UNIVERSITE BRETAGNE ECONOMIE LOIRE ET GESTION

UNIVERSITÉ DE RENNES 1

# THESE DE DOCTORAT DE

L'UNIVERSITE DE RENNES 1
COMUE UNIVERSITE BRETAGNE LOIRE

ECOLE DOCTORALE N° 597
*Sciences Economiques et sciences De Gestion*
Spécialité : Sciences Economiques

Par
**Kristýna DOSTÁLOVÁ**

## On the demand and supply of local public goods:

a quasi-experimental approach

**Thèse présentée et soutenue à Rennes, le 31 janvier 2019**
**Unité de recherche : Centre de Recherche en Economie et Management (CREM) – UMR CNRS 6211**

**Rapporteurs avant soutenance :**

| | |
|---|---|
| Etienne Farvaque | Professeur des Universités<br>Université de Lille |
| Sonia Paty | Professeure des Universités<br>Université de Lyon 2 |

**Composition du Jury :**

| | |
|---|---|
| Isabelle Cadoret | Professeure des Universités<br>Université de Rennes 1 |
| Etienne Farvaque | Professeur des Universités<br>Université de Lille |
| Sonia Paty | Professeure des Universités<br>Université de Lyon 2 |
| Benoît Le Maux<br>Directeur de thèse | Maître de Conférences, HDR<br>Université de Rennes 1 |
| **Invité(s)** | |
| Antti Moisio<br>Co-directeur de thèse | Chargé d'Etudes Economiques<br>OCDE, Paris |

This PhD thesis should not be reported as representing the views of University of Rennes 1. The views expressed are those of the author and do not necessarily reflect those of the university.

---

L'Université de Rennes 1 n'entend donner aucune approbation ni improbation aux opinions émises dans cette thèse. Ces opinions doivent être considérées comme propres à leur auteur.

# Acknowledgements

First and foremost, I would like to express my sincere gratitude to my thesis supervisor Professor Benoît Le Maux, who really is to blame for persuading me to take up this work and persevere until the very end. Without planting his initial suggestion at the end of my master's degree in Rennes, his guidance, encouragement and continuous support, the thesis literally would not be here. I deeply appreciate the time and effort invested in me and my work.

I am also grateful to my thesis co-supervisor Antti Moisio, for welcoming me in Helsinki, his aid in collecting the data and his insider input on the Finnish local public sector.

I am glad and honored that Isabelle Cadoret, Professor at University of Rennes 1, Etienne Farvaque, Professor at University of Lille, and Sonia Paty, Professor at University of Lyon 2, accepted to review the thesis and participate in the jury.

My sincere thanks also goes to Professor Fabio Padovano, whom I never expected to work with when I first met him as my lecturer years prior, for agreeing to co-author, with my thesis supervisor and I, the article based on my first chapter. Furthermore, I thank Professors Jean-Michel Josselin and Yvon Rocaboy at University of Rennes 1 for their input on the benchmarking methodology.

My daily life as a student and lecturer would have been next to impossible without the aid of the administrative faculty staff, in particular Anne L'Azou, Cécile Madoulet and Françoise Mazzoleni, and the financial support provided by the CREM that also gave me the opportunity to travel and participate in the scientific life outside the university. I thank my fellow PhD students and co-workers for keeping me company and allowing me to learn from their experience, and my students for putting up with my teaching and providing distraction when research seemed pointless and endless.

Finally, last but by no means least, I would like to thank my family back in the Czech Republic, for their support and for getting by without me in the meantime, even though it made them distinctly unhappy at times.

# Résumé général

Traditionnellement, l'existence du secteur public s'explique par le rôle prédominant du gouvernement dans la fourniture des biens et services publics et la correction des défaillances de marché. Néanmoins, tout comme les marchés peuvent échouer à fournir certains biens et services d'une manière optimale, l'intervention du gouvernement peut elle-même générer des inefficiences (*government failures* en anglais). C'est l'objet principale de l'économie publique d'analyser ces défaillances. Cette branche de l'économie contemporaine vise à appliquer la théorie microéconomique aux domaines du secteur public, notamment en se concentrant sur les questions d'efficience économique (assurer une allocation optimale des ressources) et celles d'équité (assurer une juste répartition des ressources). Une branche associée à cette discipline est celle de l'économie politique qui se concentre plus particulièrement sur les aspects institutionnels de la science politique, comme le rôle des syst'emes électoraux, celui des représentants élus, de la bureaucratie, etc. La méthodologie est en soit similaire, cependant elle mêle à la fois science politique et outils microéconomique dans le but de mieux appréhender le fonctionnement du secteur public dans son ensemble, en analysant à la fois le côté de la demande de biens et services publics mais aussi celui de l'offre. Cette approche est celle utilisée dans cette thèse de doctorat.

De manière analogue aux ressources et acteurs du secteur privé, le secteur public fonctionne avec des budgets permettant de fournir des biens et services aux citoyens. La manière dont ces budgets sont utilisés, c'est-à-dire distribués et dépensés, est à proprement parler le fruit d'une interaction entre offre et demande de biens et services publics. Dans les démocraties représentatives, la demande (les consommateurs des biens et services) est représentée par les citoyens. L'offre (les producteurs, les fournisseurs des biens et services) est quant à elle représentée par les gouvernements élus.

L'objectif de cette thèse de doctorat est de mieux appréhender les déterminants de la fourniture de biens et services publics locaux et d'évaluer l'adéquation entre l'offre de service public et la demande des citoyens. D'un point de vue théorique, les modèles analysant la demande de service public, tel le modèle de l'électeur médian ou l'hypothèse de Meltzer et Richard, démontrent que les caractéristiques de l'électorat ainsi que ses préférences sont les principaux déterminants de la fourniture de service public. Les modèles analysant l'offre de service public concluent au contraire que c'est la composition du gouvernement (nombre de partis, idéologie des élus) qui joue un rôle décisif. Cette recherche a pour objet de faire le lien entre ces deux types de théories, à travers une série d'essais empirique utilisant une approche quasi-expérimentale et se basant sur une analyse du secteur public local français et finlandais.

**Revue de la littérature**

La littérature en économie politique est riche de modèles expliquant la taille et la croissance du secteur public dans les démocraties directes et représentatives. Pour une revue exhaustive de la littérature, le lecteur peut par exemple se référer à l'ouvrage de Mueller (2003) et celui de Persson et Tabellini (2000).

Parmi les modèles de demande les plus fréquemment cités, on trouve notamment le théorème de l'électeur médian. L'idée est que la politique préférée par un électeur décisif (se trouvant au milieu de la distribution des électeurs et tournant ainsi les décisions publiques à son avantage) serait celle adoptée par un gouvernement souhaitant se faire

réélire. S'appuyant sur ce modèle, l'hypothèse de Meltzer et Richard construit ainsi un cadre simple mais pertinent pour expliquer la taille du secteur public. La manière dont les revenus sont distribués dans une économie déterminerait les préférences des électeurs en matière de redistribution. L'électeur au revenu médian serait alors décisif dans le choix de la politique redistributive.

Les modèles analysant l'offre de service public concluent au contraire que c'est la composition du gouvernement (nombre de partis, idéologie des élus, etc.) qui joue un rôle décisif. Parmi les facteurs fréquemment cités et analysés se trouvent les effets dits « partisans », c'est-à-dire résultant de l'idéologie du parti au pouvoir, amenant les politiciens à suivre un programme qui leur est propre et pouvant s'écarter des souhaits des citoyens considérés dans leur ensemble. Un autre facteur fréquemment analysé est celui de la concurrence politique telle que mesurée par exemple par le nombre de partis en compétition dans un régime politique donné. Comme pour le secteur privé, une concurrence plus intense serait mieux à même de réduire les défaillances du gouvernement, et empêcherait celui-ci de se comporter de manière monopolistique, exploitant son pouvoir et mettant en place des politiques dépensières, comportements dit de « Léviathan ». Toutefois, les modèles basés sur l'hypothèse dite de « gouvernement faible », ou *Weak Gouvernement Hypothesis* en anglais, mettent en garde contre une fragmentation politique trop aiguë : la présence d'un trop grand nombre de sujets politiques (élus, partis) peut également générer une augmentation des dépenses publiques, des déficits et un accroissement de la dette.

C'est sur les théories mentionnées ci-dessus que porte les analyses empiriques présentées dans cette thèse de doctorat. Afin de clarifier celles-ci, la littérature est présentée de manière plus détaillée ci-après.

*Les modèles de demande*

La clef de voute des modèles de demande est le théorème de l'électeur médian. Celui-ci établit que, si tous les électeurs ont des préférences unimodales sur un espace politique unidimensionnel (tel le niveau des dépenses publiques), alors l'alternative médiane sera un

vainqueur de Condorcet. En d'autres termes, le candidat proposant l'alternative médiane dans son programme sera garanti d'obtenir une majorité de voix. Les candidats politiques auront donc une incitation à proposer cette même alternative dans leur programme, quel que soit leur parti d'appartenance (Persson et Tabellini, 2000). Cette convergence des programmes peut également se généraliser à des cas où les décisions des électeurs sont erratiques. Pour Hinich (1977, 1978) et Lindbeck et Weibull (1987, 1993) par exemple, les programmes politiques convergent vers une alternative qui maximiserait une fonction de bien-être sociale.

Cette idée de convergence des politiques publiques a fait l'objet d'une attention toute particulière en économie politique, notamment en matière de fourniture de biens publics (le modèle de la demande introduit par Barr et Davis, 1966; Borcherding et Deacon, 1972; Bergstrom et Goodman, 1973) et de redistribution (l'hypothèse de Meltzer et Richard initiée par Romer, 1975; Roberts, 1977; Meltzer et Richard, 1981). Ces deux approches établissent un lien entre l'affiliation politique des électeurs et leur situation personnelle (Weatherford, 1978; Lewis-Beck et Stegmaier, 2000).

D'après le théorème de l'électeur médian, peu importe le parti au pouvoir (Besley et Case, 2003). Lors d'élections suffisamment concurrentielles, les candidats politiques maximisent le nombre de voix espérés et sont supposés tenir leurs promesses de campagne une fois élus, agissant ainsi pour le bien-être de leur électorat. Nombreuses sont les études empiriques en faveur de cette conclusion : le revenu médian, ou de manière équivalente le revenu moyen, jouent fréquemment un rôle significatif dans l'explication des niveaux de dépenses publiques, à la fois au niveau local (voir, par exemple, Ahmed et Greene, 2000; Aronson *et al.*, 2000; Guengant *et al.*, 2002) ou central (Arpaia et Turrini, 2008; Kolluri *et al.*, 2000; Magazzino, 2012; Pradhan et Bagchi, 2012; Wahab, 2004). De même, l'hypothèse de Meltzer et Richard a été récemment validée par Alesina *et al.* (2000), Borge et Rattsø (2004), Mattos et Rocha (2008), Milanovic (2000), et Mohl et Pamp (2009).

Cependant, de nombreux auteurs soutiennent que les résultats empiriques précédemment cités ne sont pas suffisant pour conclure à la suprématie des modèles de demande. Si les résultats montrent que les choix de politique publique dépendent jusqu'à

un certain degré des caractéristiques démographiques et économiques des électeurs, cela n'exclut pas la possibilité que d'autres facteurs soient également en jeu. La question de savoir à quel point les choix de politique publique reflètent les préférences des citoyens reste donc en suspens (Matsusaka, 2010).

*Les modèles d'offre*

Une part importante de la littérature en économie politique considère que c'est plutôt le côté de l'offre qui détermine la taille et la performance du secteur public. Nombreux sont les facteurs considérés, tel l'idéologie du parti au pouvoir, la formation de coalitions et leur degré de coopération, les manipulations électorales, les comportements de Léviathan ou de recherche de rente (théories de la bureaucratie), les cycles politiques, etc. Cette thèse de doctorat porte plus particulièrement sur l'existence d'effets partisans et l'impact de la fragmentation politique sur le secteur public local, comme discutés ci-dessous.

Plusieurs arguments théoriques suggèrent que l'idéologie et l'appartenance politique des élus à un parti affecte la taille du secteur public. Cette catégorie de modèles voit les politiciens comme des *citoyens-candidats* (Besley et Coate, 1997; Osborne et Slivinsky, 1996) ou des *partisans politiciens* (Persson et Tabellini, 2000) ayant des caractéristiques et des fonctions de préférences spécifiques. Dans ce contexte, les candidats ont leur propre idéologie et sont incités à révéler leurs préférences au cours des campagnes électorales. Ils sont également incités à mettre en œuvre leur programme une fois élus, la crédibilité étant un facteur crucial pour gagner les prochaines élections. Sous certaines conditions, une divergence de politique peut donc être observée d'un parti politique à l'autre, faisant ainsi de l'idéologie du gouvernement une variable pertinente pour expliquer les choix de politique publique (Padovano, 2013).

La question de l'existence d'effets partisans a été étudiée en profondeur, comme illustrée dans Jackson et Kingdon (1992), Bender et Lott (1996), et Imbeau *et al.* (2001). L'hypothèse a récemment été étayée par des preuves empiriques aux États-Unis (Besley et Case, 2003; Bjørnskov et Potrafke, 2012; Pickering et Rockey, 2013), en Espagne (Solé-Ollé,

2003), en Norvège (Borge et Rattsø, 2004), en France (Foucault *et al.*, 2008; Le Maux *et al.*, 2011), en Italie (Padovano et Petrarca, 2014; Santolini, Santolini, 2008) et dans les pays de l'OCDE (Pickering et Rockey, 2011). Dans la plupart des cas, les études montrent que les gouvernements à gauche ont tendance à augmenter les taux d'imposition et les dépenses publiques davantage que leurs homologues à droite.

Le deuxième facteur à prendre en compte est le nombre de partis en concurrence dans un régime politique donné. Le manque de concurrence politique peut augmenter la marge de manœuvre des élus qui peuvent à leur tour essayer de maximiser leur prestige, leur pouvoir ou tout autre type d'avantage. Ce type de comportement est décrit dans l'hypothèse du Léviathan (Niskanen, 1971; Brennan et Buchanan, 1980) selon laquelle le gouvernement maximise la taille du secteur public au détriment des citoyens, d'où l'importance de limiter le pouvoir des élus à travers des institutions et règles budgétaires adéquates.

Le rôle des contraintes constitutionnelles sur la taille du secteur public a par exemple été étudié empiriquement dans le contexte de restrictions d'assiette fiscale (Nelson, 1986; Shapiro et Sonstelie, 1982), de décentralisation et fédéralisme (Deacon, 1979; Marlow, 1988; Mehay, 1984; Mehay et Gonzales, 1985; Oates 1985; Schneider, 1986), ou dans un contexte où le citoyen est amené à s'informer et exercer un contrôle plus conséquent (Duncombe *et al.*, 1997; Hayes *et al.*, 1998). Récemment, cette hypothèse a fait également l'objet d'une analyse dans le cadre des politiques environnementales (voir, par exemple, Duit *et al.*, 2016, et Lundqvist, 2001, sur le « Léviathan Vert »).

Alors que la concurrence politique est généralement considérée comme un moyen de discipliner les représentants élus et d'orienter les politiques vers plus d'efficience (par exemple, Ashworth *et al.*, 2014; Ferris *et al.*, 2008; Padovano et Ricciuti, 2009), la fragmentation politique peut également conduire à des gouvernements « plus faibles », où le manque de coopération entre de trop nombreux sujets politiques pourrait amener à une hausse des dépenses publiques, des déficits publics et de la dette. Cette hypothèse initiée par Roubini et Sachs (1989a, 1989b) est décrite dans trois catégories distinctes de modèles (Ashworth *et al.*, 2005). Premièrement, *les modèles d'usure* (Alesina et Drazen, 1991; Alesina et Perotti, 1995) prévoient une augmentation de l'endettement à court terme, par

manque de coordination face aux chocs économiques. Deuxièmement, dans *les modèles basés sur une tragédie des communs*, le pluralisme politique peut amener les politiciens qui ont chacun leur stratégie et préférence propre en matière de politique publique à dépenser et taxer au-delà du niveau socialement désirable (Weingast *et al.*, 1981). Cet effet peut durer à long terme (Velasco, 2000) et s'intensifier lorsque le nombre de parties au pouvoir augmente (Olson, 1993). Enfin, *les modèles dits de dette stratégique* considèrent que la fragmentation et l'incertitude quant aux résultats électoraux sont des facteurs explicatifs de la hausse du secteur public, amenant à des comportements stratégiques de la part des élus qui se servent des politiques budgétaires et de l'endettement pour accroître leur probabilité de réélection (Alesina et Tabellini, 1990; Persson et Svensson, 1989; Tabellini et Alesina, 1990).

L'hypothèse de gouvernement faible trouve généralement un appui dans la littérature empirique : par exemple, dans Alesina et Perotti (1995), Ashworth et Heyndel (2005), Borge et Rattsø (2002), Le Maux et Zhang (2013), Padovano et Venturi (2001), Perotti et Kontopoulos (2002), ou Rattsø et Tovmo (2002), bien que mitigés (voir, par exemple, Borelli et Royed, 1995; Freitag et Sciarini, 2001; Hahm *et al.*, 1996).

## Enjeux méthodologiques

### *Endogénéité de la composition du gouvernement*

Une question est toutefois en suspens dans la littérature. Il reste à savoir si les résultats empiriques décrits précédemment mettent en avant les mécanismes tels qu'ils sont effectivement décrits dans les modèles d'offre, ou s'il s'agit d'un impact plus indirect de la demande, la manière dont sont distribués les électeurs pouvant elle-même affecter la composition du gouvernement. Bien que cette question ait été partiellement étudiée, il se pourrait en effet que la composition du gouvernement soit elle-même une variable endogène, c'est-à-dire dépendante des caractéristiques propres des citoyens et de leurs préférences en matière de politique publique. Si tel était le cas, un biais de sélection potentiel serait présent dans les données, biais qu'il serait nécessaire de résoudre. C'est un des enjeux

méthodologiques principaux auxquels la présente thèse est confrontée. Résoudre ce défi est d'autant plus important que le but primaire de la thèse est de mieux discerner les rôles joués par les mécanismes évoqués ci-dessus, ceux de l'offre et de la demande, dans l'explication des disparités de dépense publique.

La littérature théorique fournit quelques arguments en faveur de cette possible endogénéité de la composition du gouvernement. Par exemple, d'après Hinich (1978) et Lindbeck et Weibull (1987), les candidats politiques disposeraient d'une information imparfaite en ce qui concerne les préférences des électeurs. Sous certaines conditions (concurrence entre deux parties, symétrie et fonctions de vote probabiliste spécifiques), l'équilibre politique atteint serait le point utilitariste. Ce dernier caractérise une situation où l'idéologie des candidats en compétition reflète la répartition même des préférences des électeurs. Ce résultat peut être généralisé à un jeu répété où le parti au pouvoir et le parti d'opposition ont tous deux leurs réputations à préserver, obligeant le parti au pouvoir à respecter ses promesses électorales (voir, par exemple, Kreps et Wilson, 1982, ainsi que d'autres développements dans la littérature sur la concurrence imparfaite). Milesi-Ferretti *et al.* (2002), étendent cette logique à un cadre où les dépenses et services publics sont d'une nature locale. Ici, le gouvernement élu répartit les dépenses géographiquement, de manière à maximiser l'utilité des citoyens dans les diverses juridictions, afin d'accroître sa probabilité de réélection.

Les groupes d'intérêt (groupes de pression ou lobbies), peuvent également influencer les choix de politique publique, soit en influençant directement les résultats des élections (par exemple via des contributions aux campagnes électorales), soit par des pressions exercées sur les politiciens une fois ceux-ci élus (entretiens, pétitions, etc.). Dans ces modèles, plus le groupe d'intérêt est puissant, plus il tournera la politique à son avantage. Il peut notamment exister une relation entre taille du groupe (le nombre de ses membres) et son influence. Le sens de l'effet est toutefois incertain : les groupes les plus imposants peuvent d'un côté mobiliser plus de ressources mais de l'autre côté avoir plus de difficultés à se coordonner (paradoxe de l'action collective). Par exemple, en utilisant des données suisses, Stadelmann *et al.* (2013) constatent que les groupes d'intérêts joue un rôle significatif sur les

choix de politique publique, ces dernières pouvant ainsi s'écarter du point idéal médian.

