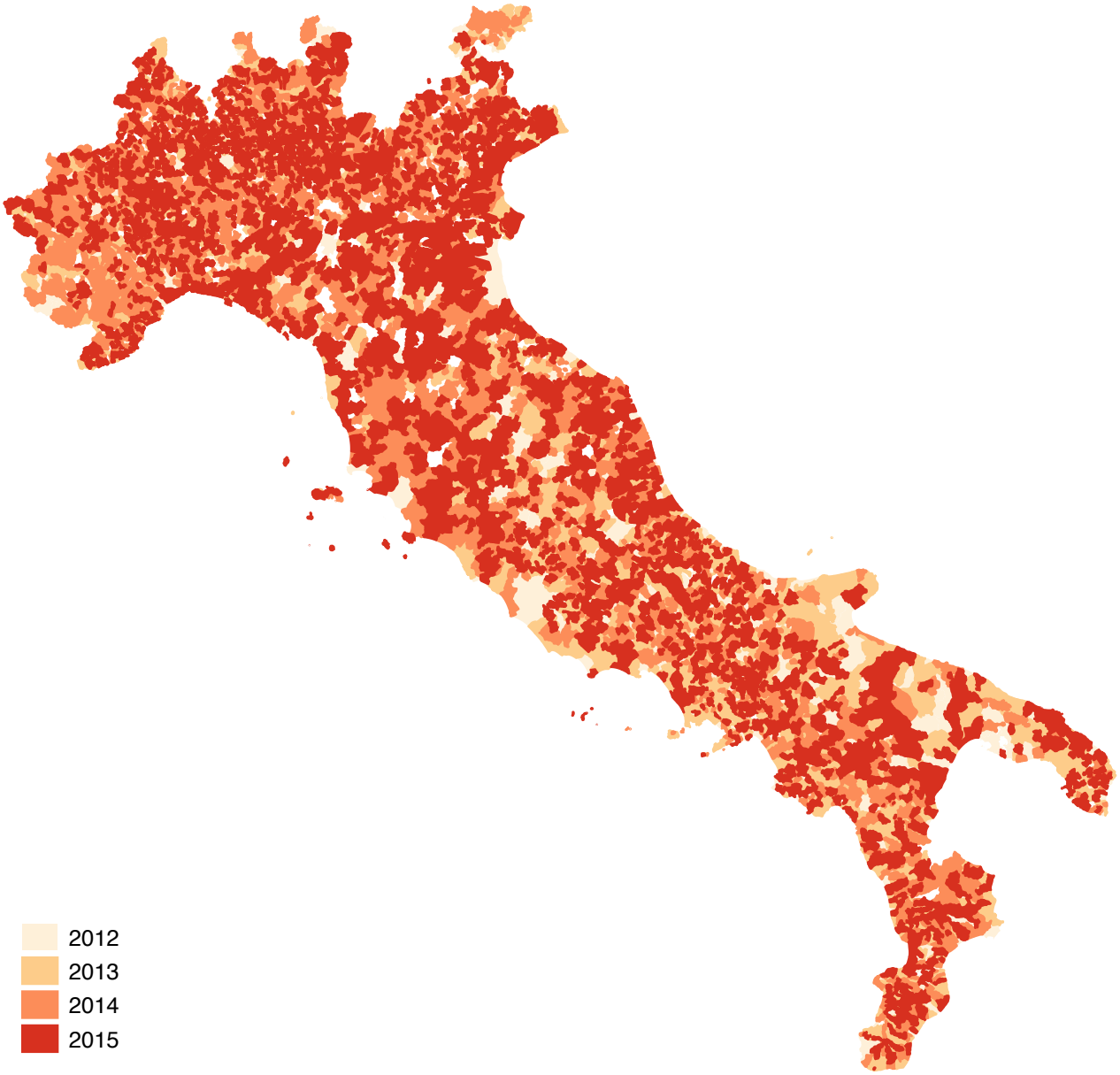
Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. All dependent variables are in per capita terms. $DebtRepayments$ and $OBSDebtRepayments$ are transformed using the inverse hyperbolic 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 initial age (in logs), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Table A19
The Effect of Random Auditor Assignment on Fiscal Sustainability, Deterrence and the Role of Spillovers

	Net Surplus		Debt Repayments		OBS Debt Repayments	
PANEL A: Estimating Spillover Effects						
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	11.01***	10.34***	0.0653***	0.0631***	-0.0208	-0.0220
Neighbor × Post	[3.848]	[3.843]	[0.0132]	[0.0131]	[0.0243]	[0.0243]
Controls	No	Yes	No	Yes	No	Yes
Dep. Var Mean	-218.5	-218.5	63.47	63.47	2.882	2.882
Observations	83424	83424	83424	83424	83424	83424
Adj. R-sq	0.699	0.699	0.769	0.769	0.405	0.405
Panel B: Accounting for Spillover Effects						
	(1)	(2)	(3)	(4)	(5)	(6)
Treated	30.94***	29.47***	0.155***	0.151***	-0.00813	-0.0110
× Post	[4.929]	[4.921]	[0.0182]	[0.0182]	[0.0362]	[0.0363]
Controls	No	Yes	No	Yes	No	Yes
Dep. Var Mean	-218.5	-218.5	63.47	63.47	2.882	2.882
Observations	62263	62263	62263	62263	62263	62263
Adj. R-sq	0.700	0.701	0.768	0.768	0.401	0.401

