STOCKHOLM SCHOOL OF ECONOMICS
Department of Economics
5350 Master's Thesis in Economics
Spring 2025

Do Gender Quotas Reshape Local Government? Evidence from France's 2014 Parity Extension

Luca Clara Dakhil (student id. 42861) and Nana Kajaia (student id. 42868)

Abstract: Gender quotas are widely studied as an indirect lever for public integrity and resource allocation, yet evidence on governance outcomes remains mixed across contexts, and effects on corruption risk and spending composition are rarely examined jointly. To address said gap in research, this thesis examines the effects of the extension of France's gender parity *zipper quota* in 2014 on both corruption and public spending composition. This policy mandated alternation between male and female candidates on electoral lists in communes of at least 1,000 inhabitants. Using corruption risk indicators and eight predefined operating expenditure accounts as outcome variables, we run sharp and fuzzy regression discontinuity designs at the population threshold. First, the reform raises women's share of council seats by five to six percentage points at the threshold but has no detectable effect on procurement corruption risk. Second, of eight expenditure accounts, six show no response, including school supplies, library media, and social subsidies, categories often associated with women's policy priorities. While the zipper quota effects on elected officials' indemnities and service contracts were statistically significant, their economic magnitude was negligible. These outcomes remain robust across placebo cutoff tests and various specifications of the regression model. Changes in council composition induced by the quota appear insufficient to reshape procurement integrity or expenditure allocation.

Keywords: Gender quotas; regression discontinuity; corruption risk; municipal expenditure; local government.

JEL: D72; D73; H72; H75; J16; K38

Supervisor: Jaakko Meriläinen
Date submitted: 18.05.2026
Date examined: 28.05.2026
Discussants: Maria Bohush and Kiryl Platonau
Examiner: Patrick Agte

Acknowledgments

We would like to express our sincere gratitude to everyone who contributed to the successful completion of this thesis. First and foremost, we would like to thank our thesis supervisor, Jaakko Meriläinen, for his guidance, insightful feedback, and encouragement throughout the research process.

We are also grateful to our fellow students and colleagues for the constructive discussions and shared experiences during our studies. In addition, we would like to thank our friends and families for their advice, late-night calls, cooking nights and encouragement throughout this journey. Nana Kajaia expresses her gratitude to the Swedish Institute for the Scholarship for Global Professionals, which made it possible for her to pursue her MSc studies.

Table of Contents

1. Introduction 1
2. Literature Review 3
3. Institutional Context: Zipper Quota in the French Municipal Elections 7
   3.1. Structure of French Municipal Government 7
   3.2. Gender Parity Legislation and the 2014 Reform 8
   3.3. Effects of the Quota on Female Representation and Governance 9
4. Data 10
   4.1. Dependent Variables: Corruption Risk and Communal Expenditures 10
   4.2. Treatment Assignment and Running Variable 13
   4.3. Municipal-Level Covariates 15
5. Empirical Method 16
   5.1. Main Specification: Sharp RDD 17
   5.2. Alternative Specification: Fuzzy RDD 18
   5.3. Identifying Assumptions 19
   5.4. Robustness Checks 19
6. Results 20
   6.1. Validity: Density Test and Covariate Balance 20
   6.2. Female Representation at the Threshold 24
   6.3. Impact of Zipper Quotas on Corruption Risk 25
   6.4. Impact of Zipper Quotas on Communal Expenditure Composition 28
   6.5. Robustness 30
7. Discussion 32
8. Conclusion 35
9. References 36

Appendices 42
   Appendix 1: Data Construction 42
   Appendix 2: Covariate Balances 58
   Appendix 3: Additional Specifications 62
   Appendix 4: Robustness Checks 64

Tables

Table 1. Descriptive statistics of the corruption risk index (CRI) in the post-treatment period....11

Table 2. Communal expenditure outcome variables....13

Table 3. Descriptive statistics of female representation in municipal councils across the 2008-2014 and 2014-2020 election periods....14

Table 4. Pre-treatment covariate balances at the 1,000-inhabitant threshold....23

Table 5. First stage results: female representation at threshold....25

Table 6. Sharp regression discontinuity results for corruption risk index....26

Table 7. Fuzzy regression discontinuity results for the corruption risk index analyzed at the commune–ruling period level....27

Table 8. Sharp regression discontinuity results on communal expenditure composition analyzed at the commune-year level....29

Figures

Figure 1. Population and zipper-quota treatment status across French communes....8

Figure 2. Density graphs and McCrary tests of the manipulation of the running variable in the CRI and expenditure samples....20

Figure 3. Regression discontinuity plots for female share in 2014 and quota bite across full municipal, CRI and expenditure samples....24

Figure 4. Sharp regression discontinuity plot for corruption risk index analyzed at the commune – ruling period level....26

Figure 5. Fuzzy Regression discontinuity plots for the corruption risk index analyzed at the commune-ruling period level....27

Figure 6. Sharp regression discontinuity plots for the selected communal expenditure categories analyzed at the commune-year level....28

1. Introduction

Women remain distinctly underrepresented in elected office worldwide. As of 1 January 2026, women hold 27.5% of national parliamentary seats and just 22.4% of cabinet positions heading ministries, with progress having slowed to its lowest rate in nearly a decade (Inter-Parliamentary Union, 2026). To narrow this gap, governments have increasingly turned to institutional remedies, and gender-parity quotas have emerged as the most widespread policy lever. More than 150 countries now apply some form of legislated or party-level gender quota, and the scope of these mandates has expanded steadily at the local-government level over the past two decades (UN Women, 2026).

Beyond their effect on female representation, quotas have motivated a substantial empirical literature asking whether female political representation matters not only for women's standing as citizens, but for the choices governments make and the integrity with which public resources are managed. The aggregate economic relevance of this governance is substantial: the World Economic Forum estimates that corruption alone absorbs more than 5% of global GDP each year, amounting to USD 2.6 trillion (2018). Complementing this, weaker governance entails not only direct losses but also distortions to the composition of public expenditure. As measured integrity deteriorates, public budgets are relocated away from socially productive uses such as education and toward infrastructure-heavy categories, which are particularly vulnerable to inflated contracts and ghost projects (Mauro, 1998; d'Agostino, Dunne, & Pieroni, 2016; Gamberoni, Garner, Giordano, & Lopez-Garcia, 2016). Integrity in public contracting and the structure of public spending are, therefore, two complementary facets of the same governance-efficiency question.

The empirical literature offers competing predictions for both governance dimensions. On the integrity side, one strand argues that increased female representation should lower corruption risk in public contracting, on the grounds that women in office tend to be less embedded in established patronage networks, often face stronger electoral sanctions for misconduct, and are perceived by voters as more honest (Dollar, Fisman, & Gatti, 2001; Swamy, Stephen, Lee, & Azfar, 2001; Goetz, 2007; Barnes & Beaulieu, 2018). Meanwhile, a more institutionalist view asserts that this association is conditional on context and tends to dissipate once women are absorbed into existing political networks or attain decision-making seniority comparable to incumbents (Esarey & Chirillo, 2013; Charron, Dahlström, Fazekas, & Lapuente, 2017). The same disagreement carries over to the structure of public spending, where a sizeable body of work documents that legislatures with higher female representation tilt budgets toward education, social services, and other gender-priority outlays (Chattopadhyay & Duflo, 2004; Clots-Figueras, 2011; Clayton & Zetterberg, 2018), whereas other studies find that increases in descriptive representation leave both the level and the composition of municipal expenditure largely unchanged in some settings (Hessami & Lopes de Fonseca, 2020).

Within this literature, the joint analysis of integrity and resource-allocation outcomes following an exogenous change in council composition has received limited attention. Local councils provide a particularly informative setting for studying these dynamics for several reasons. First, across many institutional contexts, municipalities constitute the tier of public administration most directly engaged in the provision of everyday public services, including primary schools, local roads and public spaces, sanitation, waste collection, social assistance and land-use planning. As these are the

domains through which citizens most tangibly experience the functioning of the state, the quality of municipal governance has broader implications for political trust and the perceived legitimacy of government. Secondly, local councils exercise substantial discretion over spending priorities, internal administration, and the awarding of public contracts, making them especially relevant for studying how changes in political representation affect both corruption risks and public resource allocation. Finally, these dynamics are likely to be particularly pronounced in smaller municipalities, where weaker administrative capacity and the relative absence of formal anti-corruption mechanisms increase the importance of council composition for governance outcomes (Schöberlein, 2019). At the local level, France provides a particularly suitable institutional setting for examining the impact of quotas on integrity and on the composition of public spending.

The reform central to the analysis is the electoral law of 2013, which took effect at the March 2014 municipal elections and extended the existing mandatory *zipper quota* system to municipalities with populations below 3,500 inhabitants, requiring electoral lists to alternate male and female candidates, to all communes with populations of 1,000 inhabitants or more[1]. Out of more than 36,000 communes in France, the policy extension increased coverage from 8.0% to 26.4% of all communes[2]. As treatment status is sharply determined by administrative population count, this generates a discontinuity exogenous to commune-level characteristics. This setting provides a credible basis for a regression discontinuity design: communes immediately above and below the population cutoff are comparable in their pre-treatment characteristics, treatment status is assigned mechanically by population, and the local average treatment effect of quota exposure can be identified under standard continuity assumptions.

We organize the analysis around two research questions:

i. Did the quota exposure reduce corruption risk in public procurement, and did the magnitude of any such effect scale with the actual change in council composition induced by the reform?
ii. Did the same policy-induced change in council composition shift the structure of municipal operating expenditure, focusing on spending categories the literature has linked to procurement activity and to priorities historically associated with female legislators?

Analyzed jointly, the two questions allow us to assess whether the parity policy introduced in municipal councils reshapes broader governance outcomes that span both the integrity of public spending and the way in which it is allocated.

The paper makes two concrete contributions to the literature. First, to our knowledge, it is the first to jointly analyze two dimensions of local governance, procurement integrity and the composition of operating expenditure, within a single regression discontinuity design, so that conclusions about governance quality and about resource allocation are derived from the same source of identifying

[1] *Loi n° 2013-403 du 17 mai 2013 relative à l'élection des conseillers départementaux, des conseillers municipaux et des conseillers communautaires, et modifiant le calendrier électoral*, Journal officiel de la République française, 18 May 2013, art. 29.

[2] In population terms, the reform expanded the scope of the quota system from municipalities covering approximately 41.5 million inhabitants (65.9% of total population) to those covering 53.6 million inhabitants (85.0%). Note that calculations are based on 2011 population census data (INSEE). Communes are classified by population size to eligibility thresholds: those above 3,500 inhabitants (scope of *loi sur la parité* of 2000) and those between 1,000 and 3,499 inhabitants (scope of law extension). Shares are calculated relative to the total of communes in Metropolitan France (excluding overseas territories).

variation. Second, it focuses on small communes, the part of the French municipal distribution that represents the weakest administrative oversight (AFA, 2022), a setting that has received limited attention in prior empirical work.

The remainder of the paper is structured as follows. Section 2 reviews the relevant literature, while Section 3 outlines the institutional context of the zipper quota in French municipal elections. Section 4 describes the data, including the construction of the main variables and the treatment assignment. Section 5 presents the empirical strategy, detailing both the sharp and fuzzy regression discontinuity designs, the identifying assumptions, and the robustness checks. Section 6 reports the main results, including validity tests, first-stage estimates, and the effects on corruption risk and public spending composition. Section 7 discusses the findings and their limitations, and Section 8 concludes.

2. Literature Review

This study draws on three strands of literature: (i) theoretical frameworks of political representation and policy outcomes, (ii) gender, corruption, and corruption risk in public procurement, and (iii) gender and public spending outcomes. Together, they suggest that the gender composition of political institutions can shape both the integrity of public decisions and the allocation of public resources. Gender quotas provide a source of plausibly exogenous variation in political representation, but evidence on their downstream effects on governance remains mixed. Recent work has refined measures of corruption risk in procurement. Yet, the empirical link between council composition and procurement-based risk indicators remains underexplored, especially alongside broader patterns of public spending.

Theoretical Framework: Representation and Policy Outcomes

We use political economy theories of representation as a foundational step in hypothesizing whether a policy-induced change in council gender composition can drive substantive outcomes, both in the integrity of procurement decisions and in the public spending composition. To motivate whether a change in female representation at the council level translates into changes in the policies and practices municipalities pursue, we draw on two complementary areas of political-economy theory: the median-voter framework and the citizen-candidate model. Together they bracket the range of plausible answers to our two research questions, from one in which the gender composition of the council is largely irrelevant for procurement integrity and spending choices, to one in which it is a primary determinant of both.

The median mandate tradition holds that policy in representative democracies converges toward the preferences of the pivotal voter, with the legislature translating those preferences into public policy (Downs, 1957). Cross-national evidence finds that the median voter's position is reasonably well represented overall, with distortions from electoral institutions tending to offset one another across stages and over time (Best, Burge, & McDonald, 2012; McDonald, Mendes, & Budge, 2004). Under this view, what matters is the legislature's overall alignment with median preferences rather than who individual politicians are. Thus, an exogenous change in the descriptive composition of a council need not shift policy if the median voter's position remains unchanged. Applied to our context, this benchmark anchors a null hypothesis for both research questions: the quota could

increase women's representation on the council without producing observable shifts in corruption risk or in expenditure allocation.

The citizen candidate model offers a contrasting prediction. Osborne and Slivinski (1996) develop a framework in which citizens themselves decide whether to run for office, candidates cannot credibly commit to policies other than their most preferred ones, and the winner implements their own ideal point. Building on the assumption that the descriptive characteristics of those who enter and win elections, including gender, shape their personal preferences, the composition of elected bodies may in turn influence substantive policy outcomes. Aytimur, Boukouras, and Schwager (2016) extend this logic to settings where the implemented policy reflects a compromise with unelected actors, showing that strategic delegation has limits even when entry costs are negligible. Replacing male councilors with female councilors who hold different preferences over public spending and procurement can therefore move outcomes, particularly where the new councilors gain real influence. In our setting, this makes it plausible that quota-induced changes in council composition could affect both corruption risk and the structure of public spending, the outcomes analyzed under our two research questions.

The trade-off between representation and governance further conditions these predictions. Brooks, Phillips, and Sinitsyn (2011) define this trade-off as one between representation, the breadth of preferences reflected in decision-making, and governance, the responsiveness of an institution to a pivotal voter. They show empirically that more representative governments tend to move policy away from the governance ideal. A quota reform that alters council gender composition changes the breadth of preferences represented, which can plausibly affect both procurement integrity and the allocation of operating expenditures[3].

These frameworks do not predict a single outcome. They instead bound the empirical question and identify the conditions under which a quota-induced change in council composition is likely to translate into observable shifts in governance outcomes. Our empirical analysis tests where the French case falls within those bounds.

Gender, Corruption and Corruption Risk in Public Procurement

Within these theoretical frames, one way a change in council composition can shape outcomes is through politicians' willingness to engage in or tolerate corruption. A large body of research documents a negative correlation between women's political representation and corruption. Early cross-country studies found that higher shares of women in political office are associated with lower levels of corruption (Dollar, Fisman, & Gatti, 2001; Swamy, Stephen, Lee, & Azfar, 2001; Wängnerud, 2015; Hessami & Lopes de Fonseca, 2020). Subsequent work confirms that this relationship holds across different levels of government, including local councils (Merkle, 2022; Kumar & Sarangi, 2018; Bauhr, Charron, & Wängnerud, 2019; Pereira & Fernandez-Vazquez, 2022). However, this relationship is consistently shown to be context-dependent, varying with

[3] A related literature asks whether the size of the legislature itself affects governance outcomes. Pettersson-Lidbom (2012), exploiting discontinuities in council size laws in Finland and Sweden, finds that larger legislatures are associated with smaller government, consistent with more legislators being better able to control a budget-maximizing bureaucracy. In the French setting examined here, council size is also a function of population but does not exhibit a discontinuity at the threshold relevant to our identification strategy. Descriptive statistics on council composition nonetheless confirm that council size grows with population.

institutional quality, accountability structures, and the level of government considered (Esarey & Chirillo, 2013; Merkle, 2022). Evidence from Europe further points to substantial subnational variation in both corruption and female political participation, underscoring the importance of examining these relationships at the local level (Wängnerud, 2015).

Explanations for this relationship remain mixed. Some contributions argue that women may be less likely to engage in corrupt activities or are more risk-averse, while others caution against this interpretation and emphasize the role of institutional context (Esarey & Chirillo, 2013; Charron, Dahlström, Fazekas, & Lapuente, 2017). Alternative accounts focus on structural mechanisms, suggesting that increased female representation may disrupt established patronage networks or alter patterns of political competition (Goetz, 2007; Bauhr & Charron, 2021). Women in politics may also face stronger electoral sanctions for misconduct and are often perceived as more honest by voters, potentially shaping incentives around corruption (Barnes & Beaulieu, 2014; Barnes & Beaulieu, 2018).

When analyzing corruption on the governance level, public procurement is one of the principal channels through which these dynamics become observable. Procurement processes are particularly vulnerable to irregularities such as restricted competition and procedural manipulation and have therefore become a key focus of empirical measurement (Fazekas, 2025; Fazekas & Kocsis, 2015). Rather than observing corruption directly, recent work relies on risk indicators derived from procurement data to capture patterns consistent with favoritism or reduced competition. Institutional features, including bureaucratic independence and meritocratic recruitment, have been shown to reduce such risks in procurement processes (Charron, Dahlström, Fazekas, & Lapuente, 2017). Despite advances in measurement, evidence linking political representation to procurement-based risk indicators remains scarce and is concentrated on mayors in larger municipalities rather than on council-level composition in small communes, the segment most exposed to weak oversight.

The literature, therefore, documents a recurring negative correlation between women's political representation and corruption, but differs on the underlying mechanism and offers limited causal evidence at the council level. The empirical work closest to ours, which uses procurement-based corruption risk indicators, has focused on mayors in larger municipalities rather than on the composition of councils in small communes, where oversight is weakest. This motivates our first research question: whether quota exposure reduced corruption risk in small French communes.

Gender and public spending outcomes

Beyond corruption, female political representation has also been studied through its effect on policy choice and the allocation of public expenditure, where gender quotas provide the most direct empirical entry point. Empirical studies consistently show that quota reforms increase female representation in elected office and can influence policy priorities, particularly in areas such as health and social spending (Clayton & Zetterberg, 2018; Hessami & Lopes de Fonseca, 2020; Ennser-Jedenastik, 2017). Evidence at the municipal level suggests that quotas can also affect budget composition and spending allocation (Androeli, Manzoni, & Margotti, 2022). For instance, Profeta and Accettura (2022) find that male mayors in Italian municipalities engage in more strategic pre-electoral spending on visible policies, such as road construction and maintenance, than female mayors. Meanwhile, Svaleryd (2009) finds that increased female representation in

Swedish municipalities is associated with higher spending on childcare compared to elderly care and on education compared to elderly care, but has no significant effect on the balance between childcare and education.

Closer to the present design, Dominici, Cavallini, and Masi (2025) study the 2014 Italian reform that mandated 40% gender quotas on the executive councils of municipalities above 3,000 inhabitants. Using a difference-in-discontinuities approach, they document that the policy raised female representation and reallocated municipal expenditure across functions, raising spending shares on preschools and schools while reducing the share devoted to culture, without altering total expenditure. Their results echo the broader pattern of quota-induced shifts in public spending compositions.

However, not all studies find meaningful downstream effects of quotas on policy outcomes. Using a regression discontinuity design in Spanish municipal elections, Bagues and Campa (2021) show that gender quotas in candidate lists increase women's representation among council members but do not lead to significant changes in politicians' quality, women's access to leadership positions, or the size and composition of municipal budgets. In an earlier Italian setting, Rigon and Tanzi (2011) instrument female council representation with a short-lived 1993–1995 gender quota in candidate lists and find no significant effect on the allocation of municipal expenditure across spending categories, a result they interpret as consistent with the median-voter logic that the policy preferences of female politicians may be close to those of their male counterparts.

Taken together, these findings suggest that while quotas can alter descriptive representation, they may not necessarily translate into substantive policy change, especially when women do not attain influential positions or fail to reach a critical mass (2006). The contrast between reallocation findings and null results maps onto the two theoretical frameworks set out earlier. Where descriptive representation rises, but substantive influence does not, the median voter prediction of policy stability fits the data. Where female councilors acquire real decision-making power and hold preferences that differ from those of their male counterparts, the citizen candidate logic of preference-driven policy change becomes more plausible.

Two features of this evidence shape our second research question. First, the spending categories most consistently linked to female representation, such as education, social services, and other gender priority outlays, provide a foundation for selecting the operating expenditure accounts we examine. Second, the contrast between reallocation results and null findings emphasizes that quota effects on budget composition are context-dependent and need not materialize when descriptive representation rises without a corresponding shift in decision-making influence. As public procurement is the substantial channel through which much of this spending is executed, the integrity and allocation of municipal expenditure are best examined jointly rather than in isolation, motivating the design adopted in this paper.

The three strands together reveal three persistent gaps that mirror the order of our analysis. First, causal evidence linking changes in council composition to procurement-based corruption risk remains scarce, with the existing French evidence focused on mayoral gender in larger communes. Second, the consequences of council composition for the size and composition of public spending have been studied largely in isolation from procurement integrity, despite both reflecting decisions taken by the same municipal council. Third, small municipalities, where administrative oversight is

weakest, have received limited empirical attention. This study is designed to address each of these gaps.

3. Institutional Context: Zipper Quota in the French Municipal Elections

The following section provides essential information on the French governance environment and context for the study introduced in the research questions.

3.1. Structure of French Municipal Government

France is among the most centralized states in Western Europe, with subnational authority organized into a three-tier hierarchy of *collectivités territoriales* below the national government. Mainland France is divided into 12 metropolitan regions, 96 departments, and 34,746 communes as of 1 January 2026 (INSEE, 2026). Regions handle economic development, vocational training, and upper-secondary schools (*lycées*) (Vie publique, 2025). Departments oversee social and welfare allowances, lower-secondary schools (*collèges*), and departmental roads (Vie publique, 2024). Communes are responsible for primary schools, local roads, water and sanitation, waste collection, urban planning, and social action (Vie publique, 2025).

France grants municipal councils (*conseils municipaux*) substantial authority over local public spending, including the award of public procurement contracts. Under the *Code général des collectivités territoriales*[4], the council and its executive are legally responsible for approving public contracts and may exercise broad authority over their preparation, award, execution, and settlement (CGCT, art. L2122-21; art. L2122-22)[5]. Thus, making French communes a directly relevant unit of analysis for studying governance quality and corruption risk in public procurement.

However, the formal concentration of procurement authority at the municipal level does not translate uniformly into effective oversight across communes. Administrative capacity varies sharply with municipality size: larger communes tend to maintain dedicated procurement units, internal audit functions, and compliance officers, while smaller communes often rely on generalist staff with limited expertise in public contracting (European Commission, 2016). Evidence from the French Anti-Corruption Agency surveys corroborates this pattern: formal anti-corruption mechanisms, including codes of conduct, risk mapping, and whistleblowing systems, are almost entirely absent in smaller municipalities (AFA, 2018). Awareness and reporting of corruption cases increase markedly with municipality size, with municipalities above 10,000 inhabitants being over four times more likely to report known cases than those below 3,500 (AFA, 2022). This institutional heterogeneity is directly relevant to the present study: municipalities closest to the 1,000-inhabitant threshold sit at the lower end of administrative capacity, making them both more susceptible to procurement irregularities and more sensitive to changes in council composition.