En outre, d'autres facteurs peuvent potentiellement affecter la marge de manœuvre des représentants élus, tels que le nombre de candidats en compétition, le mode de scrutin, les cycles électoraux, les caractéristiques personnelles des candidats (leur personnalité, leur charme, leur réputation, leurs caractéristiques raciales, leur religion, leur situation de famille, etc.). Tous ces facteurs peuvent ainsi impacter le pouvoir discrétionnaire des élus et les possibilités de divergence des programmes (voir, par exemple, Calvert, 1985; Hansson et Stuart, 1984; Wittman, 1973).

La question de l'endogénéité de la composition du gouvernement a également fait l'objet d'un examen approfondi en science politique. Pour un aperçu de ces théories, le lecteur peut se référer aux études de Martin et Stevenson (2001, 2010), Glasgow *et al.* (2012), Glasgow et Golder (2015) ou Laver et Benoit (2015). Concernant les gouvernements locaux, plusieurs études traitent également de ces questions : Bäck (2008), Debus et Gross (2016), Skjæveland *et al.*, (2007).

Par ailleurs, une nouvelle catégorie de modèles en économie politique suggère des interactions beaucoup plus complexes entre électeurs et politiciens. Bénabou (2000), Horstmann et Scharf (2000), Persson (1995), et Pickering et Rockey (2011) notent qu'une plus grande coopération entre riches et pauvres peut générer divers bénéfices tels une plus grande croissance économique. Cela oriente en retour les plus riches vers plus de redistribution. Le lien entre l'électorat d'un côté et l'idéologie du gouvernement de l'autre pourrait donc être associé à ce degré de coopération. On peut citer d'ailleurs à ce sujet les travaux sur l'altruisme des électeurs (*sociotropic voters*) qui pourraient voter non seulement de manière à maximiser leur propre utilité, mais aussi de manière à améliorer le bien-être de la collectivité dans son ensemble. L'idéologie du gouvernement au pouvoir dépendraient alors des préférences de l'ensemble des électeurs (Kinder et Kiewiet, 1979; Wright, 1986), et ainsi de manière endogène de l'empathie et altruisme de ces derniers.

*Usage de l'approche quasi-expérimentale*

La question de l'endogénéité de la composition du gouvernement, donnant lieu à un biais de sélection potentiel, nécessite de faire intervenir des stratégies empiriques appropriées. En effet, la différence de politique observée entre deux gouvernements élus pourrait être la somme de deux éléments: (1) les effets d'offre à proprement parler et (2) les effets de demande dues aux caractéristiques propres de l'électorat. Le problème est que le deuxième élément est un facteur de confusion pouvant affecter conjointement les choix de politique publique et la composition même du gouvernement.

Pour quantifier le pouvoir explicatif des modèles d'offre, il est donc nécessaire de corriger les biais éventuels, notamment en comparant des juridictions aux caractéristiques comparables en matière de demande, de manière à isoler les effets d'offre. L'approche utilisée dans cette thèse est celle de la quasi-expérimentation et, plus précisément, celle de la régression en discontinuité (*regression discontinuity design*, RDD) et des méthodes d'appariement sur score de propension (*propensity score matching*, PSM).

Dans le contexte de la thèse, les deux approches RDD et PSM sont similaires dans leur intuition. La méthode RDD a pour objet de comparer différentes juridictions selon un seuil prédéfini (par exemple basé sur le pourcentage de siège à gauche ou à droite). Les observations trop éloignées de ce seuil sont exclues de l'analyse. L'idée est qu'en examinant uniquement les juridictions situées à proximité du seuil, ces dernières seront similaires, ce qui permettra ainsi d'éliminer le biais de sélection évoqué. La méthode PSM consiste de manière similaire à apparier les juridictions sur la base de leurs caractéristiques observables. Elle repose sur l'estimation de scores qui résument les différences initiales et permettent la construction d'un groupe de traitement et d'un groupe de contrôle similaires (voir, par exemple, Hirano et Imbens, 2004; Imbens et Wooldridge, 2009).

Une autre méthodologie employée dans cette thèse de doctorat est celle de l'analyse comparative dite *benchmarking*. L'objectif est ici de tenir compte des différents types d'usagers concernés par les services publics et d'isoler l'effet de la demande par comparaison à une juridiction de référence appelé « benchmark ». L'idée est que selon le type d'usagers,

par exemple un chômeur ou une personne âgée, les coûts supportés par la collectivité sont différents. En isolant l'impact de la demande, on peut analyser les autres facteurs explicatifs des disparités de dépenses. L'approche permet ainsi d'évaluer la performance des collectivités locales en matière de dépenses publiques tout en tenant compte du biais de sélection préalablement cité.

**Résultats principaux**

Les analyses empiriques des secteurs publics locaux en France et en Finlande présentées dans cette thèse amènent à des conclusions diverses et variées. Dans le premier chapitre intitulé « *Ideology or voters? A quasi-experimental test of why left-wing governments spend more* », la thèse teste la présence d'effets partisans (c'est-à-dire de l'idéologie du parti au pouvoir) sur les niveaux de dépense publique en France au niveau départemental. Les résultats montrent que les gouvernements ayant une majorité de sièges à gauche et qui sont confrontés à la même situation économique que les gouvernements ayant une majorité de sièges à droite ont des niveaux de dépenses similaires, en particulier en ce qui concerne les dépenses d'aide sociale. Ce résultat exclut pour ce type de dépense la présence d'effets partisans et suggère ainsi que la demande des citoyens joue un rôle déterminant dans le choix des politiques publiques.

Dans le deuxième chapitre intitulé « *Does the size of the largest party matter? Endogenous fragmentation, political competition, and local public expenditures in Finland* », la thèse examine l'impact de la fragmentation politique (rôle de la répartition des sièges entre majorités et coalitions) sur les dépenses publiques. L'étude porte sur les municipalités finlandaises, l'unique niveau de gouvernement local en Finlande. Les résultats montrent qu'il n'y a pas de différence significative entre les dépenses publiques des gouvernements majoritaires et celles des gouvernements minoritaires. Il existe toutefois un point de rupture dans les données : lorsque le parti au pouvoir détient plus de 60 % des sièges, alors le niveau de dépense publique augmente de manière significative.

Dans le troisième chapitre, « *Performance of local governments: Benchmarking analysis of social welfare provision in France* », le rôle de la demande des citoyens sur la fourniture de bien public est examiné plus en profondeur. L'étude se fonde sur une analyse de type benchmarking et cherche à mieux comprendre les disparités de dépenses d'aide sociale des départements français. Celle-ci démontre que les niveaux de dépense par bénéficiaire dépendent non seulement du nombre total d'usagers mais aussi de la distribution des usagers entre les quatre risques que sont le chômage, l'aide à la famille, l'aide aux personnes handicapées et aux personnes âgées.

Pour conclure, la thèse montre que la demande des électeurs et l'offre de la part des gouvernements élus sont étroitement liées. D'après les données, les caractéristiques de l'électorat affectent conjointement la composition du gouvernement et, de manière plus indirecte, les politiques publiques. La demande apparaît donc comme un des facteurs principaux des disparités de dépenses publiques. Bien qu'optimiste, cette vision des démocraties représentatives locales reste à relativiser. Des inégalités peuvent en effet apparaître dans la fourniture des biens locaux, tout simplement de par l'hétérogénéité de la demande d'une juridiction à l'autre. Selon la demande, les gouvernements locaux peuvent en effet être confrontés à des contraintes budgétaires différentes comme en témoigne le cas de l'aide sociale en France étudié dans cette thèse.

# Contents

**Acknowledgements** **iv**

**Résumé général** **vii**

**Contents** **xxi**

**General introduction** **1**

**1 Ideology or voters? A quasi-experimental test of why left-wing governments spend more** **13**

1.1 Introduction . . . . . 13

1.2 The model . . . . . 16

1.3 Data description . . . . . 23

1.4 Evidence of a selection bias . . . . . 28

1.4.1 Endogenous government ideology . . . . . 28

1.4.2 Public spending and share of welfare recipients . . . . . 32

1.5 Test for partisan effects . . . . . 36

1.5.1 Regression discontinuity design (RDD) . . . . . 36

1.5.2 Propensity score matching (PSM) . . . . . 38

1.5.3 Robustness of the results . . . . . 40

1.5.4 Comparison of RDD and PSM . . . . . 45

1.6 Conclusion . . . . . 46

**2 Does the size of the largest party matter? Endogenous fragmentation, political competition, and local public expenditures in Finland 49**

2.1 Introduction . . . 49

2.2 Literature review . . . 52

2.3 Finnish local public sector . . . 55

2.3.1 The role of municipalities in Finland . . . 55

2.3.2 Changes in the structure of Finnish municipalities . . . 57

2.3.3 Municipal elections . . . 58

2.4 Identification strategy . . . 60

2.4.1 Measure of government fragmentation . . . 60

2.4.2 Assumptions and threats to identification . . . 62

2.5 Empirical results . . . 66

2.5.1 Estimation of propensity scores . . . 66

2.5.2 Propensity score matching . . . 68

2.5.3 Tests for endogenous thresholds . . . 72

2.6 Robustness checks . . . 76

2.6.1 Propensity score matching at the 59% threshold . . . 76

2.6.2 The role of largest Finnish parties . . . 77

2.6.3 Placebo tests . . . 79

2.7 Conclusion . . . 80

**3 Performance of local governments: Benchmarking analysis of social welfare provision in France 83**

3.1 Introduction . . . 83

3.2 French Departments and social welfare provision . . . 86

3.2.1 Context of the study . . . 86

3.2.2 Inter-departmental and inter-service disparities . . . 87

3.3 Population size and heterogeneity of users . . . 90

3.4 Measuring the performance of French Departments . . . . . 93
3.4.1 Benchmarking methodology . . . . . 93
3.4.2 Computing scores . . . . . 95
3.4.3 Econometric analysis of performance scores . . . . . 98
3.5 Discussion and policy implications . . . . . 102
3.5.1 Financial difficulties of the French Departments . . . . . 102
3.5.2 Urban and rural areas . . . . . 103
3.5.3 Intergovernmental grants and fiscal equalization . . . . . 107
3.6 Conclusion . . . . . 110

**General conclusion** **113**

***Appendices*** **119**

***References*** **151**

***List of Figures*** **154**

***List of Tables*** **156**

# General introduction

Traditional explanations for the existence of the public sector argue its role in providing public goods and correcting market failures. Nevertheless, just as markets may fail to provide efficient allocation of goods and services on their own, to be corrected by a government intervention, governments may in their turn fail and create inefficiencies where they would not exist in a free market. Public economics, applying microeconomic theory to the study of government policies in terms of economic efficiency (simply put, maximization of benefits from resources) and equity (fairness), is concerned with both size and performance of the public sector, including so-called government failures. A quite distinct approach is furthermore taken by public choice/political economy, focusing on the traditional topics of political science, including the role of electoral systems, politicians, bureaucracy, etc. The methodology is that of economics, however, blending politics and economics in order to better understand the public sector which encompasses both.

In the political economy literature, numerous approaches propose theoretical explanation for the size and growth of government in direct and representative democracies. The reader can for instance refer to Mueller (2003) and Persson and Tabellini (2000) for an extensive literature review. Depending on whether the key explanatory factors are presumed

to be identified among the characteristics of the electorate or, rather, of the elected governing bodies, two main groups of theories can be distinguished. So-called *demand driven models* propose that the characteristics and preferences of the electorate are the main forces driving public decisions. A pivotal assumption here is the median voter theorem, which explains how the voters'—or, more precisely, the decisive median voter's—preferred outcomes will be adopted as public policies. Building on the median voter model, the Meltzer and Richard hypothesis constructs a simple yet elegant framework that explains the size of the government using voters' incomes and their preference for further redistribution of public resources (and, analogously, taxation), depending on which will benefit them more.

So-called *supply side models* argue instead that the composition of governments (e.g., number of parties, ideology of elected officials) is the key explanatory factor. Among the principal premises is the partisan hypothesis, which sees politicians as pursuing own diverging agendas, rather than simply attempting to reflect the (median) voters' preferences. Voters then make rational choices based on the weights they place on particular issues, such as unemployment or taxation. A related concern is the role of political competition, i.e., the number of parties that compete and form the political spectrum. Analogically to private markets, competitive supply would prevent the government from behaving like a Leviathan—an ever-growing self-serving monopolist that exploits its power over the citizens, who have effectively lost all control over it. The Weak Government Hypothesis nevertheless cautions against increased government fragmentation, which would imply the cooperation of too many political subjects and lead to higher public spending, and thus deficits and debt.

Since both the theoretical and the corresponding empirical literature is quite extensive, the above-mentioned main theories are further discussed in more detail in the following text.

### *Demand effects*

The cornerstone of demand side theories, the median voter theorem, establishes that, if all voters have single-peaked preferences over a unidimensional policy space—such as the level of public spending and/or of taxation—then the median preferred alternative is a Condorcet

winner. That is to say, the candidate representing this preference would win a majority of the vote regardless of the voting system in question. Political candidates will always have an incentive to set their platforms closer to the median ideal point in order to capture a majority of votes (Persson and Tabellini, 2000). This convergence prediction can also be generalized to the case where voters' decisions are erratic: in Hinich (1977, 1978) and Lindbeck and Weibull (1987, 1993), for example, political platforms converge to a policy that maximizes a weighted social welfare function.

The idea of policy convergence has been broadly used in the political economy literature. Special attention is paid to the provision of public goods (the demand-model approach introduced by Barr and Davis, 1966; Borcherding and Deacon, 1972; Bergstrom and Goodman, 1973) and redistributive policies, where inequality of market incomes among voters is associated with higher levels of political support for redistribution (the Meltzer and Richard hypothesis initiated by Romer, 1975; Roberts, 1977; Meltzer and Richard, 1981). Both streams of studies relate the political affiliation of the voters to their personal situation, the so-called *pocketbook (egotropic) voter* (Weatherford,1978; Lewis-Beck and Stegmaier, 2000).

According to the median voter theorem, party control does not matter (Besley and Case, 2003). In sufficiently competitive elections, the political candidates maximize their expected number of votes and commit to their campaign promises once elected, thereby acting as good agents for their electorate's views. Empirical evidence lends some support to this prediction, since the median income, or equivalently the average income, has been frequently found to be a relevant variable for explaining governments' behavior both on local (see, e.g., Ahmed and Greene, 2000; Aronson *et al.*, 2000; Guengant *et al.*, 2002) and central levels (Arpaia and Turrini, 2008; Kolluri *et al.*, 2000; Magazzino, 2012; Pradhan and Bagchi, 2012; Wahab, 2004).

Likewise, the Meltzer and Richard hypothesis has recently found statistical support in Alesina *et al.* (2000), Borge and Rattsø (2004), Mattos and Rocha (2008), Milanovic (2000), and Mohl and Pamp (2009). Yet, several authors argue that those findings are not sufficient to conclude to the supremacy of the demand models. While the empirical evidence

shows that public policies respond at the margin to changes in demographic and economic variables, it does not reveal the extent to which government policies are congruent with voters' preferences (Matsusaka, 2010).

*Supply effects*

A substantial part of the political economy literature argues instead that the size and performance of the public sector depend on various supply side effects: party ideology, coalition formation and cooperation, election manipulation and logrolling, bureaucracy and rent-seeking behavior, political (business) cycles, etc. Out of the numerous streams of literature, the thesis focuses on the key topics of partisan effects and political fragmentation, as discussed below.

Several theoretical arguments suggest that the ideology and party affiliation of public officials affects the size of government. This category of models views the politicians as *citizen-candidates* (Besley and Coate, 1997; Osborne and Slivinsky, 1996) or *partisan politicians* (Persson and Tabellini, 2000) with specific characteristics and preference functions. In this setting, candidates have an incentive to reveal their preferences during the electoral campaign and to implement them if elected, since credibility is a crucial factor in winning the election. Under certain, fairly general conditions, a policy divergence can be observed between the competing political parties, thereby making the government ideology a relevant variable for explaining public policy choices (Padovano, 2013).

The question of the existence of partisan effects has been addressed extensively, as exemplified by Jackson and Kingdon (1992), Bender and Lott (1996), and Imbeau *et al.* (2001). The hypothesis has been recently supported by empirical evidence in the U.S. (Besley and Case, 2003; Bjørnskov and Potrafke, 2012; Pickering and Rockey, 2013), in Spain (Solé-Ollé, 2003), in Norway (Borge and Rattsø, 2004), in France (Foucault *et al.*, 2008; Le Maux *et al.*, 2011), in Italy (Padovano and Petrarca, 2014; Santolini, 2008), and in OECD countries (Pickering and Rockey, 2011). In most cases, the studies show that governments on the left tend to raise tax rates and public spending more than

their right-wing counterparts.

A second related consideration is the level of competition in the political arena. The lack of effective political competition may increase the margin of maneuver of elected officials, who may in turn try to maximize their prestige, power, or other type of advantage. This type of behavior is described in the Leviathan hypothesis (Niskanen, 1971; Brennan and Buchanan, 1980), where the government is assumed to maximize the size of the public sector at the expense of the citizenry, whose power is largely limited to the conception of constitutional rules. The role of constitutional constraints on the size of the public sector has been explored empirically in the context of tax base restrictions (e.g., Nelson, 1986; Shapiro and Sonstelie, 1982), decentralization and federalism (Deacon, 1979; Marlow, 1988; Mehay, 1984; Mehay and Gonzales, 1985; Oates 1985; Schneider, 1986), or citizens' incentives to be informed and exercise control (e.g., Duncombe *et al.*, 1997; Hayes *et al.*, 1998). Recently, also as a part of the environmental politics discourse (see, e.g., Duit *et al.*, 2016, and Lundqvist, 2001, on the "Green Leviathan").

While political competition is generally found to discipline politicians and provide incentives to pursue greater efficiency (e.g., Ashworth *et al.*, 2014; Ferris *et al.*, 2008; Padovano and Ricciuti, 2009), government fragmentation may lead to "weaker" governments incurring higher public spending, public deficits and debt. Three groups of models within the Weak Government Hypothesis (initiated by Roubini and Sachs, 1989a, 1989b) explain the policy effects of government fragmentation (Ashworth *et el.*, 2005). First, *war of attrition models* (Alesina and Drazen, 1991; Alesina and Perotti, 1995) anticipate increase in short-term indebtedness due to procrastinating on policy changes, such as adjustment to shocks. Second, in *common pool models*, non-cooperative politicians overtax the common resources from nationwide taxation through projects targeted at their electorate (Weingast *et al.*, 1981). The effect lasts in the long term (Velasco, 2000) and intensifies when the number of parties in the coalition increases, since it becomes harder to come to a cooperative (and socially optimal) solution (Olson, 1993). Last, *strategic debt models* link fragmentation to increased uncertainty about next term's election results, motivating incumbents' strategic behavior that raises future indebtedness (Alesina and Tabellini, 1990; Persson and Svensson,

1989; Tabellini and Alesina, 1990).

The Weak Government Hypothesis generally finds support in the empirical literature: for instance, in Alesina and Perotti (1995), Ashworth and Heyndel (2005), Borge and Rattsø (2002), Le Maux and Zhang (2013), Padovano and Venturi (2001), Perotti and Kontopoulos (2002), or Rattsø and Tovmo (2002), albeit mixed (see, e.g., Borelli and Royed, 1995; Freitag and Sciarini, 2001; Hahm *et al.*, 1996).

*Threats to identification*

One still unresolved issue is whether the empirical results emphasized above are related to a truly supply side driven process, or whether it is the indirect result of changes in voters' preferences. Put simply, government composition may be endogenous, i.e., depend on how voters are distributed, which would create a potential selection bias should we attempt to compare policy decisions directly. So far, no empirical study has been able to clearly decide.