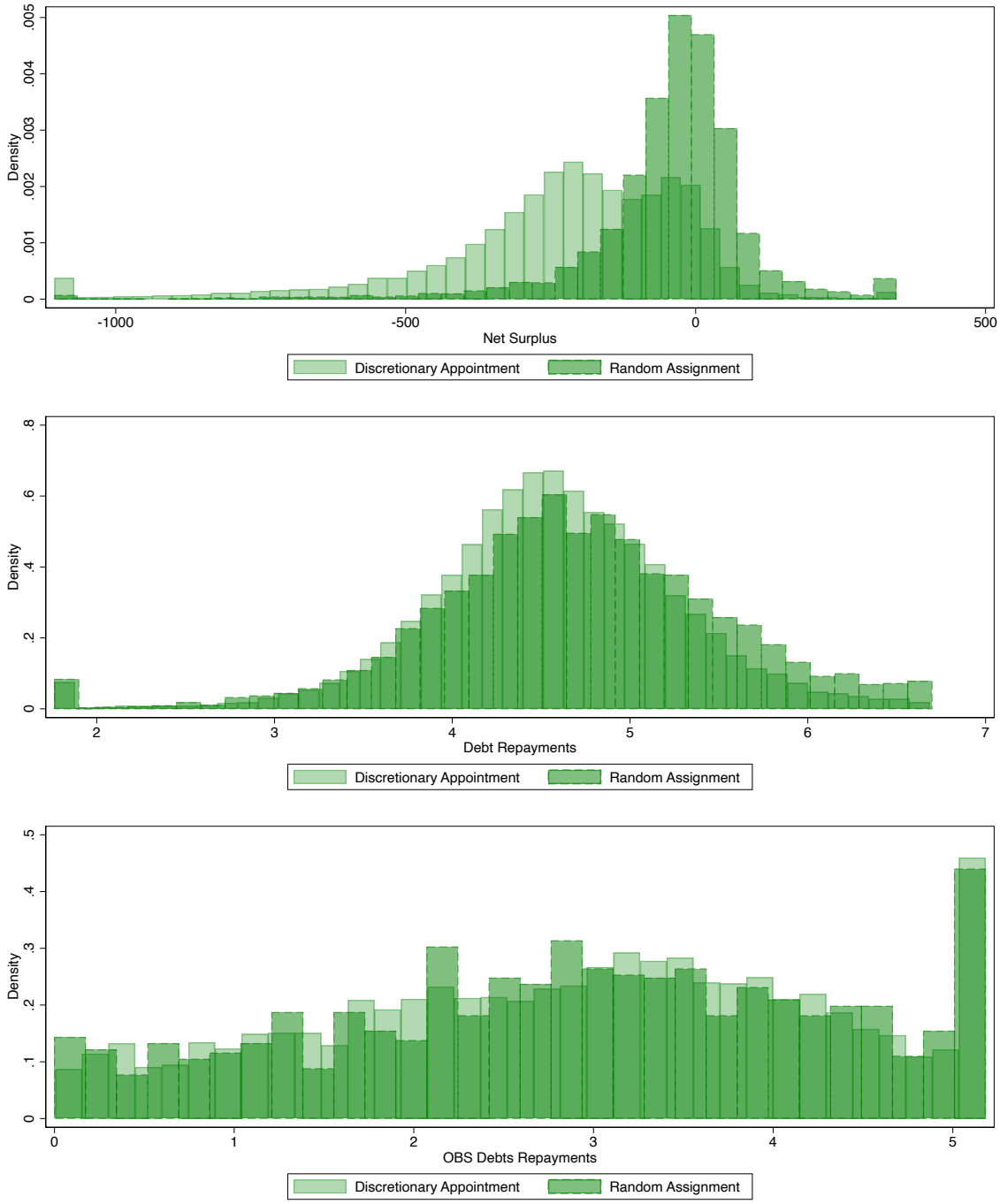
Notes: Significance: * significant at 10%; ** significant at 5%; *** significant at 1%. Standard errors clustered at the municipality level are reported in parentheses. All dependent variables are in per capita terms. *DebtRepayments* and *OBSDebtRepayments* are transformed using the inverse hyperbolic sine transformation, but their non-transformed mean is reported at the bottom of the table. All regressions include municipality, relative event-time fixed effects, election cycle fixed effects, population bins-times-year fixed effects, and the following controls: mayor's initial age (in logs), gender, term in office and a dummy equal to one if the mayor was born in the municipality. *Neighbor* is a dummy equal to 1 if a municipality is not yet treated but shares a border with an already treated municipality. Panel A excludes treated municipalities from the sample. Panel B excludes neighboring municipalities from the sample.