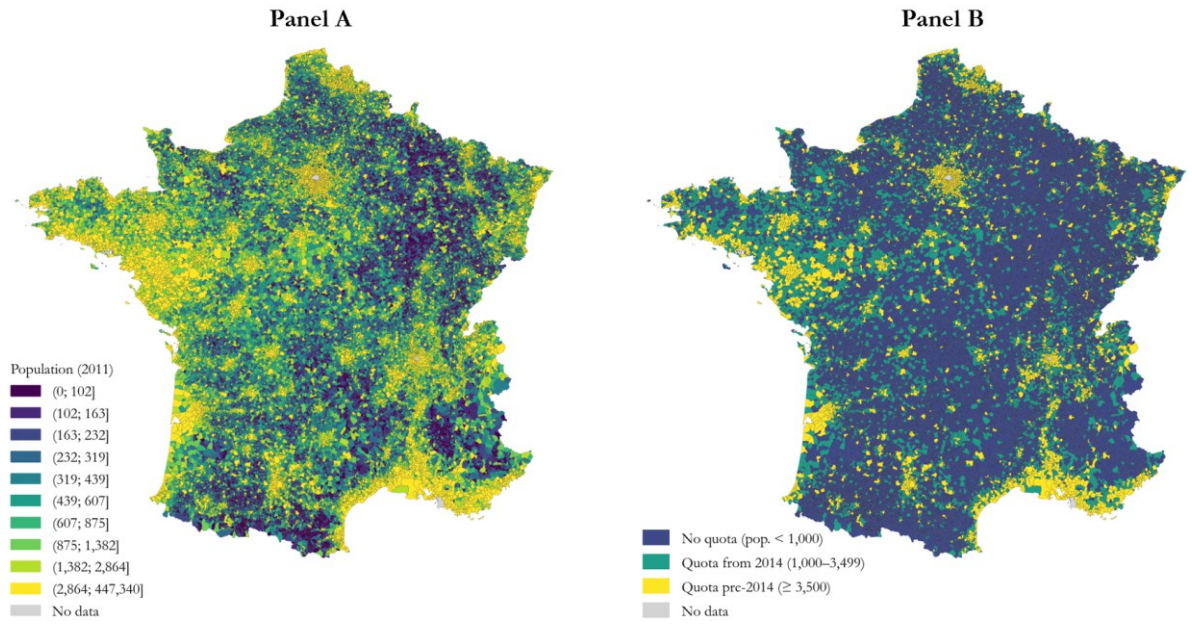
[4] *Code général des collectivités territoriales* [Legislative part promulgated on May 30, 1996, regulatory part on May 19, 2000], [consolidated version as of January 1, 2026].

[5] In addition to the council, the mayor also exercises discretion over award decisions, supplier relationships and procedure choice.

3.2. Gender Parity Legislation and the 2014 Reform

France has progressively expanded gender parity requirements in elected office through a series of legislative reforms. The foundational *loi sur la parité* of 2000 introduced parity obligations for candidate lists in municipalities above 3,500 inhabitants[6], requiring that lists alternate male and female candidates, commonly referred to as the "zipper" system (Lépinard & Lieber, 2015). Enforcement was backed by financial penalties: political parties failing to comply with parity requirements for legislative elections faced reduced public funding, establishing a credible compliance incentive.

Figure 1. Population and zipper-quota treatment status across French communes



The figure maps the 36,000-odd communes of metropolitan France. Panel A shades each commune by its population based on the 2011 census (enforced on 31st December 2013), grouped into deciles (ten equal-count bins). Each bin represents the interval of the number of inhabitants; the legend reports the population value at each decile cut. Panel B classifies communes according to their exposure to the 2014 parity ("zipper") quota for municipal council elections: communes with fewer than 1,000 inhabitants are not subject to the quota and serve as the control group (blue); communes with population between 1,000 and 3,499 were brought under the quota by the 2014 reform of the electoral code (green); communes with population of 3,500 or above had already been subject to the quota since the 2000 loi sur la parité (yellow). Population thresholds reflect the legal cutoffs as of the 2014 reform. Overseas departments and territories are excluded. Communes for which population data could not be matched due to commune mergers to commune boundaries are shown in light grey ("no data"). Sources: author's calculations using INSEE Recensement de la population (2011) and commune boundaries from IGN admin-express.

The reform central to this study was introduced through the 2013 electoral law, which took effect at the municipal elections[7] in March 2014. This legislation extended the mandatory zipper system to all communes with populations of 1,000 inhabitants or more, substantially broadening the scope of parity obligations from 8.0% to 26.4% of all communes in Metropolitan France[8](see Figure 1

[6] *Loi n° 2000-493 du 6 juin 2000 tendant à favoriser l'égal accès des femmes et des hommes aux mandats électoraux et fonctions électives,* Journal officiel de la République française, 7 Jun. 2000.

[7] The election cycles for municipal elections are six years for the office of councilors, the mayor and their deputies. The election cycles relevant to our analysis are 2008-2014 (pre-reform extension) and 2014-2020 (post-reform extension).

[8] Authors' calculation based on 2011 population census data (INSEE). Communes are classified by population size to eligibility thresholds: those above 3,500 inhabitants (scope of *loi sur la parité* of 2000) and those between 1,000 and

for the geographic distribution of population and quota exposure across communes). The extension also entailed related changes to the electoral procedure for newly covered communes, which can be understood as a necessary institutional precondition for the parity requirement: the zipper rule is only enforceable when elections are conducted based on pre-ranked closed lists rather than individual candidates. Under this system, local party branches or candidate coalitions compile and submit ranked candidate lists to electoral authorities prior to the election, with a parity requirement mandating the male-female alternation on these lists. Communes below the 1,000-inhabitant threshold were not subject to this requirement, creating a sharp institutional discontinuity in electoral rules.

The scheduled extension of parity list requirements to all communes from the 2026 municipal elections further underscores the policy's broader significance (Service Public, 2025). As mandatory gender parity becomes universal across French local government, understanding the governance implications of increased female representation becomes increasingly policy relevant. The present study contributes to this timely debate by providing causal evidence ahead of the reform's full implementation.

3.3. Effects of the Quota on Female Representation and Governance

Empirical evidence confirms that the zipper quota produced a meaningful increase in female political representation at the local level. Using a regression discontinuity design, Lassebié (2019) documents a substantial increase in the share of female candidates and elected councilors in treated municipalities. Notably, the reform had little effect on the socioeconomic, age, or experience composition of councils, suggesting that numerical representation increased without substantially altering the profiles of those entering local politics. These shared socioeconomic backgrounds may limit the extent to which increased female representation disrupts existing informal networks within local government. Nevertheless, their increased presence on councils plausibly altered deliberative dynamics and oversight capacity within municipal government.

Evidence from the French context suggests this may have governance implications, though existing work has focused predominantly on larger communes. Bauhr and Charron (2021), drawing on data from French municipal elections combined with procurement-based corruption risk indicators covering the period 2005–2016, find that newly elected women mayors are associated with lower corruption risks in public procurement. However, this effect dissipates upon re-election, with gender differences becoming negligible in municipalities where women mayors serve subsequent terms. Their analysis is conducted on larger communes, using the gender of the mayor as the key running variable, and does not extend to the effects of broader council-level gender composition in smaller municipalities.

The present study addresses this gap by examining communes near the 1,000-inhabitant threshold, largely absent from existing empirical work, focusing on the effect of an institutional policy shock directed at council composition rather than analyzing the impacts of mayoral gender composition. Studying smaller communes is analytically distinctive: these municipalities operate with weaker

3,499 inhabitants (scope of law extension). Shares are calculated relative to the total number of communes in Metropolitan France (excluding overseas territories), yielding 8.0% and 18.4% respectively.

institutional safeguards, lower administrative capacity, and less external scrutiny, conditions under which changes in council composition may carry particular weight for governance outcomes.

4. Data

Building on the institutional setting described in the previous section and data quality being essential to conducting a causal policy analysis, this section discusses the sources, construction, and limitations of the data underlying the empirical analysis. The discussion begins with the corruption risk and expenditure outcomes, proceeds to the running variable and treatment assignment, and finally to the pre-treatment municipal-level covariates used in the balance tests.

4.1. Dependent Variables: Corruption Risk and Communal Expenditures

Corruption Risk

The corruption risk outcome variables are derived from the Global Public Procurement Dataset (GPPD), compiled by the Government Transparency Institute (2024). The GPPD harmonizes contract-level procurement records across countries[9] and constructs standardized corruption risk indicators at the tender level. Following the framework developed by Fazekas and Kocsis (2015), corruption risks in procurement are proxied through a set of "red flag" indicators capturing restrictions on competition and irregular procurement patterns, including single-bidder tenders, non-open procedures, and unusually short submission or decision periods. The authors aggregate these indicators into a composite corruption risk index (CRI)[10] that summarizes the extent to which procurement processes exhibit characteristics associated with favoritism or manipulation.

Although these indicators do not measure corruption directly, they provide a consistent behavioral proxy for corruption risk in government contracting and have been widely validated in empirical research (Fazekas & Kocsis, 2020; Broms, Dahlström, & Fazekas, 2019; Bauhr, Czibnik, de Fine Licht, & Fazekas, 2019; Coviello & Mariniello, 2014). Relative to perception-based indices[11], procurement-based indicators offer several advantages: they rely on administrative transaction data rather than subjective assessments, allow corruption risks to be measured at a granular level and aggregated over time, and are directly observable in the domain through which a substantial share of municipal public resources is allocated (Fazekas, 2025).

The dataset contains more than three million contract-level observations for France covering the period 2006–2020, enabling corruption risk indicators to be observed both before and after the 2014 reform. Because the empirical analysis focuses on corruption risks at the municipal level, procurement contracts are linked to communes through a multi-stage cleaning and geocoding

[9] Procurement data are generally more systematically collected and publicly available at higher levels of government, particularly at the national level, while subnational procurement data remain uneven and fragmented across many countries. This limitation is especially pronounced in the Global South, where corruption risks in public procurement are often considered more severe, yet administrative capacity, digitization, and transparency standards frequently constrain the availability and comparability of procurement records.

[10] The CRI itself was constructed by calculating the average over all available indicators (Fazekas & Kocsis, 2015).

[11] The most widely used perception-based measure is Transparency International's Corruption Perceptions Index (CPI), which aggregates expert assessments and business surveys to rank 180 countries according to perceived levels of public-sector corruption.

procedure[12]. The final matched dataset covers 14,259 communes, representing 39.0% of all communes in mainland France and approximately 2.27 million procurement observations, following the removal of exact duplicates and the exclusion of overseas departments. For the regression discontinuity analysis, communes affected by post-2014 mergers, splits, or *commune nouvelle* reorganizations are subsequently excluded so that each commune is observed under a fixed treatment status with respect to the 1,000-inhabitant threshold around which identification is constructed. This reduces the analysis sample by 602 communes (8,757 contract-level observations). Further filtering of the dataset to communes with inhabitants below 3,500 narrows down the analysis-eligible sample to 10,535 unique communes, which covers 33.6% of all relevant communes[13]. Contract-level corruption indicators are then aggregated to the commune–ruling period level to construct the municipal outcome measures used in the empirical analysis.

Table 1. Descriptive statistics of the corruption risk index (CRI) in the post-treatment period

Corruption risk index	Treatment = 0	Treatment = 1	Total
Mean	0.473	0.441	0.458
Median	0.500	0.500	0.500
Standard deviation	0.080	0.101	0.092
Min.	0.000	0.000	0.000
Max.	0.900	0.875	0.900
Number of observations	3,676	3,471	7,147
Frequency	51.43%	48.57%	100.00%

Descriptive statistics for the CRI in the post-treatment period, grouped by treatment assignment. Treatment equals 1 if a municipality exceeds the population threshold of 1,000 inhabitants. The CRI ranges from 0 to 1, with higher values indicating greater corruption risk. Frequencies denote the share of observations in each group. Contract-level corruption indicators are aggregated to the commune–ruling period level to construct the municipal outcome measures used in the empirical analysis.

Table 1 reports descriptive statistics for the post-policy sample of the corruption risk index cross-tabulated by treatment assignment. The sample is relatively balanced across treatment groups, with 51.4% of observations in the control group and 48.6% in the treated group. The average corruption risk is marginally higher in the control group than in the treated group, and the dispersion of CRI is somewhat larger among treated units, indicating slightly higher heterogeneity across the total treatment group.

Three identified features of the procurement data: (i) sparse tender coverage among small communes, (ii) the tendency of larger communes to file more detailed and more frequent tenders, and (iii) the resulting asymmetric coverage above and below the 1,000-inhabitant threshold, would obstruct identification under cross-sectional or panel comparisons of communes of different sizes. They are addressed in the present analysis using the regression discontinuity design, which identifies the local average treatment effect at the cutoff and requires only that observability and outcomes

[12] Location information in the raw dataset is inconsistently reported across establishment identifiers, postal codes, and free-text city names. To address this, each procurement observation is assigned to an official INSEE commune code, prioritizing establishment identifiers linked to the SIRENE business registry (2026), followed by deterministic matching based on postal codes and standardized municipality names. For further information, please refer to Appendix 1.1.

[13] For further overview of sample representativeness, please refer to Appendix 1.3. Sample Representativeness, Procurement-based Corruption Risk Sample.

vary smoothly with the population around the threshold. Detailed coverage statistics and sample representativeness checks are reported in Appendix 1.

Communal Expenditures

The outcome variables for the local public spending analysis are constructed from the *balances comptables des communes*, the annual municipal budget-execution files published by the *Direction générale des Finances publiques* (DGFiP) of the French Ministry of Finance. These files record account-level debit and credit flows for the universe of French communes under four nationally regulated public-accounting nomenclatures (M14 and M14A[14] for ordinary and small communes; M57 and M57A[15] as their modernized successors). This provides an almost universal (99.5%) coverage of eligible communes for the analysis[16]. The analysis focuses on Class 6 accounts, which capture operating expenditures, and is restricted to communes' main budgets (*budgets principaux*) to abstract from annex and auxiliary budgets that are unevenly distributed across communes. For each commune–year and each account, realized expenditure is measured as the budgetary debit flow net of cancellations less the corresponding net credit flow, and net of internal "budgetary order" operations. The realized flow is then aggregated to selected two-, three-, and four-digit account groups defined by the M14 / M57 hierarchies[17] and expressed both as a share of total Class 6 spending in the same commune–year and on a per-capita basis using the legally binding 2011 census population.

Eight accounts are retained as outcome variables for the regression discontinuity analysis as presented in Table 2. The selection is anchored in the political-economy literature on female political representation and the level and composition of public spending, as well as in the link between operating expenditure and procurement-related corruption risk. Two accounts capture the political rent and procurement mechanisms studied here. Elected officials' indemnities and allowances (account 653) cover the indemnities, mission costs, training, and pension contributions of municipal councilors. We interpret 653 as a measure of legal rent extraction, discretionary expenditure that the council sets, in part, for itself, and include it as a primary outcome under the hypothesis that gender composition may shift the size and use of these in-office benefits. Service contracts (account 611) capture the operating-expenditure counterpart to the external-procurement records used to construct the corruption-risk indicators, thereby directly linking the expenditure analysis to the procurement-corruption analysis.

Three further accounts are included as gender-priority spending categories based on a literature linking the gender composition of legislatures to higher spending on social and human-capital infrastructure: library books (6065) and school supplies (6067), which proxy investment in education-related public goods, and operating subsidies granted (657), which fund associations and inter-municipal bodies delivering locally provided social services (Chattopadhyay & Duflo, 2004; Clots-Figueras, 2011; Clayton & Zetterberg, 2018; Hessami & Lopes de Fonseca, 2020).

[14] *L'instruction M14 - Comptabilité des communes, Gérer les finances publiques locales, Direction générale des Finances publiques, Le portail des collectivités locales.*
[15] *Le référentiel budgétaire et comptable M57, Gérer les finances publiques locales, Direction générale des Finances publiques, Le portail des collectivités locales.*
[16] For further overview of sample representativeness, please refer to Appendix 1.3. Sample Representativeness, Communal Expenditure Sample.
[17] For further information, please refer to Appendix 1.2.

Maintenance and repairs (615) is a second procurement-related outcome covering buildings, roads, networks, and movable property; existing work documents that female legislators tend to substitute away from visible infrastructure spending toward less salient social spending (Clayton & Zetterberg, 2018; Hessami & Lopes de Fonseca, 2020), which yields a non-positive prediction for this category. Employer social security contributions (645) are included as a mechanical baseline, since their level is set by employer obligations rather than by political discretion. Non-storable supplies (6061; water, electricity, heating) are included as a falsification test, since spending on these utility flows is largely fixed by infrastructure stock and consumption patterns rather than by political choices.

Table 2. Communal expenditure outcome variables

Code	Description	Motivation
653	Elected officials' indemnities and allowances	Legal rent extraction
611	Service contracts	Procurement related
615	Maintenance and repairs	Literature indicates gender-based differences in spending
645	Employer social-security contributions	Mechanical baseline
657	Operating subsidies granted	Literature indicates gender-based differences in spending
6061	Non-storable supplies (water, electricity, heating)	Falsification check
6065	Library books and media	Literature indicates gender-based differences in spending
6067	School supplies	Literature indicates gender-based differences in spending

All outcomes are constructed from Class 6 accounts in communes' main budgets (budgets principaux) under the M14, M14A, M57, and M57A nomenclatures. For each outcome, the analysis reports both a share: The account's spending as a fraction of total Class 6 spending in the same commune–year and a per-capita measure normalized by 2011 INSEE population.

The dataset spans the years 2010–2019, covering four pre-reform and six post-reform years over which expenditure outcomes are reliably observed. The analytical sample is then restricted to communes within the [500, 1,500] population bandwidth used for the regression discontinuity design and to communes present in every year of the analysis period; 306 communes affected by post-2014 mergers, splits, or *commune nouvelle* reorganizations are excluded so that each commune is observed under a fixed treatment status throughout the panel. The resulting balanced panel covers 9,259 communes observed in each year, on which the eight expenditure variables play a dual role: in the pre-reform period (2010–2013), they serve as pre-treatment covariates in the balance tests, and in the post-reform period (2014–2019), they serve as outcomes in the regression discontinuity estimation. The full sample-construction pipeline, descriptive statistics and sample representativeness are reported in Appendix 1.

4.2. Treatment Assignment and Running Variable

Population

The running variable in the regression discontinuity design is the official municipal population recorded in the 2011 census, made legally binding on 31 December 2013 (INSEE, 2013). These

figures were collected through a rolling survey methodology conducted independently of and prior to policy considerations, providing no basis for manipulation by local actors around the threshold. Treatment status is accordingly defined as a binary indicator equal to one for communes with a population at or above 1,000 inhabitants[18].

Council Composition

To verify that crossing the threshold translated into actual changes in council composition, the analysis draws supplementarily on electoral data from the *Base Révisée des Élu·es de France* (BREF), a longitudinal administrative dataset covering elected representatives across the Fifth Republic (Février, Labatut, Volpi, & Marrel, 2025). Female representation is calculated as the share of women among elected municipal councilors at the start of the 2008 and 2014 electoral terms, respectively. Several additional variables are constructed to capture the intensity of the reform's impact at the commune level. The primary measure is a continuous quota bite, defined as the change in female representation between 2008 and 2014. The analytical sample is restricted to communes with populations below 3,500 inhabitants and to communes observed in both election years to ensure valid pre- and post-comparisons. This electoral data further serves as the basis for robustness checks examining whether estimated effects on procurement outcomes vary with the intensity of the reform's impact on council composition.

Table 3. Descriptive statistics of female representation in municipal councils across the 2008-2014 and 2014-2020 election periods

Variable	Control (pop<1,000)	Treated (pop≥1,000)	Difference
Number of communes	25,020	6,259	-
Avg. total councilors (2008)	11.435	17.331	5.896
Avg. total councilors (2014)	11.507	18.254	6.747
Avg. Female Share (2008)	0.308	0.355	0.047
Avg. Female Share (2014)	0.343	0.478	0.135
Avg. Quota Bite (2014–2008)	0.035	0.123	0.088
Positive Bite Incidence	0.515	0.873	0.358

The Table reports raw means for the 2014–2020 ruling period at the commune level for all communes affected by the zipper quota introduced in 2014 (i.e., communes with pop≤3500), excluding merged communes. Quota Bite represents the change in the share of female councilors between the 2008 and 2014 mandates, while positive bite incidence shows the share of communes where female representation increased in 2014 compared to 2008.

To provide context, we begin with descriptive evidence on treatment intensity, comparing communes above and below the threshold in terms of female representation and its evolution between the 2008 and 2014 elections. This step serves two purposes. First, it documents the extent to which treated communes experienced larger increases in female representation, thereby confirming that the reform had a substantive impact on council composition. Second, it motivates the use of a continuous treatment measure of quota bite in the subsequent fuzzy regression discontinuity design.

[18] For the overview of population distribution across samples of analysis, refer to Appendix 1.3.

The data presented in Table 3 confirms a substantial descriptive gap in gender composition between the two groups. In treated communes, the average share of female councilors reached 47.9%, compared to 34.3% in untreated communes. Treated communes saw an average increase of 12.3 percentage points, while untreated communes experienced a much smaller shift of 3.5 percentage points. Positive bite incidence was 35.8 percentage points higher in treated communes compared to the control group. Together, these patterns provide clear descriptive support for a strong first-stage effect of the reform at the threshold, indicating that crossing the population cutoff was associated with a markedly greater feminization of municipal councils.

4.3. Municipal-Level Covariates

The selection of pre-treatment covariates follows best practices in regression discontinuity designs, which recommend including variables that could plausibly affect the outcome and may be correlated with the running variable, in order to assess the smoothness assumption (Lee & Lemieux, 2010). Our covariates are sourced from five French administrative sources. Demographic and household variables are drawn from INSEE's population census,[19] which provides commune-level counts of adults, households, and property ownership. Income, fiscal capacity, and labor force measures are taken from the DGFiP fiscal records[20] and the INSEE labor force survey[21], which gives household-level mean income and the per-capita size of the active and salaried labor force. Pre-treatment municipal expenditure variables are constructed from the *balances comptables des communes*, which record annual operating expenditure by Class 6 chart-of-accounts category at the commune level. Firm-level variables are derived from the SIRENE business registry[22] and divided by population to obtain sectoral firm density (services, commerce, construction, industry) and the rate of new firm creation. Finally, the number of councilors is sourced from historical records of the *Répertoire National des Élus*, which present the full roster of locally elected officials in each commune.

We assess balance on a wide set of pre-treatment characteristics that fall into five substantive groups. First, population structure is captured by the per-capita number of adults, household density, and homeownership measured both as the per-capita count of owner-occupiers and as a share of dwellings. Second, local economic resources are summarized by mean household income, commune fiscal capacity, and the per capita size of the active and salaried labor force. Third, productive structure is proxied by per-capita establishment counts in services, commerce, construction, and industry, together with the rate of new firm creation. Fourth, to account for pre-existing differences in municipal spending patterns, pre-treatment expenditure composition is tested directly on the Class 6 expenditure accounts from the *balances comptables*, measured both as a share of total operating expenditure and as real per-capita spending. Finally, to test for pre-existing differences at the threshold for council size, we analyze the number of communal councilors on either side of the threshold.

---

[19] *Populations légales 2009*, *Institut national de la statistique et des études économiques (Insee), Recensement de la population*, published Dec. 31, 2011, enforced since Jan. 1, 2012.