In the theoretical literature, numerous arguments propose that the parameters measuring the ideology space of each candidate can be endogenous. In Hinich (1978) and Lindbeck and Weibull (1987), for instance, political candidates have imperfect information about voter preferences. Under certain conditions (namely, two-party competition, symmetry and monotone probabilistic voting function) the political equilibrium outcome would be the utilitarian point, i.e., a situation where the size of the ideological spaces reflects the distribution of voters' ideological preferences. This result can be generalized to a repeated-game setting where both the ruling and the opposition parties have their reputation to protect, which forces the incumbent to fulfill their electoral promises (see, e.g., Kreps and Wilson, 1982, as well as other developments in the literature on imperfect competition). Milesi-Ferretti *et al.* (2002) extend this logic to a setting where expenditures on public goods and services have a local nature. In this setting the incumbent government distributes expenditures geographically, to maximize the joint utility of the various jurisdictions and so its probability of reelection.

Voters' interests may also be propagated through pressure groups, either by influencing

elections (via campaign contributions or indirect lobbying) or through pressures exerted on the politicians once in office (interviews, petitions, etc.). In these models, the candidates' policy space depends on the lobby's ability to map support to the elected politicians. The logic of collective action literature generally assumes that the size of the lobby groups negatively affects their influence. The direction of the effect is however unclear, since larger groups may also mobilize more resources and thus map more support to the targeted politician. For instance, using Swiss data, Stadelmann *et al.* (2013) find that special interest groups are able to make politicians diverge from policy positions that the median voter had already endorsed in a referendum.

Moreover, numerous other factors could also endogenize the candidates' policy space, such as the number of candidates who choose to run, electoral campaigns, incumbency, the multidimensionality of platforms, personal characteristics of the candidates—their personality, charm, reputation, racial characteristics, religion, family situation, etc. This may affect the margin of maneuver of politicians, i.e., modify their opportunity to implement their favorite policy without electoral sanctions (see, e.g., Calvert, 1985; Hansson and Stuart, 1984; Wittman, 1973).

The fact that the government composition can be endogenous has been discussed extensively in the political science, rather than the political economy literature. For an overview of government/coalition formation theories through the optics of (comparative) political science see, e.g., Martin and Stevenson (2001, 2010), Glasgow *et al.* (2012), Glasgow and Golder (2015), or Laver and Benoit (2015). Concerning the local government level in particular, several studies deal with political actors' incentives for government formation: e.g., Bäck (2008), Debus and Gross (2016), Skjæveland *et al.*, (2007).

Theory also points out another still unsolved issue in the relationship between voters, ideology and policy decisions: nonlinearity. A new class of political economy models argues that the interactions among voters and politicians that affect the candidates ideological or policy spaces may be quite complex. Bénabou (2000), Horstmann and Scharf (2000), Persson (1995) and Pickering and Rockey (2011) suggest that cooperation among the rich and the poor may generate benefits such as economic growth, which increase the willingness

of the rich to support redistributive policies. In such a case the relationship between the distribution of voters and the candidates' endogenous ideological stance might well be non-monotonic, depending on the degree of interclass cooperation.

This literature has roots in the models of *sociotropic voting* and *altruistic voting*, where voters base their decisions not only on their personal degree of altruism, but also on the candidate's ability to improve the well-being of the whole community. The candidates' policy spaces would then depend on the distribution of voters' preferences (Kinder and Kiewiet, 1979; Wright, 1986) and the weights of the government's social welfare function would depend endogenously on the empathy/altruism of voters.

*Quasi-experimental approach*

While the issue of nonlinearity is fairly easy to address empirically, the problem of endogeneity, i.e., the selection bias discussed previously, requires more sophisticated empirical strategies. In theory, the observed difference between policy decisions could be the sum of two components: (1) the direct effect of the supply side's composition (ideology, fragmentation, etc.) and (2) the indirect effect created by the distribution of voters, which determines the supply side's composition and thereby, jointly, the policy choices. In order to be able to decide in favor of supply side explanations, the latter component has to be removed from the equation, or at least, as much as possible, its impact explicitly expressed and separated in the results ("controlled for").

Ideally, to simply and clearly determine whether one jurisdiction would have changed its policies had its government been different, it would need to be observed under various conditions in the same time period. An impossible task. Unfortunately, neither can citizens be randomly assigned to specific government compositions, so as to conduct a controlled experiment that would compare the treated and control groups. Economic research in general rarely allows for a true experimental approach (randomized, replicable), due to practical, ethical or legal obstacles. Yet, collected data can still be interpreted *as if* indeed the subjects have passed through an experimental intervention, the impact of which can then be estimated.

A suitable tool is presented by so-called *quasi*-experimental designs, adapted to deal with confounding factors affecting the analysis. This type of approach allows for a hypothetical alternative outcome, the *counterfactual*, to be calculated and used as a counter to the real data, thereby enabling the seemingly impossible comparison of the same jurisdiction under different conditions.

Although imperfect, since only group averages, and thus difference in the mean outcome caused by the intervention, can be considered, quasi-experimental designs deal with the misleading bias that would otherwise plague the analysis. First, the omitted variable bias, which would result from naively comparing the same individual pre and post intervention, e.g., the election, all the while disregarding the evolution and impact of other factors as time passes. Second, there is the concern of the selection bias, which occurs when comparing subjects with heterogeneous characteristics. With a naive comparison, the role played by confounding factors, such as the characteristics of the subjects, would be ignored, which could bias the estimation of the effect we are interested in. Hence, in the absence of random treatment assignment, there is a need to compare groups with similar characteristics.

Based on how the counterfactual is constructed, quasi-experimental designs can be applied through the introduction of instrumental variables into econometric regression models, or via methods comparing directly the treated and comparison groups: the simplest difference-in-differences (DID, DD), or regression discontinuity design (RDD) and propensity score matching (PSM)—two techniques employed by us.

While both RDD and PSM are designed to control for selection bias, they approach the same problem from different angles. RDD is based on a comparison around a threshold in the variable of interest, which is assumed to depend on the characteristics creating the selection bias. The threshold defines the contrast between the treatment and comparison groups: for instance, the share of seats won by leftist politicians, depending on the characteristics of the electorate, leads (or not) to a left-wing majority being elected. Gradually decreasing bandwidths are plotted around this cutoff, creating subsamples of observations that, being so similar in the values of the treatment variable, should also be increasingly similar in correlated characteristics, thus eliminating the selection bias. RDD then estimates an

econometric model around the discontinuity, measuring differences in the outcome variable, e.g., public spending, between the two groups.

PSM, meanwhile, is not based on the observed values of one variable denoting treatment, but rather on econometrically estimated probability of belonging to the treatment group due to a set of relevant characteristics—i.e., the propensity score. The idea is to select and pair ("match") treated and non-treated observations with similar propensity scores, assuming them to be near identical in all underlying characteristics as well. Limitations may be placed on the difference in scores, further enforcing their likeness. Treatment effects can then be calculated as differences between group averages in the outcome variable. Note that, customarily, PSM is limited to a binary treatment setting, although progress has been made into accommodating multiple treatment options (see, e.g., Hirano and Imbens, 2004; Imbens and Wooldridge, 2009).

Another technique that may be applied with an analogical purpose is benchmarking. Capable of accounting for multiple inputs transformed into a variety of combinations, depending on the specification of users' needs, benchmarking methodology chiefly allows to highlight the impact demand structure has on the costs of the services a facility provides. The cost comparison is performed with respect to a relevant reference point (the benchmark), which may be national, regional or whatever is suitable, depending on the institutional context. The benchmarking methodology is mainly used as follow-up evaluation tool in the healthcare and education sector. In our context, controlling for the impact of the demand structure allows us to examine the remaining effects of other factors on policy choices.

*Thesis outline*

The empirical analysis of local public sectors in France and Finland, presented in the thesis, takes advantage of the aforementioned quasi-experimental techniques and benchmarking methodology in order to examine whether government composition (e.g., ruling party's ideology) affects the structure and level of public spending, *ceteris paribus*. On data from French Departments, we show that left-wing governments do not exhibit higher social

spending levels than their right-wing counterparts once we control for demand factors. Similarly, the largest Finnish parties do not prove to play a significant role in determining policies. The demand side, i.e., the electorate and its socio-economic characteristics, thus seems to be an important factor that affects policy choices. Its role is studied in more detail in the context of social welfare provision in France. The research is presented in the form of three empirical essays, as follows.

In the first chapter, titled *"Ideology or voters? A quasi-experimental test of why left-wing governments spend more"*, the thesis studies the impact of partisan effects (leftist ideology) on local public finance in France, at the departmental level. The results show that left-wing governments facing the same economic situation as right-wing ones do not spend more, particularly in the case of social expenditures, which are the spending item most exposed to ideological pressures. This result rules out the partisan-politicians hypothesis and lends support to demand driven policy selection processes.

In the second chapter, *"Does the size of the largest party matter? Endogenous fragmentation, political competition, and local public expenditures in Finland"*, the thesis examines the determinants of political fragmentation (distribution of seats among majorities and coalitions) and its effect on public spending levels. The focus is on Finnish municipalities, the single level of local government in Finland. The results show that there is no significant difference in public spending between single-party majorities and minority governments. Instead, using sample-splitting techniques, the analysis demonstrates that a significant breakpoint exists in our data when the largest party holds a supermajority of approximately three-fifths of the seats: at this threshold, spending levels are found to increase significantly.

In the third chapter, *"Performance of local governments: Benchmarking analysis of social welfare provision in France"*, we further explore how the composition of the demand side, i.e., the social inequalities observed in the population, influences the provision of social services by the French Departments. Electorate's heterogeneity may not only determine the composition of the government, but also the costs of public good provision. Employing benchmarking analysis, particularly suitable for dealing with heterogeneous

users of services, in order to isolate the effects of the size and the composition of the welfare-dependent population, we examine what drives the observed differences in social spending.

Overall, the thesis demonstrates how closely the demand (voters) and the supply (politicians) sides of public policies are connected. According to the data, the characteristics of the electorate jointly affect the configuration of the political arena and policy choices. The demand side thus appears as an important determinant of local public good provision. If this yields an optimistic prediction as to how local representative democracies are functioning, the fact remains that the demand for local public goods can be higher in some jurisdictions than in others. Hence, local governments are faced with different budget constraints depending on the characteristics of the electorate. This in return can yield unequal access to public services, as evidenced in the case of social welfare provision in France.

# Chapter 1

# Ideology or voters? A quasi-experimental test of why left-wing governments spend more

[†]

## 1.1 Introduction

Politics is partisan: the theoretical political economy literature hold this as a stylized and generally accepted fact. Moreover, the empirical literature seems to support this "partisan hypothesis", since a considerable number of studies find—in a large array of samples and institutional contexts—that governments with a left-wing ideology tend to increase public expenditure and taxation. Conversely, right-wing governments tend to reduce spending, or at least curb the expansion of the public sector (Borcherding, 1997; Hansen and Stuart, 2003; Tridimas and Winer, 2005). This view has recently found support in the U.S. (Besley and Case, 2003; Bjørnskov and Potrafke, 2012; Pickering and Rockey, 2013), Spain (Solé-Ollé, 2003), Norway (Borge and Rattsø, 2004), France (Foucault *et al.*, 2008; Le Maux *et al.*, 2011), Italy (Padovano and Petrarca, 2014; Santolini, 2008), as well as in OECD countries (Pickering and Rockey, 2011).

[†]This research is the result of cooperation with Benoît Le Maux (University of Rennes 1, CREM-CNRS, Condorcet Center for Political Economy) and Fabio Padovano (University of Rennes 1, CREM-CNRS, Condorcet Center for Political Economy).

This paper instead claims that such evidence of policy divergence between left-wing and right-wing governments is insufficient to settle the issue in favor of the partisan hypothesis, for two reasons. The first is that government ideology may itself be endogenous: a left-wing government may spend more than a right-wing one because a leftward shift of voters' preferences results in the election of, precisely, a left-wing government—which in turn expands the budget to satisfy voters' preferences. In other words, what may appear as an ideology, or supply side driven process, is in fact a demand driven one. This "selection bias" generates a problem of observational equivalence in empirical analysis. So far, no empirical study has been able to ascertain whether policy changes depend on a truly ideologically driven, supply side process, or on a demand side one, which indirectly determines policy outcomes by affecting the probability that an ideology constrained candidate is elected.

Secondly, a new class of political economy models suggests a more complex relationship between the distribution of voters' preferences, government ideology and policy decisions. Bénabou (2000), Horstmann and Scharf (2000), Persson (1995) and Pickering and Rockey (2011) suggest that cooperation among high and low income voters may result in larger welfare expenditure and faster economic growth at the same time. This increases the willingness of high income voters to support redistribution. Low cooperation instead reduces interclass externalities and increases the ideological polarization of political parties. Hence there exists an optimal level of interclass cooperation that makes the relationship between voters' political preferences and candidates' ideology non-monotonic. Testing the impact of government ideology on policy choices is therefore more complicated than what the empirical literature usually assumes.

To address these two problems, we first need a unified theoretical structure able to generate policy equilibria through both supply side and demand side driven processes. In our model, three different channels are considered. First, voting behavior directly affects the candidates' platforms, thus endogenizing the ideology of the government in office, and inducing a first form of selection bias. Second, a change in the distribution of voters also modifies the total tax base and, consequently, affects the willingness to cooperate of different classes of taxpayers, hence inducing a second form of selection bias. Third, the ideology of

the incumbent government can also reflect the candidates' political preferences toward one class of voters, holding voters' preferences constant, which is the actual partisan hypothesis.

Next, we test the partisan hypothesis by means of two quasi-experimental techniques, namely, regression discontinuity design (RDD) and propensity score matching (PSM). Several studies have applied RDD to explain the relationship between policy choices of the incumbent government and electoral processes (see, e.g., Folke, 2014; Freier and Odendahl, 2015; Pettersson-Lidblom, 2008). The approach is efficient at taking care of unobservables; yet, in our context, RDD can also present shortcomings. The close elections on which it relies can introduce other types of bias: first, because a close two-party electoral contest can result in issue convergence (see, e.g., Padovano, 2013; Sigelman and Buell, 2004); second, because close elections can severely reduce the political clout of the incumbent government, forcing compromises (Le Maux *et al.*, 2011; Le Maux and Rocaboy, 2016). In both cases close contests might mitigate ideological differences. To deal with this concern, we extend the analysis to a much larger array of not necessarily close elections and resort to propensity score matching (PSM). To the best of our knowledge, PSM has never been used in this subject matter before.

Both RDD and PSM strategies consist in selecting a group of right-wing jurisdictions and making them resemble left-wing ones in all features except for their ideology. Our empirical strategy thus keeps constant the demand driven characteristics, i.e., all the proxies normally used to capture the distribution of voters' preferences in the policy space. This way we single out the impact of government ideology on policy choices, holding all other conditioning factors constant. In this analytical framework, a statistically significant coefficient of the government ideology indicators provides conclusive evidence in favor of the partisan explanation, precisely because we control for changes in the distribution of voters. We select the French Departments (*Départements*) as a convenient testing ground for the research question at hand.

Overall, the tests of the model yield two main results. In the first step of the analysis, the estimation of a vote function and of a spending equation confirms the existence of a selection bias and suggests a strong relationship between voters' support for public spending

and actual spending. In the second step, the resort to quasi-experimental techniques allows to control for the socio-economic characteristics of the electorate; we can then show that left-wing governments facing the same socio-economic situation as right-wing ones do not actually spend more, particularly on social expenditures, which are the spending item most exposed to ideological pressures. This evidence conflicts with the partisan hypothesis, but it is consistent with the demand driven vision that voters' preferences endogenously determine the ideology of the government.

The remainder of the paper is organized as follows. Section 1.2 provides a unified microeconomic theoretical framework that leads to three propositions. Section 1.3 justifies the choice of the French Departments as a testing ground and describes the data. Section 1.4 provides suggestive evidence of a selection bias. Section 1.5 develops the second step of the empirical analysis, where the selection bias is solved through quasi-experimental techniques and the partisan effects are verified controlling for demand driven processes. Finally, Section 1.6 highlights the main findings of the analysis.

## 1.2 The model

Consider $N$ voters living in a representative democracy and a government providing a public service in quantity $Z$. For simplicity, $Z$ is assumed to be produced by a firm acting in a competitive market, with price equal to a constant marginal cost of production $c$. To capture the essential features of the demand side explanations, we exogenously divide the voting population into two classes of different size: class 1 is composed of $n_1$ welfare-dependent individuals, while class 2 of $n_2$ taxpayers. Again for simplicity, every member of a given class has the same income $y_i$, with $y_2 > y_1$. Moreover, all members of class $i$ have a utility function given by $U_i(x_i, z)$, where $x_i$ denotes the quantity of a composite private good consumed (the *numéraire* of the economy) and $z$ the quantity of public service available to every individual class member. $U_i$ is strictly concave, twice differentiable, increasing in both variables, and its cross partials are assumed to be non-negative to rule out inferior goods.

A parameter $\alpha$ captures possible congestion effects, to account for the taxonomy of public policies:

$$z = \frac{Z}{N^{\alpha}}. \tag{1.1}$$

If the service is privately consumed, then $\alpha = 1$; if, instead, it is a pure public good, $\alpha = 0$. Importantly, this parameter can also be viewed as an inverse proxy for the degree of cooperation, empathy or altruism of voters (Andreoni, 2007). When $\alpha < 0$, each voter achieves a higher utility when sharing the publicly provided good, thus generating a social network effect. If instead $\alpha = 1$, an increase in the group size reduces the benefit for each individual class member.

An important difference between the two classes of voters is that income $y_2$ is taxable, whereas $y_1$ is not. The individual budget constraint for a member of class 1 therefore is $y_1 = x_1$, while it equals $(1-r)y_2 = x_2$ for a member of class 2, where $r$ is a proportional tax rate assumed to apply identically to all taxpayers in that class. The government is assumed to balance its budget, hence:

$$n_2(rY_2) = cZ. \tag{1.2}$$

By substituting equations (1.1) and (1.2) into the budget constraint of class 2, we obtain $y_2 = x_2 + p_2 z$, where $p_2 = \frac{cN^{\alpha}}{n_2}$ denotes the tax price paid by the wealthier class. This tax price plays a crucial role in the model, as it represents the share of the cost $c$ that each voter of class 2 finances. The larger the size of class 2, the lower the average tax burden on its members. Substituting the budget constraint into the utility functions gives the following reduced forms of preferences: $U_1(y_1, z)$ for class 1 and $U_2(y_2 - p_2 z, z)$ for class 2.

The electoral process is characterized as follows. There are two politicians (parties), $j = A, B$, who maximize their probability of being elected into office. They are assumed to be ideologically oriented, in the sense that candidate $A$ is ideologically oriented towards class 1, while candidate $B$ towards class 2. The timeline of the model is in two steps: first, both candidates simultaneously announce their political platforms, respectively $z^A$ and $z^B$. In

the second step, elections are held. The voting game is solved by backward induction.

Voters' decisions are based on the candidates' policy positions. An individual $k$ from class 1 will vote for $A$ iff:

$$(1+\delta)U_1(y_1,z^A)-(1-\delta)U_1(y_1,z^B)>\sigma_1^k. \tag{1.3}$$

Likewise, an individual $k$ from class 2 will vote for $A$ if and only if:

$$(1-\delta)U_2(y_2-p_2z^A,z^A)-(1+\delta)U_2(y_2-p_2z^B,z^B)>\sigma_2^k, \tag{1.4}$$

where $\sigma_i^k$ is an individual-specific parameter uniformly distributed over $[-\frac{1}{2\phi},+\frac{1}{2\phi}]$ and with density $\phi$. As in Persson and Tabelini (2000, p. 52), those parameters represent voter's $k$ ideological bias toward the political candidates. For simplicity, these distributions have the same density and are common knowledge. Coefficient $\delta$ is a positive parameter representing the candidates' ability to influence the vote of each class.[1] Candidate $A$ (respectively candidate $B$), who is assumed to be ideologically oriented toward class 1 (respectively class 2), is more popular in that class.[2]

The voter of class 1 who is indifferent between both candidates has an ideology parameter equal to:

$$\sigma_1^*=(1+\delta)U_1(y_1,z^A)-(1-\delta)U_1(y_1,z^B). \tag{1.5}$$

[1]Note that by assuming quasi-linear preferences, $U_i(x_i,z)=x_i+v(z)$, and assigning the weights $(1+\delta)$ and $(1-\delta)$ on $v(z)$ only, the model can be extended to any continuous income distribution. The left-hand sides of equations (1.3) and (1.4) would become $(1+\delta)v_1(z^A)-(1-\delta)v_1(z^B)$ and $(1-\delta)v_2(z^A)-p_2z^A-(1+\delta)v_2(z^B)+p_2z^B$, respectively. Differences in income would thus matter only via the tax price, i.e., the results would not fundamentally change.