Figure A1
Staggered Treatment, Geographic Variation



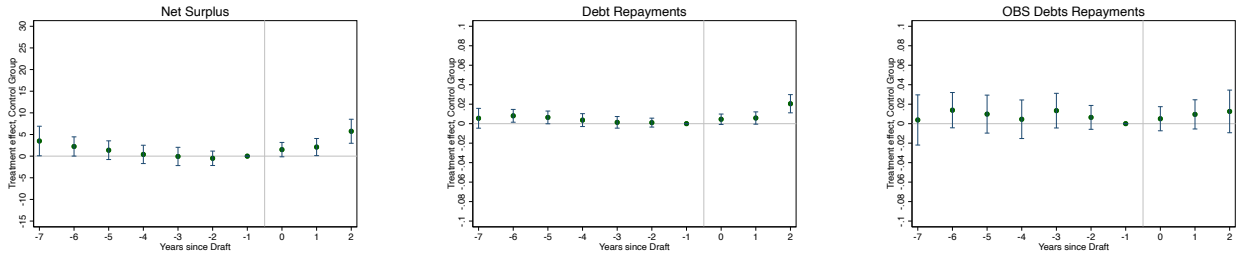
Notes: The figure shows the geographic variation in treatment timing. Darker gradation reflects later treatment timing.

Figure A2
Distribution of Outcomes, by Treatment Status



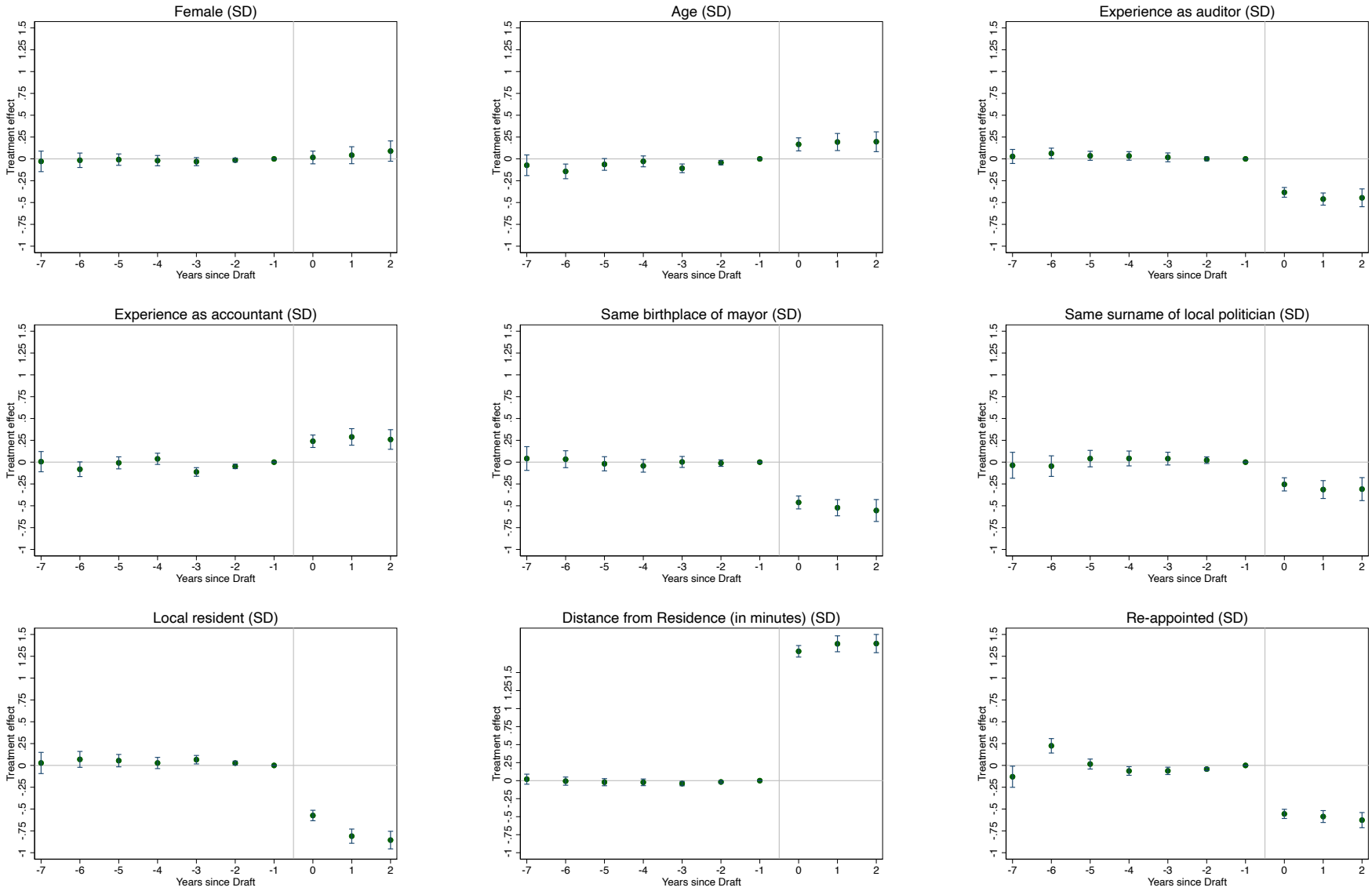
Notes: The figure shows histograms of the three outcome variables, $NetSurplus$, $DebtRepayments$ and $OBSDebtRepayments$, by treatment status. All dependent variables are in per capita terms, $DebtRepayments$ and $OBSDebtRepayments$ are transformed using the inverse hyperbolic sine transformation. For the bottom panel, the graph only shows the distribution of the variable for municipalities that have positive $OBSDebtRepayments$, which happens in less than 20% of the sample.