[20] *Revenus fiscaux des ménages and the comptes des collectivités locales* and *Balances comptables des collectivités locales, Direction générale des Finances publiques (DGFiP).*

[21] *Enquête Emploi, Institut national de la statistique et des études économiques* (*Insee*), collected on an annual basis.

[22] *Système d'identification du répertoire des entreprises et des établissements (SIRENE)*, Insee, National registry continuously updated on a daily basis

The selection is motivated by a well-established body of cross-national evidence linking modernization indicators, including income, education, and economic development, to lower corruption levels (Holmberg & Rothstein, 2014; Bauhr & Charron, 2021). Homeownership is included as a proxy for community stability and civic engagement: DiPasquale and Glaeser (1999) show that homeowners participate more in local civic and political life, which may translate into greater electoral monitoring of local government and lower tolerance for integrity irregularities. The inclusion of female political representation (Sundström & Stockemer, 2015) is motivated by its established link to the level and composition of public spending (Borge & Rattsø, 2008; Hessami & Lopes de Fonseca, 2020).

The choice of expenditure accounts is further motivated by evidence on the spending priorities associated with female legislators in comparable institutional contexts. Studies of Nordic and Western European local governments find that higher female council representation is associated with increased spending on social services, childcare, and education, and reduced spending on physical infrastructure and general administration (Androeli, Manzoni, & Margotti, 2022; Ennser-Jedenastik, 2017; Cavallini, Dominici, & Masi, 2025; Hessami & Lopes de Fonseca, 2020). This account-level selection therefore, allows us to test whether the French reform produced spending shifts consistent with the patterns documented in the broader literature on female legislative representation.

As for council size, we test for pre-treatment differences to account for any existing issues caused by the potential impact of council size on equilibrium governance outcomes (Brooks, Phillips, & Sinitsyn, 2011; Pettersson-Lidbom, 2012). Analyzing balance for these variables helps ensure that any estimated discontinuity at the threshold is not confounded by underlying socioeconomic or fiscal differences across communes that independently predict both council composition, corruption and public expenditure outcomes.

5. Empirical Method

Having the data in place, this section sets out our research design. Estimating the causal effect of gender quota policies on governance outcomes is complicated by the potential endogeneity of female political representation, which may correlate with socioeconomic, cultural, and institutional characteristics that independently influence procurement outcomes. The 2014 zipper quota generates a single, sharp, deterministic discontinuity at the 1,000-inhabitant cutoff of legal population at the time of policy introduction. This discontinuity can, in principle, be exploited for both regression discontinuity design (RDD) and difference-in-discontinuities (DiDisc) design[23], utilizing variation across both the cutoff and the reform date (Grembi, Nannicini, & Troiano, 2016).

A DiDisc estimator is necessary for identification, however, only when a second policy or institutional feature switches at the same cutoff and risks confounding the treatment of interest. In our setting, no such confounder exists during the period of analysis. A review of national policy

[23] DiDisc design combines the logic of regression discontinuity and difference-in-differences approaches by comparing the size of a discontinuity at a policy threshold before and after a reform. The method is used in settings when other policies jump discontinuously at the same cutoff, thereby isolating the reform-induced effect from any other discontinuities around the threshold.

confirms that the principal post-2014 anti-corruption instruments[24] do not apply at the threshold of 1,000 inhabitants. The most natural mechanical confound is municipal council size, which is a step function of population; however, the bins do not fall around the 1,000-inhabitant threshold. The balance tests in Section 6.1 empirically confirm that council size and other observable covariates evolve smoothly across the cutoff. Therefore, a standard sharp RDD directly and more efficiently identifies the local effect of the quota than a DiDisc design.

5.1. Main Specification: Sharp RDD

Estimation throughout follows current best practice for local-polynomial RDD. Estimation is implemented in the rdrobust package, using a local linear estimator with a triangular kernel and a data-driven MSE-optimal bandwidth h (Calonico, Cattaneo, & Titiunik, 2014; Calonico, Cattaneo, & Farell, 2017). Following the most recent empirical recommendations (De Magalhães, et al., 2025), we report the conventional local-linear point estimate together with robust confidence intervals, standard errors, and p-values. The pilot bandwidth used for bias correction is set equal to the main bandwidth ($\rho \equiv h/b = 1$), so that bias estimation is performed on the same neighborhood as point estimation. The preferred specification uses a local linear estimator with mean-squared error optimal (MSE-optimal) bandwidth selection, following the data-driven procedure implemented in the rdrobust package in Stata and R. In panel specifications, standard errors are clustered at the commune level to account for within-commune correlation over time, whereas cross-sectional ruling-period specifications employ heteroskedasticity-robust standard errors.

Building on this motivation for the empirical method, we first specify the equation for the corruption risk. For the primary outcome variable, the baseline estimating equation takes the following form:

$$Y_i = \alpha + \tau D_i + \beta_1(X_i - c) + \beta_2 D_i(X_i - c) + \varepsilon_i \quad (1)$$

where Y_i is the corruption risk outcome in the commune i, α is capturing the intercept, $D_i = 1\{\text{Xi} \geq \text{c}\}$ the treatment indicator, equal to one if the commune's population X_i exceeds the threshold c of 1,000 inhabitants and zero otherwise. The term $(X_i - c)$ is the running variable centered at the cutoff[25]. β_1 measures the slope of the relationship between the running variable and the outcome for communes below the cutoff, while β_2 allows this slope to differ above the cutoff. ε_i is an error term capturing unobserved determinants of corruption risk.

The coefficient of interest τ captures the local average treatment effect at the threshold. This effect has a standard local average treatment effect (LATE) interpretation, as is identified from variation among communes near the cutoff, and does not generalize to the broader population of French municipalities. Notably, as the reform raised women's numerical share without essentially altering the socioeconomic, age, or experience profile of elected councilors (Lassébie, 2019), the

[24] Review covered information available through *Haute Autorité pour la transparence de la vie publique*; the French Anti-Corruption agency (AFA) guidelines; Law on transparency, the fight against corruption and the modernization of economic life (also known as Sapin II law)

[25] The running variable X_i(commune population) is centered at the cutoff c (1,000 inhabitants) so that the threshold corresponds to zero. This normalization is standard in regression discontinuity designs, as it allows the coefficient τ to be interpreted as the treatment effect exactly at the cutoff and simplifies the interpretation of the polynomial terms on either side of the threshold.

discontinuity isolates a change in gender representation from correlated shifts in councilor characteristics.

The estimating equation for the municipal-expenditure outcomes introduced in the data section follows the same sharp-RDD logic as for corruption. However, it takes advantage of the high time-frequency of the *balances comptables*: communes report their accounts annually, so the panel runs at the commune-year level over 2010–2019 rather than at the ruling-period level used for the procurement analysis. We exploit this time-variant data density by augmenting the baseline specification with year fixed effects.

For each chosen account j, we estimate the local treatment effect on two complementary outcomes: the share of total Class 6 operating expenditure in commune i devoted to account j (capturing expense distribution *across* categories) and real per-capita spending in euros (capturing changes in *levels*). The estimating equation extends the baseline RDD of the equation to the panel structure of the communal accounting balances:

$$Y_{i,t,j} = \alpha_j + \tau_j D_i + \beta_{1,j}(X_i - c) + \beta_{2,j} D_i (X_i - c) + \gamma_{t,j} + \varepsilon_{i,t,j} \quad (2)$$

where i indexes communes, t indexes years, j indexes the expenditure account, $Y_{i,t,j}$ is the expense j outcome variable. X_i is the 2011 INSEE census population, $c = 1{,}000$ is the cutoff, $D_i = \mathbf{1}\{X_i \geq c\}$ is treatment status, and $\gamma_{t,j}$ are year fixed effects estimated outcome-by-outcome. The coefficient of interest τ_j is identified from cross-sectional variation around the cutoff within year and is not contaminated by aggregate trends in municipal finances. Similar to the previous specification, the estimation strategy uses a linear polynomial, a triangular kernel and the MSE-optimal bandwidth selection, with standard errors clustered at the commune level[26].

5.2. Alternative Specification: Fuzzy RDD

A complementary approach is used to examine the role of actual changes in council composition in corruption risk outcomes. We use a fuzzy RDD specification, with the population threshold as an instrument for the magnitude of the change in female representation, as captured by the quota bite variable described in Section 4.2. This approach allows the reduced-form estimates to be examined in relation to the intensity of the reform's impact on council composition, providing additional insight into the mechanism through which the policy may have affected procurement outcomes.

The first stage estimates the discontinuity in treatment intensity at the cutoff:

$$quota_bite_i = \alpha + \pi D_i + \beta_1 (X_i - c) + \beta_2 D_i (X_i - c) + \varepsilon_i \quad (3)$$

where i indexes communes. $quota_bite_i$ measures treatment intensity, defined as the increase in the share of female councils after 2014 in the commune i. α is capturing the intercept. $D_i = \mathbf{1}\{X_i \geq c\}$ is an indicator equal to one if the commune's population X_i exceeds the cutoff $c = 1{,}000$, and

[26] Each commune contributes multiple observations across years and expenditure categories, which may be correlated due to persistent local characteristics such as administrative capacity, political leadership, fiscal preferences, revenue base and tax capacity, demographic composition, and long-standing institutional or budgeting practices.

zero otherwise. β_1 measures the slope of the relationship between population and treatment intensity below the threshold, while β_2 allows this slope to differ above the cutoff. ε_i is an error term. The coefficient π captures the discontinuous change in female representation induced by the reform.

The second stage estimates the causal effect of increased female representation on corruption risk:

$$Y_i = \alpha + \tau \widehat{quota_bite}_i + \beta_1(X_i - c) + \beta_2 D_i(X_i - c) + \varepsilon_i \quad (4)$$

where $\widehat{quota_bite}_i$ denotes the predicted value from the first stage. The coefficient of interest τ represents the local average treatment effect of an increase in female representation on corruption risk for municipalities at the threshold.

5.3. Identifying Assumptions

The validity of the RDD rests on two core assumptions (Cunningham, 2021). The first is the continuity assumption, which requires that potential outcomes evolve smoothly as a function of the running variable at the threshold, such that the only change occurring at the point of discontinuity is the shift in treatment status, with communes just below and just above the cutoff being essentially identical in all pre-determined characteristics in the absence of treatment. This assumption is assessed empirically through formal balance tests that estimate RDD models for each pre-treatment covariate using the optimal bandwidth and the same kernel as in the main specification, with estimated coefficients and confidence intervals reported in a covariate balance table presented in Section 6.1[27].

The second assumption is the absence of manipulation of the running variable. Although the 2011 census population figures were collected independently of and before the reform, this is nonetheless assessed empirically through a formal density test (McCrary, 2008) and a heaping test examining whether anomalous clustering around round numbers introduces measurement misspecification near the cutoff.

5.4. Robustness Checks

The robustness of the baseline estimates is assessed through a series of sensitivity and falsification checks, applied separately to each of the three sets of outcomes examined in this thesis: the first-stage discontinuities in female representation and quota bite, the corruption-risk results, and the expenditure outcomes.

First, alternative fixed bandwidths, set at 0.75 and 1.25 times the MSE-optimal bandwidth, are examined to verify that results are not sensitive to the bandwidth choice. Second, placebo cutoff tests are conducted by re-estimating the main specifications at artificial population thresholds of 750 and 1,250 inhabitants, where no institutional discontinuity exists. Under the null of no treatment effect at placebo thresholds, no significant discontinuity should be detected. Third, placebo pre-policy tests examine whether comparable discontinuities are detectable in the years preceding the 2014 reform, using the 2008 election cohort for the first stage and the 2011–2013 period for the corruption-risk and expenditure outcomes. The absence of a pre-reform

[27] For further information please refer to Appendix 2.

discontinuity supports the validity of the identifying assumption that communes on either side of the cutoff would have evolved similarly in the absence of the policy.

6. Results

The following chapter discusses the results of the validity checks, the covariate balances and the effectiveness of the quota on female share within the council. It continues with reporting the results of the models described in Section 5 and the robustness checks.

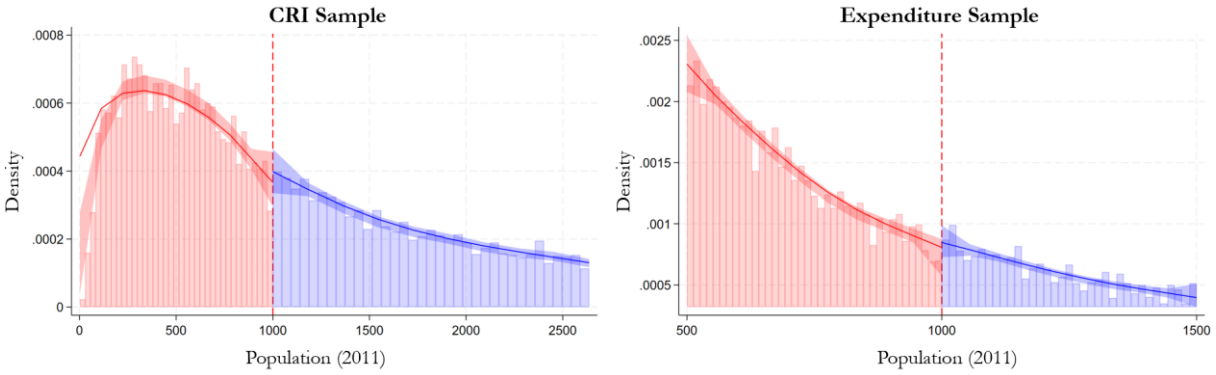
6.1. Validity: Density Test and Covariate Balance

Before presenting the main results, the validity of the regression discontinuity design is assessed through a series of manipulation and balance checks.

Density and Heaping Tests

A central concern in any regression discontinuity design is whether units can precisely manipulate their position relative to the threshold. In the present setting, treatment status is determined by the 2011 INSEE population count, which was fixed 3 years before the 2014 election to which the zipper quota first applied. Because the quota was introduced after these population counts had already been recorded, and because municipal populations evolve slowly and are not under the discretionary control of any individual actor, there is no plausible mechanism through which communes could have sorted themselves above or below the 1,000-inhabitant threshold in anticipation of the policy. Manipulation is therefore not expected on theoretical grounds. The assumption is nonetheless tested empirically through both visual and formal procedures.

Figure 2. Density graphs and McCrary tests of the manipulation of the running variable in the CRI and expenditure samples



The figure displays the distribution of communes around the population threshold (1,000 inhabitants) for both the CRI and expenditure samples. The solid lines represent kernel density estimates on either side of the cutoff, while the McCrary test assesses the continuity of the running variable at the threshold. The absence of a statistically significant discontinuity indicates no evidence of manipulation of the population variable, supporting the validity of the regression discontinuity design.

The density of the running variable is first visually inspected using histograms of the 2011 population, with narrow bin widths around the cutoff. This step is intended to detect anomalous spikes at round numbers, which would suggest rounding or reporting conventions rather than genuine population counts, as well as any visible discontinuity at the threshold itself. The histograms reveal a smooth density consistent with the natural distribution of commune sizes in

France, with no spikes at 1,000 inhabitants, no excess mass immediately to one side of the cutoff, and no evidence of heaping at round values.

Formally, we test this utilizing the McCrary (2008) density test, under which the null hypothesis of no manipulation implies a continuous density at the cutoff. The test is implemented separately for each of the two samples used in the main specifications: the corruption risk index sample of communes with procurement observations between 2014 and 2019, and the expenditure sample of communes with available budget data over the same period. The test yields p-values of 0.685 in the CRI sample and 0.220 in the expenditure sample, providing no evidence against the continuity of the density at the cutoff in either case (see Figure 2). Therefore, the formal manipulation test combined with the institutional features of the setting supports the assumption that communes did not sort around the 1,000-inhabitant threshold in a way that would invalidate the regression discontinuity design.

Covariate Balance

The identifying assumption of the research discontinuity design is that pre-treatment commune characteristics evolve smoothly with population and do not change discontinuously at the 1,000-inhabitant threshold. We assess this assumption by estimating a separate local-linear regression discontinuity for each pre-treatment covariate, using the same MSE-optimal bandwidth-selection procedure and triangular kernel similar to our main outcome specifications. Table 4 reports balance tests in two panels: socioeconomic and productive characteristics in Panel A, and municipal expenditure-related covariates from the accounting balances in Panel B.

Panel A shows that most of the selected demographic and economic variables relevant to outcome variables are all balanced at the cutoff, indicating that the observable local economic and productive base evolves smoothly across the 1,000-inhabitant threshold and that any estimated treatment effect cannot easily be attributed to differences in income, fiscal resources, or business density. Two demographic variables (households per capita and the homeownership rate) and one firm-density measure (construction firms per capita) return statistically significant coefficients. The magnitudes, however, are economically insignificant. The household-density coefficient of 0.009 ratio points represents a movement of roughly two percent of the sample-average ratio; the homeownership shift of −3.96 percentage points is small relative to the very high baseline homeownership in small French communes; and the construction-firm-density jump of 0.001 establishments per capita is negligible. None of these differences is large enough to plausibly drive a first-order shift in corruption or expenditure composition outcomes, and the pattern is internally consistent with a modest urbanity gradient at the cutoff rather than structural discontinuities.

Panel B explores the covariate smoothness across the pre-treatment expenditure distribution. We test for smoothness of pre-treatment expenditure covariates across the Class 6 accounts. The selected accounts presented in Panel B cover: four two-digit aggregate totals (purchases, external services, personnel costs, and other current management charges), three three-digit detailed sub-accounts (non-stocked supplies, professional fees and intermediaries, and other social charges), and two four-digit detailed sub-accounts (permanent and non-permanent staff remuneration). The remaining expenditure-account tests are reported in Appendix 2. Seven of the nine covariates are smooth at the threshold. Two accounts, other social charges and non-permanent staff remuneration return statistically significant pre-treatment coefficients. However, the magnitudes

are economically negligible (approximately 0.07 and 0.72 percentage points of total operating expenditure, respectively, on accounts that represent small budget shares). The aggregate spending base and broad expenditure structure, therefore, evolve smoothly through the cutoff, supporting the claim that pre-existing fiscal patterns cannot account for the treatment effects estimated in our main specifications.

In addition to the above-mentioned covariates, a natural concern specific to this setting is that the number of municipal councilors might itself jump at the 1,000-inhabitant cutoff. The legal council size of French communes is a step function of population, but the 1,000-inhabitant threshold falls strictly inside the 500–1,499 bracket, so council size should not move mechanically at the cutoff[28]. We verify this empirically by running the same local-linear RD specification on council size in 2008, the last municipal-election cycle prior to the zipper quota. The pre-treatment point estimate is 0.021 councilors and statistically insignificant, indicating no detectable discontinuity in council size at the threshold prior to the policy. As a complementary check, we also estimate the same RD on the 2014 council size, i.e., on the first post-treatment cycle. Here, the coefficient is 0.17 councilors and is statistically significant; however, the size of the change is negligible. Because identification of our main effects rests on the smoothness of pre-treatment characteristics, we read the 2008 result as the relevant balance test and treat the 2014 estimate as a negligible post-treatment feature[29].

A related concern is that communes on either side of the threshold may differ systematically in baseline levels of female political representation. Such differences could threaten identification if they existed prior to the reform and were correlated with subsequent outcomes of interest. To assess this possibility, we estimate the same local-linear RD specification using the share of female councilors as the outcome. The pre-treatment test uses the 2008 municipal-election cycle. The estimated discontinuity is small in magnitude and statistically indistinguishable from zero, indicating no evidence of sorting or pre-existing differences in female representation around the threshold. This finding supports the continuity assumption underlying the RD design and suggests that municipalities on either side of the cutoff were comparable in terms of women's political representation before the reform[30].

Overall, we find that the observable demographic, economic, fiscal, and institutional characteristics of communes evolve smoothly across the 1,000-inhabitant threshold, supporting the identifying assumption of the regression discontinuity design and the causal interpretation of the estimates reported in the next section.

[28] *Code général des collectivités territoriales,* Article L.2121-2 [Modifié par LOI n° 2013-403 du 17 mai 2013 - art. 28], [Consolidated version as of January 1, 2026].
[29] Full estimates and bandwidth diagnostics are reported in Appendix 2.
[30] Full estimates and bandwidth diagnostics are reported in Appendix 2.

Table 4. Pre-treatment covariate balances at the 1,000-inhabitant threshold

Panel A. Population structure, local economic resources, and productive structure								
	(1) Adults p.c.	(2) Households p.c.	(3) Homeownership rate (pp)	(4) Mean household income (€)	(5) Fiscal capacity	(6) Active labor force p.c.	(7) Service firms p.c.	(8) Construction firms p.c.
RD estimate	0.000	0.009**	-3.961**	-29.799	-1.889	-0.003	0.001	0.001*
Robust SE	(0.004)	(0.003)	(1.532)	(84.143)	(1.948)	(0.005)	(0.001)	(0.001)
Robust 95% CI	[-0.005, 0.013]	[0.004, 0.016]	[-6.336, -0.331]	[-202.512, 127.322]	[-4.780, 2.857]	[-0.016, 0.002]	[-0.002, 0.004]	[0.000, 0.001]
MSE-opt. bandwidth	281.4	202.0	190.8	405.8	393.6	354.3	223.8	352.3
Eff. obs. left of cutoff	2,807	1,812	1,693	4,717	4,512	3,854	2,064	3,824
Eff. obs. right of cutoff	1,737	1,357	1,293	2,302	2,256	2,079	1,469	2,074

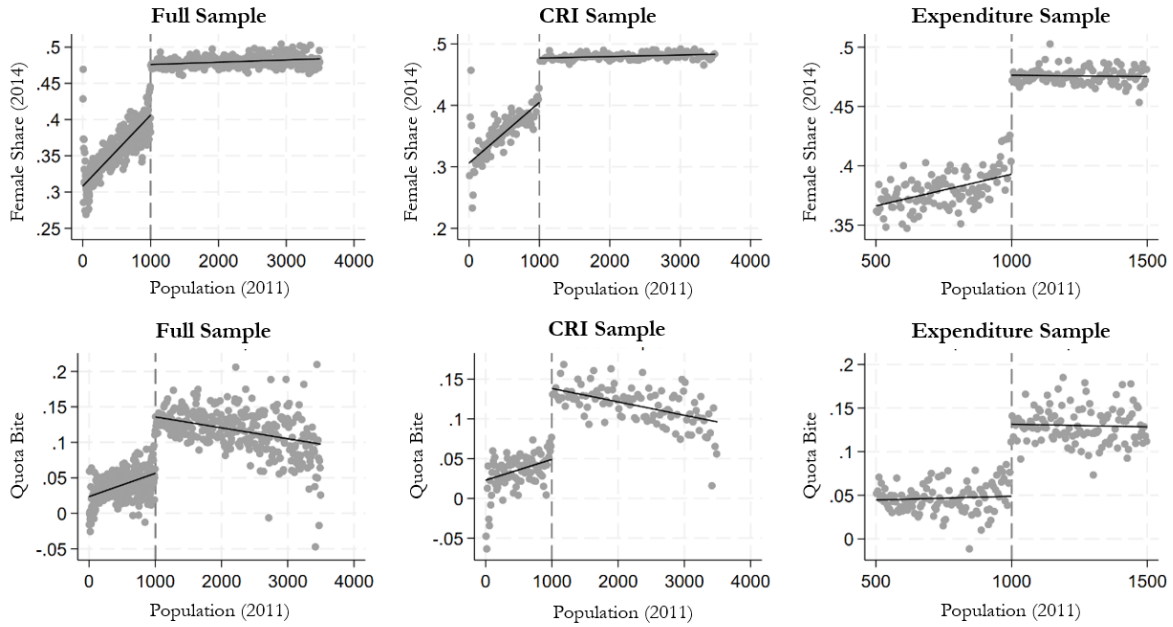
Panel B. Pre-treatment expenditure covariates (share of Class 6 operating expenditure, percentage points)									
	(1) Total purchases (60)	(2) Total external services (61)	(3) Total personnel costs (64)	(4) Total other charges (65)	(5) Non-stocked supplies (606)	(6) Professional fees (622)	(7) Other social charges (647)	(8) Permanent staff remun. (6411)	(9) Non-permanent staff remun. (6413)
RD estimate	-0.102	0.121	0.609	-0.474	-0.042	0.028	0.067**	-0.113	0.718**
Robust SE	(0.402)	(0.477)	(0.939)	(1.025)	(0.383)	(0.068)	(0.033)	(0.707)	(0.383)
Robust 95% CI	[-0.761, 0.558]	[-0.674, 0.917]	[-0.951, 2.169]	[-2.213, 1.264]	[-0.666, 0.583]	[-0.087, 0.143]	[0.009, 0.125]	[-1.265, 1.040]	[0.051, 1.385]
MSE-opt. bandwidth	160.0	137.5	119.7	119.2	157.4	157.5	116.0	127.2	103.7
Eff. obs. left of cutoff	5,464	4,580	4,004	4,004	5,376	5,342	3,371	4,284	3,142
Eff. obs. right of cutoff	4,600	3,972	3,486	3,496	4,548	4,522	3,087	3,704	2,855

Each column reports a separate local-linear regression discontinuity estimate at the 1,000-inhabitant cutoff using MSE-optimal bandwidths (Calonico, Cattaneo, & Titiunik, 2014)*, a triangular kernel, and ρ = 1. Reported coefficients are conventional, and confidence intervals are robust. In Panel A, the homeownership rate is measured in percentage points, the mean household income in euros, and all other variables are per-capita ratios. In Panel B, dependent variables are budget shares (percentage points of total Class 6 operating expenditure); estimates are based on the panel of pre-treatment commune-years, and the number of communes varies across accounts depending on reporting coverage. Significance: *** p<0.001, ** p<0.01, * p<0.05.*

6.2. Female Representation at the Threshold

The 2014 extension of the gender-parity zipper quota to communes with at least 1,000 inhabitants generates the policy variation on which our research design rests. Before turning to outcomes, we examine whether the quota produces a discontinuous increase in female representation at the population threshold, essential for our analysis (see Figure 3). We estimate sharp RD specifications on two outcomes: the share of female councilors elected in the 2014 municipal election, and the quota bite estimated separately on the full universe of eligible non-merged communes (N=31,279), the CRI analysis sample (N=8,111), and the expenditure analysis sample (N=9,261)[31].