[2]This so-called party identification, providing voters with a shortcut for making voting decisions, can be explained either through the citizens' social group identities and perceptions of the social groups that support each party (e.g., Campbell *et al.*, 2011; Green *et al.*, 2002), or through their evaluation of the parties' policies and ideological orientation (e.g., Downs, 1957; Abramowitz and Saunders, 2006). The extent to which the elements of an election reflect the voters' summary judgment about parties should condition the strength and character of partisan voting (Campbell *et al.*, 2011).

Similarly, the swing voter for class 2 has an ideology parameter:

$$\sigma_2^* = (1-\delta)U_2(y_2 - p_2 z^A, z^A) - (1+\delta)U_2(y_2 - p_2 z^B, z^B). \tag{1.6}$$

The share of votes for candidate $A$ coming from the voters in class $i$ is:

$$\pi_i^A = \phi\left(\sigma_i^* + \frac{1}{2\phi}\right) \qquad (i = 1,2). \tag{1.7}$$

The overall vote shares for candidates $A$ and $B$ are therefore:

$$\pi^A = \frac{n_1}{N}\pi_1^A + \frac{n_2}{N}\pi_2^A \quad \text{and} \quad \pi^B = 1 - \pi^A. \tag{1.8}$$

Replacing $\pi_i^A$ and $\sigma_i^*$ with their expression, we finally obtain:

$$\begin{aligned} \pi^A &= \frac{n_1}{N}\phi\left[(1+\delta)U_1(y_1, z^A) - (1-\delta)U_1(y_1, z^B) + \frac{1}{2\phi}\right] \\ &\quad + \frac{n_2}{N}\phi\left[(1-\delta)U_2(y_2 - p_2 z^A, z^A) - (1+\delta)U_2(y_2 - p_2 z^B, z^B) + \frac{1}{2\phi}\right]. \end{aligned} \tag{1.9}$$

Political candidates choose $z$ to maximize their expected number of votes. From equation (1.9), the platform announced by candidate $A$ must satisfy the following first-order condition:

$$(1+\delta)\frac{n_1}{N}\frac{\partial U_1}{\partial z^A} + (1-\delta)\frac{n_2}{N}\left[\frac{\partial U_2}{\partial z^A} - p_2\frac{\partial U_2}{\partial x_2}\right] = 0. \tag{1.10}$$

Similarly, candidate $B$ chooses $z^B$ so that:

$$(1-\delta)\frac{n_1}{N}\frac{\partial U_1}{\partial z^B} + (1+\delta)\frac{n_2}{N}\left[\frac{\partial U_2}{\partial z^B} - p_2\frac{\partial U_2}{\partial x_2}\right] = 0. \tag{1.11}$$

When $\delta = 0$, i.e., when candidates share the same ideology, both political platforms converge to the utilitarian optimum, a standard result in political economy.[3] Should $\delta > 0$, i.e.,

[3]Note that the utilitarian optimum obtained in a probabilistic setting is unlikely to correspond to the median

political candidates do not share the same ideology, then their policy platforms diverge. Candidate $A$ weights more the utility of class 1, while candidate $B$ favors class 2. In such a case one obtains $z^A > z^B$.

Overall, candidates' behavior can be characterized as a maximization of the following objective function with a control variable $z$:

$$\max_z \quad \Omega = \theta_1^j U_1(y_1, z) + \theta_2^j U_2(y_2 - p_2 z, z) \qquad (j = A, B) \tag{1.12}$$

with

$$\theta_1^A = (1+\delta)\frac{n_1}{N}; \theta_2^A = (1-\delta)\frac{n_2}{N}; \theta_1^B = (1-\delta)\frac{n_1}{N}; \theta_2^B = (1+\delta)\frac{n_2}{N}, \tag{1.13}$$

where $\theta_1^j$ and $\theta_2^j$ are two parameters accounting for both the demand ($n_1/N$ and $n_2/N$) and supply ($\pm\delta$) sides of the policy process.

To simplify the exposition, let us assume that the solution to (1.12) satisfies the (work) incentive compatibility constraint $U_1 < U_2$. The first order conditions can be rewritten as:

$$\theta_1^j \frac{\partial U_1}{\partial z} + \theta_2^j \left[\frac{\partial U_2}{\partial z} - p_2 \frac{\partial U_2}{\partial x_2}\right] = 0 \qquad (j = A, B), \tag{1.14}$$

where $\partial U_1/\partial z$ and $\partial U_2/\partial z - p_2 \partial U_2 \partial x_2$ represent the net marginal benefit that class 1 and 2, respectively, obtain from the public policy. While class 1 always derives a positive marginal utility from the public policy $z$, members of class 2 must pay taxes; hence they suffer a utility loss when $z$ is too high.

More generally, let $z^*$ denote the solution to equation (1.14). From the policy $z^*$, we obtain the following comparative static derivatives (see Appendix A):

$$\frac{\partial z^*}{\partial y_1} > 0, \quad \frac{\partial z^*}{\partial y_2} > 0, \quad \frac{\partial z^*}{\partial p_2} < 0, \quad \frac{\partial z^*}{\partial \alpha} < 0, \quad \frac{\partial z^*}{\partial c} < 0, \quad \frac{\partial z^*}{\partial \theta_1^j} > 0, \quad \frac{\partial z^*}{\partial \theta_2^j} < 0. \tag{1.15}$$

voter's ideal point, an outcome that is instead usually obtained in a deterministic setting. Both frameworks however predict full policy convergence.

The value of $z^*$ rises as $y_1$ and $y_2$ increase and $p_2$ (the tax price paid by class 2) decreases. In the same vein, $z^*$ decreases as $c$ (the marginal cost of $z$) and $\alpha$ (the degree of rivalry in consumption of $z$) increase. In particular, with respect to $\alpha$, demand will be higher for a pure public good ($\alpha = 0$) than for a subsidy ($\alpha = 1$), because $p_2$ tends to 0 as $\alpha$ approaches 0. All these results are consistent with the demand side view that the demand for public services should react as that of any normal private good, even when partisan effects are involved.

The novelty of our approach is that the weights $\theta_1^j$ and $\theta_2^j$ in the government's objective function for aggregating the utilities of each class play a determinant role in setting $z^*$. In equilibrium, the higher the influence of class 2 ($\theta_2^j$), the lower will be the demand for $z$. Reciprocally, the higher the influence of class 1 ($\theta_1^j$), the higher the demand for $z$. This result generates three main propositions:

(P1) *Endogenous government ideology*. The weights $\theta_1^j$ and $\theta_2^j$ in the government objective function (1.12) depend on the relative size of the two classes of voters ($n_1$ and $n_2$). The greater the size of class 1 relative to class 2, the higher the influence of welfare recipients, and the higher the level of $z^*$. This result is driven by the fact that the distribution of voters affects the ideology of government. This induces a first form of selection bias.

(P2) *Tax base erosion effect*. The tax price $p_2$ depends negatively on the share of taxpayers $n_2/N$. The lower $n_2/N$ or, equivalently, the greater the share of welfare recipients $n_1/N$, the higher will be $p_2$ and the lower the taxpayers' demand for public policy. This induces a second form of selection bias.

(P3) *Partisan effects*. The weights $\theta_1^j$ and $\theta_2^j$ may not only reflect the demand of the voters, but also the candidates' political bias toward one of the classes, i.e., $1 \pm \delta$. Hence $z^*$ is also conditional on which political candidate or party is elected.

The first two driving forces in P1 and P2 offset each other. On the one hand, welfare recipients demand more of the public goods they heavily rely on. On the other, taxpayers have to pay higher taxes if the tax base erodes. To illustrate, consider the derivative of $z^*$ with

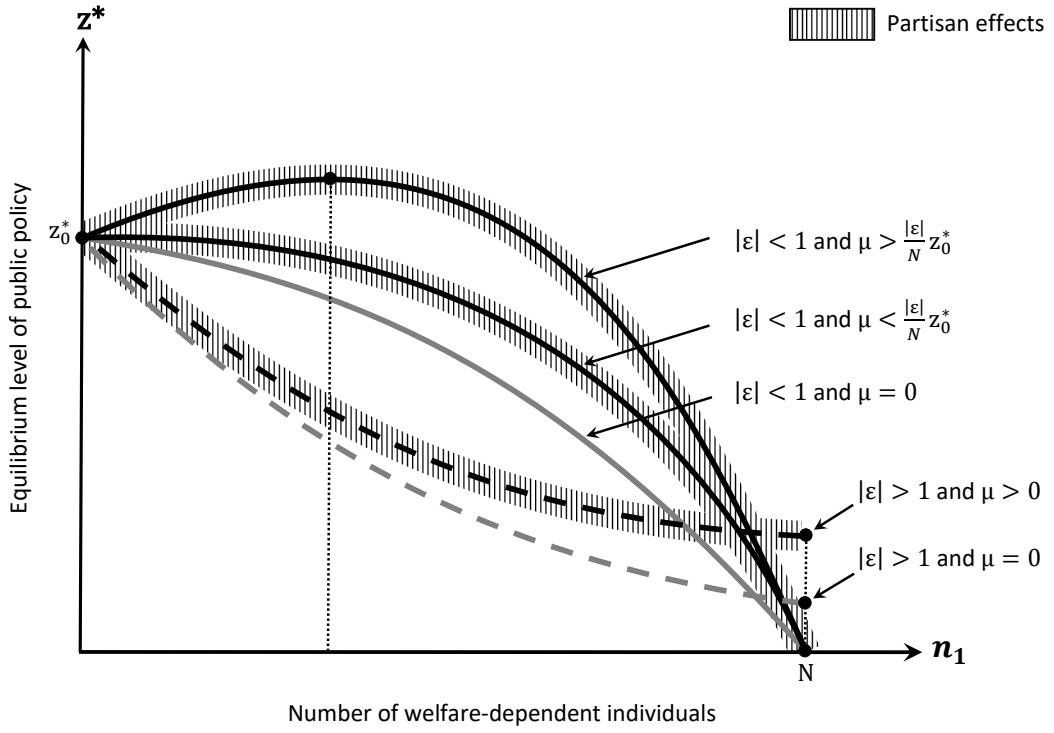



Figure 1.1. Distribution of voters and public good provision.

respect to $n_1$. Given that $\theta_1^j = \frac{n_1}{N}(1 \pm \delta)$, and assuming that the marginal voting influence of class 1, $\frac{\partial z^*}{\partial \theta_1^j}$, is constant and equal to $\mu$, we have:

$$\frac{\partial z^*}{\partial n_1} = \frac{\partial z^*}{\partial p_2}\frac{\partial p_2}{\partial n_1} + \frac{\partial z^*}{\partial \theta_1^j}\frac{\partial \theta_1}{\partial n_1} = \varepsilon \times \frac{z^*}{N - n_1} + \frac{\mu}{N}(1 \pm \delta), \tag{1.16}$$

where $\varepsilon = \frac{\partial z^*}{\partial p_2}\frac{p_2}{z^*}$ denotes the (tax) price elasticity of demand ($\varepsilon < 0$). The second derivative is:

$$\frac{\partial^2 z^*}{\partial n_1^2} = \varepsilon(1 + \varepsilon) \times \frac{z^*}{N - n_1} + \varepsilon\frac{\mu}{N}(1 \pm \delta). \tag{1.17}$$

Consider now the case where $n_1$ has an impact on $z^*$ through the tax price only, i.e., $\mu = 0$. If the absolute price elasticity of demand is smaller than 1, the demand curve for $z$ will be concave and always decreasing with $n_1$, as the solid gray curve in Figure 1.1 shows. In this case, the highest level of provision, hereafter denoted $z_0^*$, is obtained when $n_1 = 0$ (proposition P2).

When instead $\mu > 0$, the demand curve shifts upwards, as the black solid curves indicate. The voting influence of class 1 raises the levels of provision of $z^*$ with respect to the $\mu = 0$ case (proposition P1). In particular, if $\mu > \frac{z_0^* \times |\varepsilon|}{1 \pm \delta}$, the demand curve first increases as the size of class 1 grows. At some point, there are so many welfare-dependent people, and still a sufficient number of taxpayers, that there is a near unanimous support for the policy. Beyond such level, the decrease in the relative number of taxpayers more than offsets the demand for the public good expressed by the welfare-dependent class. Note also that if the absolute price elasticity of demand is higher than 1, the demand curve will be convex (gray dashed line). At high levels of $n_1$, the government is more reluctant to reduce $z^*$, especially as $\mu$ rises (black dashed line).

Lastly, if political candidates do not share the same ideology, a partisan effect could affect the policy in one direction or another, depending on the value and sign of $\delta$ (proposition P3). This effect is illustrated by the short vertical lines surrounding the curves.

To conclude, we argue that two factors need to be controlled for when testing the partisan hypothesis: first, a greater share of welfare recipients reinforces support for the left wing; second, this increases the burden on those who pay taxes, which puts a bound on what the left can do. Solving this potential selection bias is at the heart of the paper.

## 1.3 Data description

We select the French Departments as a testing ground for the model above for two reasons. First, welfare expenditures are the main responsibility and source of public outlays for the Departments, which ensures a close representation of the endogenous variable $z^*$ of the theoretical model. Second, the social and political contexts of the Departments show a considerable degree of variability, both cross-sectionally and over time. These features allow a good representation of the exogenous factors of the model as well.

France is divided into 96 metropolitan and 5 overseas Departments (Guadeloupe, Guyane, Martinique, Mayotte and Réunion). Because of their unique characteristics, Paris,

the two Corsican and the five overseas Departments are excluded from the analysis, as is customary for this sample. This leaves us with 93 Departments, observed over the period of 11 years between 1998 and 2008. The Departments are governed by a council elected in the so-called "cantonal elections", which are held approximately every three years.[4]

Our main interest lies in the operating expense categories of the Departments, namely:

A) **Social expenditures** (*social*), which include social aid to the unemployed (through a specific healthcare program and, since 2004, also through an unemployment benefit), social assistance to families with dependent children (through prevention, protection, aid to family, etc.), assistance to the disabled (through housing subsidies, direct payments, housing modifications for accessibility, etc.) and, lastly, to pensioners and the elderly (through direct payments and home subsidies).

B) **Non-social expenditures** (*nonsocial*), which cover principally the provision of transport services for all students in the Department and the maintenance of roadway and waterway networks, the management of ports, airports and public buildings. Moreover, since 1986, French Departments are responsible for building and maintaining schools for students aged between 11 and 15 years (*collèges*). As the great majority of French families have children attending these schools, this responsibility entrusts the Departments with an important role in the French educational system.

Since public expenditures do not consist only of targeted aids, but also of in-kind services, using both measures of social and of non-social expenditures seems a convenient approach to test our theory. Note that pensions are excluded from the empirical analysis because they are not a competence of the French Departments; Departments also do not hold responsibility for

[4]Voters directly elect the departmental councilors for a six-year term through a two-ballot, uninominal majority voting procedure. An important feature of this system is that only one-half of the councilors are renewed at each election, with one councilor per constituency. A constituency is a grouping of municipalities referred to as a *canton*—a subset of the Department. In a given constituency, a candidate who secures at least 25% of the registered voters and more than 50% of the total number of votes is elected. If no political candidate satisfies these conditions, a second electoral round is held one week later. The two candidates who obtain the largest number of votes in the first round proceed to the second round, plus any other candidate who received at least 10% of the votes in the constituency. In the second round, the most voted candidate is elected.

immigration programs. We furthermore exclude unemployment benefits, since their amount is defined nationally and Departments have no discretionary power there.

Table 1.1 offers the description and summary statistics of all the variables under consideration; the GDP deflator ($2010 = 100$) is used to compare financial variables over time. Note that both types of expenditures are mainly financed through the Departments' own tax revenues and, to a minor extent, through two types of grants from the central government, the DGF and the DGD (see Dollery and Lorenzo, 2008). The *Dotation globale de fonctionnement* (DGF) is a lump-sum grant aimed at reducing fiscal differences among Departments; the *Dotation générale de décentralisation* (DGD) has been designed to compensate for the transfer of responsibilities to the sub-national governments in the 1980s.

The welfare-dependent population, targeted by the social expenditures, is defined as:

$$WD = families + elder + unemp + disabled, \tag{1.18}$$

where $families$, $elder$, $unemp$, and $disabled$ stand, respectively, for the number of families, elderly, unemployed and disabled people receiving social assistance, divided by the total population.[5] Since the number of welfare recipients in a Department depends on eligibility criteria defined by the national law, these variables are not functions of the Department's public policies; they can be considered as truly exogenous in the model. Conversely, with the exception of unemployment benefits, the *amount* of individual social aid and of non-social expenditures is a discretionary choice of the Department. The considerable differences between Departments in per capita and per beneficiary amounts of aid show that the Departments take full advantage of their margin of maneuver with respect to spending decisions; if the central government was to entirely set the the amounts of aid through formulas and regulated standards, which the Departments were only to apply, cross-departmental differences in aid per beneficiary would not appear.

Although imperfect because of possible, but quite limited, overlaps between the

[5]The variable *unemp* does not correspond to a measure of the rate of unemployment, but instead to the share of the unemployed people in the population who benefit from the departmental program.

Table 1.1. Summary statistics

| Variable | Description | N | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|---|---|
| $leftshare$ | Share of seats on the left. $L$ equals 1 when $leftshare > 0.5$. | 372 | 0.475 | 0.201 | 0.088 | 0.955 |
| *social* | Per capita social expenditures (in euros), excluding unemployment benefits. | 1,023 | 274.861 | 65.602 | 128.709 | 483.756 |
| *nonsocial* | Per capita non-social expenditures (in euros). | 1,023 | 196.523 | 61.734 | 59.479 | 483.074 |
| $families$ | Share of families with children in the population receiving social assistance (in %). | 1,023 | 0.188 | 0.059 | 0.060 | 0.384 |
| *elder* | Share of elderly people in the population receiving social assistance (in %). | 1,023 | 1.556 | 0.876 | 0.196 | 4.134 |
| *unemp* | Share of the unemployed in the population receiving social assistance (in %). | 1,023 | 1.519 | 0.620 | 0.552 | 3.824 |
| *disabled* | Share of disabled people in the population receiving social assistance (in %). | 1,023 | 0.388 | 0.094 | 0.184 | 0.919 |
| $WD$ | Share of welfare-dependent people in the population, measured as a sum of $families$, *elder*, *unemp* and *disabled* (in %). | 1,023 | 3.652 | 1.159 | 1.273 | 7.033 |
| *population* | Number of inhabitants (in thousands). | 1,023 | 620.439 | 450.579 | 73.507 | 2,565.257 |
| *density* | Number of inhabitants per km$^2$. | 1,023 | 330 | 1,178 | 14 | 8,825 |
| *income* | Taxable income per capita (in euros). | 1,023 | 8,868 | 1,855 | 6,075 | 20,036 |
| *grants* | Grants per capita received by the Department (in euros). | 1,023 | 180.740 | 78.775 | 30.403 | 644.996 |

*Note: DEC* represents a time-dummy variable indicating the transfer of competences after 2002. Variables *social* and *nonsocial* are used as averages calculated over each term of office (*social_average*, *nonsocial_average*) where indicated.