Figure A3
The Dynamic Effect of Random Auditor Assignment on Fiscal Sustainability, Anticipation Effects in the Control Group



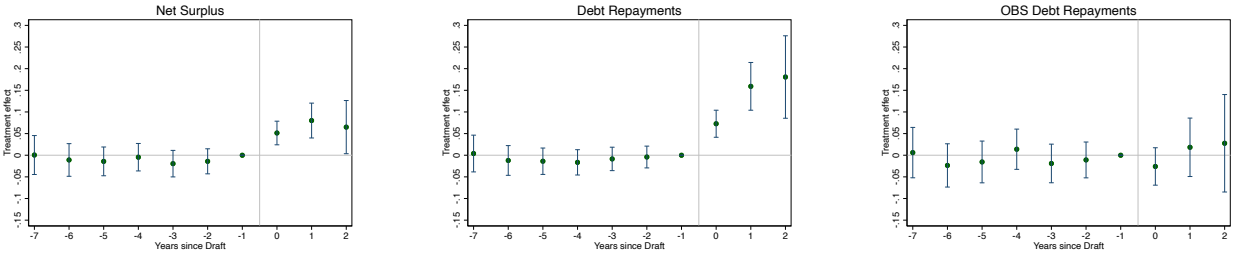
Notes: The graphs report the β_k coefficients and 95% confidence intervals estimated according to Equation 2. Standard errors clustered at the municipality level. All dependent variables are in per capita terms, $DebtRepayments$ and $OBSDebtRepayments$ are transformed using the inverse hyperbolic sine transformation. All regressions include municipality, population-bins-by-year fixed effects, and election cycle fixed effects, as well as the following controls: mayor's age (in logs) gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Figure A4
The Effect of Random Auditor Assignment on the Selection and Allocation of Auditors, Event-Study Estimates



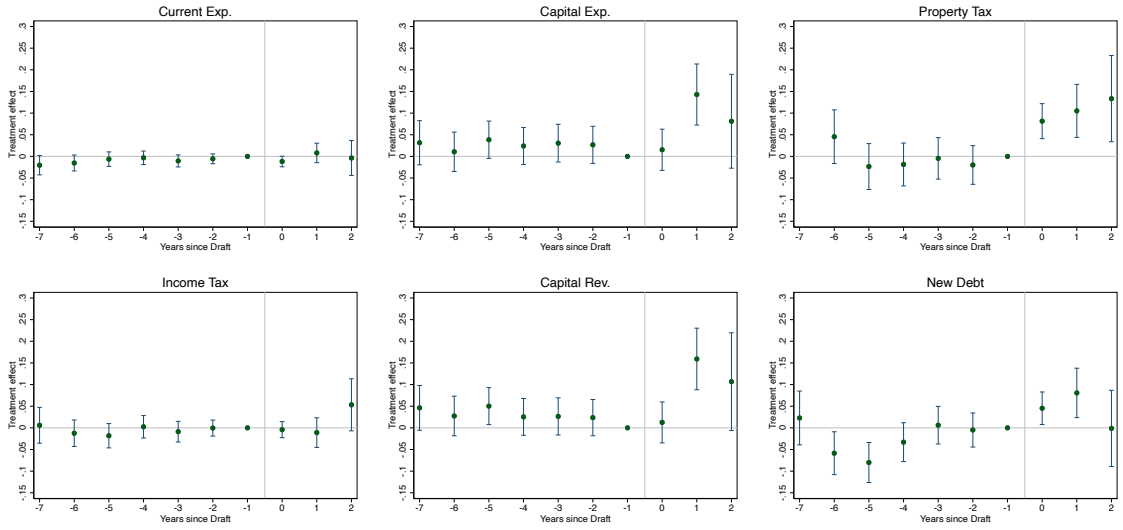
Notes: The graphs report coefficients and 95% confidence intervals estimated according to Equation 2. Standard errors are clustered at the municipality level. All dependent variables are standardized. All regressions include municipality, population-bins-by-year fixed effects, relative time fixed effects 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.

Figure A5
The Dynamic Effect of Random Auditor Assignment on Fiscal Sustainability, Standardized Outcomes



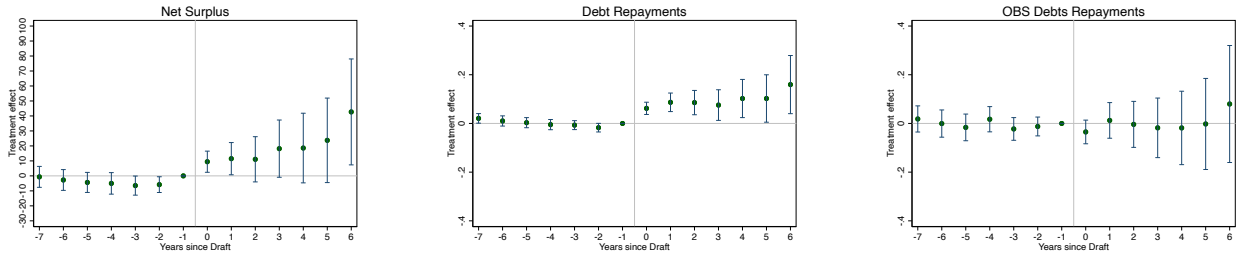
Notes: The graphs report coefficients and 95% confidence intervals estimated according to Equation 2. Standard errors clustered at the municipality level. All dependent variables are in per capita terms, $DebtRepayments$ and $OBSDebtRepayments$, transformed using the inverse hyperbolic sine transformation and standardized. 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 (in logs) gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Figure A6
The Dynamic Effect of Random Auditor Assignment on Aggregate Spending and Revenue Choices, Standardized Outcomes