Figure 3. Regression discontinuity plots for female share in 2014 and quota bite across full municipal, CRI and expenditure samples



The figure presents regression discontinuity plots at the 1,000-inhabitant threshold for the share of female councilors in 2014 and the change in female representation between 2008 and 2014 (quota bite), shown separately for the full municipal, CRI, and expenditure samples. Points represent binned averages of the outcome variable, while fitted lines correspond to local polynomial regressions estimated on either side of the cutoff. Bins are selected through the default rdplot bin-selection method binselect (esmv), through a data-driven procedure mimicking the variance evenly spaced method using spacing estimators. The vertical line indicates the population threshold at which the zipper quota becomes binding. The visible discontinuities at the cutoff provide graphical evidence of a strong first-stage effect of the reform on council gender composition across all samples.

All six specifications produce a positive and tightly estimated jump at the threshold (see Table 5). The 2014 female share rises by 5.96 percentage points at the cutoff in the full sample, 5.10 points in the CRI sample, and 5.88 points in the expenditure sample. The quota bite jumps by 5.32, 5.82 and 4.83 percentage points across the three samples. All estimates are significant at the 1% level on robust bias-corrected inference.

The near-identical magnitudes across samples confirm that restricting the universe to communes with non-missing procurement or expenditure data does not weaken the treatment intensity at the

[31] Please note that the restricted sample of expenditure analysis included communes with inhabitants between 500 and 1,500, hence the size of the final sample. For further information on sample representativeness please refer to Appendix 1.3.

threshold: the analysis samples inherit the same first-stage variation as the underlying universe of treated communes.

Table 5. First stage results: female representation at threshold

	(1)	(2)	(3)	(4)	(5)	(6)
Sample	Full	Full	CRI	CRI	Expenditure	Expenditure
Outcome	Female share	Quota bite	Female share	Quota bite	Female share	Quota bite
RD estimate	0.0596***	0.0532***	0.0510***	0.0582***	0.0588***	0.0483***
Robust SE	(0.0087)	(0.0129)	(0.0163)	(0.0207)	(0.0095)	(0.0162)
Robust 95% CI	[0.0444; 0.0697]	[0.0310; 0.0678]	[0.0251; 0.0692]	[0.0235; 0.0822]	[0.0417; 0.0712]	[0.0206; 0.0691]
Robust p-value	0.0000	0.0000	0.0000	0.0004	0.0000	0.0003
MSE-opt. Bandwidth	129.3	165.8	142.5	182.1	105.2	112.3
Eff. obs. left of cutoff	1090	1431	462	616	868	935
Eff. obs. right of cutoff	949	1193	431	557	779	829
Total observations	31,279	31,279	8111	8111	9261	9261

The table reports local linear regression discontinuity estimates of the effect of crossing the 1,000-inhabitant threshold on female representation in municipal councils. Outcomes include the share of female councilors in 2014 and the change in female representation between 2008 and 2014 (quota bite), shown for the full municipal sample as well as the CRI and expenditure subsamples. Estimates are obtained using a triangular kernel with ρ=1, a first-order polynomial, and mean squared error–optimal bandwidths (MSERD). Robust bias-corrected confidence intervals and p-values are reported alongside conventional estimates. Effective observations correspond to those within the selected bandwidth on either side of the cutoff. Standard errors are clustered at the municipality level.
**** p<0.001, ** p<0.01, * p<0.05*

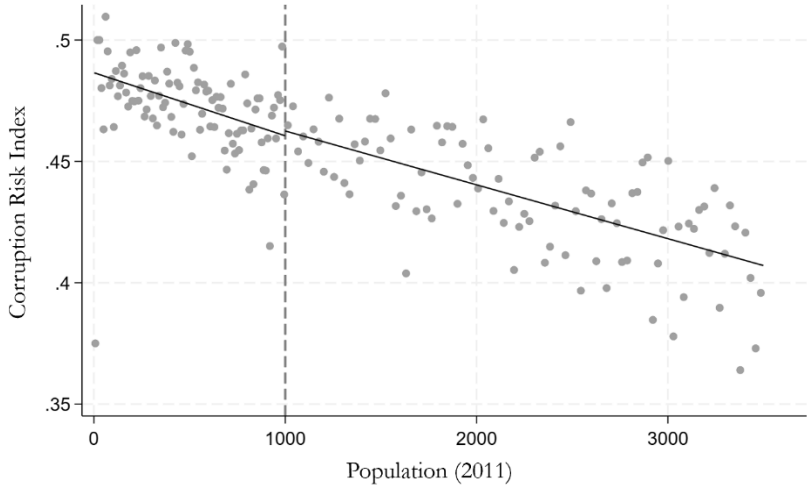
6.3. Impact of Zipper Quotas on Corruption Risk

The main analysis estimates the effect of the zipper quota reform on the corruption risk index (CRI). RDD specifications use a local linear estimator with MSE-optimal bandwidth selection and a triangular kernel, which is the default recommendation in the RDD literature as it assigns greater weight to observations closer to the threshold. Inference is based on robust bias-corrected confidence intervals. The specification aggregates outcomes to a single observation per commune over the full 2014–2020 ruling period. By averaging over the entire post-reform mandate, this specification produces the most stable corruption risk measure but at the cost of within-period variation[32].

Table 6 reports the sharp regression discontinuity estimates of the effect of crossing the 1,000-inhabitant threshold on the corruption risk index (CRI). The estimated coefficient is small in magnitude and not statistically different from zero (−0.0017, p = 0.68). This suggests that treatment at the threshold did not translate into a detectable change in corruption risk as measured by the CRI (see Figure 4).

[32] We also estimate analogous specifications at the tender level and at the commune–year level. Due to data limitations and concerns regarding measurement precision, these results are not included in the main text. However, they are reported in Appendix 3 and are consistent with the main findings.

Figure 4. Sharp regression discontinuity plot for corruption risk index analyzed at the commune – ruling period level



The figure displays a sharp regression discontinuity plot of the corruption risk index (CRI) at the 1,000-inhabitant threshold, using commune–ruling-period-level observations. Points represent binned averages of the CRI, while fitted lines correspond to local polynomial regressions estimated separately on either side of the cutoff. Bins are selected through the default rdplot bin-selection method binselect (esmv), through a data-driven procedure mimicking the variance evenly spaced method using spacing estimators. The vertical line marks the population threshold at which the quota reform becomes binding.

Table 6. Sharp regression discontinuity results for corruption risk index

	Corruption risk index
RD estimate	-0.0017
Robust SE	(0.0158)
Robust 95% CI	[-0.0374, 0.0244]
Robust p-value	0.680
Conventional SE	0.0103
MSE-opt. bandwidth	264.7
Eff. obs. left of cutoff	837
Eff. obs. right of cutoff	692
Total observations	7,147

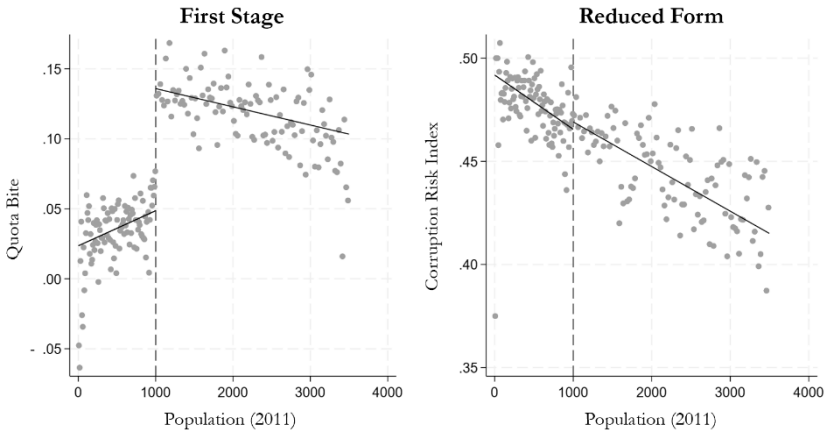
*The table reports local linear sharp regression discontinuity estimates of the effect of crossing the 1,000-inhabitant threshold on the corruption risk index (CRI), measured at the commune–ruling period level. Estimates are obtained using a triangular kernel, a first-order polynomial, and mean squared error–optimal bandwidths (MSERD), with $\rho=1$. Conventional estimates are reported alongside robust bias-corrected confidence intervals and p-values. Effective observations correspond to those within the selected bandwidth on either side of the cutoff. *** $p<0.001$, ** $p<0.01$, * $p<0.05$*

Fuzzy Regression Discontinuity Design: Treatment Intensity

While the baseline analyses treat quota eligibility as a binary policy exposure, the reform's impact varies across communes depending on the extent to which female representation increased. To account for this variation, we implement a fuzzy regression discontinuity design that uses the 1,000-inhabitant threshold as an instrument for the change in female representation[33]. The analysis is conducted at the commune-ruling period level, similar to the previous specification.

[33] As an additional robustness check, we also estimate sharp regression discontinuity specifications restricting the sample to treated communes that experienced a positive increase in female representation following the reform. This

Figure 5. Fuzzy Regression discontinuity plots for the corruption risk index analyzed at the commune-ruling period level



The figure presents fuzzy regression discontinuity plots at the 1,000-inhabitant threshold for the commune–ruling period level analysis. The left panel shows the first-stage relationship between crossing the threshold and the treatment intensity (quota bite), while the right panel displays the reduced-form relationship between the threshold and the corruption risk index (CRI). Points represent binned averages, and fitted lines correspond to local polynomial regressions estimated separately on either side of the cutoff. Bins are selected through the default rdplot bin-selection method binselect (esmv), through a data-driven procedure mimicking the variance evenly spaced method using spacing estimators. The vertical line marks the population threshold at which the quota reform becomes binding.

The first stage estimates indicate a statistically significant increase in female representation at the threshold. The preferred specification yields a discontinuity corresponding to a 5.8 percentage-point increase in the share of female councilors at the cutoff (see Figure 5). This confirms that the reform generated meaningful variation in council composition among municipalities near the threshold.

Table 7. Fuzzy regression discontinuity results for the corruption risk index analyzed at the commune–ruling period level

	First Stage	Reduced Form	Fuzzy RD
	(1)	(2)	(3)
RD estimate	0.058***	0.006	0.079
Robust SE	(0.013)	(0.009)	(0.120)
Robust 95% CI	[0.0167, -0.0952]	[-0.0246, 0.0284]	[-0.2910, 0.4179]
Robust p-value	0.000	0.887	0.726
MSE-opt. bandwidth	181.79	285.44	309.10
Eff. obs. left of cutoff	612	1034	1145
Eff. obs. right of cutoff	555	824	888
Total observations	8,111	8,111	8,111

*The table reports fuzzy regression discontinuity estimates of the effect of increased female representation on the corruption risk index (CRI) at the 1,000-inhabitant threshold, using commune–ruling-period-level data. Column (1) presents the first-stage effect of crossing the threshold on treatment intensity (quota bite), column (2) reports the reduced-form effect on CRI, and column (3) shows the corresponding fuzzy RD (Wald) estimate. Estimates are obtained using local linear regressions with a triangular kernel and mean squared error–optimal bandwidths (MSERD) and ϱ=1. Conventional estimates are reported, alongside robust bias-corrected confidence intervals and p-values. Effective observations correspond to those within the selected bandwidth on either side of the cutoff, and standard errors are clustered at the municipality level. *** p<0.001, ** p<0.01, * p<0.05*

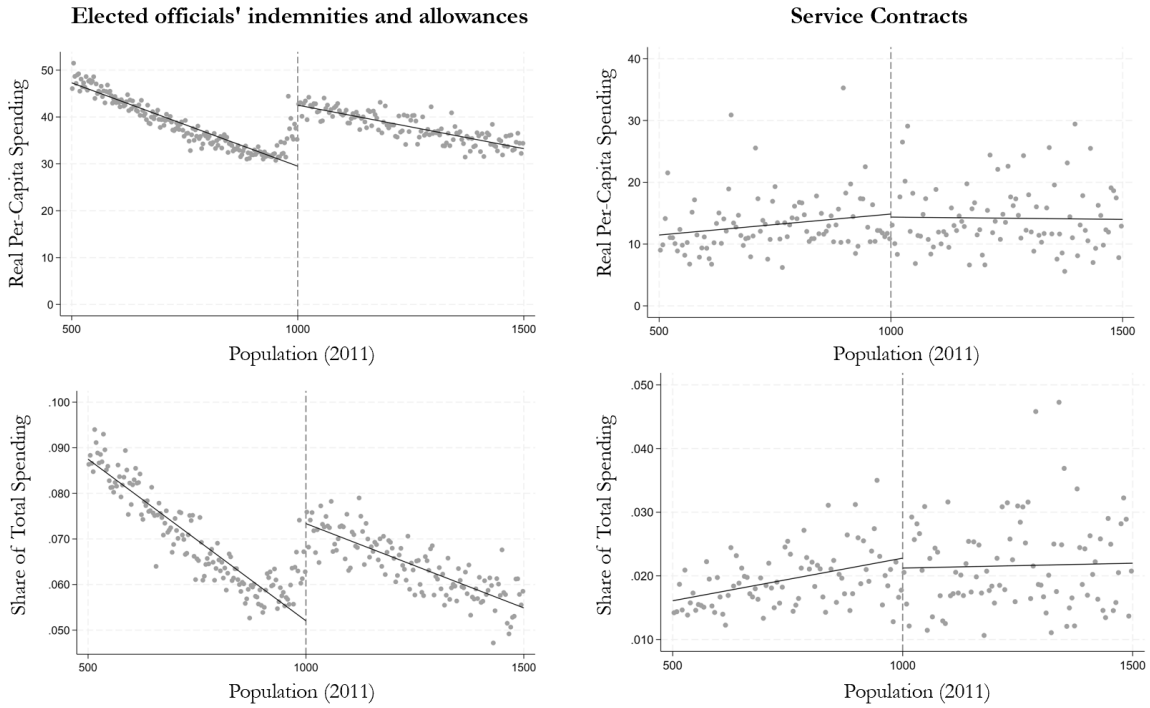
restriction isolates municipalities that responded to the quota mandate and yields qualitatively similar results, with no statistically significant effects on corruption risk. The corresponding estimates are reported in Appendix 3.

6.4. Impact of Zipper Quotas on Communal Expenditure Composition

Having found no detectable effect of the zipper quota on procurement corruption, neither under the sharp design nor in the fuzzy specification that exploits variation in quota bite, we now ask whether the policy left a measurable footprint on a different margin of municipal governance: how communes allocate their operating budgets. Following the strategy outlined in Section 5.1, we estimate 16 sharp RDD specifications (see Table 8), one for each combination of the 8 expenditure accounts introduced in Section 4.1: the share of Class 6 operating expenditure devoted to that account (capturing reallocation across categories) and real per-capita spending (capturing changes in levels).

Of the sixteen estimated treatment effects, only two accounts, elected officials' indemnities and allowances and service contracts, return coefficients distinguishable from zero at conventional significance levels (see Figure 6). As we discuss below, neither result can be interpreted as evidence of a behavioral response.

Figure 6. Sharp regression discontinuity plots for the selected communal expenditure categories analyzed at the commune-year level



The figure presents sharp regression discontinuity plots at the 1,000-inhabitant threshold for selected categories of communal expenditures at the commune–year level. Only expenditure items for which statistically significant effects are detected, namely elected officials' indemnities and allowances, and service contracts, are displayed. Points represent binned averages of spending outcomes, while fitted lines correspond to local polynomial regressions estimated separately on either side of the cutoff. Bins are selected through the default rdplot bin-selection method binselect (esmv), through a data-driven procedure mimicking the variance evenly spaced method using spacing estimators. The vertical line indicates the population threshold at which the quota reform becomes binding.

Table 8. Sharp regression discontinuity results on communal expenditure composition analyzed at the commune-year level

Account	(1) Elected officials' indemnities	(2) Service contracts	(3) Maintenance & repairs	(4) Social-security contributions	(5) Operating subsidies granted	(6) Non-storable supplies	(7) Library books & media	(8) School supplies
Panel A. Outcome: share of Class 6 operating expenditure (percentage points)								
RD estimate	0.887***	0.671**	-0.169	0.184	-0.376	-0.013	-0.006	-0.002
Robust SE	(0.251)	(0.374)	(0.351)	(0.324)	(0.387)	(0.174)	(0.023)	(0.049)
Robust 95% CI	[0.453, 1.321]	[0.030, 1.312]	[-0.740, 0.403]	[-0.362, 0.730]	[-1.029, 0.276]	[-0.298, 0.273]	[-0.045, 0.032]	[-0.087, 0.082]
Robust p-value	0.001	0.035	0.566	0.401	0.360	0.926	0.798	0.772
Bandwidth	95.5	113.2	145.7	115.3	125.2	189.8	191.0	109.7
Eff. obs. left of cutoff	4,680	3,502	7,326	5,745	6,283	10,042	4,208	4,133
Eff. obs. right of cutoff	4,264	3,097	6,288	5,047	5,464	7,867	3,888	3,701
Total observations	55,528	32,977	55,554	55,433	55,236	55,474	21,539	37,676
Panel B. Outcome: real per-capita spending (euros per inhabitant)								
RD estimate	4.056**	6.499***	0.314	2.021	-0.063	0.505	0.052	-0.019
Robust SE	(1.497)	(2.837)	(3.245)	(4.000)	(5.133)	(1.719)	(0.188)	(0.297)
Robust 95% CI	[1.376, 6.736]	[1.558, 11.440]	[-4.977, 5.604]	[-4.601, 8.643]	[-8.779, 8.653]	[-2.376, 3.386]	[-0.254, 0.359]	[-0.513, 0.475]
Robust p-value	0.020	0.008	0.808	0.564	0.955	0.612	0.820	0.881
Bandwidth	64.7	111.4	164.6	137.7	180.3	179.2	157.4	115.2
Eff. obs. left of cutoff	2,958	3,470	8,514	6,863	9,486	9,459	3,348	4,341
Eff. obs. right of cutoff	2,998	3,046	7,044	5,932	7,580	7,513	3,378	3,884
Total observations	55,528	32,977	55,554	55,451	55,236	55,474	21,539	37,676

Each column is an independent sharp regression-discontinuity specification estimated by local linear regression on the post-treatment sample 2014–2019. The running variable is 2011 INSEE census population, centered at the 1,000-inhabitant cutoff. Estimates use a triangular kernel and the MSE-optimal bandwidth selector (Calonico, Cattaneo, & Titiunik, 2014)*, $\rho=1$; robust bias-corrected standard errors are reported in parentheses below each conventional coefficient and are clustered at the commune level. Year fixed effects are partial led out by Frisch–Waugh residualization. Coefficients in Panel A are reported in percentage points (×100 of the underlying budget share); coefficients in Panel B are real euros per inhabitant. Bandwidths and effective sample sizes (left + right of the cutoff) are reported separately for the two panels because the MSE-optimal bandwidth is selected outcome-by-outcome. *** p<0.001, ** p<0.01, * p<0.05.*

For elected officials' indemnities and allowances, we estimate a statistically significant discontinuity of +0.89 percentage points in the budget share and of €4.06 in real per-capita spending. We interpret this jump as an economically negligible shift rather than a behavioral response of newly composed councils. Notably, this marginal jump is not attributable to council size changes, as we find no significant discontinuity at the threshold, despite council size being broadly linked to the population-band schedule[34]. For a representative commune of about 1,000 inhabitants, the per-capita coefficient implies an annual increase of roughly €4,000 in absolute terms, accounting for only 0.5% of total average expenditure. This magnitude is economically negligible within the broader operating budget. The Account 653 discontinuity should therefore be read as reflecting the statutory council-size schedule rather than as evidence that the quota changed incumbents' propensity to engage in legal rent extraction.

For service contracts, the estimated effect is +0.67 percentage points on the budget share and €6.50 on real per-capita. Despite the statistical significance, the magnitudes are economically modest. In absolute value, the coefficient implies on the order of €6,500 in additional service-contracting expenditure per year for a 1,000-inhabitant commune, which is equivalent of 0.8% of total average expenditures. This quantity can be attributed to routine year-to-year variation in contract renewals and is small relative to typical baseline outlays on this account. Additionally, service contracts fail the pre-treatment test on both outcomes. The discontinuity in service-contract spending therefore predates the 2014 reform, indicating that we are observing a pre-existing difference in service-contracting practice between communes immediately above and below the cutoff rather than a behavioral response to the quota. We accordingly do not interpret the post-period coefficient on service contracts as a causal effect of the policy.