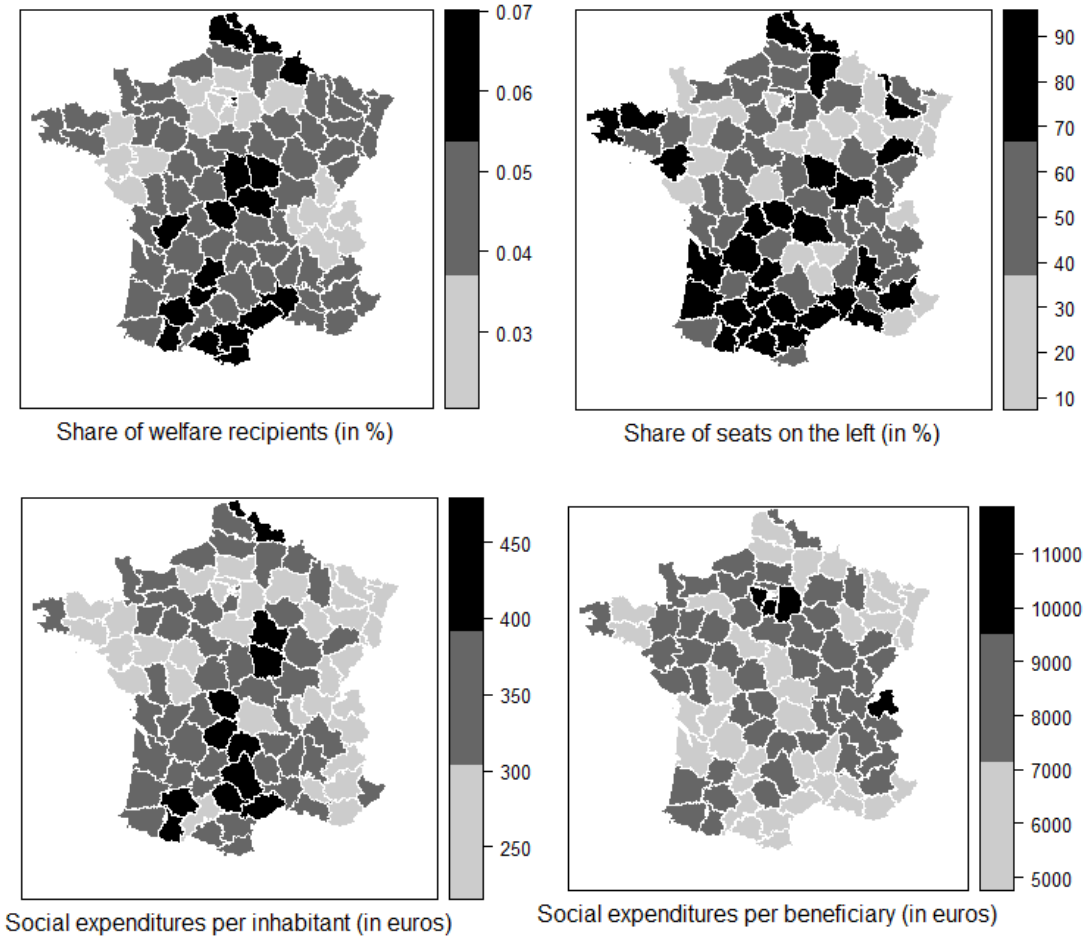



Figure 1.2. Cross-sectional comparison – 2008.

various subcategories of welfare recipients, the $WD$ index offers a good proxy to evaluate the share of welfare-dependent individuals in the population of the Department. More importantly, this measure allows to directly test the theory of Section 1.2, as it is the empirical proxy for the theoretical variable $n_1$. Figure 1.2 illustrates the situation of the Departments in 2008. In the top-left panel, we observe an unequal distribution of beneficiaries of social assistance. Difficult social situations are particularly evident in the South and in the North of France. These regions are characterized by a high amount of social expenditures per inhabitant (bottom-left panel), which is consistent with P1, but also by a lower amount of social aid per beneficiary (bottom-right panel), which lends support to P2. Meanwhile, as shown in the top-right panel, these Departments are usually governed by left-wing coalitions. Assessing whether partisan effects are at play, as stated in P3, requires further and more rigorous analysis.

## 1.4 Evidence of a selection bias

### 1.4.1 Endogenous government ideology

The descriptive analysis in the previous section has revealed some *prima facie* evidence of a correlation between the share of welfare recipients and left-wing votes. In this section we estimate a vote-popularity function to better quantify this relationship. The model assesses the impact of the socio-economic characteristics of a Department on the probability that left-wing representatives hold more than 50% of the seats. This stage is essential since it also determines the propensity scores used in the implementation of the matching process in Section 1.5. That is why we prefer to estimate a binary outcome model, rather than an equation predicting the seat shares.

The vote function model is specified as follows:

$$L_{i,t} = \alpha_0 + \alpha_1 \ln WD_{i,t} + \alpha_2 \ln income_{i,t} + \alpha_3 \ln density_{i,t} + \alpha_4 \ln population_{i,t} + \varepsilon_{i,t}, \quad (1.19)$$

Table 1.2. Geographical dummies

| Area | Regions |
|---|---|
| Center | Île-de-France |
| North-West | Basse-Normandie, Bretagne, Centre-Val de Loire, Haute-Normandie, Pays de la Loire |
| North-East | Alsace, Bourgogne, Champagne-Ardenne, Franche-Comté, Lorraine, Nord-Pas-de-Calais, Picardie |
| South-West | Aquitaine, Limousin, Midi-Pyrénées, Poitou-Charentes |
| South-East | Auvergne, Languedoc-Roussillon, Provence-Alpes-Côte d'Azur, Rhône-Alpes |

where $i$ and $t$ stand for Department $i$ and year $t$, respectively. Four election years are examined: 1998, 2001, 2004 and 2008. The endogenous variable $L$ is the government ideology, a dummy equal to 1 if the governing party or coalition is leftist or far-leftist ("left-wing" hereafter), and 0 otherwise ("right-wing"). Appendix B explains the methodology adopted to construct this variable. The variable $WD$ measures the share of welfare recipients, as defined in Section 1.3; *income* denotes the mean taxable income. The variables *density* and *population* stand for the density and the size of the population, respectively. They are incorporated in the model to account for additional tax price effects. For instance, larger local government entities could be more efficient than their smaller counterparts due to economies of scale.

A logistic regression estimates the coefficients for the binary dependent variable $L$, in order to circumvent the well-known shortcomings of a linear probability model applied to binary outcomes. The ideology $L$ of the governments exhibits a low degree of variation over time: among the 93 Departments, 65 have kept the same partisan affiliation throughout the entire sample period. As such, fixed individual effects could remove much of the time variation needed to obtain good estimates, especially on those coefficients (Beck, 2001). To avoid this potential bias, additional random effects and a pooled model with geographical dummies are estimated. The geographical dummies regroup the Departments into five areas (Table 1.2), providing a compromise between the pooled and the fixed-effects estimator, while at the same time taking into account the potential impact of regional specificities.

Table 1.3. Vote function – estimation results[a]

| | *Dependent variable:* | | | | | |
|---|---|---|---|---|---|---|
| | *L* | | | | | |
| | Logit Pooled | Logit RE | Logit Geogr. dummies | Logit Pooled | Logit RE | Logit Geogr. dummies |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Constant | 27.477*** | −5.310 | 23.042*** | 20.280** | 18.632 | 23.846** |
| | (7.824) | (22.324) | (7.873) | (10.140) | (27.912) | (10.305) |
| log(WD) | 3.721*** | 9.377*** | 3.582*** | | | |
| | (0.423) | (1.870) | (0.463) | | | |
| log(families) | | | | 0.852* | 3.859** | 2.022*** |
| | | | | (0.439) | (1.663) | (0.557) |
| log(unemp) | | | | 2.592*** | 8.146*** | 1.976*** |
| | | | | (0.418) | (2.496) | (0.483) |
| log(elder) | | | | 1.218*** | 2.541*** | 1.205*** |
| | | | | (0.273) | (0.788) | (0.308) |
| log(disabled) | | | | −0.887 | −1.074 | −0.885 |
| | | | | (0.677) | (2.113) | (0.736) |
| log(income) | −2.090*** | 1.963 | −1.951*** | −0.347 | 5.712* | −0.571 |
| | (0.735) | (2.146) | (0.746) | (0.984) | (3.035) | (1.059) |
| log(density) | 0.459** | −1.561 | 0.240 | 0.294 | 0.628 | 0.317 |
| | (0.200) | (1.129) | (0.366) | (0.220) | (0.676) | (0.363) |
| log(population) | 0.131 | 1.965* | 0.461 | −0.142 | −0.740 | 0.113 |
| | (0.292) | (1.033) | (0.427) | (0.300) | (1.000) | (0.417) |
| NORTHWEST | | | −1.353 | | | −1.084 |
| | | | (1.046) | | | (0.954) |
| NORTHEAST | | | −0.914 | | | −0.971 |
| | | | (1.036) | | | (0.925) |
| SOUTHWEST | | | 0.180 | | | 0.747 |
| | | | (1.062) | | | (1.004) |
| SOUTHEAST | | | −0.666 | | | 0.102 |
| | | | (1.011) | | | (0.959) |
| Observations | 372 | 372 | 372 | 372 | 372 | 372 |
| Log-Likelihood | -206.236 | -137.747 | -196.451 | -199.474 | -134.780 | -186.883 |

[a] Arellano's method (1987) was used to compute a robust covariance matrix, allowing a fully general structure with respect to heteroskedasticity and serial correlation in panel data. In the context of a logistic regression, the approach was not possible for the logit RE model.

***, ** and * indicate a significance level of 1%, 5% and 10%, respectively. Standard errors in brackets.

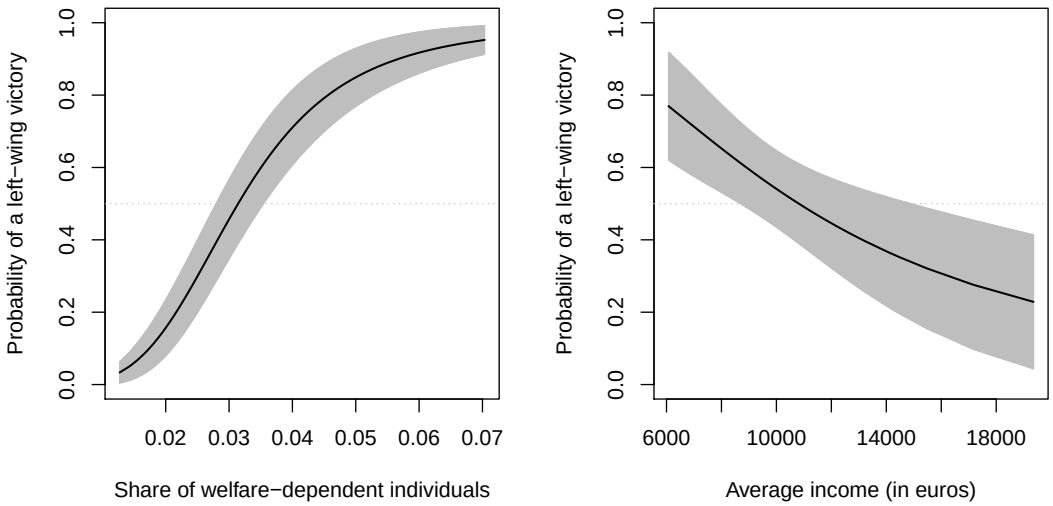



Figure 1.3. Vote function: simulating the probability of being on the left.

The overall results of the estimates (Table 1.3) are consistent with what we expected. A higher value of $WD$ increases the chances that a Department is governed by a left-wing coalition, which lends empirical support to P1. As for the impact of *income*, wealthier jurisdictions are less likely to vote for the left. To better assess the importance of these effects, the simulated probability of the victory of a left-wing party has been computed using an "average Department", coupled with the pooled logit model in column 1. The results for the share of welfare recipients and the mean taxable income are plotted in Figure 1.3, with 95% confidence intervals in gray. In both cases, the $x$-axis is specified with the minimum and maximum values of the related variable (see Table 1.1). As can be seen, the probability of a left-wing victory is highly sensitive to the $WD$ measure, yielding a differential shift of over 90% in the probability of the victory of a left-wing party (left-hand side). The differential shift amounts to 54% when it comes to the variation in income (right-hand side).

Columns 4–6 of Table 1.3 provide estimates where different groups of welfare recipients are examined in greater detail. The coefficients for $families$, *unemp*, and $elder$ appear with a positive and significant sign, with the greatest impact found for *unemp*. These results confirm the existence of a relationship between, on the one hand, electoral support for left-wing parties and, on the other, welfare-dependency and, to some extent, per capita income.

### 1.4.2 Public spending and share of welfare recipients

Next, we examine how the share of welfare recipients affects the level of public spending. The empirical specification of the spending equation is as follows:

$$\begin{aligned} \ln E_{i,t+1} &= \beta_0 + \beta_1 \ln WD_{i,t} + \beta_2 WD_{i,t} + \beta_3 \ln income_{i,t} + \beta_4 \ln density_{i,t} \\ &\quad + \beta_5 \ln population_{i,t} + \beta_6 \ln grants_{i,t} + \beta_7 DEC_t + \eta_{i,t}, \end{aligned} \tag{1.20}$$

where $i$ and $t$ again stand for Department $i$ and year $t$, respectively, now analyzed over the entire period of eleven years: from 1998 to 2008. The endogenous variable $E_{i,t+1}$ denotes the level of public expenditures per inhabitant; it represents either the per capita social expenditures of the Departments (*social*) or the per capita non-social expenditures (*nonsocial*). Expenditures are lagged one year to take into account the budget process: in French Departments the budget of year $t+1$ is proposed and approved between September of year $t$ and January of year $t+1$, i.e., by the government of year $t$. $WD$ is included both in a linear and a logarithmic form, allowing for the possibility of an inverted U-shaped relationship.

Two additional covariates are included in the specification of the vote function: *grants* and *DEC*. The variable *grants* denotes the transfers that the Departments receive from the central government; *DEC* is a time-dummy variable indicating the transfers of competences that occurred after 2002, when the central government endowed the Departments with two additional tasks: in 2002, a new welfare program targeting the elderly, called APA (*Allocation personnalisée d'autonomie*); then, with the *Decentralization Act* of 2004, an unemployment benefit using their own funds, the RMI (*Revenu minimum d'insertion*), replaced by the RSA (*Revenu de solidarité active*) starting from 2009.

The results are shown in Table 1.4. In line with what we expected, the estimated coefficients reveal a statistically significant relationship between $WD$ and $E$. This result appears consistent with P1. We once again use an "average Department" to better assess

Table 1.4. Spending equation – estimation results[a]

| | *Dependent variable:* | | | | | |
|---|---|---|---|---|---|---|
| | log(*social*) | | | log(*nonsocial*) | | |
| | Pooled | RE | Geogr. dummies | Pooled | RE | Geogr. dummies |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Constant | 4.010*** | 4.131*** | 4.237*** | 3.623* | 3.106*** | 4.753*** |
| | (0.730) | (0.533) | (0.612) | (1.864) | (1.121) | (1.512) |
| log(WD) | 0.140 | 0.161*** | 0.167* | 0.191 | 0.253 | 0.427* |
| | (0.097) | (0.062) | (0.093) | (0.291) | (0.180) | (0.221) |
| WD | 7.934*** | 6.588*** | 7.589*** | -8.409 | −17.884*** | −16.415*** |
| | (2.691) | (1.697) | (2.646) | (7.874) | (5.548) | (6.021) |
| log(income) | 0.152*** | 0.236*** | 0.136*** | 0.608*** | 0.779*** | 0.556*** |
| | (0.041) | (0.019) | (0.038) | (0.104) | (0.059) | (0.089) |
| log(density) | 0.001 | −0.015 | −0.023 | 0.129*** | 0.110*** | 0.098** |
| | (0.018) | (0.020) | (0.021) | (0.031) | (0.039) | (0.042) |
| log(population) | −0.020 | −0.034 | −0.001 | −0.307*** | −0.332*** | −0.270*** |
| | (0.032) | (0.033) | (0.030) | (0.053) | (0.056) | (0.044) |
| log(grants) | 0.111*** | 0.001 | 0.108*** | 0.124* | 0.099*** | 0.159*** |
| | (0.035) | (0.013) | (0.032) | (0.073) | (0.031) | (0.061) |
| DEC | 0.098*** | 0.168*** | 0.097*** | −0.218*** | −0.130*** | −0.220*** |
| | (0.019) | (0.010) | (0.018) | (0.036) | (0.020) | (0.032) |
| NORTHWEST | | | −0.083 | | | −0.199 |
| | | | (0.057) | | | (0.128) |
| NORTHEAST | | | −0.119** | | | −0.162 |
| | | | (0.055) | | | (0.122) |
| SOUTHWEST | | | −0.073 | | | −0.136 |
| | | | (0.059) | | | (0.142) |
| SOUTHEAST | | | −0.147*** | | | 0.085 |
| | | | (0.052) | | | (0.132) |
| Observations | 1,023 | 1,023 | 1,023 | 1,023 | 1,023 | 1,023 |
| $R^2$ | 0.772 | 0.909 | 0.793 | 0.312 | 0.291 | 0.428 |
| Adjusted $R^2$ | 0.766 | 0.902 | 0.784 | 0.309 | 0.289 | 0.423 |

[a] Arellano's method (1987) was used to compute a robust covariance matrix, correcting for heteroskedasticity and serial correlation in panel data.

***, ** and * indicate a significance level of 1%, 5% and 10%, respectively. Standard errors in brackets.

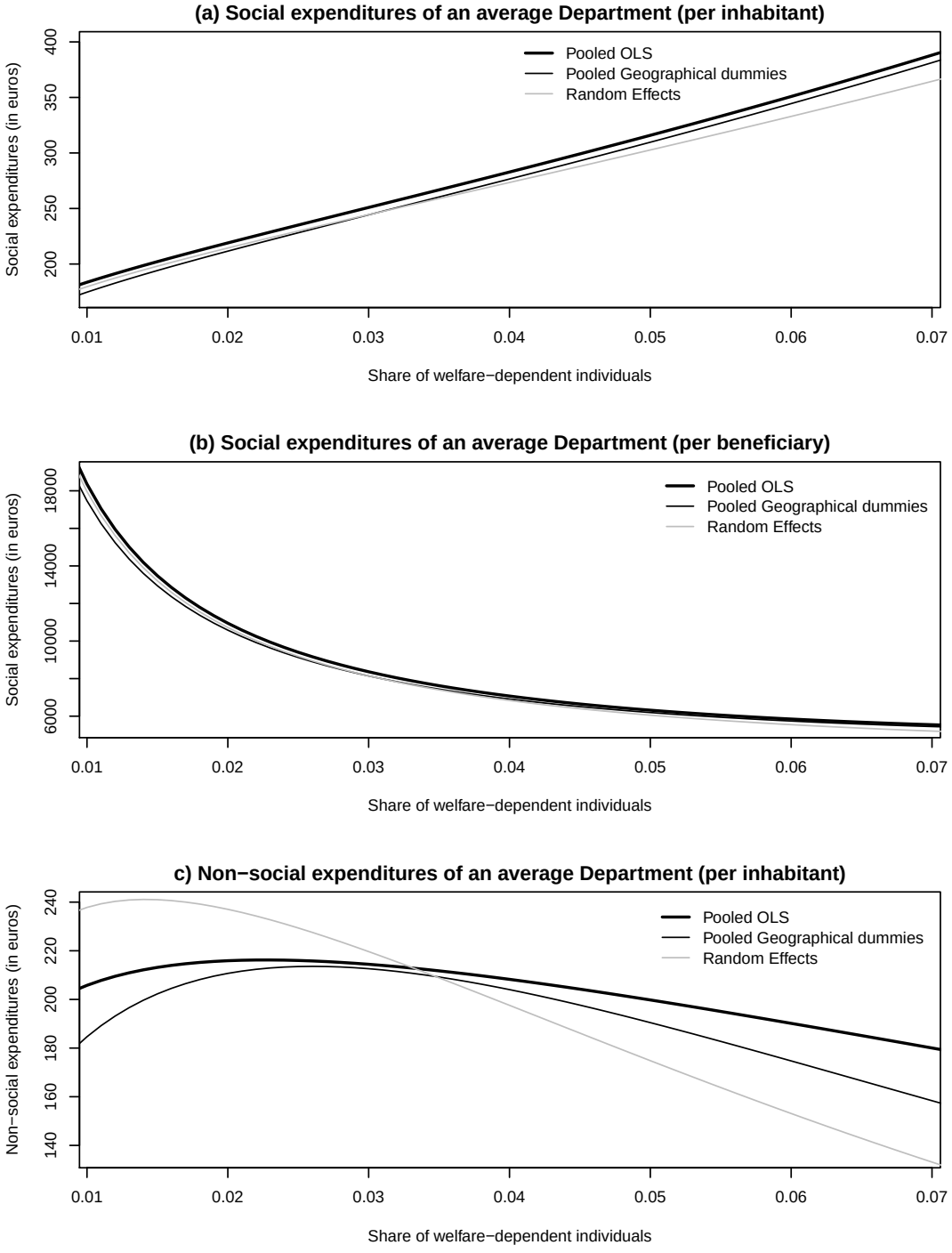



Figure 1.4. Spending and inequalities: simulations of an average Department (for graphical convenience, confidence intervals are not drawn).

the importance of the relationship, depicted in Figure 1.4. In panel (a), as we move from the minimum to the maximum share of welfare recipients in the electorate, average social expenditures per inhabitant double, from less than 200 euros to almost 400 euros. This is not surprising, since the endogenous variable may increase automatically with the number of welfare recipients. More importantly, panel (b) shows that the larger is the share of welfare recipients, the smaller is the amount of aid they individually receive from the Department. This lends support to P2, since a higher level of $WD$ is also synonymous with a higher tax price for taxpayers. Therefore, despite a higher number of welfare-dependent people (and, from Section 1.4.1, stronger political support for left-wing parties), the Departments marginally decrease the amount of aid per beneficiary. This effect appears to be convex: in panel (b), the larger is the share of welfare recipients, the lower will be the marginal decrease in social expenditures per beneficiary.