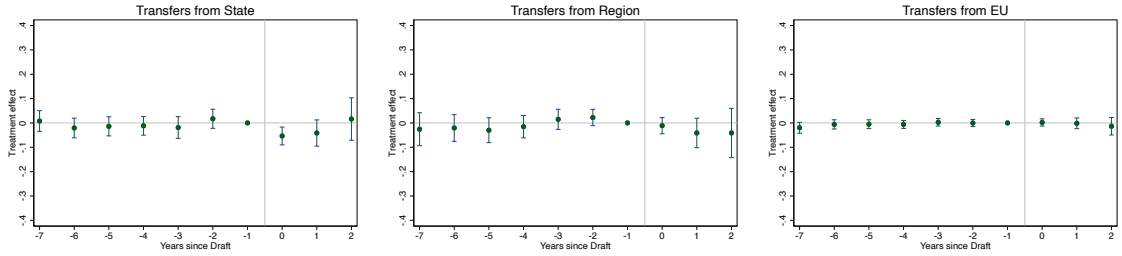
Notes: The graphs report coefficients and confidence intervals estimated according to Equation 2. Standard errors clustered at the municipality level. All dependent variables are in per capita terms and transformed using the inverse hyperbolic sine transformation. All regressions include municipality, year, and election cycle fixed effects, and the following controls: mayor's initial age (in logs), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Figure A7
The Dynamic Effect of Random Auditor Assignment on Fiscal Sustainability, Persistent Effects in the Long Run



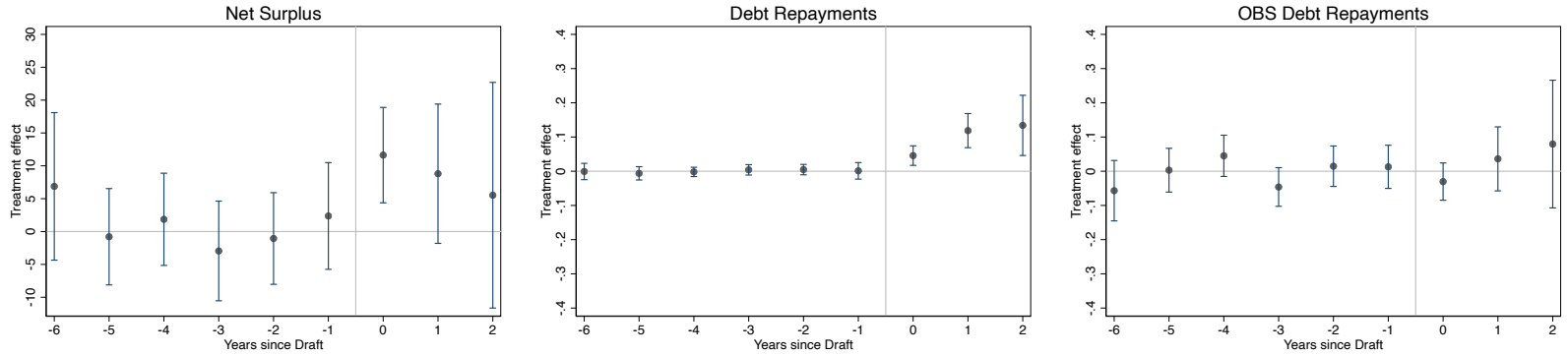
Notes: The graphs report the β_k coefficients and 95% confidence intervals estimated according to Equation 4. Standard errors clustered at the municipality level. All dependent variables are in per capita terms, $DebtRepayments$ and $OBSDebtRepayments$ are transformed using the inverse hyperbolic sine transformation. All regressions include municipality, population-bins-by-year fixed effects, and election cycle fixed effects, as well as the following controls: mayor's age (in logs) gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Figure A8
The Dynamic Effect of Random Auditor Assignment on Transfers from Other Levels of Government