The remaining six accounts return null effects, including the three female-priority accounts emphasized in the gender-and-spending literature (operating subsidies, library media, school supplies), which show no detectable response to the policy-induced increase in female representation. This pattern of nulls is consistent with recent quasi-experimental evidence from other developed-country settings: Bagues and Campa (2021) report similarly null spending effects for Spain's gender-quota reform, and the broader review by Hessami and Lopes da Fonseca (2020) finds limited evidence of female-induced spending shifts outside contexts of severe baseline gender exclusion.

6.5. Robustness

The following sub-section summarizes the most notable patterns from the robustness exercises described in Section 5.4. The full set of results, across alternative bandwidths, placebo population cutoffs, and pre-policy placebos, is reported in Appendix 4.

Female Representation

The first-stage discontinuities in female representation and quota bite are highly stable across both bandwidth options. At alternative bandwidths, all six coefficients remain significant at the 1% level, with point estimates close to the baseline (between 0.0472 and 0.0606 for the female share and

[34] For further information please refer to Section 6.1. Validity: Density Test and Covariate Balance; and Appendix 2: Covariate Balances.

between 0.0493 and 0.0628 for the quota bite). At the lower placebo cutoff of 750 inhabitants, no discontinuity is detected for either outcome in any of the three samples (full, CRI, expenditure). At the upper placebo cutoff of 1,250, two of the six coefficients are marginally significant at the 10% level: the female-share estimates in the full sample and the expenditure sample (0.0055* and 0.0073*). However, their magnitudes are approximately one-tenth the size of the effects observed at the true cutoff.

Moreover, this marginal significance does not extend to the CRI sample. Taken together, these isolated findings are more plausibly interpreted as statistical noise rather than evidence of a systematic discontinuity at 1,250. Finally, the pre-period (2008 elections) placebo yields small, statistically insignificant coefficients across all three samples, consistent with the absence of a pre-existing discontinuity in female representation at the 1,000-inhabitant threshold.

Corruption Risk

The null effect on the Corruption Risk Indicator is robust to all three classes of checks. The coefficient remains small and statistically insignificant at 0.75 (-0.0042, $p = 0.75$) and 1.25 (-0.0008, $p = 0.75$) at the MSE-optimal bandwidths. No discontinuity is detected at either placebo cutoff (-0.0008 at 750, $p = 0.82$; 0.0026 at 1,250, $p = 0.70$). The pre-policy placebo over 2011–2013 likewise yields a small and insignificant coefficient (-0.0039, $p = 0.58$). Together, these results indicate that the post-reform null is neither a product of the chosen specifications nor an offsetting pre-existing discontinuity at the threshold.

Communal Expenditure Structure

At the two-digit level of communal expenditure accounts, the positive baseline effect on elected officials' allowances (account 653) is stable across bandwidth choices, remaining significant at the 5% level, with point estimates that are economically insignificant, similar to the main specification results. The baseline effect on service contracts (account 611), although significant in the main specification, is more fragile: at alternative bandwidths, the point estimates are statistically insignificant in both the real-per capita and share-in-total-expenditures specifications. Account 611 also exhibits a sizeable and statistically significant pre-period discontinuity (8.70*** in real per-capita spending, $p < 0.01$; 0.0075** in the share, $p < 0.05$), indicating that communes on either side of the 1,000-inhabitant threshold differed in service-contract spending before the reform. The combined loss of significance under bandwidth alternatives and the pre-existing differential at the threshold imply that the baseline effect on 611 should be interpreted cautiously.

The remaining two-digit accounts (615 maintenance and repairs, 645 social-security contributions, 657 operating subsidies) are null across alternative bandwidths, placebo cutoffs, and the pre-period placebo, with one isolated marginal significance at the 10% level (account 645 at the 1,250 placebo, point estimate 0.0071*) that does not survive across the remaining checks. At the three-digit level (accounts 6061 non-storable supplies, 6065 library books and media, 6067 school supplies), the picture is similarly null: pre-period placebos and the bandwidth alternations all yield small, statistically insignificant coefficients.

Taken together, the robustness exercises support both the validity of the regression discontinuity design and the substantive interpretation of the main results. The first stage shift in female

representation is sharply identified at the 1,000-inhabitant threshold and survives every robustness check. The null effect on corruption risk is not a product of specification choice or of pre-existing differences across the threshold. Among the expenditure outcomes, the effect on elected officials' allowances, which Section 6.4 already attributes to the statutory council size schedule rather than to a behavioral response of newly composed councils, fails the robustness checks. The effect on service contracts is both bandwidth-sensitive and confounded by a pre-period discontinuity at the threshold. The robustness analysis therefore reinforces, rather than overturns, the overall pattern of results observed in the main specifications.

7. Discussion

The reform delivers a clear shift in council composition at the 1,000-inhabitant threshold, yet shows no detectable effect on procurement corruption risk and only a narrow imprint on expenditure composition. This section describes how we read the two patterns, situates them against the closest causal evidence in the French setting, and addresses the limitations of the empirical setting.

Interpreting the null on corruption risk

The cross-country and cross-municipal correlational literature would have predicted a reduction in procurement-based corruption risk under a quota that raises female representation. The closest causal evidence in the French institutional setting tells a different story. Bauhr and Charron (2021), using regression discontinuity and first difference designs to study the impact of gender of mayors on procurement-based corruption risk indicators throughout 2005–2016, show that the integrity gain associated with female representation operates at the executive margin: newly elected female mayors are linked to lower procurement corruption risks, and that differential dissipates upon re-election. Two features of their results provide a basis for discussion for our design.

First, the relevant treatment in their work is the mayor's gender. The 2014 zipper-quota extension does not change this. The reform mandates parity in the rank-ordering of electoral lists. However, the head of the list, and therefore the modal mayoral appointment, is not directly affected by the parity rule. Our null on procurement corruption risk reflects the causal effect of the reform itself, which alters council composition and not necessarily the gender of mayors. Procurement decisions are ratified and overseen by the council, but they are signed by the mayor, who also exercises discretion over award decisions, supplier relationships and procedure choice. If the integrity differential Bauhr and Charron (2021) document operates specifically through the mayoral office, a reform that directly impacts council composition would not be expected to deliver a similar reduction in procurement corruption risk, as our results show. This follows the citizen-candidate logic from the literature review: the politician with the highest implementation authority is the one whose preferences shape outcomes. In our setting, the mayor is not necessarily affected by policy.

Second, the dissipation upon re-election in their results points to a mechanism in which the disruption of established patronage and supplier networks, rather than gender preferences per se, drives the integrity gain. Our post-reform window covers a single ruling period (2014–2020), within which most affected communes are observing their first parity-compliant council. This works in favor of detecting a similar effect, making the null more informative: council-level disruption of incumbency does not appear to reproduce the integrity gains that Bauhr and Charron (2021) document for executive-level disruption in larger municipalities. Our design tests the same

mechanism on the council rather than the mayor. Our post-reform window covers a single ruling period (2014–2020), during which most affected communes are observing their first parity-compliant council, a period during which a marginalization effect at the council level should be most visible. The null we estimate suggests that the marginalization mechanism does not extend straightforwardly to newly elected women councilors in smaller French communes.

There is also an institutional dimension that may impact our results. Communes near the 1,000-inhabitant cutoff sit at the lower end of administrative capacity in France. Surveys by the Agence française anticorruption show that formal anti-corruption mechanisms, such as codes of conduct, risk mapping and whistleblowing systems, are largely absent in municipalities of this size, and that detection and reporting of procurement irregularities scale steeply with commune size (2022; 2018). In this context, a change in the gender composition of the council may translate into observable differences in corruption risk only through channels that depend on the council having both the information and the procedural standing to challenge award decisions, scrutinize single-bidder procedures, or demand supplier rotation. Where these channels are weak, the marginal female councilor brought in under the quota may have limited operational leverage over procurement outcomes, regardless of any underlying preferences. Our null can therefore be read as speaking to the small-commune procurement environment as much as to the behavior of female councilors, and may refine rather than overturn the Bauhr and Charron (2021) result: their evidence is consistent with a "new mayor" channel for integrity gains. At the same time, ours suggests this channel may not extend straightforwardly to council composition in smaller municipalities, where the parity reform was binding. However, the institutional preconditions for translating representation into procurement oversight are less developed.

Interpreting the results on municipal expenditures

The expenditure analysis examines three groups of pre-specified accounts: procurement-adjacent operating items; spending categories that the literature on female representation has identified as gendered (Svaleryd, 2009; Profeta & Accettura, 2022; Hessami & Lopes de Fonseca, 2020); and elected officials' indemnities and allowances, where elected officials can engage in legal rent extraction within statutory ceilings.

The most informative finding is the absence of any response in the gendered categories. The increase in female representation produced by the quota does not translate into a detectable reallocation of the operating budget toward the spending lines that the comparative literature would have predicted. Several mechanisms are compatible with this pattern. Smaller French communes have rigid budget structures, with most operating expenditure pre-committed to staffing, statutory transfers, and ongoing service contracts, leaving narrow discretionary scope at the annual budget vote. The quota also moves council composition toward, but generally not above, parity, so the post-reform council is closer to being balanced than to female-majority. The result also echoes Rigon and Tanzi (2011), who find a similar null effect in the Italian setting, with male and female councilors holding preferences close to the same median.

The two non-null coefficients fall outside the gendered-spending group. The results of both allowances and service contracts are economically insignificant. However, the result on service contracts is more delicate to interpret, as the same account shows a pre-treatment imbalance of a similar size. Following Marshall (2022), we treat the post-reform estimate as a compound of the

quota-induced shift in council composition and the pre-existing differences at the cutoff, and we do not attribute it to the parity rule.

The results of both research questions are consistent with the representation-governance trade-off introduced in the literature review (Brooks, Phillips, & Sinitsyn, 2011): widening the set of voices represented on the council does not automatically change who effectively decides, and without that shift, neither procurement integrity nor expenditure composition responds.

Limitations

The regression discontinuity design addresses most of the comparability concerns that would otherwise affect comparisons across communes of different sizes. Identification rests on the local comparison of communes immediately above and below the 1,000-inhabitant cutoff, and the covariate balance tests support smoothness of the variables for which we observe pre-treatment values. The limitations discussed here therefore primarily affect external validity and the precision of the corruption-risk measure rather than the internal validity of the estimates.

The first limitation concerns the coverage of procurement data near the cutoff. Communes in our analysis bandwidth conduct relatively few procurement procedures annually, and the GPPD's coverage of these communes reflects variation in reporting practices, in digitization quality, and in whether lower-value contracts pass the data collection threshold[35]. The RDD design absorbs the part of this variation that is symmetric around the cutoff, but commune-level corruption risk measures are noisier than they would be in larger municipalities. Aggregation to the commune-year and ruling-period levels, which we use for this reason, trades off some of that noise against within-period variation. Our estimates, therefore, capture the integrity dimensions visible in the procurement-level corruption risk indicators. They are less informative about forms of corruption that leave no contract-award footprint, such as informal favoritism in supplier selection or post-award execution irregularities, hindering the external validity of the results.

The second limitation concerns expenditure-side measurement. The DGFiP balances permit a clean compositional analysis at the chart of accounts level. However, they do not, for the communes in our sample, support a fine-grained classification of spending by policy area. The functional activity codes introduced from 2016 onwards would, in principle, have allowed us to identify spending on social services, culture, sport, and other policy areas directly, providing closer proxies for the categories most often emphasized in the gender-and-spending literature. In the relevant communes, these codes are predominantly populated for non-principal budgets, so coverage in main budgets is too sparse to support a separate analysis. Our outcomes are therefore necessarily defined at the accounting-category level, which may understate or mask reallocations across policy areas that share an accounting parent.

The third limitation concerns the availability of covariates for smaller communes. Several variables that would ideally enter the covariate balance tests, including detailed fiscal indicators, fine-grained

[35] The current sample of our analysis represents 33.6% of all unique communes, excluding merged communes. Among communes with fewer than 1,000 inhabitants, the sample covers 24.5%, while for communes with between 1,000 and 3,500 inhabitants, representation increases to 69.9%. However, over the entire analysis period (2011-2020), the number of tenders per commune is low, with a mean of 8.995 tenders and a median of 2 tenders per each unique commune.

labor market data, and political competition measures, are either unavailable or of limited quality in the official administrative sources at this size. The smoothness assumption is therefore validated on the subset of pre-treatment variables we can observe. We have no reason to believe that the unobserved characteristics behave differently at the cutoff, but the balance evidence is necessarily partial.

A fourth limitation concerns the compound nature of the treatment at the 1,000-inhabitant threshold. The 2013 electoral law extended the zipper quota alongside a shift in the electoral procedure for newly covered communes, specifically the move toward a closed list system under which voters choose between pre-ranked party or coalition lists rather than individual candidates. This procedural change serves as the institutional precondition for the parity rule to be operable rather than as an independent policy intervention with separate governance implications. Our design identifies the effect of crossing the threshold, which in principle conflates this procedural shift with the zipper quota itself. Our intent-to-treat estimates can therefore be interpreted as the effect of the full set of institutional changes at the 1,000-inhabitant cutoff, of which the zipper quota is the primary and most novel component.

8. Conclusion

This paper examined the effect of France's 2014 extension of the gender-parity zipper quota on two dimensions of local governance in communes near the 1,000-inhabitant threshold: corruption risk in public procurement and the composition of communal expenditure. Exploiting the threshold as a regression-discontinuity cutoff, we find that the reform substantially raised the share of women on municipal councils but did not produce detectable changes in either outcome.

The reform meets its first-stage objective: a council-targeted parity rule reliably and substantially increases female representation, an outcome with independent democratic value. What our findings refine is the further claim that compositional change at the council level translates into integrity and resource-allocation effects. First, raising the share of women on the council is not sufficient to reduce procurement-based corruption risk in smaller French communes; existing causal evidence has located female-representation effects on procurement integrity at the executive mayoral level. Second, the broader claim that female representation shifts public spending toward gendered categories is not evident in the small-commune setting we examine, characterized by rigid budget structures and post-reform councils that approach parity rather than a female majority. Descriptive change at the council does not, on its own, deliver substantive change in either of the two governance outcomes we study.

Three directions for future research would help interpret these nulls more sharply. The first is to combine variation in council composition with independent variation in mayoral characteristics, which our setting does not allow. The second is to extend the analysis beyond the contract-award record to dimensions of council behavior the corruption-risk indicator does not capture, including the use of negotiated procedures, contract fragmentation, and council responsiveness to constituents. The third is to revisit the same questions following the 2026 extension of parity requirements to all French communes.

9. References

AFA. (2018). *Survey on the Prevention of Local Government Corruption.* https://www.agence-francaise-anticorruption.gouv.fr/files/files/EN%20Survey%20local%20government%20antibribery%20policy.pdf: Agence Française Anticorruption.

AFA. (2022). *Prevention and Detection of Corruption in Local Governments and Public Entities.* https://www.agence-francaise-anticorruption.gouv.fr/files/files/summary%20of%20the%20national%20statistical%20survey%20on%20the%20prevention%20and%20detection%20of%20corruption%20in%20local%20governments%20and%20public%20entities%20April%202022.pdf: Agence Française Anticorruption.

Androeli, F., Manzoni, E., & Margotti, M. (2022). Women at work: Gender quotas, municipal elections and local spending. *European Journal of Political Economy, vol. 75*, https://doi.org/10.1016/j.ejpoleco.2021.102175.

Aytimur, R. E., Boukouras, A., & Schwager, R. (2016). The citizen-candidate model with imperfect policy control: Strategic delegation and polarization. *The Canadian Journal of Economics / Revue canadienne d'Economique, Vol. 49, Issue 3*, 997-1015.

Bagues, M., & Campa, P. (2021). Can gender quotas in candidate lists empower women? Evidence from a regression discontinuity design. *Journal of Public Economics, Vol. 194*.

Barnes, T., & Beaulieu, E. (2014). Gender stereotypes and corruption: How candidates affect perceptions of election fraud. *Politics and Gender, vol. 10, Issue 3*.

Barnes, T., & Beaulieu, E. (2018). Women Politicians, Institutions, and Perceptions of Corruption. *Comparative Political Studies, vol. 52, Issue 1*, https://doi.org/10.1177/0010414018774355.

Bauhr, M., & Charron, N. (2021). Will Women Executives Reduce Corruption? Marginalization and Network Inclusion. *Comparative Political Studies*, Vol. 54(7) 1292 –1322, DOI: 10.1177/0010414020970218.

Bauhr, M., Charron, N., & Wängnerud, L. (2019). Exclusion or interests? Why females in elected office reduce petty and grand corruption. *European Journal of Political Research 58, no. 4*, 1043-1065.

Bauhr, M., Czibnik, A., de Fine Licht, J., & Fazekas, M. (2019). Lights on the shadows of public procurement: Transparency as an antidote to corruption. *Governance 33(5)*, DOI:10.1111/gove.12432.

Best, R. E., Burge, I., & McDonald, M. D. (2012). Representation as a median mandate: Taking cross-national differences seriously. *European Journal of Political Research , Vol.51 , Issue 1* , 1-23.

Bideau, G., & Ysebaert, R. (2022). Les communes nouvelles françaises (2012-2022) : une méthode pour l'analyse de données à l'échelon municipal selon des limites évolutives. *Cybergeo: European Journal of Georgraphy*, https://doi.org/10.4000/cybergeo.39387.

Borge, L.-E., & Rattsø, J. (2008). Young and Old Competing for Public Welfare Services. *CESifo Working Paper No. 2223* .

Broms, R., Dahlström, C., & Fazekas, M. (2019). Political Competition and Public Procurement Outcomes. *Comparative Political Studies*, 1-34.

Brooks, L., Phillips, J., & Sinitsyn, M. (2011). The Cabals of a Few or the Confusion of a Multitude: The Institutional Trade-Off between Representation and Governance. *American Economic Journal: Economic Policy, Vol. 3, Issue 1*, 1-24.

Calonico, S., Cattaneo, M. D., & Farell, M. (2017). rdrobust: Software for regression-discontinuity designs. *The Stata Journal, 17, Number 2*, 372-404.

Calonico, S., Cattaneo, M. D., & Titiunik, R. (2014). Robust Nonparametric Confidence Intervals for Regression-Discontinuity Designs. *Econometrica, Vol. 82, Issue 6*, 2295-2326.

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

Cavallini, F., Dominici, A., & Masi, O. (2025). Gender Quotas in Municipal Executives: Reallocating Public Spending in Italy. *International Tax and Public Finance (ITAX), 5.12 Working Paper*, https://dx.doi.org/10.2139/ssrn.4981662.

CGCT, art. L2122-21; art. L2122-22. (n.d.). *Code général des collectivités territoriales (CGCT), art. L2122-21*. Retrieved from Le service public de la diffusion du droit, Légifrance: https://www.legifrance.gouv.fr/codes/section_lc/LEGITEXT000006070633/LEGISCTA000006192256/#LEGISCTA000006192256

Charron, N., Dahlström, C., Fazekas, M., & Lapuente, V. (2017). Careers, Connections, and Corruption Risks. *The Journal of Politics, vol.79, no. 1*, 89-104.

Chattopadhyay, R., & Duflo, E. (2004). Women as Policy Makers: Evidence from a Randomized Policy Experiment in India. *Econometrica, Vol 72., No. 5* , 1409-1443.

Clayton, A., & Zetterberg, P. (2018). Quota Shocks: Electoral Gender Quotas and Government Spending Priorities Worldwide. *The Journal of Politics, vol. 80, Issue 3*, 916-932.

Clots-Figueras, I. (2011). Women in politics: Evidence from the Indian States. *Journal of Public Economics, Vol. 95, Issues 7-8*, 664-690.

Coviello, D., & Mariniello, M. (2014). Publicity Requirements in Public Procurement: Evidence from a Regression Discontinuity Design. *Journal of Public Economics, vol. 109*, 76–100.

Cunningham, S. (2021). *Causal Inference: The Mixtape.* Yale University Press, https://mixtape.scunning.com/.

d'Agostino, G., Dunne, P., & Pieroni, L. (2016). Government Spending, Corruption and Economic Growth. *World Development* , 190-205.

Dahlerup, D. (1989). From a Small to a Large Minority Women in Scandinavian Politics. *Scandinavian Political Studies, Vol. 11, Issue 2*.

Dahlerup, D. (2006). The Story of the Theory of Critical Mass. *Politics & Gender, Vol.2, Issue 4*.

De Magalhães, L., Hangartner, D., Hirvonen, S., Meriläinen, J., Ruiz, N. A., & Tukiainen, J. (2025). When Can We Trust Regression Discontinuity Design Estimates from Close Elections? Evidence from Experimental Benchmarks. *Political Analysis, Vol. 33, Issue 3*, 258-265.

DGFiP. (2026). *Balances comptables des communes*. Retrieved from data.economie.gouv.fr: https://data.economie.gouv.fr/explore/?sort=modified&q=Balances+comptables+des+communes

DiPasquale, D., & Glaeser, E. L. (1999). Incentives and Social Capital: Are Homeowners Better Citizens? *Journal of Urban Economics*.

Dollar, D., Fisman, R., & Gatti, R. (2001). Are Women Really the 'Fairer' Sex? Corruption and Women in Government. *Journal of Economic Behavior & Organization 46, no. 4* , 423–29.

Downs, A. (1957). *An Economic Theory of Democracy*. New York: Harper & Row.

Ennser-Jedenastik, L. (2017). How Women's Political Representation affects Spending on Family Benefits. *Journal of Social Policy, vol 46, Issue 3*, 563-581.

Esarey, J., & Chirillo, G. (2013). 'Fairer Sex' or Purity Myth? Corruption, Gender, and Institutional Context. *Politics & Gender 9, no. 04*, 361-389.

European Commission. (2016). *Public procurement – Study on administrative capacity in the EU, France Country Profile*. Brussels: European Union.

Fazekas, M. (2025, May 21). From transparency to metrics: Measuring corruption in government contracts. *The U4 Anti-Corruption Resource Centre*, pp. https://www.u4.no/blog/from-transparency-to-metrics-measuring-corruption-in-government-contracts.

Fazekas, M., & Kocsis, G. (2015). Uncovering High-Level Corruption: Cross-National Corruption Proxies Using Government Contracting Data. *Government Transparency Institute*.

Fazekas, M., & Kocsis, G. (2020). Uncovering high-level corruption: Cross-national objective corruption risk indicators using public procurement data. *British Journal of Political Science, vol. 50*, 155-164.

Fazekas, M., Tóth, B., Abdou, A., & Al-Shaibani, A. (2024). Global contract-level public procurement dataset. *Data in Brief*, vol. 54. https://doi.org/10.1016/j.dib.2024.110412 .

Févrat, N., Labatut, V., Volpi, É., & Marrel, G. (2025). A dataset of the representatives elected in France during the fifth republic. *Data in Brief, vol. 60* .

Gamberoni, E., Garner, C., Giordano, C., & Lopez-Garcia, P. (2016). Is corruption efficiency-enhancing? A case study of nine Central and Eastern European countries. *The Competitiveness Research Network, European Central Bank*.

Gelman, A., & Imbens, G. (2018). Why High-Order Polynomials Should Not Be Used in Regression Discontinuity Designs. *Journal of Business & Economic Statistics , vol. 37, Issue 3*, 447-456.

Goetz, A. M. (2007). Political Cleaners: Women as the New Anti-Corruption Force? *Development and Change, vol. 38, Issue1*, 87-105.

Grembi, V., Nannicini, T., & Troiano, U. (2016). Do Fiscal Rules Matter? *American Economic Journal. Applied Economics, Vol. 8, Issue 3*, 1-30.