In panel (c) of Figure 1.4, non-social expenditures show first a diminishing marginal per capita increase, then they decline when the share of welfare recipients exceeds 3% of the population. This identifies an "optimal" level of $WD$ where the number of welfare recipients is sufficiently large to support an increase in public spending, but also sufficiently low not to erode taxpayers' support. Hence the two types of expenditures, *social* and *nonsocial*—chiefly targeted public spending the former, more universally spent the latter—are at first complementary goods, but then become potentially substitutes when the share of welfare recipients grows too large.

The log-log functional form allows to interpret the other estimated parameters as elasticities. The income elasticities and the grant elasticities are always positive and highly significant; their estimated coefficients range from 0.1 to 0.24 for social expenditures and from 0.1 to 0.8 for non-social ones (Table 1.4). Departments that are both richer and receive more in the form of grants tend to spend more, *ceteris paribus*. This suggests that the demand for public spending reacts as any normal good. Finally, Departments with a larger population and a lower population density are associated with lower non-social spending per inhabitant, which is evidence of important scale economies. Some of the estimated variations of public spending are also related to the mechanical effects of the new transfer of competences in

2002 and 2004: the *DEC* dummy has a significant positive impact on social expenditures, but a negative impact on non-social ones.

## 1.5 Test for partisan effects

So far the previous analysis has shown that left-wing parties tend to be elected in areas characterized by greater inequalities; because of that, they do not face the same demand for public goods as right-wing ones. To deal with this potential selection bias, generated by the voting process itself, we apply two different approaches: (1) regression discontinuity design (RDD) and (2) propensity score matching (PSM).

### 1.5.1 Regression discontinuity design (RDD)

The RDD approach exploits the discontinuity in the share of seats held by the incumbent political coalitions. Departments with a share of left-wing seats ($leftshare$) just below 50%, which are governed by the right wing, are compared with Departments where left-wing parties have obtained a number of seats just above the majority and have therefore formed the government. The two groups of Departments are likely to be very similar in their exogenous characteristics. Differences in policy choices found around the 50% threshold are therefore likely to be generated by ideological differences rather than by differences in voters' preferences, which eases the concerns of a selection bias. If such differences are found, the partisan hypothesis is validated.

We focus on the election years (1998, 2001, 2004, 2008) and on the average spending per term of office, labeled *social_average* and *nonsocial_average*. Following the usual methodology, a bandwidth around the threshold is selected, in order to examine differences in the proximity of the 50% seat share. The estimations will use only the observations inside the bandwidth, excluding Departments with a very large electoral margin. The choice of the appropriate bandwidth is therefore crucial. Equally important is ensuring that the forcing

Table 1.5. Regression discontinuity design – estimation results: LATE

| | *social_average* | | *nonsocial_average* | |
|---|---|---|---|---|
| | IK | CCT | IK | CCT |
| Conventional RD | −4.536 | −10.646 | 25.967** | 32.369* |
| | (14.485) | (19.360) | (11.906) | (17.417) |
| Bias-corrected (robust) | 58.527 | −8.836 | 23.930 | 32.024 |
| | (44.266) | (22.559) | (54.811) | (20.269) |
| Observations | 275 | 161 | 342 | 152 |
| BW RD point-estimator ($h$) | 0.245 | 0.137 | 0.343 | 0.131 |
| BW bias-correction estimator ($b$) | 0.143 | 0.200 | 0.166 | 0.203 |

***, ** and * indicate a significance level of 1%, 5% and 10%, respectively. Standard errors in brackets.

variable ($leftshare$) is not subject to manipulation around this threshold, which the McCrary (2008) density test refutes.[6]

Of course, the smaller the bandwidth, the lower will be the likelihood of a selection bias; but there is a trade-off between the size of the bandwidth and the number of observations. The analysis employs bandwidths selected according to the method of Imbens and Kalyanaraman (2012, henceforth IK), which ensures an asymptotically optimal bandwidth under squared error loss, and of Calonico *et al.* (2014, CCT), to further guard against the use of too large bandwidths.

Estimates of the local average treatment effects (LATE) at the cutoff are reported in Table 1.5. The main bandwidths ($h$) are used to construct local linear regression point-estimators: the IK bandwidths treat respectively 275 and 342 observations out of 372, the smaller CCT bandwidths 161 and 152. The bias-corrected bandwidths ($b$) offer a bias-correction using local quadratic regression with a robust variance estimation. For *social_average*, the estimators reveal a non-significant difference between left-wing and right-wing jurisdictions. For *nonsocial_average*, the conventional RDD suggests a significant difference, of about 26 to 32 euros on average; with a robust variance estimation, however, there is no significant impact of ideology on spending.

[6]For each election year, the McCrary density test yields the following $p$-value: 0.142 in 1998, 0.770 in 2001, 0.861 in 2004, and 0.559 in 2008.

### 1.5.2 Propensity score matching (PSM)

Previous RDD results are limited to close elections; the method identifies treatment effects only locally, disregarding other electoral outcomes. This might pose a threat to identification, since policy convergence is more likely to be observed as elections get close, as first exemplified by the median voter theorem: if two main political blocs compete for office, one might expect that they converge to the position of the median voter, to maximize expected votes. In a multidimensional context, issue convergence can also be observed, as suggested by Sigelman and Buell (2004), who point out that the attention of competitors in the U.S. presidential campaigns converges on the same issues when elections are close (Padovano, 2013, provides a review). Moreover, using data from the French Departments, Le Maux *et al.* (2011) and Le Maux and Rocaboy (2016) show that political power is not only a matter of the absolute majority of seats, but also of their relative shares. In a close electoral contest, a powerful opposition holding a sufficient number of seats would still be able to influence policy outcomes. For all these reasons, we resort to propensity score matching (PSM) as an alternative identification strategy.

PSM allows to consider a larger variety of jurisdictions with various degrees of political competition. The idea is to select a comparison group of right-wing jurisdictions, also referred to as a "control" group in the literature, so as to make them resemble the left-wing ones in all respects, except for the ideology of the government in office. Once the matched groups are formed, the average treatment effect (ATE) is estimated, simply by computing the difference in means between the two groups.

In performing the matching, one would ideally find for each left-wing jurisdiction a right-wing one that is identical in all respects. Since the list of possible variables is too large to allow an exact match to be achieved, the focus shifts onto propensity scores, generally defined as the probability of treatment assignment. In our case, the propensity scores correspond to the fitted values of our vote function (column 6 of Table 1.3).[7]

[7]The random effects model, with the greatest log-likelihood, performs very poorly in terms of solving the selection bias, as it mostly yields standardized biases of over 25 to over 100; for that reason it has been excluded. A pooled model with geographical dummies is a good proxy for the correct specification of individual fixed

Table 1.6. Propensity score matching – results of matching algorithms: ATE

| | | | | *social_average* | | *nonsocial_average* | |
|---|---|---|---|---|---|---|---|
| | Observations | Treated | Controls | Mean diff. | *t*-stat | Mean diff. | *t*-stat |
| Before matching | 372 | 160 | 212 | 46.085*** | 6.859 | 16.972** | 2.576 |
| | Matched observations | Matched unweighted | Loss in % | *social_average* ATE | *t*-stat | *nonsocial_average* ATE | *t*-stat |
| Nearest neighbor | 372 | 394 | 0 | 14.219 | 1.211 | 18.169 | 1.445 |
| Caliper = 0.1 | 322 | 344 | 13.44 | 8.748 | 1.107 | 18.383** | 2.206 |
| Caliper = 0.05 | 293 | 315 | 21.24 | 6.638 | 0.948 | 15.056** | 2.051 |
| Caliper = 0.025 | 257 | 279 | 30.91 | 5.248 | 0.877 | 17.929*** | 2.815 |

***, ** and * indicate a significance level of 1%, 5% and 10%, respectively.

Different matching algorithms are applied. First, nearest neighbor matching (NNM) links any two left-wing and right-wing Departments based on the most similar propensity scores, until all observations have been matched. Replacement, i.e., a repeated use of the same controls, is allowed, which should ensure a higher average quality of matching and avoid problems of dependence on the order in which the matches are made (Caliendo and Kopeinig, 2008). NNM may, however, associate Departments with very different propensity scores, and hence characteristics, if no closer match is available. For that reason, a maximum permitted distance is determined next, through a caliper matching algorithm, in our case 0.1, 0.05 and 0.025 standard deviations of the propensity score.

Before performing the matching, we could conclude that left-wing and right-wing councils implement significantly different social and non-social expenditure policies. Per capita social expenditures averaged 308.5 euros in leftist Departments, compared to 262.4 euros for their right-wing counterparts, i.e., a difference of 46 euros; non-social expenditures stood at 214.8 euros and 197.9 euros, respectively, i.e., a difference of 17 euros. After the matching, however, the test for the difference between the means leads to an altogether different conclusion: left-wing governments do *not* spend more on social expenditures than right-wing governments. None of the matching algorithms report a statistically significant ATE for *social_average* (column 5 in Table 1.6), which means that

effects, which are impossible to estimate due to insufficient inter-departmental variation in our case.

no partisan effects are observed for social expenditures. In contrast, for *nonsocial_average* (column 7), all caliper matching algorithms show evidence that leftist governments actually spend more, by about 18 euros per capita. Such a difference is nevertheless relatively small when compared with the data range, given the minimum of 59.5 and the maximum of 483 euros (see Table 1.1).

### 1.5.3 Robustness of the results

An important feature of the RDD and PSM analysis is the possibility to verify whether the question of selection bias has been resolved, i.e., whether the treatment group and the control group are balanced. To formally assess the quality of the comparison, two-sample $t$-tests of exogenous variables are recommended (see Tables 1.7 and 1.9). Before applying the methods, we can observe highly significant differences in the control variables between the left-wing and the right-wing group (columns 2–3), with the exception of the average income and population density. With a decreasing RDD bandwidth, Table 1.7 indicates an increasing quality of the comparison at the threshold, showing the diminishing statistical difference in the key socio-economic variables. With PSM (Table 1.9), the differences between the compared groups of Departments are reduced merely to the share of families with children at risk.

PSM also offers additional popular approaches to assess the quality of the match, such as the calculation of the average standardized bias (SB), defined by Rosenbaum and Rubin (1985) as the difference of sample means for each covariate in the treated group and in a matched control group, divided by a square root of the average of sample variances in both groups. While there are no formal rules, SB between 3% and 5% is usually seen as sufficient (Caliendo and Kopeinig, 2008), while SB of 20% after matching is considered large (Rosenbaum and Rubin, 1985). The tighter the caliper, the higher the quality of the match, as demonstrated by the fewer significant differences in the control variables between the left-wing and the right-wing group (Table 1.9), and also by the reduced value of the SB in Table 1.8 (e.g., by 99 percentage points for the key variable $WD$).

Table 1.7. RDD quality indicators – difference in exogenous variables: two-sample $t$-tests

| | | | *Bandwidth:* | | | | | |
|---|---|---|---|---|---|---|---|---|
| | Whole sample | | 0.343 | | 0.245 | | 0.131 | |
| Variable | Right | Left | Right | Left | Right | Left | Right | Left |
| WD | 0.031 | 0.041*** | 0.031 | 0.040*** | 0.031 | 0.040*** | 0.031 | 0.037*** |
| families | 0.002 | 0.002*** | 0.002 | 0.002*** | 0.002 | 0.002*** | 0.002 | 0.002* |
| unemp | 0.013 | 0.018*** | 0.013 | 0.018*** | 0.013 | 0.017*** | 0.014 | 0.016*** |
| elder | 0.012 | 0.017*** | 0.012 | 0.016*** | 0.012 | 0.017*** | 0.012 | 0.015** |
| disabled | 0.004 | 0.004** | 0.004 | 0.004 | 0.004 | 0.004* | 0.004 | 0.004 |
| income | 8,815 | 9,015 | 8,816 | 9,048 | 8,823 | 9,111 | 8,715 | 9,296* |
| density | 277 | 398 | 290 | 437 | 325 | 310 | 227 | 474 |
| population | 569,951 | 685,241** | 573,693 | 727,072*** | 574,978 | 708,794** | 604,438 | 716,028 |

***, ** and * indicate a significant difference at a 1%, 5% and 10% level, respectively.

Table 1.8. Matching quality indicators – standardized bias (SB) in percentage

| | | *Matching algorithm:* | | | |
|---|---|---|---|---|---|
| Variable | Before matching | Nearest neighbor | Caliper 0.1 | Caliper 0.05 | Caliper 0.025 |
| WD | 98.58 | 12.38 | 3.58 | 0.71 | −0.50 |
| families | 33.54 | −5.73 | −18.99 | −17.18 | −21.74 |
| unemp | 93.67 | 21.25 | 14.01 | 7.56 | 2.81 |
| elder | 55.43 | 2.05 | −2.22 | −1.52 | −0.21 |
| disabled | 24.28 | 9.82 | −0.07 | −3.31 | −2.81 |
| income | 10.45 | −7.68 | −7.87 | −11.26 | −7.67 |
| density | 10.29 | 0.74 | −2.78 | −6.99 | −7.53 |
| population | 25.11 | 15.00 | 7.98 | 2.28 | −0.63 |

Table 1.9. Matching quality indicators – difference in exogenous variables: two-sample $t$-tests

| | | | *Matching algorithm:* | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Before matching | | Nearest neighbor | | Caliper 0.1 | | Caliper 0.05 | | Caliper 0.025 | |
| Variable | Right | Left | Right | Left | Right | Left | Right | Left | Right | Left |
| WD | 0.031 | 0.041*** | 0.034 | 0.036* | 0.036 | 0.036 | 0.035 | 0.035 | 0.035 | 0.035 |
| families | 0.002 | 0.002*** | 0.002 | 0.002 | 0.002 | 0.002** | 0.002 | 0.002** | 0.002 | 0.002** |
| unemp | 0.013 | 0.018*** | 0.014 | 0.015*** | 0.014 | 0.015* | 0.014 | 0.015 | 0.014 | 0.014 |
| elder | 0.012 | 0.017*** | 0.015 | 0.015 | 0.015 | 0.015 | 0.015 | 0.015 | 0.015 | 0.015 |
| disabled | 0.004 | 0.004** | 0.004 | 0.004 | 0.004 | 0.004 | 0.004 | 0.004 | 0.004 | 0.004 |
| income | 8,815 | 9,015 | 8,891 | 8,748 | 9,007 | 8,855 | 9,060 | 8,839 | 9,024 | 8,874 |
| density | 277 | 398 | 254 | 261 | 276 | 248 | 295 | 223 | 317 | 234 |
| population | 569,951 | 685,241** | 539,793 | 596,279** | 548,814 | 578,341 | 553,390 | 561,799 | 555,568 | 553,333 |

***, ** and * indicate a significant difference at a 1%, 5% and 10% level, respectively.

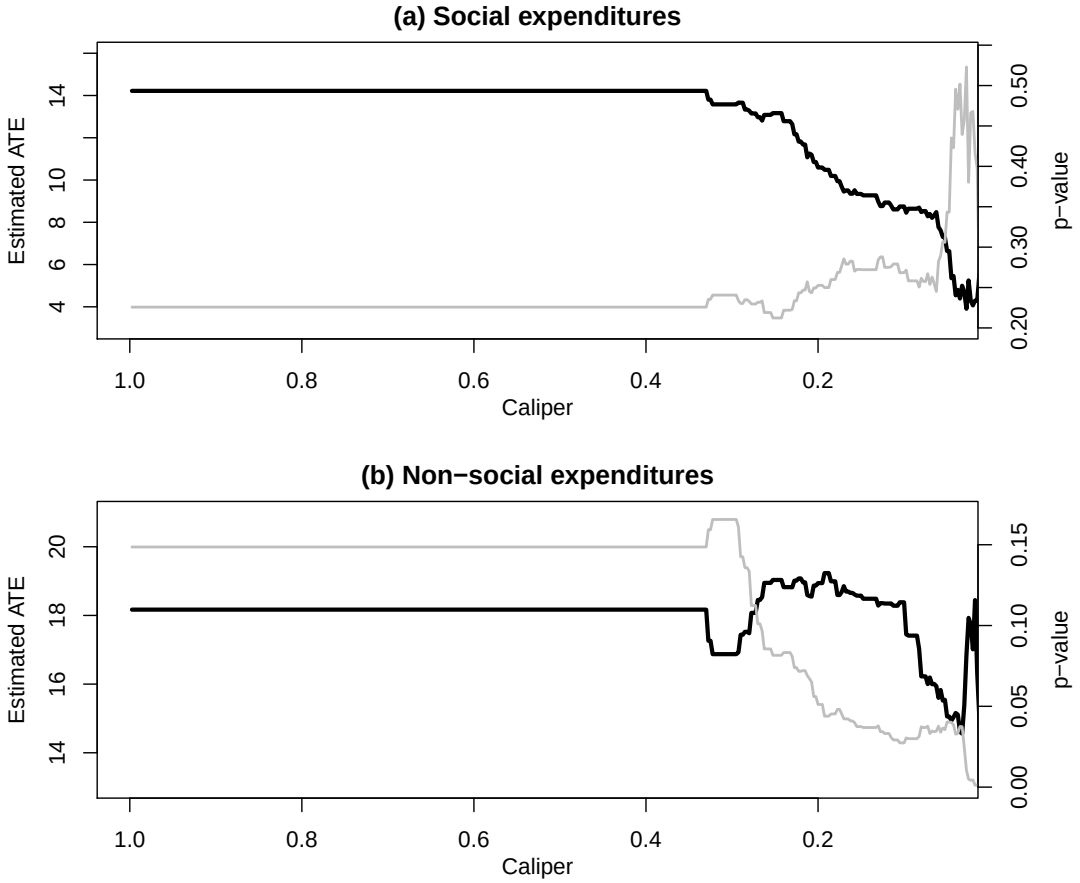



Figure 1.5. Propensity score matching: ATE (in black) and $p$-value (in gray) as a function of caliper size.