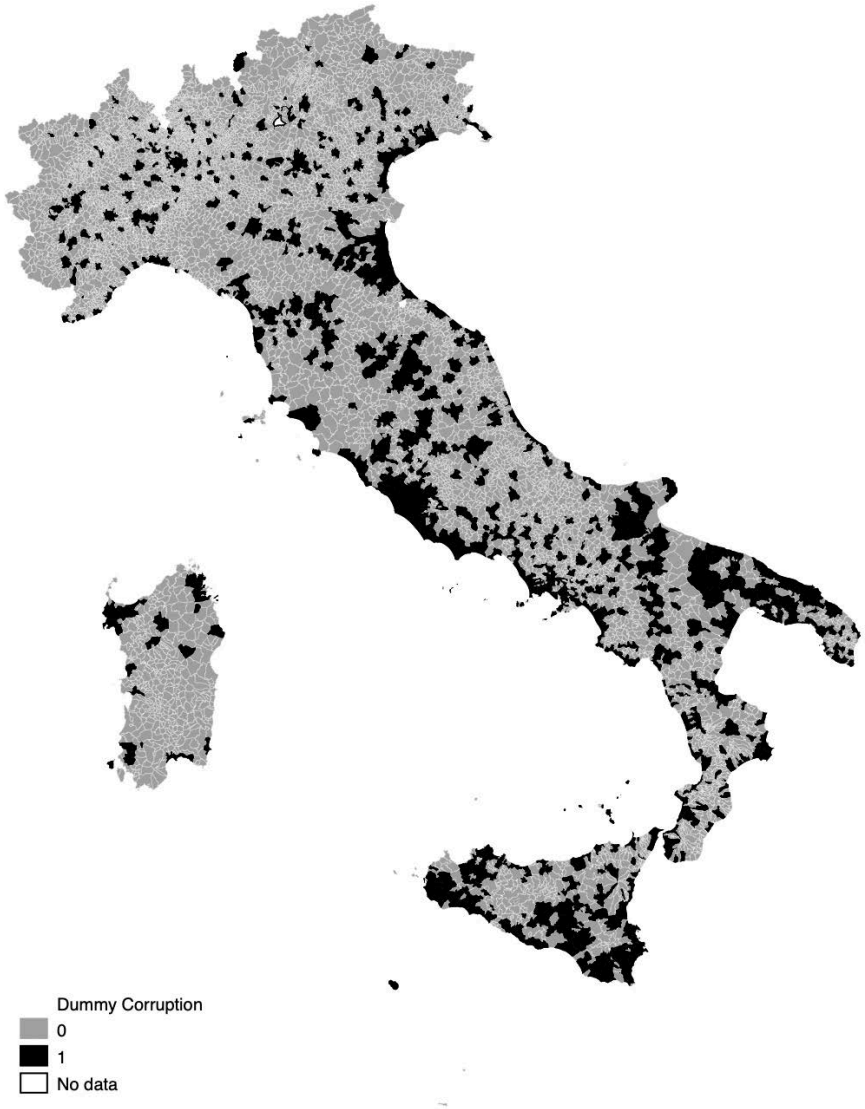
Notes: The graphs report coefficients and confidence intervals estimated according to Equation 2. Standard errors clustered at the municipality level. All dependent variables are in per capita terms and transformed using the inverse hyperbolic sine transformation. All regressions include municipality, population-bins-by-year fixed effects, region-by-year fixed effect, relative time fixed effects, and election cycle fixed effects, as well as the following controls: mayor's age (in logs) gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Figure A9
The Dynamic Effect of Random Auditor Assignment on Fiscal Sustainability, using de Chaisemartin and D'Haultfœuille (2020) Methodology



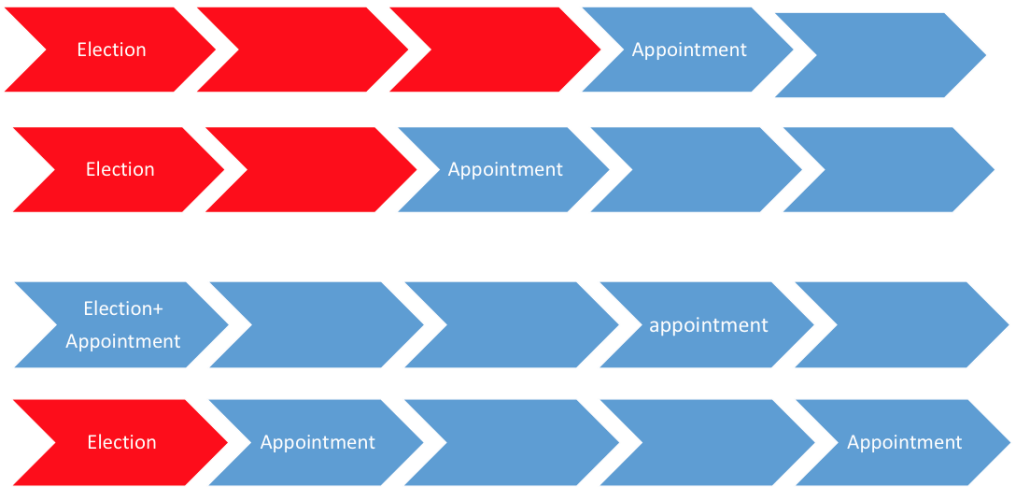
Notes: The graphs report coefficients and confidence intervals of the DID_M estimators estimated according to de Chaisemartin and D'Haultfœuille (2020) methodology using the did_multipleGT command in Stata. Standard errors clustered at the municipality level. All dependent variables are in per capita terms. *DebtRepayments* and *OBSDebtRepayments* are transformed using the inverse hyperbolic sine transformation. All regressions include municipality, election cycle fixed effects, population bins-times-year fixed effects, and the following controls: mayor's initial age (in logs), gender, term in office and a dummy equal to one if the mayor was born in the municipality.