Hessami, Z., & Lopes de Fonseca, M. (2020). *Female political representation and substantive effects on policies: A literature review*. https://docs.iza.org/dp13125.pdf: IZA: Institute of Labor Economics.

Holmberg, S., & Rothstein, B. (2014). Correlates of Corruption. *The Quality of Government Institute*, https://www.gu.se/sites/default/files/2020-05/QoGWP_2014_17_Rothstein_Holmberg.pdf.

INSEE. (2013, 31 12). *Populations légales 2011*. Retrieved from https://www.insee.fr/fr/statistiques/2119745?sommaire=2119751

INSEE. (2026, 02). *Business register identification system - SIRENE*. Retrieved from https://www.insee.fr/en/metadonnees/source/serie/s1020

INSEE. (2026, January 1). *Code officiel géographique au 1er janvier 2026* . Retrieved from https://www.insee.fr/fr/information/8740222

Inter-Parliamentary Union. (2026). *Women in parliament in 2025: The year in review*. Inter-Parliamentary Union.

Inter-Parliamentary Union. (2026). *Women in Politics: 2026*. Retrieved from https://www.ipu.org/resources/publications/infographics/2026-03/women-in-politics-2026

Kumar, J. C., & Sarangi, S. (2018). Women and corruption: What positions must they hold to make a difference? *Journal of Economic Behavior & Organization, vol 151*, 219-233.

Kutlu, L., & Mao, X. (2023). The effect of corruption control on efficiency spillovers. *Journal of Institutional Economics, Vol. 19, Issue 4*, 564-578.

Lassébie, J. (2019). Gender quotas and the selection of local politicians: Evidence from French municipal elections. *Elsevier*, https://doi.org/10.1016/j.ejpoleco.2019.101842.

Lee, D., & Lemieux, T. (2010). Regression Discontinuity Designs in Economics. *Journal of Economic Literature, vol. 48, no. 2*, 281-355.

Lépinard, E., & Lieber, M. (2015). *The Policy on Gender Equality in France: In-depth analysis for the FEMM Committee.* https://www.europarl.europa.eu/RegData/etudes/IDAN/2015/510024/IPOL_IDA(2015)510024_EN.pdf: European parliament, Policy Department C: Citizens' Rights and Constitutional Affairs, Committee on Women's Rights and Gender Equality.

Marshal, J. (2022). Can Close Election Regression Discontinuity Designs Identify Effects of Winning Politician Characteristics? *American Journal of Political Science, Vol. 68, Issue 2*, 494-510.

Mauro, P. (1998). Corruption and the composition of government expenditure. *Journal of Public Economics, Vol. 69, Issue 2*, 268-279.

McCrary, J. (2008). Manipulation of the running variable in the regression discontinuity design: A density test. *Journal of Econometrics, vol.142, Issue 2*, 698-714.

McDonald, M. D., Mendes, S. M., & Budge, I. (2004). What Are Elections for? Conferring the Median Mandate. *British Journal of Political Science, Vol. 34, Issue 1*, 1–26.

Merkle, O. (2022). *Anti-corruption and gender: the role of women's political participation.* The Westminster Foundation for Democracy anti-corruption and integrity series.

Osborne, M. J., & Slivinski, A. (1996). A Model of Political Competition with Citizen-Candidates. *The Quarterly Journal of Economics, Vol. 111, Issue. 1* , 65-93.

Pereira, M. M., & Fernandez-Vazquez, P. (2022). Does Electing Women Reduce Corruption? A Regression Discontinuity Approach. *Legislative Studies Quarterly, vol. 48, Issue 4*, 731-763.

Pettersson-Lidbom, P. (2012). Does the size of the legislature affect the size of government? Evidence from two natural experiments. *Journal of Public Economics, Vol. 96, Issues 3-4*, 269-278.

Profeta, P., & Accettura, C. (2022). Gender Differences in Political Budget Cycles. https://papers.ssrn.com/sol3/papers.cfm?abstract_id=4014355.

Rigon, M., & Tanzi, G. M. (2011). Does gender matter for public spending? Empirical evidence from Italian municipalities. *MPRA Paper 34845, University Library of Munich, Germany*.

Schöberlein, J. (2019). Lessons learned from anti-corruption efforts at municipal and city level. *Transparency International*.

Service Public. (2025, August 25). Municipales 2026 : le mode de scrutin évolue dans de nombreuses communes. *Direction de l'information légale et administrative (Premier ministre)*, pp. https://www.service-public.gouv.fr/particuliers/actualites/A18432.

Smith, A. (2000). *Regional Government in France.* University College London, https://www.ucl.ac.uk/social-historical-sciences/sites/social_historical_sciences/files/64.pdf.

Sundström, A., & Stockemer, D. (2015). What determines women's political representation at the local level? A fine-grained analysis of the European regions. *International Journal of Comparative Sociology, vol. 56, Issue 3-4* , https://doi.org/10.1177/002071521559569.

Svaleryd, H. (2009). Women's representation and public spending. *European Journal of Political Economy, vol 25, Issue 2*, 186-198.

Swamy, A., Stephen, K., Lee, Y., & Azfar, O. (2001). Gender and corruption. *Journal of Development Economics, vol. 64, Issue 1*, 25-55.

Treisman, D. (2000). The causes of corruption: a cross-national study. *Journal of Public Economics, vol.76*, 399-457.

Treisman, D. (2007). What Have We Learned About the Causes of Corruption From Ten Years of Cross-National Empirical Research? *Annual Review of Political Science, vol. 10 (1).*

UN Women. (2026, March 11). *Facts and figures: Women's leadership and political participation*. Retrieved from https://www.unwomen.org/en/articles/facts-and-figures/facts-and-figures-womens-leadership-and-political-participation#_edn13; https://localgov.unwomen.org/

United Nations. (2018, September 10). *Global Cost of Corruption at Least 5 Per Cent of World Gross Domestic Product, Secretary-General Tells Security Council, Citing World Economic Forum Data*. Retrieved from https://www.un.org/sg/en/content/sg/statements/2018-09-10/secretary-generals-remarks-the-security-council-corruption-conflict-delivered

Vie publique. (2024, October 25). *Quel est le rôle de la région ?* Retrieved from https://www.vie-publique.fr/infographie/280077-infographie-quelles-sont-les-competences-de-la-region

Vie publique. (2025, November 3). *Quelles sont les compétences des départements ?* Retrieved from Vie publique: https://www.vie-publique.fr/fiches/19620-quelles-sont-les-competences-exercees-par-les-departements

Vie publique. (2025, August 7). *Qu'est-ce qu'une commune ?* Retrieved from https://www.vie-publique.fr/fiches/19613-quest-ce-quune-commune

Wängnerud, L. (2015). *The Link Between Gender and Corruption in Europe.* Gothenburg: Quality of Government Institute, ANTICORRP; https://anticorrp.eu/wp-content/uploads/2015/10/D5.3-Case-studies-on-gender-and-corruption_Lena-W%C3%A4ngnerud.pdf.

Appendices

Appendix 1: Data Construction

1.1. Procurement Data Cleaning and Commune Matching

The procurement records used in this study are drawn from the Global Public Procurement Dataset (GPPD), a contract-level dataset compiled by harmonizing tender publications from national procurement authorities and the European Tenders Electronic Daily (TED) portal. The publishers harmonize variable definitions across publication channels, apply filters intended to remove invalid or duplicate records, and prepare the data for further analysis. Several limitations of the source data are worth noting. Many of the underlying fields (notably buyer and city names) are free-text entries filled in by the tender filer and exhibit common formatting inconsistencies, including the use of the postal-routing suffix CEDEX in city names. The dataset also contains tender publications that were never executed; the GPPD filters intended to drop such cases are not exhaustive. So-called "umbrella tenders", single tender notices that bundle multiple subcontracts, pose an additional challenge for unit-of-analysis decisions. These caveats are noted here and addressed at successive stages of the cleaning pipeline.

The raw GPPD France extract contains 3,075,410 contract-level observations covering the 2006-2021 period. Initial preprocessing applies three restrictions. First, the dataset is restricted to contracts with French buyers. Second, cancelled procedures are removed. Third, the dataset is limited to procurement procedures that contain information on competition, including open tenders with losing bids: A restriction motivated by the requirements of the corruption risk indicators, which depend on competition-related fields. During preprocessing, key string variables (buyer names, commune names, and postal codes) are standardized to ensure consistent formatting across records.

Procurement records do not systematically report INSEE commune codes (geo_code). To enable aggregation to the municipality level, observations are linked to communes through a multi-stage matching procedure that prioritizes the most reliable identifiers and progressively applies geographic-matching rules to observations that remain unmatched at each step. The seven matching stages are summarized in Table A.1.1.1.

At Stage 1, observations carrying a valid SIRET establishment identifier are matched to the French SIRENE business registry (a reduced version of sirene_buyers_reduced.parquet), which provides the official commune code (codeCommuneEtablissement) associated with each establishment. This stage produces the most reliable assignments because SIRET identifiers uniquely identify establishments in the administrative registry. At Stage 2, observations without usable SIRET identifiers are matched on postal code, restricted to the 1,764 postal codes with a unique one-to-one mapping to a commune. Stages 3 and 4 combine postal codes with normalized commune names: Stage 3 applies to postal codes that are unique to a single commune after applying the norm_city() routine (which removes accents, punctuation, and administrative suffixes such as CEDEX), while Stage 4 uses an expanded long-format lookup that lists all postal codes associated with each commune. Stage 5 matches on commune name alone, restricting matches to commune names that are unique nationally; for Paris, Lyon, and Marseille, arrondissement-level records are aggregated to the city level under synthetic geo-codes with summed populations. Stages 6 and 7

combine standardized commune names with department identifiers derived, respectively, from NUTS3 codes (using a department crosswalk constructed from already-matched rows where the modal NUTS3 share is at least 95%) and from the first two digits of the postal code, with deterministic lookups restricted to (department, commune) pairs with a unique geo-code. Finally, a small set of known non-commune strings (for example, "La Défense" → Nanterre, "Sophia Antipolis" → Valbonne) is resolved deterministically through an alias table.

Table A.1.1.1. Multi-stage commune-matching procedure.

Stage	*Identifier used*	*Observations matched*	*Share of dataset*
1	SIRET → SIRENE registry (codeCommuneEtablissement)	553,055	18.0%
2	Postal code with unique commune mapping	393,108	12.8%
3	Postal code + normalized city (single-postcode communes)	425,005	13.8%
4	Postal code + normalized city (multi-postcode communes)	89,310	2.9%
5	Nationally unique commune name	≈1,300,000	≈42%
6	NUTS3-derived department + normalized city	≈120,000	≈3.9%
7	Postal-derived department + normalized city	≈80,000	≈2.6%
Total	Across all stages	2,990,483	97.2%
Unmatched	Missing or inconsistent geographic identifiers	84,927	2.8%

Note. Shares are computed relative to the 3,075,410 contract-level observations in the cleaned French subset of GPPD. Shares for Stages 5–7 are reported to the nearest tenth of a percentage point; the corresponding observation counts are reported as the order-of-magnitude figures used in the matching pipeline.

Across all seven stages, geo-codes are successfully assigned to 2,990,483 observations (97.2% of the cleaned dataset). The remaining 84,927 observations (2.8%) cannot be reliably matched due to missing or inconsistent geographic identifiers; these observations are retained in the underlying dataset but excluded from any commune-level aggregation.

Following the matching procedure, contract-level corruption risk indicators are aggregated to the commune-year level. The resulting file is then merged with commune-level population dataset, sourced from the INSEE 2011 census, effective at the policy introduction. Overseas departments are excluded because of inconsistencies in the corresponding population and fiscal datasets. The merged analytical dataset contains 2,272,385 procurement observations across 14,259 communes in mainland France.

Table A.1.1.1. summarizes the successive sample-construction filters applied to obtain the final estimation sample. The first step drops communes that fail to match the 2011 INSEE population frame, removing 65,328 observations. The second step restricts the sample to communes with population_2011 below 3,500, the upper bound of the broader RDD comparison frame around the 1,000-inhabitant threshold; this step substantially reduces the dataset because larger communes contribute disproportionately many procurement records relative to their share of the population.

The third step restricts the sample to tender years 2010–2019 to align with the period over which both treatment and outcome variables are reliably observed. The fourth step excludes communes affected by post-2014 mergers (communes nouvelles): a commune that merges with one or more neighbors during the sample period changes its boundaries, its population, and potentially its treatment status under the 1,000-inhabitant threshold, so its procurement records cannot be unambiguously assigned to a stable treatment status. This exclusion drops 602 unique communes and 8,757 observations, leaving 136,857 procurement observations in the final analysis sample.

Table A.1.1.2. reports the number of procurement observations, unique tenders, and unique communes in the final analysis sample, broken down by tender year. Coverage in 2010 is markedly thinner than in subsequent years, with only 224 observations across 166 communes; this likely reflects incomplete backfilling of pre-publication data in the source. Due to coverage issues, pre-policy analysis covers the 2011-2013 period, excluding 2010.

Table A.1.1.2. Annual coverage of the procurement analysis sample (post-merger-exclusion).

Tender year	*Procurement rows*	*Unique tenders*	*Unique communes*
2010	224	224	166
2011	17,589	15,397	4,541
2012	15,026	12,730	4,099
2013	15,071	12,383	3,894
2014	12,333	9,655	3,108
2015	16,685	10,495	3,447
2016	14,929	8,800	3,164
2017	14,061	7,963	2,756
2018	14,751	8,196	2,739
2019	16,188	8,724	2,874

Table A.1.1.3. reports the distribution of unique tenders per commune across the 2011–2019 sample period. The median commune posts only two unique tenders over this nine-year window, while the 90th percentile posts nineteen and the maximum exceeds one thousand. The resulting commune-by-year panel is therefore highly unbalanced and the within-commune distribution of activity is heavily right-skewed. This pattern is taken into account in the choice of estimator: the selected specification uses an RDD on commune-year aggregates with MSE-optimal bandwidth choices and clustered standard errors at the commune level.

Table A.1.1.3. Distribution of unique tenders per commune, 2011–2019.

Statistic	*Value*
Communes (N)	10,535
Mean unique tenders per commune	8.96
Median unique tenders per commune	2
90th percentile	19
Maximum	1,075

Note. Statistics are computed over communes that appear at least once in the 2011–2019 procurement sample after the population, year, and merger filters of Table B.3.

The Corruption Risk Indicator (CRI) used in this study is constructed as the average of a set of binary red-flag indicators, each capturing a distinct dimension of procurement risk (e.g., single-bidder outcomes, short submission windows). For tenders in which only a small number of components are valid, that is, observable in the underlying contract record, the CRI takes a smaller number of possible values and is mechanically more likely to land at 0.5. Table A.1.1.4. documents this pattern: when only three components are valid, 94% of tenders take the value 0.5; when four or more components are valid, the CRI is concentrated away from 0.5. The frequency with which 0.5 occurs at three valid components is therefore not informative about underlying corruption risk and reflects a mechanical feature of the indicator. The main analysis addresses this in two ways. First, the baseline outcome restricts attention to tenders with at least four valid CRI components, so that the underlying flag distribution is not dominated by mechanical 0.5 values. Second, robustness checks reported in the main text re-estimate the headline specification using each individual red-flag component as the outcome variable.

Table A.1.1.4. Frequency of CRI = 0.5 by number of valid CRI components.

Valid CRI components	*CRI ≠ 0.5*	*CRI = 0.5*
3	1,698	26,132
4	13,990	335
5	3,510	300
6	17	1

Note. Tabulation is over the matched analytical dataset prior to the population-bandwidth restriction

1.2. Construction of communal expenditure covariates

The source data are the annual balances comptables des communes, published by the Direction générale des Finances publiques (DGFiP) of the French Ministry of Finance. These files report annual budget execution for all French communes and contain account-level debit and credit flows alongside commune identifiers, budget-type indicators, and accounting-framework codes. Our extract spans the years 2010–2023; the empirical analysis uses the 2010–2019 sub-period.

Four standardized public-accounting nomenclatures apply to French municipalities during the sample period. M14 is the long-standing nomenclature for communes; M14A is its simplified variant for communes with fewer than 500 inhabitants. M57 is a modernized, unified framework introduced from 2015 on a voluntary basis and scheduled to become mandatory in 2024; M57A is its simplified variant. Account codes are largely consistent across the four frameworks, but a small number of codes do not fully overlap; see Section C.4 and Table C.5.

The functional activity codes introduced from 2016 onwards, which would in principle allow expenditure to be classified by policy area, are not used in this study, because in the relevant communes they are predominantly populated for non-principal budgets and the resulting coverage in main budgets is too sparse to support cross-commune comparisons. We therefore use the account-group hierarchy (defined by the leading digits of the account code) as a proxy for the functional composition of spending.

Table A.1.2.1 summarizes the cleaning and sample-construction pipeline. Communes are identified by a constructed geo_code (the combination of the department code NDEPT and the INSEE municipal code), which is matched to the 2011 INSEE population reference. Observations that fail to match this reference are dropped: this category includes communes created after the 2011 reference year, communes that subsequently merged into communes nouvelles, vintage mismatches between the accounting and population files, and budget codes (notably annex budgets) that do not correspond to a stable 2011 administrative unit. The exclusion is conservative by design, since identification under the 2014 reform requires that each commune be observed at a fixed treatment status throughout the panel.

After matching, we restrict the dataset to budgets principaux: Communes' main budgets, by dropping observations with budget-type code CTYPE = 410 (annex and auxiliary budgets). Annex budgets are associated with specific services such as water, waste, or social housing, and are populated unevenly across communes; their inclusion would distort cross-commune comparisons of operating spending, particularly for smaller communes. We then restrict the dataset to Class 6 accounts (account numbers beginning with the digit 6), which capture booked operating expenditures.

For the regression discontinuity analysis, the dataset is further restricted to communes with population_2011 within the [500, 1,500] bandwidth around the 1,000-inhabitant treatment threshold, and to communes that are present in every year of the analysis period. A non-trivial set of communes (approximately 400) appears intermittently in the panel, primarily because of post-2014 mergers and splits associated with the commune nouvelle reform; among these, several absorbed neighboring communes during the sample period and exhibit anomalously large year-on-year increases in expenditure (in some cases exceeding twenty-five-fold from 2015 to 2016). We

drop a total of 306 communes affected by mergers or splits, leaving a balanced panel of 9,259 communes that are observed in all analysis years.

Table A.1.2.1. Sequential filters applied to the communal expenditure dataset.

Step	*Restriction applied*	*Result*
(0)	Annual balances comptables des communes, 2010–2019, all entries	Raw extract
(1)	Restrict to Class 6 accounts (account number begins with 6)	Operating expenditure entries
(2)	Construct GEOCODE = NDEPT + INSEE; exclude overseas departments	Mainland France only
(3)	Match to 2011 INSEE population frame; drop unmatched entries	Stable 2011 commune universe
(4)	Drop annex / auxiliary budgets (CTYPE = 410); retain budgets principaux	Main budgets only
(5)	Convert French decimal commas to numeric points; aggregate to commune-year	Clean panel
(6)	Restrict to communes with population_2011 < 3,500	33,325 communes
(7)	Restrict to communes with population_2011 in [500, 1,500] (RDD bandwidth)	approx. 9,967 communes per year
(8)	Drop communes with zero recorded 2011 population	Active 2011 communes
(9)	Exclude communes affected by post-2014 mergers, splits, or commune nouvelle reorganizations (306 communes); restrict to balanced panel (present in every analysis year)	9,259 communes

Note. Steps are applied cumulatively. Counts in the rightmost column refer to the resulting universe at each step; commune counts at steps (7)–(9) refer to the [500, 1,500] bandwidth used in the main analysis.

Within each Class 6 account, we measure realized annual expenditure using the standard accounting flows reported in the balances comptables. Three intermediate variables are constructed at the commune-year-account level:

- *spend_net = OBNETDEB – OBNETCRE: net budgetary flow: debit flows of the year (net of cancellations) less the corresponding credit flows (net of cancellations);*
- *spend_order = OOBDEB – OOBCRE: net "budgetary order" operations: technical/internal accounting movements that do not correspond to economic spending;*
- *spend_real = spend_net − spend_order: realized expenditure, which is the variable used throughout the analysis.*

We retain observations for which spend_real is negative due to in-period rebates: only a negligible number of commune-account-year cells (approximately 2) have negative realized values, and excluding them would introduce a small selection bias toward communes whose accounts are reversed in the same fiscal year. We accordingly interpret spend_real as a net flow rather than a strictly non-negative spending measure.

Notably, the Class 6 nomenclature is hierarchical: each account code can be read as a sequence of digits, with successive digits indicating progressively finer levels of disaggregation. We aggregate spend_real to three levels.

- Level 1 (L1): Two-digit family totals: The sum of all Class 6 accounts whose code begins with a given two-digit prefix. Five families are reported as covariates and outcomes (60_total, 61_total, 62_total, 64_total, 65_total).
- Level 2 (L2): Selected three-digit accounts within these families, chosen with reference to the literature on local procurement and gender-of-council composition (606, 611, 613, 615, 622, 625, 628, 641, 645, 647, 653, 657, 658).
- Level 3 (L3): Selected four-digit sub-accounts within 606 and 641 (6061, 6062, 6063, 6065, 6067 and 6411, 6413).

All remaining account codes are collapsed into the residual category and recorded as missing, since the analysis is restricted to the pre-specified set of theoretically motivated accounts. The full account list, with parent families and scope, is reported in Table C.4.

Table A.1.2.2. Account groups used as covariates and outcomes.

Level	*Code*	*Type*	*Parent family*	*Expense description*
L1	60	2-digit family total	60	Purchases & inventory changes
L1	61	2-digit family total	61	External services
L1	62	2-digit family total	62	Other external services
L1	64	2-digit family total	64	Personnel costs
L1	65	2-digit family total	65	Other current management charges
L2	606	Selected 3-digit account	60	Non-stocked purchases of materials and supplies
L2	611	Selected 3-digit account	61	Service contracts
L2	613	Selected 3-digit account	61	Rentals
L2	615	Selected 3-digit account	61	Maintenance and repairs
L2	622	Selected 3-digit account	62	Fees and intermediary services
L2	625	Selected 3-digit account	62	Travel, missions, and receptions
L2	628	Selected 3-digit account	62	Miscellaneous external services
L2	641	Selected 3-digit account	64	Staff remuneration
L2	645	Selected 3-digit account	64	Employer social security contributions

Level	*Code*	*Type*	*Parent family*	*Expense description*
L2	647	Selected 3-digit account	64	Other social charges
L2	653	Selected 3-digit account	65	Elected officials' indemnities and allowances
L2	657	Selected 3-digit account	65	Operating subsidies granted
L2	658	Selected 3-digit account	65	Miscellaneous current management charges
L3	6061	Selected 4-digit sub-account	606	Non-storable supplies (water and sanitation, electricity, heating)
L3	6062	Selected 4-digit sub-account	606	Non-stocked supplies (fuels, food, treatment products)
L3	6063	Selected 4-digit sub-account	606	Maintenance and small equipment supplies
L3	6065	Selected 4-digit sub-account	606	Books and media (libraries, media libraries)
L3	6067	Selected 4-digit sub-account	606	School supplies
L3	6411	Selected 4-digit sub-account	641	Permanent staff remuneration
L3	6413	Selected 4-digit sub-account	641	Non-permanent staff remuneration

Note. Codes follow the M14/M57 nomenclature. Where an account code does not appear in all four nomenclatures, the spend_real variable is computed only over the years in which it is defined; see Table C.5.