To further validate our findings, we have performed two robustness checks in the form of variations of the original PSM quasi-experimental approach. First, we have addressed the fact that we have exogenously defined the calipers, showing only three possible outcomes of the analysis. To avoid this limitation, Figure 1.5 plots the estimates of the average treatment effect (ATE) provided by PSM as a function of a decreasing caliper width for both (a) *social_average* and (b) *nonsocial_average* expenditures. The evolution of the corresponding $p$-value (gray line) emphasizes the boundary of statistical significance for the ATE, which for *social_average* is never below 10%. Both the size and the significance of the effect continue to decrease steadily. In contrast, the treatment effect for *nonsocial_average* grows in statistical significance with a tightening caliper, surpassing the 5% and 1% level once the caliper size drops below 0.2 and 0.03, respectively. The difference in non-social expenditures remains stable at around 18 euros per capita, corroborating the small partisan

Table 1.10. PSM – cross-sectional analysis and robustness of the results with respect to election years

| | *social* | | | *nonsocial* | | |
|---|---|---|---|---|---|---|
| | Before matching | Caliper 0.1 | Caliper 0.05 | Before matching | Caliper 0.1 | Caliper 0.05 |
| **1998** | 22.972*** | -13.514** | -12.631** | 17.360* | 21.350** | 19.870*** |
| 1999 | 21.241*** | -19.837*** | -17.502*** | 21.902** | 27.233*** | 25.061*** |
| 2000 | 25.475*** | -9.719 | -6.954 | 19.575* | 15.816* | 15.529** |
| **2001** | 26.002*** | 2.105 | -1.934 | 21.409** | 31.514*** | 22.651*** |
| 2002 | 37.152*** | 5.558 | 1.807 | 15.015 | 25.152** | 16.977* |
| 2003 | 41.125*** | 5.841 | -2.641 | 14.713 | 34.910*** | 25.759*** |
| **2004** | 30.746*** | 6.751 | 12.021 | 0.288 | 5.883 | -2.778 |
| 2005 | 33.887*** | 9.618 | 16.206* | 0.265 | 8.244 | -0.478 |
| 2006 | 31.593*** | 9.778 | 14.100 | 4.010 | 2.408 | -4.820 |
| 2007 | 30.019*** | 7.968 | 13.115 | 0.031 | 12.359 | 5.148 |
| **2008** | 26.379** | 3.878 | 0.069 | 11.662 | 6.428 | 15.595 |

***, ** and * indicate a significance level of 1%, 5% and 10%, respectively.
Caliper 0.025 was disregarded due to a small remaining sample size ($N = 23$).

effects found in this type of departmental outlays in Section 1.5.2.

Second, we have performed a cross-sectional analysis (Table 1.10). This modification is especially important for PSM, since it does not automatically distinguish between the time series of each Department when observations get matched. Before matching (columns 2 and 5 of Table 1.10), a series of *t*-tests reveals a statistically significant impact of ideology on social expenditures, in the sense that left-wing governments tend to spend more, although rarely on non-social expenditures. As before, no significant difference is found with regard to social expenditures in matched Departments (columns 3–4), with the exception of the years 1998–1999 where the ATE is actually negative, implying that left-wing governments in fact spent less.

For non-social expenditures, the cross-sectional PSM again corroborates the result that left-wing governments spend more (columns 6–7), although only until the 2004–2007 legislature. Overall, the range of results is quite similar to what has been measured with the whole sample in Section 1.5.2. The change observed after 2004 may be due to the transfer of competences that introduced new mandatory social expenditures, which in turn reduced the discretionary power of departmental councils and induced pressure on their budget, particularly for left-wing governments facing a large share of welfare-dependent people.

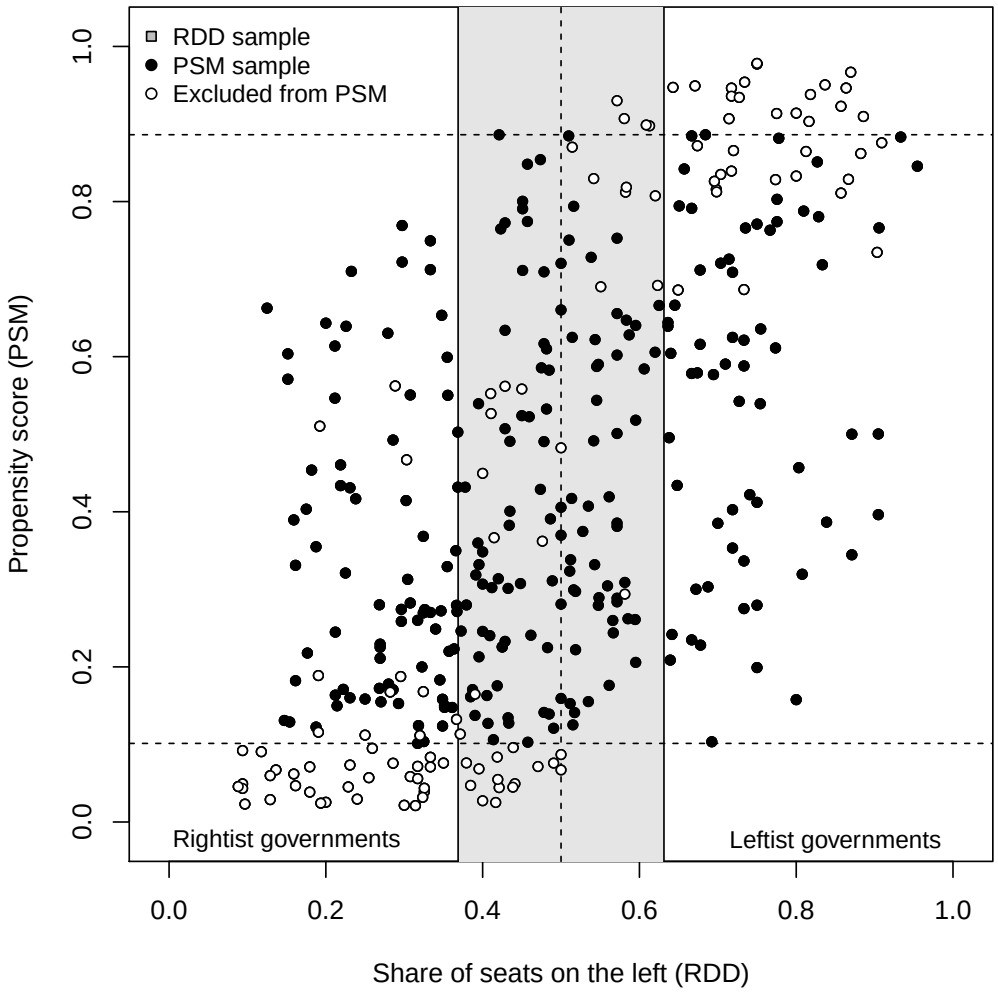



Figure 1.6. Comparison of observations taken into account by RDD and PSM.

### 1.5.4 Comparison of RDD and PSM

Although both regression discontinuity design and propensity score matching are quasi-experimental methods designed to control for selection bias, they approach the same problem from two very different angles. Neither treats the same subsample, nor do they provide exactly the same results. RDD plots gradually decreasing bandwidths around a threshold at a 50% share of seats, measured by the results of actual elections. PSM instead treats the entire sample, excluding *a priori* only the extreme observations that exceed the

common support area, i.e., an overlap in propensity scores between the two groups, based on an estimated probability of winning the election.

Despite these differences, both RDD and PSM concur that partisan effects are not found in social expenditures, while they appear to be either non-significant (24–32 euros, RDD) or significant (18 euros, PSM) in non-social expenditures. Figure 1.6 reiterates our results from Sections 1.5.1 and 1.5.2 to show that, while within the smallest bandwidth (CCT bandwidth of 13.1%; gray area) RDD deals with only 152 observations, under the tightest caliper (equal to 0.025) PSM concentrates on 257 (black points). Both methods therefore always take into account only the observations marked in black inside the gray field.

The obvious advantage of PSM is that it does not focus only on close elections, which allows comparing more Departments than RDD. Furthermore, by controlling the matching directly, we are able to reduce the observed differences in the exogenous variables between the left-wing and the right-wing Departments much more accurately. The assessment of the quality of the RDD and the PSM analysis in Section 1.5.3 shows that the comparison of the groups is better balanced for PSM (Tables 1.8 and 1.9), particularly with respect to the key variable $WD$ and its components, which is not the case for RDD (Table 1.7).

## 1.6 Conclusion

Despite numerous theoretical explanations and a large empirical literature, the issue of what ultimately drives policy choices in representative democracies—politicians' ideology or voters' preferences—is far from being settled. In the political economy literature, the standard theories stem from two alternative characterizations of the political process: an "opportunistic" vision, where reelection-seeking politicians move freely along the entire policy space to satisfy voters' demands; and a "partisan" vision, where elected politicians implement their policy platforms and are therefore bounded by an ideological space. Explanations based on the opportunistic vision relate policy choices to the distribution of voters' preferences. Under this view, politics is competitive and politicians will always

satisfy their constituency, regardless of their personal convictions, for fear of losing the election. Conversely, partisan models attribute a more active role to the supply side of politics, as they consider politicians' and parties' ideology an important driver of policy choices. Politics is partisan, inasmuch as candidates who want to prevail over their opponents must present different political platforms and the probability of being elected depends on their credibility in implementing the promised policies. Empirical tests have exploited this idea, finding evidence that left-wing governments behave differently than their right-wing counterparts. Most of the tests, however, do not analyze what drives the emerging differences, i.e., whether they depend on voters' preferences or on candidates/parties' ideologies.

In this study we have innovated on the literature by considering both the demand and the supply side of the policy-making process. To begin with, we have provided a theoretical model that summarizes all the possible interactions between the distribution of voters' preferences, the ideology of the party in office, and the policy choices that are made. We show that the interactions between these three items are quite complex, even accepting the hypothesis that the distribution of voters eventually determines the policy choices. Furthermore, the nature of this relationship may involve non-constant marginal effects. More specifically, the mechanisms described in two of our propositions offset each other: while a greater share of welfare recipients may reinforce support for public spending, it may also increase the corresponding tax price, which puts a constraint on the scope of the government's policy choices. Hence the relationship between the share of welfare recipients and the amount of public spending is non-monotonic.

Our empirical analysis confirms that the distribution of voters significantly impacts the probability of an electoral victory of a left-wing party. In particular, we have uncovered a positive relationship between government left-wing ideology and categories of welfare recipients (families with children at risk, unemployed or elderly people). Moreover, the levels of both social and non-social expenditures have been found to be significantly related to the distribution of voters. For instance, we find that higher shares of welfare recipients correlate with lower marginal decrease in social assistance per recipient.

However, the resort to quasi-experimental techniques suggests that the partisan effects disappear in the case of social expenditures once the socio-economic characteristics of each Department, that proxy voters' preferences, are controlled for. The demand side process thus seems to entirely determine the policy decisions in this expenditure domain. Only in the case of non-social expenditures—curiously, those that are believed to be less ideologically sensitive—the situation appears more complex: left-wing governments are still observed to spend slightly more on average than their right-wing counterparts and socio-economic variables do not seem to represent a sufficient explanatory factor.

Of course, these results are obtained for the particular case of French Departments and cannot be easily generalized to other types of government levels and/or countries. Nevertheless, our findings do provide support to the idea that the distribution of voters is a strong determinant of both government ideology and public spending.

# Chapter 2

# Does the size of the largest party matter? Endogenous fragmentation, political competition, and local public expenditures in Finland

---

†

> *"The party with 51 percent of the vote in one election can do very little; that with 65 percent in two consecutive elections can do considerably more. At least as a first approximation, an economic firm with 49 percent of output exerts 49/51 as much influence on price as its larger rival. The situation in political parties is not much different."*
>
> *(Stigler, 1972, p. 99)*

## 2.1 Introduction

How does government composition affect the level of public spending? The political economy literature generally offers three main answers to this question: the Weak Government Hypothesis, the Partisan Theory and the Leviathan Government Hypothesis.

---

†This research is the result of cooperation with Benoît Le Maux (University of Rennes 1, CREM-CNRS, Condorcet Center for Political Economy) and Antti Moisio (OECD Centre for Entrepreneurship, SMEs, Regions and Cities).

For all three, institutional settings are sufficiently permissive to allow elected politicians to maximize their own utility functions at the expense of citizens' preferences. Political parties matter in the sense that they may divert public expenditures from their socially desired level. While sharing similarities, these hypotheses also depart significantly from each other. For instance, the Weak Government Hypothesis suggests a positive relationship between government fragmentation and public spending, irrespective of the ideology of the government in office. For the Partisan Theory, governments differ in their policy choices because political parties have different preference functions. Last, the Leviathan Government Hypothesis suggests that public spending will be higher the lower is the intensity of political competition.

Finding which theory is best suited to explain the size of government implies the need for an accurate identification strategy. This is particularly the case since several confounding factors may blur our understanding of the phenomenon. First, as suggested in Ashworth *et al.* (2014), competition for political office can reduce the potential for opportunism but, at the same time, if competition is too fierce, it may result in more fragmented governments and inefficient spending levels. A distinction thus has to be made between the degree of government fragmentation (e.g., whether or not there is a single-party majority) and the intensity of political competition (e.g., the number and size of parties). Supporting this idea, Ferris *et al.* (2008) suggest that the intensity of political competition should be considered as the primary, if not the only, important political factor explaining government size (see also Caplan, 2001; Le Maux *et al.*, 2011; Padovano and Ricciuti, 2009).

Second, an important issue that has not been given much attention in the empirical literature is that the government composition itself can be endogenous.[1] Depending on the electoral system, preexisting social cleavages and social heterogeneity may affect the

[1]The question of endogenous government composition has been discussed extensively in the political science, rather than the political economy literature. See, e.g., Martin and Stevenson (2001, 2010), Glasgow *et al.* (2012), Glasgow and Golder (2015), and Laver and Benoit (2015), for an overview of government/coalition formation theories, or Bäck (2008), Debus and Gross (2016), and Skjæveland *et al.*, (2007), concerning the local government level in particular.

number of parties running for office (see, e.g., Duverger, 1954; Lijphart, 1990; Ordeshook and Shvetsova, 1994), which may in turn affect the composition of the government (Folke 2014; Le Maux and Rocaboy, 2016; Pettersson-Lidblom, 2008). The characteristics of the electorate can therefore act as a potential confounding factor, which jointly affects the configuration of the political arena and policy choices.

To address these issues, our study not only controls for the socio-economic characteristics of the municipalities, but also extends the analysis to different types of elections with a varied intensity of competition. Our data set comprises of an 8-year panel of more than three hundred Finnish municipalities, from 2000 to 2007, covering two municipal elections in 2000 and 2004. The following empirical strategy is used. First, we provide suggestive evidence of a selection bias by showing that the socio-economic characteristics of the municipalities influence not only their levels of public spending, but also the composition of the government. Second, we resort to propensity score matching (PSM) and regression discontinuity design (RDD) methods to test for the impact of political fragmentation on the level of public spending, *ceteris paribus*. As a matter of fact, we show that there is no significant difference in public spending between single-party majorities and minority governments once the selection bias is controlled for.

Third, and more importantly, we find that fiscal choices are dependent on the intensity of political competition. Using sample-splitting techniques, we demonstrate that a significant breakpoint exists in our data when the largest party holds a supermajority of approximately three-fifths of the seats. This is confirmed by propensity score matching and regression discontinuity design: at this threshold, spending levels are found to increase significantly by no less than 233–249 euros per capita. Hence, while our approach is different from that of Ashworth *et al.* (2014), Ferris *et al.* (2008), and Padovano and Ricciuti (2009), our results are in line with their findings: political competition is found to discipline politicians.

The paper is organized as follows. Section 2.2 presents a review of the literature.

Section 2.3 introduces the political context in Finland and highlights several stylized facts. Section 2.4 shows evidence of a selection bias and discusses the identification strategy. Section 2.5 provides the empirical results. Section 2.6 offers several robustness checks. Last, Section 2.7 concludes.

## 2.2 Literature review

The Weak Government Hypothesis states that government fragmentation leads to higher public spending, public deficits and debt (Buchanan and Tullock, 1962; Weingast *et al*., 1981; Roubini and Sachs, 1989a, 1989b). Several empirical studies give support to the phenomenon: Ashworth and Heyndel (2005), Borge and Rattsø (2002), Le Maux and Zhang (2013), Padovano and Venturi (2001), and Rattsø and Tovmo (2002), among others. This link is mostly explained by the theory of common pool resources. Elected representatives are assumed to defend their own interests by, for instance, expanding a particular item of public spending. Since the cost of this policy will be financed by the public budget, a non-cooperative politician could opt for an increase in public spending that is higher than efficient. The phenomenon is all the more likely to be observed (1) if the number of political parties in office is high (the "law of 1/n" of Bradbury and Crain, 2001), (2) if the parties exchange favors to secure support for a policy (the "theory of logrolling" by Tullock, 1959), or (3) if interest groups influence the political agenda through the parties (see, e.g., Hyytinen *et al*., 2018, for an application in Finland).

There is also a sizeable literature on the effects of party ideology on the level of public spending. This is usually exemplified by left-wing governments spending or raising tax rates more than their right-wing counterparts (see, e.g., Bjørnskov and Potrafke, 2012; Padovano and Petrarca, 2014; Santolini, 2008). The theory assumes that politicians have incentives to reveal their preferences about the level of spending during the electoral campaign and to

implement their announced platform if they want to be reelected (Besley and Coate, 1997; Osborne and Slivinski, 1996; Persson and Tabellini, 2000). A policy divergence can thus be observed from one government to the other, making the ideology of the government a relevant variable for explaining policy choices (Padovano, 2013).

Regression discontinuity design has recently been used for testing the Weak Government Hypothesis and the Partisan Theory. The identifying assumption is that observations around a seat threshold are similar in all respects, excepting government composition (e.g., a discontinuity in government ideology at the 50% threshold). If this condition holds, any observed differences in policy outcomes are attributed to the differences in government composition. For instance, applying this approach to a set of Swedish local governments, Pettersson-Lidblom (2008) demonstrates that left-wing governments spend and tax 2–3 percent more than right-wing governments. Using data from a sample of Finnish municipalities for the years 1997–2012, Meriläinen (2018) finds that coalitions tend to spend more than single-party governments. Folke (2014) shows that changes in the representation of anti-immigration and green parties in Swedish municipalities have a causal effect on the key policies for these parties. Using data from German municipalities in the state of Bavaria, Freier and Odendahl (2015) suggest that the value of property taxes depends on the ideology of the ruling party.

Yet, the Weak Government Hypothesis and the Partisan Theory are not the only explanations of government's growth. In particular, Brennan and Buchanan (1980) depict political parties as Leviathan entities, whose main interest is the maximization of government revenues and spending. In this context, electoral constraints are supposed to play an important role: the higher the likelihood of an electoral victory, the lower the political costs of a growing budget (e.g., Caplan, 2001). Empirical evidence of this effect is however mixed. For instance, using data on seat margins, Dubois *et al.* (2007) and Solé-Ollé (2006) refute this hypothesis for the French Departments and the Spanish municipalities, respectively.

Ferris *et al.* (2008), on the other hand, examine the evolution of public expenditures by the Government of Canada over 130 years and conclude that less competition leads temporarily to a larger-sized government. Similarly, using data from 308 Flemish municipal governments, Ashworth *et al.* (2014) corroborate that competition is associated with lower productive inefficiency. Last, examining a sample of 15 Italian Regions from 1980 to 2002, Padovano and Ricciuti (2009) find evidence that higher political competition improves economic performance, through the choice of more efficiency-oriented policies.

The literature thus offers three different insights into the role government composition may play in public spending. While the Weak Government Hypothesis focuses on the number of parties/politicians independently of their ideology, the Partisan Theory treats the ideology of the government as a key determinant which is likely to influence the level of public spending in one direction or another. The Leviathan Hypothesis, meanwhile, considers the intensity of political competition a key variable that may drive policy choices in the optimal direction. In the past years, few studies have attempted to consider these explanations in a single empirical framework. Exceptions are Ashworth *et al.* (2014), Dubois *et al.* (2007), Le Maux *et al.* (2011), Osterloh and Debus (2012), and Solé-Ollé (2006), but with contradictory results and examining only two hypotheses at a time.

Another issue is that political power is not merely a matter of absolute majority of seats. For instance, according to Le Maux *et al.* (2011) and Le Maux and Rocaboy (2016), the electoral margin of the majority and the number of parties in both coalitions are key variables that determine their effective political power. A powerful opposition can thus impose large costs upon the majority. This idea can be traced back to Wildavsky's (1986) theory of budgetary incrementalism and Stigler's (1972) theory of party competition. What actually matters is party strength and political influence. A similar view has been expressed in Ashworth *et al.* (2014). Their study makes a distinction between (1) the degree of government fragmentation, e.g., through a dummy variable indicating whether or not there

is a single-party government, and (2) the intensity of political competition, e.g., using a measure of political volatility or the size of the ruling party.