Figure A10
Municipal Corruption Flags



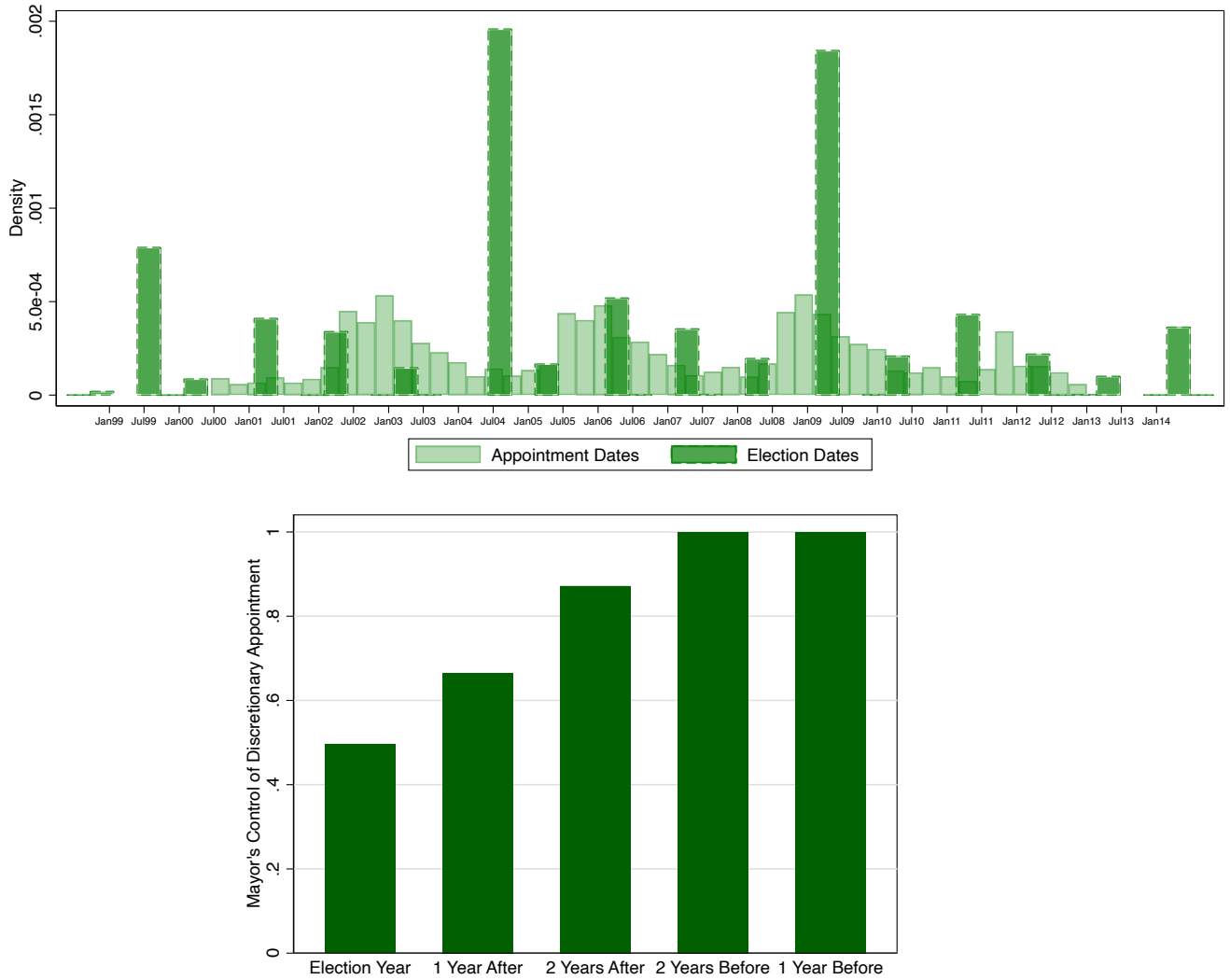
Notes: The figure displays a map of the municipalities having the *Corrupt* dummy equal to 1. To construct this variable, I exploit restricted-access data from the *Sistema D'Indagine Interforze* (SDI), a centralized investigation archive that contains reports of all individuals investigated by any of the Italian police forces. The final data contains information, for each municipality, on the total number of investigations for all types of crimes in the years 2004-2013. I construct an indicator variable flagging whether, in any given municipality, there was at least one investigation for corruption-related crimes in the entire period.

Figure A11
Auditors' Appointment Control, Example



Notes: The figure shows an example of how the interplay between the audit and election cycle across municipalities creates variation in the control of appointment by the mayor. Each of the rows represents a different municipality. In the first case (top row), the auditor was just appointed when the election takes place, so the newly elected mayor gets to appoint the auditor only after 3 years. In the third row, the opposite happens: the auditor cycle ends right after the election, and so the mayor gets to appoint a new auditor immediately after being elected.

Figure A12
Auditors' Appointment Dates and Municipalities' Election Cycles



Notes: The top panel of the figure displays the distribution of auditors' appointment dates in light green, as well as election dates in dark green. The bottom panel of the figure shows the share of mayors who had control of the auditor's appointment, by the moment of the election cycle. This is the variation used for identification in Table A15. Municipalities are on different electoral cycles as well as auditors' cycles for historical reasons, unrelated to the reform.