Twelve account codes do not appear in every nomenclature applied during the sample period. To preserve consistency across communes and years, we collapse these codes to their nearest common parent group prior to aggregation; the affected codes are listed in Table A.1.2.3.

Table A.1.2.3. Account codes that do not fully overlap across the M14, M14A, M57, and M57A nomenclatures.

Coverage	*Account codes*
Present in only one nomenclature	642, 643, 646, 656
Present in two nomenclatures	671, 674, 678, 687
Present in three of four nomenclatures	652, 665, 666, 667

Table A.1.2.4. reports the number of account-level rows, the number of unique communes, and the total realized Class 6 expenditure (in euro billions) at the wider <3,500-inhabitant cut, prior to the [500, 1,500] bandwidth restriction. Coverage is essentially complete in 2010–2014 (over 33,000 communes per year) and declines modestly in later years, primarily because cumulative post-2014 mergers reduce the number of independent communes in the population frame.

Table A.1.2.4. Annual coverage of the expenditure dataset, communes with population_2011 < 3,500.

Year	*Rows*	*Unique communes*
2010	1,589,532	33,320
2011	1,590,450	33,320
2012	1,599,184	33,323
2013	1,612,029	33,321
2014	1,622,007	33,321
2015	1,621,130	33,292
2016	1,644,132	32,514
2017	1,618,986	32,071
2018	1,614,792	32,017
2019	1,621,046	31,642

Table A.1.2.5. Annual coverage of the expenditure dataset, communes with population_2011 in [500, 1,500].

Year	Unique communes (500–1,500)
2010	9,967
2011	9,967
2012	9,966
2013	9,966
2014	9,966
2015	9,961
2016	9,771
2017	9,664
2018	9,651
2019	9,568

Table A.1.2.5. restricts attention to communes within the RDD bandwidth ([500, 1,500] inhabitants). Coverage fell from 9,967 communes in 2010 to 9,568 in 2019; the gap reflects communes that exited the panel because of mergers and reorganizations rather than data loss.

Following filtering and the panel-balance restriction, the final commune-year analysis dataset contains 9,259 unique communes observed in each year from 2010 to 2019. For each commune-year and each account group defined in Table A.1.2.2., we construct two outcome variables:

- *share_it = (commune i's spending in account group g, year t) / (commune i's total Class 6 spending in year t);*
- *spend_real_pc = (commune i's spending in account group g, year t) / population_2011, expressed in euros per capita.*

The share is the primary outcome because it is invariant to commune-level differences in overall operating expenditure, thereby isolating compositional shifts. The per-capita measure is reported as a robustness check; it relies on the assumption that the 2011 INSEE population is a sufficient denominator for normalization, which is plausible at the commune level over a ten-year window in which population changes are modest.

The same set of variables serves both as pre-treatment covariates in the balance analysis (over the 2010–2013 sub-period) and as outcomes in the post-reform regression discontinuity analysis (over the 2014–2019 sub-period). The full 2010–2019 panel is used for placebo and falsification checks reported.

For further insights on expenditure data, we also present the descriptive statistics for the post-reform period in Table A.1.2.6.

Table A.1.2.6. Descriptive statistics of post-reform expenditure outcomes, communes with population_2011 in [500, 1,500].

Panel A. Share of total commune spending

	Treatment = 0	***Treatment = 1***	***Total***
653 *Elected officials' indemnities and allowances*			
Mean	0.0730	0.0653	0.0707
Median	0.0704	0.0638	0.0681
Standard deviation	0.0241	0.0197	0.0231
Min.	0.0003	0.0052	0.0003
Max.	0.2275	0.1620	0.2275
Number of observations	6,488	2,771	9,259
Frequency	70.07%	29.93%	100.00%
611 *Service contracts*			
Mean	0.0189	0.0216	0.0198
Median	0.0084	0.0098	0.0088
Standard deviation	0.0266	0.0296	0.0276
Min.	0.0000	0.0000	0.0000
Max.	0.3473	0.3031	0.3473
Number of observations	4,609	2,213	6,822
Frequency	67.56%	32.44%	100.00%
615 *Maintenance and repairs*			
Mean	0.0812	0.0745	0.0792
Median	0.0744	0.0699	0.0729
Standard deviation	0.0365	0.0305	0.0350
Min.	0.0112	0.0070	0.0070

	Treatment = 0	*Treatment = 1*	*Total*
653 *Elected officials' indemnities and allowances*			
Max.	0.3258	0.2522	0.3258
Number of observations	6,488	2,771	9,259
Frequency	70.07%	29.93%	100.00%
645 *Employer social-security contributions*			
Mean	0.1283	0.1392	0.1316
Median	0.1306	0.1415	0.1343
Standard deviation	0.0322	0.0282	0.0314
Min.	0.0011	0.0004	0.0004
Max.	0.2934	0.2720	0.2934
Number of observations	6,486	2,769	9,255
Frequency	70.09%	29.91%	100.00%
657 *Operating subsidies granted*			
Mean	0.0365	0.0426	0.0384
Median	0.0242	0.0297	0.0257
Standard deviation	0.0407	0.0418	0.0411
Min.	0.0002	0.0011	0.0002
Max.	0.4536	0.3752	0.4536
Number of observations	6,484	2,770	9,254
Frequency	70.07%	29.93%	100.00%
6061 *Non-storable supplies (water, electricity, heating)*			
Mean	0.0542	0.0576	0.0552
Median	0.0518	0.0555	0.0529
Standard deviation	0.0195	0.0193	0.0195
Min.	0.0000	0.0003	0.0000
Max.	0.2057	0.1758	0.2057
Number of observations	6,483	2,769	9,252
Frequency	70.07%	29.93%	100.00%
6065 *Library books and media*			
Mean	0.0019	0.0021	0.0020
Median	0.0013	0.0016	0.0014
Standard deviation	0.0021	0.0021	0.0021
Min.	0.0000	0.0000	0.0000
Max.	0.0403	0.0314	0.0403
Number of observations	3,049	1,804	4,853
Frequency	62.83%	37.17%	100.00%
6067 *School supplies*			
Mean	0.0084	0.0081	0.0083
Median	0.0082	0.0079	0.0081
Standard deviation	0.0047	0.0038	0.0044
Min.	0.0000	0.0000	0.0000

	Treatment = 0	*Treatment = 1*	*Total*
653 *Elected officials' indemnities and allowances*			
Max.	0.0459	0.0349	0.0459
Number of observations	4,450	2,348	6,798
Frequency	65.46%	34.54%	100.00%

Panel B. Real per-capita spending (EUR)

	Treatment = 0	**Treatment = 1**	**Total**
653 *Elected officials' indemnities and allowances*			
Mean	40.00	38.47	39.55
Median	39.39	38.38	39.06
Standard deviation	8.72	8.32	8.63
Min.	0.42	5.78	0.42
Max.	98.28	97.02	98.28
Number of observations	6,488	2,771	9,259
Frequency	70.07%	29.93%	100.00%
611 *Service contracts*			
Mean	12.90	14.21	13.32
Median	4.75	5.92	5.08
Standard deviation	32.55	23.23	29.85
Min.	0.00	0.00	0.00
Max.	1139.19	372.60	1139.19
Number of observations	4,609	2,213	6,822
Frequency	67.56%	32.44%	100.00%
615 *Maintenance and repairs*			
Mean	49.89	47.90	49.30
Median	41.56	41.42	41.49
Standard deviation	40.15	34.37	38.52
Min.	4.73	3.90	3.90
Max.	1314.60	805.56	1314.60
Number of observations	6,488	2,771	9,259
Frequency	70.07%	29.93%	100.00%
645 *Employer social-security contributions*			
Mean	78.63	89.51	81.88
Median	71.71	82.82	75.16
Standard deviation	44.78	44.26	44.90
Min.	-1.35	-1.69	-1.69
Max.	1045.63	883.03	1045.63
Number of observations	6,486	2,769	9,255
Frequency	70.08%	29.92%	100.00%
657 *Operating subsidies granted*			
Mean	26.98	32.61	28.66

	Treatment = 0	Treatment = 1	Total
653 *Elected officials' indemnities and allowances*			
Median	13.46	17.43	14.60
Standard deviation	83.67	89.50	85.49
Min.	0.12	0.45	0.12
Max.	2907.15	2341.97	2907.15
Number of observations	6,484	2,770	9,254
Frequency	70.07%	29.93%	100.00%
6061 *Non-storable supplies (water, electricity, heating)*			
Mean	32.65	36.31	33.74
Median	28.91	33.11	30.29
Standard deviation	20.34	18.52	19.88
Min.	0.02	0.22	0.02
Max.	429.37	252.80	429.37
Number of observations	6,483	2,769	9,252
Frequency	70.07%	29.93%	100.00%
6065 *Library books and media*			
Mean	1.16	1.33	1.22
Median	0.77	0.99	0.85
Standard deviation	1.33	1.41	1.36
Min.	0.00	0.00	0.00
Max.	25.60	20.74	25.60
Number of observations	3,049	1,804	4,853
Frequency	62.83%	37.17%	100.00%
6067 *School supplies*			
Mean	4.98	4.90	4.95
Median	4.95	4.80	4.89
Standard deviation	2.82	2.18	2.61
Min.	0.00	0.00	0.00
Max.	37.73	16.89	37.73
Number of observations	4,450	2,348	6,798
Frequency	65.46%	34.54%	100.00%

Notes: Sample: post-reform commune-year observations at the commune level. Treatment = 1 if the 2011 INSEE population ≥ 1,000. Each commune-year is collapsed to the commune level by averaging within the commune over the post-reform years. "Number of observations" reports the number of communes with non-missing values for the corresponding outcome. "Frequency" reports the share of those communes falling on each side of the 1,000-inhabitant threshold. Shares are reported as proportions of total commune expenditure; per-capita amounts are deflated to constant euros.

1.3. Sample Representativeness

This appendix documents the representativeness of the two estimation samples used in the empirical analysis: (i) the procurement-based corruption sample, used to construct contract-level corruption indicators, and (ii) the communal expenditure sample, used to estimate effects on local fiscal outcomes. Both samples are evaluated against the eligible commune universe defined by the 2014 zipper-quota reform, excluding the merged communes.

Procurement-based corruption risk sample

Three data sources are combined to construct the corruption sample: (i) a GPPD dataset with commune-level corruption indicators aggregated from contract awards, (ii) a population reference file providing 2011 census population counts used to determine eligibility under the 2014 reform, and (iii) an exclusion file identifying communes affected by post-2014 mergers ("communes nouvelles"), therefore distorting the commune composition.

The sample is constructed in three steps. First, commune identifiers are harmonized into a consistent five-digit INSEE Geo Codes. Second, merged communes are excluded, and the population frame is restricted to communes with fewer than 3,500 inhabitants, defining the RDD bandwidth around the 1,000-inhabitant cutoff. Third, the procurement data are matched to this restricted population frame using the commune identifier.

Table A.1.3.1. reports the coverage of the corruption risk sample relative to the eligible commune universe, excluding merged communes. Every commune in the main analysis sample is contained within the eligible universe (overlap rate of 100%), so the procurement data introduce no out-of-frame observations. However, the sample is not a census of the universe: only 33.6% of eligible communes are observed in the procurement file, leaving 20,787 communes in the restricted population frame absent from the analysis sample.

Table A.1.3.1. Coverage of the corruption risk sample against the universe of eligible communes, excluding mergers.

Measure	*Value*
Unique communes	10,535
Coverage rate	33.6%
Coverage rate, communes below 1,000	24.5%
Coverage rate, communes at or above 1,000	69.9%

Coverage above the threshold (69.9%) is roughly three times higher than below it (24.5%). This asymmetry is consistent with the underlying structure of the procurement data: larger communes contract more frequently and are more often subject to publication thresholds, so they are mechanically more likely to appear in the procurement-based indicators.

Table A.1.3.2. reports population summary statistics for the corruption risk sample and for the corresponding eligible communes benchmark. The procurement sample is shifted toward larger communes throughout the distribution: the median commune in the sample has 1,725 inhabitants, compared with 379 in the restricted universe overall. Within each side of the threshold, the mean

population in the sample also exceeds the universe mean, although the gap narrows considerably above 1,000 inhabitants.

Table A.1.3.2. Population summary statistics, corruption risk sample versus the universe of eligible communes, excluding mergers

Sample	*N*	*Mean*	*Median*	*Std. dev.*
Eligible communes (all)	31,322	638.8	379.0	683.0
Corruption risk sample (all)	10,535	1,740.6	1,725.0	946.7
Eligible communes, < 1,000	25,037	350.9	283.0	246.7
Corruption risk sample, < 1,000	6,142	568.3	591.0	244.9
Eligible communes, ≥ 1,000	6,285	1,785.7	1,593.0	661.1
Corruption risk sample, ≥ 1,000	4,393	2,180.8	2,213.0	707.5

Two implications follow for the interpretation of estimates obtained on this sample. First, the procurement sample is a strict subset of the eligible commune universe rather than a census of it, and it overrepresents larger communes; particularly above the 1,000-inhabitant threshold. Second, because the regression discontinuity design identifies treatment effects locally at the cutoff, the relevant concern is not that the sample is unrepresentative of all small French communes in absolute terms. However, rather than selection into the procurement sample varying discontinuously across the threshold, selection into the procurement sample could itself vary discontinuously across the threshold. This concern is addressed in the main text through balance and density checks of observability around the cutoff; estimates from the procurement sample should be interpreted as local effects for the subpopulation of communes whose corruption risk is observable.

Communal Expenditure Sample

The communal expenditure sample is derived from a commune-level fiscal panel restricted to communes with populations between 500 and 1,500 inhabitants, based on 2011 INSEE census data. To construct a comparable representativeness benchmark, the eligible communes reference file is restricted to the same 500–1,500 bandwidth.

Table A.1.3.3. Coverage of the expenditure sample against the universe of eligible communes, excluding mergers (500–1,500 bandwidth).

Measure	*Value*
Unique communes	9,967
Coverage rate	99.5%
Coverage rate, communes below 1,000	99.5%
Coverage rate, communes at or above 1,000	99.3%

Coverage of the expenditure sample is reported in Table A.1.3.3. Unlike the procurement sample, the expenditure data provide near-census coverage of the eligible bandwidth: 9,967 of the 10,022 communes in the 500–1,500 range are included in the analysis file, corresponding to a coverage rate of 99.5%.

This near-complete coverage is essentially symmetric around the 1,000-inhabitant cutoff (99.5% below versus 99.3% above), so missingness is unlikely to be driving any discontinuity at the threshold.

Table A.1.3.4. reports population summary statistics. The expenditure sample tracks the restricted-population benchmark almost exactly, with means, medians, and standard deviations agreeing to the first decimal place both overall and within each side of the threshold.

Table A.1.3.4. Population summary statistics, expenditure sample versus the universe of eligible communes, excluding mergers (500–1,500 bandwidth).

Sample	*N*	*Mean*	*Median*	*Std. dev.*
Restricted population (all)	10,022	858.8	796.0	274.2
Expenditure sample (all)	9,967	858.7	796.0	274.2
Restricted population, < 1,000	7,013	704.9	682.0	140.4
Expenditure sample, < 1,000	6,979	705.0	683.0	140.5
Restricted population, ≥ 1,000	3,009	1,217.5	1,199.0	143.6
Expenditure sample, ≥ 1,000	2,988	1,217.6	1,199.0	143.7

Taken together, the evidence indicates that the expenditure sample is highly representative of the eligible commune universe within its 500–1,500 bandwidth, and that selection into the sample is essentially balanced across the 1,000-inhabitant cutoff. Differential observability is therefore not a first-order concern for estimates on the expenditure outcomes.

Implications for Identification

The two samples place different demands on the identification strategy. For the expenditure outcomes, near-complete and balanced coverage of the eligible universe means that the local-randomization interpretation of the regression discontinuity design proceeds with minimal additional assumptions: the sample on either side of the threshold reflects the population on either side of the threshold. For the procurement-based corruption risk outcomes, lower and uneven coverage means that estimated effects should be interpreted as effects for the subset of communes whose procurement activity is observable. The main text reports balance and continuity checks designed to assess whether observability varies discontinuously at the cutoff, and probes the sensitivity of estimates to alternative bandwidth choices and inclusion rules.

Appendix 2: Covariate Balances

This appendix reports the complete set of pre-treatment covariate balance tests summarized in Section 6.1. Each row corresponds to a separate local-linear regression discontinuity at the 1,000-inhabitant cutoff, estimated with MSE-optimal bandwidth, triangular kernel, and errors clustered at the commune level.

Table A.2.1. Demographic, economic, and productive covariates

Variable	*Coef.*	*Robust SE*	*P-value*	*Robust 95% CI*	*MSE-opt. bandwidht*	*N (L)*	*N (R)*
Minors p.c.	-0.012***	0.005	0.004	[-0.024, -0.005]	227	2,106	1,489
Adults p.c.	0.000	0.004	0.844	[-0.007, 0.009]	281	2,807	1,737
Households p.c.	0.009***	0.003	0.002	[0.004, 0.016]	202	1,812	1,357
Homeowners p.c.	0.001	0.003	0.721	[-0.004, 0.006]	275	2,728	1,703
Homeownership rate (pp)	-3.96***	1.19	0.000	[-6.77, -2.10]	191	1,693	1,293
Mean household income (€)	-29.8	69.0	0.718	[-160.0, 110.3]	406	4,717	2,302
Fiscal capacity	-1.89	1.63	0.298	[-4.90, 1.50]	394	4,512	2,256
Active labor force p.c.	-0.003	0.004	0.394	[-0.011, 0.004]	354	3,854	2,079
Salaried workers p.c.	-0.007	0.005	0.113	[-0.018, 0.002]	266	2,604	1,657
Service firms p.c.	0.0015	0.0011	0.167	[-0.0007, 0.0038]	224	2,064	1,469
Commerce firms p.c.	0.0012	0.0011	0.285	[-0.0009, 0.0032]	300	3,049	1,816
Construction firms p.c.	0.0006**	0.0003	0.012	[0.0002, 0.0013]	352	3,824	2,074
Industry firms p.c.	0.0003	0.0004	0.310	[-0.0004, 0.0011]	315	3,250	1,902
New firm creations p.c.	0.0004	0.0002	0.107	[-0.0001, 0.0009]	301	3,076	1,823
Educ. & health institutions p.c.	-0.292	0.698	0.681	[-1.655, 1.081]	131	1,090	958

Notes: Source: harmonized commune-level dataset combining INSEE's population census (population structure, household, ownership variables), DGFiP (mean household income, fiscal capacity, labor-force measures), and SIRENE (sectoral firm density, new firm creations, educational and health institutions). Per-capita variables are defined as the count divided by total population; the homeownership rate is the share of owner-occupied dwellings (in percentage points). Bandwidths are MSE-optimal (Calonico, Cattaneo, & Titiunik, 2014) *with ρ=1 and a triangular kernel; reported coefficients are conventional, and the confidence intervals are robust. Significance: *** *** p<0.001, ** p<0.01, * p<0.05.*

Table A.2.2. Pre-treatment expenditure accounts

Code	*Description*	*Outcome*	*Coef.*	*Robust SE*	*P-value*	*Robust 95% CI*	*MSE-opt. bandwidth*	*N (L)*	*N (R)*
60	Purchases and changes in inventories	*Share (pp)*	-0.102	0.402	0.702	[-0.761, 0.558]	160	5,464	4,600
60	Purchases and changes in inventories	*Per cap. (€)*	0.440	4.028	0.910	[-6.205, 7.086]	153	5,232	4,412
61	External services	*Share (pp)*	0.121	0.477	0.630	[-0.674, 0.917]	138	4,580	3,972
61	External services	*Per cap. (€)*	1.963	4.719	0.516	[-5.930, 9.855]	151	5,092	4,324
62	Other external services	*Share (pp)*	0.281	0.387	0.367	[-0.351, 0.913]	181	6,400	5,072
62	Other external services	*Per cap. (€)*	2.557	2.761	0.270	[-2.012, 7.127]	154	5,232	4,412
64	Personnel costs	*Share (pp)*	0.609	0.939	0.378	[-0.951, 2.169]	120	4,004	3,486
64	Personnel costs	*Per cap. (€)*	6.099	11.180	0.531	[-12.305, 24.503]	136	4,497	3,926
65	Other current management charges	*Share (pp)*	-0.474	1.025	0.445	[-2.213, 1.264]	119	4,004	3,496
65	Other current management charges	*Per cap. (€)*	-7.130	13.386	0.537	[-28.755, 14.495]	165	5,676	4,696
606	Non-stocked supplies	*Share (pp)*	-0.042	0.383	0.868	[-0.666, 0.583]	157	5,376	4,548
606	Non-stocked supplies	*Per cap. (€)*	0.995	3.847	0.786	[-5.339, 7.330]	149	5,004	4,272
613	Rentals	*Share (pp)*	-0.180*	0.122	0.097	[-0.382, 0.023]	161	4,166	3,674
613	Rentals	*Per cap. (€)*	-1.126	1.102	0.222	[-2.922, 0.670]	160	4,122	3,644
622	Professional fees and intermediaries	*Share (pp)*	0.028	0.068	0.503	[-0.087, 0.143]	158	5,342	4,522
622	Professional fees and intermediaries	*Per cap. (€)*	0.447	0.648	0.431	[-0.622, 1.515]	134	4,450	3,885
625	Travel and missions	*Share (pp)*	-0.010	0.048	0.860	[-0.088, 0.069]	150	3,225	2,793
625	Travel and missions	*Per cap. (€)*	-0.041	0.256	0.916	[-0.462, 0.381]	145	3,089	2,706
628	Miscellaneous external services	*Share (pp)*	0.031	0.233	0.984	[-0.349, 0.411]	165	5,406	4,458
628	Miscellaneous external services	*Per cap. (€)*	0.699	1.235	0.558	[-1.316, 2.713]	188	6,237	4,917
641	Staff remuneration	*Share (pp)*	0.332	0.641	0.459	[-0.731, 1.396]	121	4,032	3,513
641	Staff remuneration	*Per cap. (€)*	3.957	7.695	0.576	[-8.701, 16.616]	135	4,477	3,893
647	Other social charges	*Share (pp)*	0.067**	0.033	0.013	[0.009, 0.125]	116	3,371	3,087
647	Other social charges	*Per cap. (€)*	0.485***	0.195	0.003	[0.139, 0.830]	92	2,671	2,536
658	Various current management charges	*Share (pp)*	0.143**	0.072	0.027	[0.027, 0.259]	78	507	518

Code	*Description*	*Outcome*	*Coef.*	*Robust SE*	*P-value*	*Robust 95% CI*	*MSE-opt. bandwidth*	*N (L)*	*N (R)*
658	Various current management charges	*Per cap. (€)*	0.869**	0.467	0.034	[0.162, 1.575]	75	484	499
6062	Non-stocked supplies (food)	*Share (pp)*	-0.096	0.274	0.635	[-0.545, 0.352]	183	6,481	5,099
6062	Non-stocked supplies (food)	*Per cap. (€)*	0.224	1.949	0.898	[-2.957, 3.404]	147	4,933	4,211
6063	Maintenance & small equipment	*Share (pp)*	0.094	0.150	0.494	[-0.150, 0.337]	133	4,428	3,848
6063	Maintenance & small equipment	*Per cap. (€)*	1.051	1.190	0.373	[-0.904, 3.006]	138	4,580	3,972
6411	Permanent staff remuneration	*Share (pp)*	-0.113	0.707	0.795	[-1.265, 1.040]	127	4,284	3,704
6411	Permanent staff remuneration	*Per cap. (€)*	0.849	6.919	0.994	[-10.638, 12.336]	135	4,474	3,892
6413	Non-permanent staff remuneration	*Share (pp)*	0.718**	0.383	0.026	[0.051, 1.385]	104	3,142	2,855
6413	Non-permanent staff remuneration	*Per cap. (€)*	4.841**	2.611	0.026	[0.278, 9.405]	113	3,448	3,103