A2 Additional Details on Data and Sample Construction

I supplement my main data with further data sources to account for municipal characteristics that might affect fiscal sustainability. First, I use detailed data on the universe of municipal elections provided by the Ministry of the Interior, including information about election dates, results, and the mayor's characteristics (gender, age, and education). Local electoral data allows me to control for election cycle fixed effects as well as the mayor's term. Second, I include data on the universe of municipal council dismissals, which have been put together by a non-profit research entity, OpenPolis, by digitizing each of the dismissals' decrees. Italian municipal councils can be dismissed for several different reasons, the predominant one being mafia infiltration, but also for a failure to approve the budget. Council dismissals have proven to affect economic activity (Fenizia et al. 2020), and this is likely to be reflected in the municipal budgets. While I prefer not to include this as a control in my main specifications, as council dismissal could in principle be itself an outcome, all results are robust to controlling for or excluding dismissed municipalities (165 municipalities, which amount to less than 3% of my sample). Finally, I use data on municipal geographic and socio-demographic characteristics from the Italian Statistical Census (ISTAT). My analysis sample is based on the 6,627 municipalities located in Ordinary regions. To keep a balanced panel during the analysis period, I exclude all those municipalities that were dissolved or newly created between 2008 and 2015. Furthermore, I exclude 265 municipalities that were affected by earthquakes during the analysis period, as they were granted several fiscal exceptions following the catastrophic events. I also exclude 341 municipalities that voluntarily participated in the experimental phase, between 2013 and 2015, of the so-called "harmonized accounting system" that became binding from 2016 onwards. This new system significantly changed both the structure of the financial reports as well as the set of fiscal rules. The final sample consists of 5603 municipalities observed between 2007 and 2015.

A3 Alternative Estimators

To corroborate my analysis, I apply the alternative estimator recently proposed by de Chaisemartin and D'Haultfœuille (2020), which is robust to negative-weighting issues.[52] I start by estimating a "naive" two-way fixed effects model outlined in Equation of the following form:

$$Y_{mt} = \alpha_m + \delta_t + \beta Treated_{mt} + X'_{mt}\zeta + \epsilon_{mt} \tag{5}$$

[52]For further details, see, in particular, section 5.2 of de Chaisemartin and D'Haultfœuille (2020).

Results are displayed in Table A7. Reassuringly, the estimates are all very similar to the ones in Table 3. I then estimate the weights attached to each of the average treatment effects (ATTs) to compute the overall β_{fe} estimate.[53] Results are displayed in the bottom part of Table A7. The β_{fe} coefficients in Table A7 are obtained as a weighted average of 10,747 ATTs, of which 3175 receive a negative weight, but overall the negative weights only sum up to -0.106, suggesting that the relative importance of ATTs receiving negative weights is limited. To have a better sense of whether this represents a problem, one can look at the two diagnostic measures to assess the robustness of the $\hat{\beta_{fe}}$ estimate to treatment effect heterogeneity. The first one $\underline{\sigma}_{fe}$, corresponds to the ratio between the $\hat{\beta}_{fe}$ and the standard deviation of the weights, which is a proxy of the (unobserved) degree of heterogeneity in ATEs across treated groups and periods and reflects the minimal value under which it would be possible to have a β_{fe} which is of the opposite sign of the true ATT. The second one,$\underline{\underline{\sigma}}_{fe}$, is a proxy of the minimal amount of treatment effect heterogeneity under which it would be possible to obtain a β_{fe} which is of the opposite sign of all the ATEs. This second measure is defined only if at least one of the weights is negative, otherwise, it would be impossible to obtain a β_{fe}, which is a weighted average of all the ATEs, of opposite sign to them. Furthermore, across all outcomes, $\underline{\sigma}_{fe}$, which reflects the minimal value under which it would be possible to have a β_{fe} which is of the opposite sign of the true ATT, is as large as the $\hat{\beta}_{fe}$ estimate per se, thus implying that a substantial amount of treatment effect heterogeneity across $municipalityXyear$ cells would be required to invalidate the naive estimates. I then estimate the alternative Wald estimator proposed by de Chaisemartin and D'Haultfœuille (2020), the $DIDm$ estimator, which is robust to treatment effect heterogeneity across groups and periods. In my case, the $DIDm$ is estimated only by comparing "joiners", i.e. units whose treatment status changes between periods, to untreated stable units, i.e. units who remain untreated between periods.[54] Similarly, one can also compute $DIDm$ placebo estimators by looking at outcomes' evolution in pre-treatment periods, as well as the different dynamic treatment effects over time. I present results for the full dynamic specifications in Appendix Figure A9. Results look very similar to the ones in Figure 2.

As an additional robustness test, in Table A8 I also repeat my analysis using the estimator developed by Callaway and Sant'Anna (2021). Their approach amounts to estimate all the different group-time average treatment effects on the treated, $ATT(g,t) = E[Y_{it}(g) - Y_{it}(\infty)|G_i = g]$, which gives the average treatment effect at time t for the cohort first treated in time g. Under standard assumptions of parallel trends and no anticipation,

[53] I use the most recent versions of *twowayfeweights* and *did_multiplegt* commands, developed by the authors, and available in STATA repository.

[54] So, for example, to estimate the treatment effect for the first year of treatment (t = 0), the $DIDm$ compares the evolution of outcomes between $t-1$ and t for units that become treated at time t to units that are not yet treated in t.

each of the $ATT(g,t)$ can be identified by comparing the expected change in outcome for cohort g between periods g-1 and t to that for a control group not-yet treated at time t., i.e. any group g' such that $g' > t$, or an average over some set of comparisons G such that $g' > t$ for all $g' \in G$. Callaway and Sant'Anna (2021) consider two possibilities for G, one using only never-treated units and the second uses all not-yet-treated units. In my case, I can only use the second option, given the absence of a never-treated group in my sample. In this case, this estimator is analogous to the one proposed by Sun and Abraham (2021).