Notes: Tested for both outcome definitions: budget share (percentage points of total Class 6 operating expenditure) and real per-capita spending. 2-digit codes denote top-level chart-of-accounts categories; three and four-digit accounts are progressively detailed sub-accounts. Source: balances comptables des communes (Class 6 operating expenditure). Number of communes covered varies across accounts because reporting of detailed sub-accounts is uneven for smaller communes (ranging from 3,595 communes for account 658 to 9,259 for the two-digit categories). Bandwidths are MSE-optimal (Calonico, Cattaneo, & Titiunik, 2014) *with ϱ=1 and a triangular kernel; reported coefficients are conventional, and the confidence intervals are robust. Significance: *** p<0.001, ** p<0.01, * p<0.05.*

Table A.2.3. Discontinuity in council size at the 1,000-inhabitant threshold, pre- and post-policy

	(1) n_councillors (2008)	*(2) n_councillors (2014)*
RD coefficient	0.021	0.172*
Robust SE	(0.083)	(0.080)
P-value	0.812	0.033
95% CI	[-0.097, 0.1243]	[-0.0324, 0.2830]
MSE-opt. bandwidth	140.92	154.93
Eff. N (left)	1,174	1,316
Eff. N (right)	1,023	1,112

Notes. Each column reports a separate local-linear regression discontinuity estimate of the number of municipal councilors at the 1,000-inhabitant cutoff. Column (1) uses the 2008 council size, the last municipal-election cycle prior to the introduction of the zipper quota, and constitutes the pre-treatment balance test relevant for identification. Column (2) uses the 2014 council size, the first post-treatment cycle, and is reported as a complementary check. Specifications mirror the main outcome regressions: bandwidths are MSE-optimal (Calonico, Cattaneo, & Titiunik, 2014) *with ϱ=1 and a triangular kernel; reported coefficients are conventional, and the confidence intervals are robust. Significance: *** p < 0.01, ** p < 0.05, * p < 0.10.*

Table A.2.4. Discontinuity in share of female councillors at the 1,000-inhabitant threshold, pre- and post-policy

	(1) Share of female councillors (2008)	*(2) Share of female councillors (2014)*
RD coefficient	0.0026	0.978***
Robust SE	(0.0073)	(0.1343)
P-value	0.0079	0.000
95% CI	[-0.0082, 0.0228]	[0.6793, 1.2060]
MSE-opt. bandwidth	411.175	145.171
Eff. N (left)	4,763	1,222
Eff. N (right)	6,268	1,1057

Notes. Each column reports a separate local-linear regression discontinuity estimate of the share of female councilors at the 1,000-inhabitant cutoff. Column (1) uses the 2008 council data, the last municipal-election cycle prior to the introduction of the zipper quota, and constitutes the pre-treatment balance test relevant for identification. Column (2) uses the 2014 council data, the first post-treatment cycle, and is reported as a complementary check. Specifications mirror the main outcome regressions: bandwidths are MSE-optimal (Calonico, Cattaneo, & Titiunik, 2014) *with ϱ=1 and a triangular kernel; reported coefficients are conventional, and the confidence intervals are robust. Significance: *** p < 0.01, ** p < 0.05, * p < 0.10.*

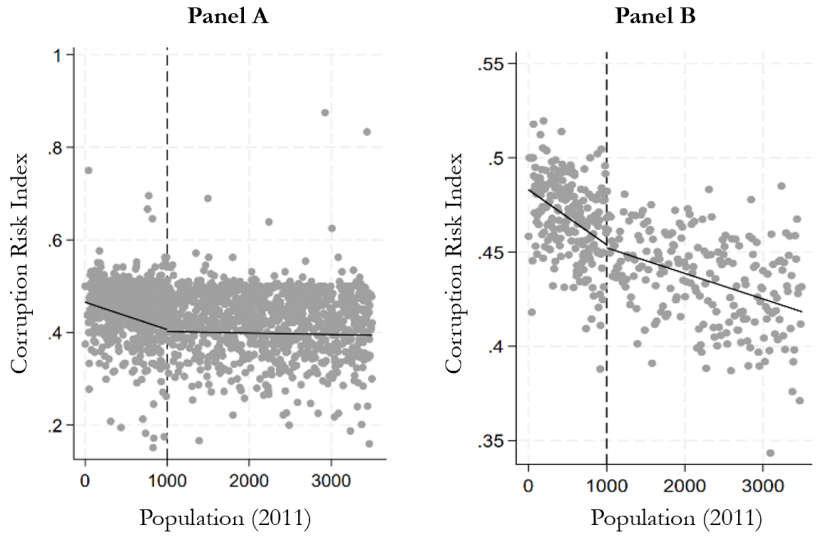
Appendix 3: Additional Specifications

Table A.2.1. Sharp regression discontinuity results for corruption risk index at tender and commune-year levels

	(1)	*(2)*
RD coefficient	0.009	0.0019
Robust SE	(0.040)	(0.0107)
P-value	0.793	0.805
95% CI	[-0.0657, 0.0925]	[-0.0183, 0.0236]
MSE-opt. bandwidth	388.287	274.003
Eff. N (left)	13,569	2,324
Eff. N (right)	14,517	2,191

Notes. Each column reports a separate local-linear regression discontinuity estimate corruption risk index at different levels of granularity. Column (1) uses the post-policy tender level data, clustering errors at commune level. Column (2) uses the post-policy commune-year level data, clustering tender level data and using the time fixed effects. Specifications mirror the main outcome regressions: bandwidths are MSE-optimal (Calonico, Cattaneo, & Titiunik, 2014) *with ϱ=1 and a triangular kernel; reported coefficients are conventional, and the confidence intervals are robust. Significance: *** p < 0.01, ** p < 0.05, * p < 0.10.*

Figure A.2.1. Sharp Regression Discontinuity Plots for Corruption Risk Index at tender and commune-year levels



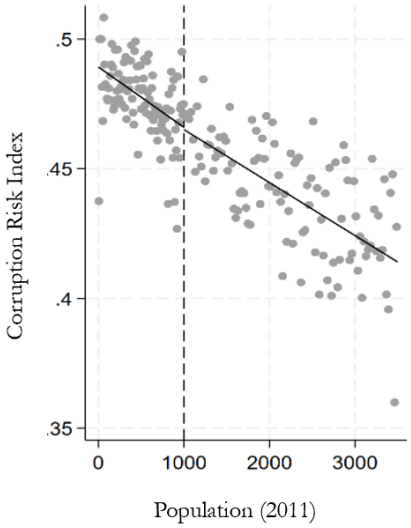
The figure displays a sharp regression discontinuity plot of the corruption risk index (CRI) at the 1,000-inhabitant threshold, using tender level and commune-year level observations. Points represent binned averages of the CRI, while fitted lines correspond to local polynomial regressions estimated separately on either side of the cutoff. Bins are selected through the default rdplot bin-selection method binselect (esmv), through a data-driven procedure mimicking the variance evenly spaced method using spacing estimators. The vertical line marks the population threshold at which the quota reform becomes binding.

Table A.2.2. Sharp regression discontinuity results for corruption risk index on restricted sample at commune-ruling period level

	CRI
RD coefficient	0.0007
Robust SE	(0.0088)
P-value	0.936
95% CI	[-0.0215, 0.0198]
MSE-opt. bandwidth	267.095
Eff. N (left)	1,014
Eff. N (right)	837

Notes. The table reports the local-linear regression discontinuity estimate of the effect of the zipper quota reform on the corruption risk index (CRI) for the subsample of treated communes that experienced a positive increase in female representation following the reform. The outcome is measured as the average CRI over the 2014–2020 ruling period. The running variable is municipal population measured relative to the 1,000-inhabitant threshold at which the reform becomes binding. Specifications mirror the main outcome regressions: bandwidths are MSE-optimal (Calonico, Cattaneo, & Titiunik, 2014) *with ϱ=1 and a triangular kernel; reported coefficients are conventional, and the confidence intervals are robust. Significance: *** p < 0.01, ** p < 0.05, * p < 0.10.*

Figure A.2.1. Sharp Regression Discontinuity Plot for Corruption Risk Index on restricted sample at commune-ruling period level



The figure displays a sharp regression discontinuity plot of the corruption risk index (CRI) at the 1,000-inhabitant population threshold for the subsample of treated communes that experienced a positive increase in female representation following the reform. Points represent binned averages of the CRI, while fitted lines correspond to local polynomial regressions estimated separately on either side of the cutoff. Bins are selected through the default rdplot bin-selection method binselect (esmv), through a data-driven procedure mimicking the variance evenly spaced method using spacing estimators. The vertical line marks the population threshold at which the quota reform becomes binding.

Appendix 4: Robustness Checks

Table A.4.1. Female representation robustness checks

Specification	Full sample		CRI sample		Expenditure sample	
	Female share	*Quota bite*	*Female share*	*Quota bite*	*Female share*	*Quota bite*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
A1: Fake cutoff = 750						
Estimate	-0.0069	-0.0055	-0.0144	-0.0056	-0.0044	-0.0068
Robust SE	(0.0111)	(0.0130)	(0.0193)	(0.0224)	(0.0123)	(0.0141)
Robust p	[0.5396]	[0.7829]	[0.4017]	[0.8186]	[0.7598]	[0.6678]
95% CI	[-0.0285; 0.0149]	[-0.0290; 0.0219]	[-0.0541; 0.0217]	[-0.0490; 0.0387]	[-0.0279; 0.0204]	[-0.0336; 0.0215]
BW	89.9	77.4	100.1	81.7	74.5	67.4
N eff. (L / R)	1235 / 992	1027 / 866	443 / 407	349 / 339	982 / 831	874 / 758
A2: Fake cutoff = 1250						
Estimate	0.0055*	0.0093	0.0022	0.0084	0.0073*	0.0082
Robust SE	(0.0040)	(0.0144)	(0.0062)	(0.0182)	(0.0046)	(0.0174)
Robust p	[0.0878]	[0.3476]	[0.6727]	[0.4348]	[0.0788]	[0.7132]
95% CI	[-0.0010; 0.0145]	[-0.0147; 0.0416]	[-0.0096; 0.0149]	[-0.0214; 0.0498]	[-0.0009; 0.0170]	[-0.0277; 0.0406]
BW	160.3	127.0	146.5	158.2	102.0	85.9
N eff. (L / R)	980 / 762	753 / 622	410 / 367	443 / 399	580 / 501	474 / 424
B1: Bandwidth = 0.75 × h						
Estimate	0.0589***	0.0496***	0.0472**	0.0569**	0.0606***	0.0493***
Robust SE	(0.0099)	(0.0152)	(0.0189)	(0.0250)	(0.0109)	(0.0189)
Robust p	[0.0000]	[0.0013]	[0.0103]	[0.0128]	[0.0000]	[0.0100]
95% CI	[0.0437; 0.0824]	[0.0191; 0.0787]	[0.0114; 0.0854]	[0.0133; 0.1113]	[0.0404; 0.0832]	[0.0117; 0.0858]
BW	96.9	124.3	106.9	136.5	78.9	84.2
N eff. (L / R)	790 / 726	1045 / 916	335 / 334	435 / 414	615 / 594	676 / 637
B2: Bandwidth = 1.25 × h						
Estimate	0.0603***	0.0578***	0.0526***	0.0628***	0.0596***	0.0499***
Robust SE	(0.0078)	(0.0115)	(0.0141)	(0.0181)	(0.0086)	(0.0143)
Robust p	[0.0000]	[0.0001]	[0.0011]	[0.0042]	[0.0000]	[0.0011]
95% CI	[0.0435; 0.0740]	[0.0238; 0.0687]	[0.0186; 0.0740]	[0.0163; 0.0872]	[0.0424; 0.0762]	[0.0186; 0.0746]
BW	161.6	207.2	178.1	227.6	131.6	140.4
N eff. (L / R)	1391 / 1171	1889 / 1436	601 / 549	808 / 674	1095 / 951	1171 / 1015
C: Pre-period placebo (2008)						

Specification	Full sample		CRI sample		Expenditure sample	
	Female share	*Quota bite*	*Female share*	*Quota bite*	*Female share*	*Quota bite*
	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*
Estimate	0.0027	-	-0.0088	-	0.0098	-
Robust SE	(0.0062)	-	(0.0104)	-	(0.0114)	-
Robust p	[0.4704]	-	[0.4562]	-	[0.3354]	-
95% CI	[-0.0077; 0.0166]	-	[-0.0280; 0.0126]	-	[-0.0114; 0.0333]	-
BW	410.3	-	395.8	-	142.1	-
N eff. (L / R)	4745 / 2422	-	1565 / 1099	-	1191 / 1030	-

Notes: Robustness of the first-stage discontinuities in female representation and quota bite at the 1,000-inhabitant threshold. Panels A1–A2 report estimates at placebo cutoffs of 750 and 1,250 inhabitants. Panel B1-B2 report estimates at 0.75 × h and 1.25 × h, where h is the MSE-optimal bandwidth. Bandwidths are MSE-optimal (Calonico, Cattaneo, & Titiunik, 2014) *with ϱ=1 and a triangular kernel; reported coefficients are conventional, and the confidence intervals are robust. Panel C reports the pre-period placebo (2008 election cohort). *** p<0.001, ** p<0.01, * p<0.05.*

Table A.4.2. Corruption risk robustness checks

Specification	*CRI*
A1: Fake cutoff = 750	
Estimate	-0.0008
Robust SE	(0.0157)
Robust p	[0.8229]
95% CI	[-0.0344; 0.0273]
BW	97.6
N eff. (L / R)	373 / 347
A2: Fake cutoff = 1250	
Estimate	0.0026
Robust SE	(0.0199)
Robust p	[0.6952]
95% CI	[-0.0312; 0.0468]
BW	139.3
N eff. (L / R)	343 / 319
B1: Bandwidth = 0.75 × h	
Estimate	-0.0042
Robust SE	(0.0190)
Robust p	[0.7470]
95% CI	[-0.0435; 0.0312]
BW	198.5
N eff. (L / R)	600 / 527
B2: Bandwidth = 1.25 × h	
Estimate	-0.0008
Robust SE	(0.0138)
Robust p	[0.7502]
95% CI	[-0.0314; 0.0226]
BW	330.9
N eff. (L / R)	1084 / 840
C: Pre-policy placebo (2011–2013)	
Estimate	-0.0039
Robust SE	(0.0098)
Robust p	[0.5771]
95% CI	[-0.0246; 0.0137]
BW	377.1
N eff. (L / R)	1340 / 990

Notes: Robustness of the commune-ruling-period RDD on the Corruption Risk Indicator (CRI). Panels A1–A2 report placebo cutoffs at 750 and 1,250 inhabitants. Panels B1–B2 report 0.75 × h and 1.25 × h bandwidth variations relative to the MSE-optimal bandwidth h. Panel C reports the pre-policy placebo using procurement data over 2011–2013. Bandwidths are MSE-optimal (Calonico, Cattaneo, & Titiunik, 2014) *with ϱ=1 and a triangular kernel; reported coefficients are conventional, and the confidence intervals are robust. *** p<0.001, ** p<0.01, * p<0.05.*

Table A.4.3. Expenditure composition robustness checks

Specification	L2 accounts					L3 accounts		
	653	*611*	*615*	*645*	*657*	*6061*	*6065*	*6067*
Panel 1 Outcome: Real per-capita spending								
A1: Fake cutoff = 750								
Estimate	-0.3747	-1.0197	2.6947	0.9718	16.5676	-0.3121	0.0587	0.6434
Robust SE	(0.9437)	(3.0920)	(6.0699)	(5.4920)	(12.1308)	(2.3305)	(0.2018)	(0.4037)
Robust 95% CI	[-1.9776; 1.7217]	[-7.3465; 4.7738]	[-9.6910; 14.1024]	[-8.5119; 13.0162]	[-4.7560; 42.7957]	[-4.7407; 4.3949]	[-0.3454; 0.4456]	[-0.1568; 1.4255]
A2: Fake cutoff = 1250								
Estimate	1.2256	3.9409	3.5364	-1.8860	-4.8512	4.4197	-0.3515	0.9085**
Robust SE	(1.4062)	(4.8436)	(13.1961)	(10.6486)	(31.9264)	(4.3455)	(0.4986)	(0.4709)
Robust 95% CI	[-1.4085; 4.1037]	[-5.0108; 13.9755]	[-23.7156; 28.0122]	[-23.7280; 18.0139]	[-69.6980; 55.4511]	[-3.8576; 13.1763]	[-1.3711; 0.5836]	[0.1011; 1.9471]
B1: Bandwidth = 0.75 × h								
Estimate	4.0336***	3.9373	-0.7124	0.5823	0.4878	0.1261	0.0478	0.1589
Robust SE	(2.2146)	(3.3992)	(4.5181)	(5.2811)	(4.7538)	(2.2888)	(0.2820)	(0.4959)
Robust 95% CI	[1.7676; 10.4487]	[-6.6978; 6.6270]	[-8.0163; 9.6944]	[-9.2116; 11.4897]	[-8.2572; 10.3775]	[-5.5870; 3.3850]	[-0.4616; 0.6438]	[-0.3489; 1.5950]
B2: Bandwidth = 1.25 × h								
Estimate	4.4016**	4.5354	0.6372	2.1270	-1.8096	0.4227	0.0157	0.0269
Robust SE	(1.7737)	(2.7209)	(3.6673)	(4.4194)	(5.4479)	(1.9323)	(0.2155)	(0.3835)
Robust 95% CI	[0.3798; 7.3327]	[-1.1894; 9.4765]	[-7.6557; 6.7199]	[-7.9226; 9.4012]	[-8.9618; 12.3936]	[-3.4230; 4.1514]	[-0.4092; 0.4353]	[-0.5041; 0.9994]
C: Pre-period placebo								
Estimate	1.1438	8.6976***	-0.3424	1.7423	-0.8588	-0.1618	0.0650	-0.1407
Robust SE	(1.1631)	(2.9453)	(2.9113)	(3.6957)	(5.1049)	(1.6293)	(0.1730)	(0.2912)
Robust 95% CI	[-1.6097; 2.9496]	[3.9532; 15.4987]	[-5.2695; 6.1426]	[-4.9985; 9.4882]	[-9.7679; 10.2430]	[-3.1430; 3.2437]	[-0.2566; 0.4215]	[-0.6860; 0.4555]
Panel 2 Outcome: Share of total spending								
A1: Fake cutoff = 750								
Estimate	-0.0015	-0.0007	-0.0015	0.0005	0.0024	-0.0004	0.0001	0.0011
Robust SE	(0.0025)	(0.0046)	(0.0041)	(0.0043)	(0.0072)	(0.0027)	(0.0003)	(0.0008)
Robust 95% CI	[-0.0063; 0.0037]	[-0.0107; 0.0073]	[-0.0108; 0.0054]	[-0.0074; 0.0094]	[-0.0130; 0.0153]	[-0.0058; 0.0049]	[-0.0005; 0.0007]	[-0.0006; 0.0025]
A2: Fake cutoff = 1250								
Estimate	-0.0016	0.0053	0.0029	-0.0071*	0.0030	0.0012	-0.0002	0.0007
Robust SE	(0.0038)	(0.0075)	(0.0055)	(0.0048)	(0.0084)	(0.0035)	(0.0004)	(0.0008)
Robust 95% CI	[-0.0097; 0.0053]	[-0.0087; 0.0207]	[-0.0082; 0.0132]	[-0.0177; 0.0010]	[-0.0132; 0.0197]	[-0.0056; 0.0081]	[-0.0010; 0.0006]	[-0.0008; 0.0024]

Specification	L2 accounts					L3 accounts		
	653	*611*	*615*	*645*	*657*	*6061*	*6065*	*6067*
B1: Bandwidth = 0.75 × h								
Estimate	0.0072*	0.0047	-0.0010	0.0019	-0.0040	0.0002	-0.0001	0.0004
Robust SE	(0.0040)	(0.0046)	(0.0053)	(0.0049)	(0.0051)	(0.0026)	(0.0004)	(0.0009)
Robust 95% CI	[-0.0004; 0.0154]	[-0.0085; 0.0095]	[-0.0068; 0.0138]	[-0.0079; 0.0113]	[-0.0130; 0.0068]	[-0.0058; 0.0046]	[-0.0009; 0.0007]	[-0.0006; 0.0028]
B2: Bandwidth = 1.25 × h								
Estimate	0.0098**	0.0032	-0.0012	0.0004	-0.0040	0.0001	-0.0002	0.0001
Robust SE	(0.0031)	(0.0038)	(0.0041)	(0.0039)	(0.0042)	(0.0020)	(0.0003)	(0.0007)
Robust 95% CI	[0.0009; 0.0131]	[-0.0020; 0.0128]	[-0.0091; 0.0069]	[-0.0049; 0.0104]	[-0.0121; 0.0044]	[-0.0040; 0.0039]	[-0.0008; 0.0004]	[-0.0007; 0.0019]
C: Pre-period placebo								
Estimate	0.0038	0.0075**	0.0003	0.0018	-0.0044	-0.0001	0.0000	-0.0004
Robust SE	(0.0023)	(0.0034)	(0.0038)	(0.0034)	(0.0044)	(0.0020)	(0.0003)	(0.0006)
Robust 95% CI	[-0.0014; 0.0076]	[0.0021; 0.0154]	[-0.0065; 0.0086]	[-0.0042; 0.0093]	[-0.0131; 0.0040]	[-0.0040; 0.0038]	[-0.0005; 0.0005]	[-0.0015; 0.0007]

Notes: Robustness of the commune-year RDD on communal expenditures, reported jointly for L2 accounts (653 elected officials' allowances; 611 service contracts; 615 maintenance and repairs; 645 social-security contributions; 657 operating subsidies) and L3 accounts (6061, 6065, 6067). Panel 1 reports estimates with real per-capita spending as the outcome; Panel 2 reports estimates with the account share of total spending as the outcome. Panels A1–A2 report placebo cutoffs at 750 and 1,250 inhabitants. Panels B1–B2 report 0.75 × h and 1.25 × h bandwidth variations. Panel C reports the pre-policy placebo test using 2011-2013 communal expenditure data. Bandwidths are MSE-optimal (Calonico, Cattaneo, & Titiunik, 2014) *with ϱ=1 and a triangular kernel; reported coefficients are conventional, and the confidence intervals are robust. *** p<0.001, ** p<0.01, * p<0.05.*

AI disclosure statement

This thesis has made limited use of generative artificial intelligence (AI), specifically tools such as ChatGPT and Claude, in accordance with institutional guidelines. AI was primarily used to support coding tasks in R, Python, and Stata, particularly in the context of extensive data cleaning, data matching, and visualization. All AI-assisted outputs were carefully reviewed and verified line by line to ensure accuracy and reliability.

In addition, AI tools were used for minor language support, including grammar checks and improvements to written clarity. No AI tools were used to generate substantive academic content. All ideas, analyses, results, interpretations, and conclusions presented in this thesis are entirely the authors' own original work. The use of AI has not influenced the intellectual contributions of this research.