In order to capture the different aspects of political interactions, our analysis will rely on propensity score matching, regression discontinuity design, and several test for endogenous thresholds. More specifically, we address the question of the optimal seat share for a ruling party to occupy a comfortable position. We thus depart from the existing empirical literature (e.g., Folke, 2014; Freier and Odendahl, 2015; Meriläinen, 2018; Pettersson-Lidblom, 2008) in that we do not examine close elections only.

## 2.3 Finnish local public sector

### 2.3.1 The role of municipalities in Finland

Finland has two tiers of government: the central and the local level. The local government in Finland is currently formed by a little over 300 municipalities, following a massive series of mergers since 2004 and particularly 2008, induced by the *Act on Restructuring Local Government and Services (169/2007)* which came into effect in March 2007.[2] The purpose of the restructuring process, still progressing and now including a debate on introducing a regional administrative level, is to create a solid structural and financial basis for local-government services equally in all parts of Finland.

Finnish municipalities are self-governing entities by Constitution. This means that central government cannot assign new responsibilities to municipalities without first passing legislation to this effect. Nevertheless, the provision of many public services has been delegated from central government to the municipal sector, leaving Finnish

[2]E.g., in 2012 there were 320 registered municipalities, 399 in 2008, but as many as 436 as of 2000, most of them under 5,000 residents. The number of municipalities has already gone down by municipal mergers in the past: after the Second World War, Finland counted 558 local units, reduced to merely 460 by 1990.

municipalities with exceptionally wide-ranging responsibilities by international comparison. They are responsible for providing social welfare and healthcare services, as well as most education and culture, basic environment and technical infrastructure services. Due to the many assigned tasks, the overall economic importance of municipalities is considerable: municipality spending as a share of GDP reaches around 18% and municipalities employ roughly 20% of the total Finnish workforce.

According to their self-governing role, municipalities have independent taxation rights and they decide their own budgets. Although central government administers tax collection, municipalities have full autonomy in determining local earned income tax rates, which is their most important source of revenue. In addition to the local income tax, central government levies its own income tax. Municipalities are the sole receivers of property taxes, although their importance as a revenue source is small. Municipalities also receive a share of corporate tax revenues which is a central government tax. The main part of total municipal sector revenues is raised through own source revenues (e.g., in 2010 only 19% of the revenues came from grants). However, it must be noted that most Finnish municipalities are small and therefore heavily reliant on the grant system: in 2010 the grants covered more than a half of the budget for one-fourth of the municipalities.

Between 2000 and 2008, municipal finances have developed favorably with an average annual growth of taxable income at around 5%. Grants from central government have been increased during this period as well. However, the economic situation of Finnish municipalities has not always been so prosperous. Between 1990 and 1994, Finland faced a severe economic slump, during which GDP fell cumulatively by more than 10%, and the recession drove the public sector into a serious deficit. The economic situation started to improve after 1994, but municipal finances were still tight for many years, partly due to grant reductions during the years 1993–1998 (Moisio *et al.*, 2010), and municipalities reacted to the decreasing income tax base by raising tax rates, increasing fees for healthcare and social

welfare services, by borrowing, holding back investments and restraining the healthcare and social welfare expenditures.

The fact that nearly a half of the Finnish municipalities have a population of less than 5,000 inhabitants means that many municipalities are too small to organize all their services alone. As an answer to the economies of scale problem, the smallest municipalities have actively sought and participated in cooperative arrangements. The most common form of cooperation has been the joint authority—set up by two or more municipalities mainly for tasks that require a larger population base. The most important joint authorities include hospital districts, basic health care (health centers), districts for care for the disabled, vocational education and regional councils. Membership in a joint authority is voluntary with few exceptions, the most important being hospital services, where each municipality is obliged by law to belong to a hospital district joint authority.

### 2.3.2 Changes in the structure of Finnish municipalities

The present study uses a panel of Finnish municipalities between the years 2000–2007, covering two election years 2000 and 2004. The time period had to be reduced considerably due to the unavailability of reliable detailed data before 2000 and the recent frequent changes in the structure of the Finnish local public sector, as mentioned above.

To illustrate, Figure 2.1 overviews the recent developments, depicting on the left the total number of registered municipalities in each year (gray area) and their share having remained unchanged since 2000 (black line). The individual changes (municipalities divided or partially merged, created as new, or disestablished) are plotted in the right panel. By 2009, less than one-third of the municipalities registered at that point were the same administrative units with regard to their population, area, etc. (see, e.g., Saarimaa and Tukiainen, 2015, for an in-depth analysis of common pool problems related to municipal mergers).

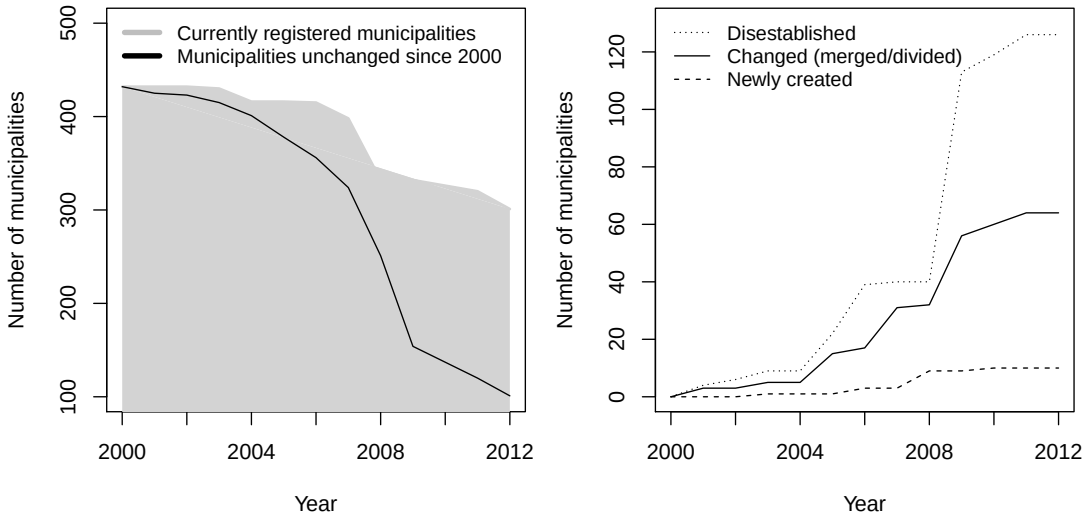



Figure 2.1. Changes in the structure of Finnish municipalities.

Unfortunately, likely due to the predominantly small size of the units concerned, data for the last and/or first year following mergers is not reported consistently, frequently reusing the same administrative code for either the original or the modified unit, or both. For all these reasons, and because we do not want to introduce a selection bias by focusing solely on the municipalities that did not merge, our study relies on the period 2000–2007, before the massive series of mergers took place.

### 2.3.3 Municipal elections

*Finnish Local Government Act* stipulates that each municipality must have a municipal council, a municipal board, an auditing committee, a committee for organizing elections, and a municipal manager, a civil servant elected by the municipal council. Municipal council is the highest decision-making body of the municipality. Its members are local politicians who have been elected in local elections. The council decides the yearly budget, the tax rates and the main objectives of the municipality.

The municipal board, selected by the municipal council, is responsible for municipal

administration and financial management. The municipal manager is a top civil servant who works under the municipal board as the head of municipal administration, financial management and other functions. This arrangement differs from most other European countries, where municipality managers are elected mayors, who are also chairs of local councils or boards or both. In Finland, the municipality managers hold their position either for a fixed term or the positions are permanent. However, in the largest cities the city managers are often *de facto* politicians who have risen up to this position from city councils or national politics (Moisio *et al*., 2010).

Municipal elections are held approximately every four years (e.g. 1996, 2000, 2004, 2008, 2012, and most recently, 2017). All persons who have residence in the municipality and are entitled to vote in local elections, and who are not legally incompetent, are eligible as candidates. The candidates are nominated by political parties or by at least a group of ten people entitled to vote in the municipality. The number of councilors that will be elected in each municipality depends on the population size. Each party may nominate a number of candidates equaling the number of councilors to be elected multiplied by one and a half. Parties may form electoral alliances, but the number of candidates nominated by an alliance may not exceed the maximum number of candidates for a single party.

The municipal elections are organized as direct secret voting, so that voters vote directly for the candidate. Municipal elections are proportional, i.e., each party gains seats in relation to the votes cast for it compared with the votes cast for other groups; the seats are allocated to parties based on the party vote shares in accordance with competitive indices set by the d'Hondt method. After the council elections, the members of the municipal board are chosen by the municipal council. The composition of the municipal board is based on the political makeup of the council: the parties represented in the council get seats on the municipal board according to their share of council seats.

Table 2.1. Summary statistics

| Variable | Description | N | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|---|---|
| *largest* | The largest share of seats held by a single party in each council (in %). | 736 | 49.005 | 14.954 | 19.608 | 100 |
| *MAJ* | A dummy equal to 1 if there is a single-party majority ($largest > 50\%$). | 736 | 0.454 | 0.498 | 0 | 1 |
| *HHI* | Herfindahl-Hirschman index (HHI) of the government, computed from shares of seats by party. | 736 | 0.372 | 0.143 | 0.163 | 1 |
| *Y* | Per capita total current operating expenditures (in euros). | 2,944 | 3,311 | 506 | 1,942 | 5,416 |
| *children* | Share of inhabitants aged 0–6 years in the population (in %). | 2,944 | 7.378 | 1.762 | 2.954 | 14.855 |
| *young* | Share of inhabitants aged 7–14 years in the population (in %). | 2,944 | 10.267 | 1.732 | 3.984 | 18.218 |
| *elderly* | Share of inhabitants over 65 years of age in the population (in %). | 2,944 | 19.354 | 4.637 | 5.162 | 35.563 |
| *unemp* | Unemployment rate (in %). | 2,944 | 11.322 | 4.425 | 1.900 | 28.935 |
| *population* | Number of inhabitants. | 2,944 | 12,643 | 37,273 | 237 | 568,531 |
| *density* | Number of inhabitants per $km^2$ (without water areas). | 2,944 | 57 | 207 | 0.2 | 3,040 |
| *income* | Mean taxable income (in euros). | 2,944 | 9,428 | 2,089 | 5,229 | 27,528 |

## 2.4 Identification strategy

This section discusses the identification strategy of the average treatment effect of government composition on public spending. Table 2.1 provides a summary of our dataset; the GDP deflator ($2000 = 100$) is used to deflate all the financial variables.

### 2.4.1 Measure of government fragmentation

Following the existing empirical literature (e.g., Folke, 2014; Freier and Odendahl, 2015; Meriläinen, 2018; Pettersson-Lidblom, 2008), we will test whether single-party majority governments spend more on average compared to minority governments. To do so, we use a simple dummy variable stating whether the government of municipality *i*

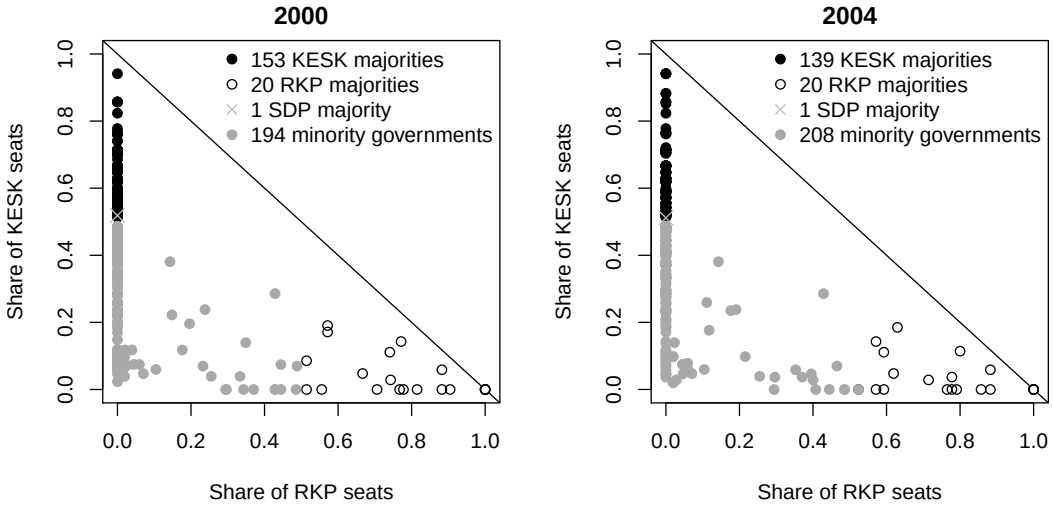



Figure 2.2. Distribution of seats after the 2000 and 2004 elections.

is a single-party majority ($MAJ = 1$ for a single-party majority in power and $MAJ = 0$ otherwise).

Figure 2.2 describes the dataset. In the 2000 and 2004 elections, the main political parties at the municipal level are namely KESK (the *Center Party*), SDP (the *Social Democratic Party of Finland*), KOK (the *National Coalition Party*) and RKP (the *Swedish People's Party*).[3] While KESK, SDP and KOK are large bodies averaging 15 to more than 40% of seats, the fourth party, RKP, holds the notable position of looking after the interests of the Swedish-speaking population in Finland, thus being consistently fairly popular in southern and western regions of the country. In the two election years, 334 single-party majority governments ($MAJ = 1$) were elected: 174 (i.e., in 47.3% cases) in 2000, 160 (43.5%) in 2004. Only three parties obtained a single-party majority: predominantly KESK in 153+139 (approximately 87.5%) and RKP in 20+20 (12%) instances, with SDP also winning 1 council majority in each election.

[3]KESK is a centrist moderately reformist party supporting social ideals and values, also known to defend the interests of rural areas and agriculture, and to protect natural assets. SDP is a socialist party that supports social democracy, welfare services provided by the state and municipalities, and opposes uncontrolled capitalism and conservative values. KOK is a moderately liberal and reformist party in support of individualism and entrepreneurship, emphasizing social responsibility of people and companies. Last, RKP, is a centrist non-socialist party promoting social liberalism and the position of the Swedish language.

Table 2.2. Evidence of a selections bias – two-sample $t$-tests

| | Single-party majorities $MAJ = 1$ | Minority governments $MAJ = 0$ |
|---|---|---|
| *children* (%) | 7.153 | 7.565*** |
| *young* (%) | 10.337 | 10.209** |
| *elderly* (%) | 20.822 | 18.135*** |
| *unemp* (%) | 11.875 | 10.862*** |
| *population* | 4,566 | 19,355*** |
| *density* | 9.111 | 95.937*** |
| *income* | 8,313 | 10,354*** |
| Nb. obs. | 1,608 | 1,336 |

***, ** and * indicate a significant difference at a 1%, 5% and 10% level, respectively.

The dummy *MAJ* is a convenient measure of government fragmentation (see, e.g., Meriläinen, 2018). The budget proposed by a single-party majority should rarely be amended, as it holds a majority of seats, while minority governments have to negotiate with other parties, form coalitions, and adjust policies to get enough votes to pass a project. In this consideration, however, we exclude the possibility that political power is also determined by the size of the ruling party. One might also argue that the size of the ruling party is negatively associated with the intensity of political competition (e.g., Ashworth *et al.*, 2014). Moreover, the distribution of voters' preferences may also affect the number of parties running for office. Accounting for these effects is the main challenge to our identification strategy.

### 2.4.2 Assumptions and threats to identification

In order to investigate whether government composition is related to the distribution of voters' preferences, we first provide a simple comparison between single-party majorities and minority governments. Table 2.2 provides results of two-sample $t$-tests and compares councils with a majority share of seats held by a single party ($MAJ = 1$) and those necessitating a multi-party ruling coalition ($MAJ = 0$). The highly statistically significant

differences in all the electorate attributes imply that single-party majorities and minority governments are elected under very different conditions, i.e., with councils facing different demands for public goods. In other words, our caution regarding the impact of the demand side seems pertinent here.

Propensity score matching is first used to overcome this selection bias. To explain the pros and cons of the method, let $i$ index the Finnish municipalities, $Y_{i1}$ denote the spending levels when the ruling party holds a majority of seats ($MAJ_i = 1$) and $Y_{i0}$ the spending levels when a coalition is required to pass a budget ($MAJ_i = 0$). We would like to estimate the average treatment effect:

$$\tau \equiv E[Y_{i1} - Y_{i0}]. \tag{2.1}$$

Estimating $\tau$ is however complicated by the fact that a municipality cannot be both exposed and unexposed to government fragmentation. A proper identification strategy thus implies finding a good reference point for comparison. For this purpose, propensity score matching constructs a statistical comparison group, aka control group in this literature, that is based on the estimation of a propensity score $\Pr(MAJ_i = 1|X_i)$, defined for a unit $i$ as the probability of belonging to the exposed group conditional on its observed characteristics $X_i$. Provided that there is a sufficient overlap between the estimated scores (aka common support condition), a matching based on the distance between point estimates of the propensity scores should eliminate all confounding $X_i$ (aka unconfoundedness assumption). Once the matched groups are formed, the average treatment effect is estimated by simply computing the difference in mean outcomes between the two groups (Abadie and Imbens, 2006, 2016; Rosenbaum and Rubin, 1983).

In our analysis, the propensity scores are obtained by regressing the single-party

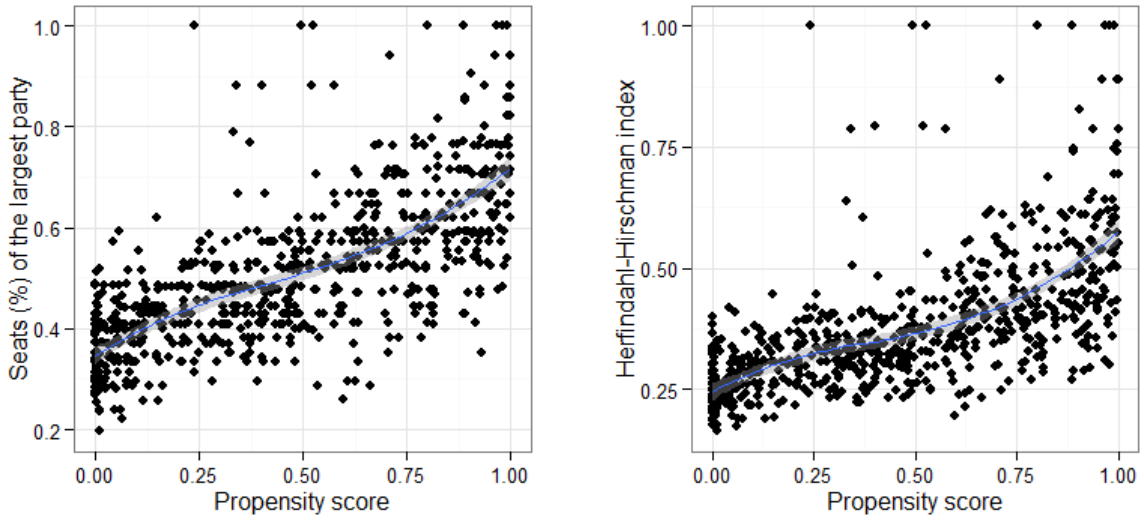



Figure 2.3. Comparison of electoral results and estimated propensity scores.

majority dummy *MAJ* on a set of socio-economic control variables:

$$MAJ_{i,t} = \gamma_0 + \mathbf{X}'_{i,t}\iota + \delta_{i,t}. \tag{2.2}$$

Following the empirical literature, $\mathbf{X}_{i,t}$ is a vector of control variables commonly used in local public finance research: population size (*population*), population density (*density*), share of inhabitants aged between 0–6 years (*children*), between 7–14 years (*young*), over 65 years (*elderly*), the unemployment rate (*unemp*), and the mean income (*income*).

Since the estimated scores represent the probability that a single party holds a majority in municipality $i$ in year $t$, the measure can be considered a proxy for the intensity of electoral competition. As shown in Figure 2.3 (see Section 2.5 for a description of the estimated scores), this measure is strongly related to the size of the largest party (left panel), but also to the total number of parties and their size, as exemplified by the Herfindahl-Hirschman index[4] (right panel). In other words, depending on the electorate attributes, it may be easier for the

[4]The Herfindahl-Hirschman index index is computed as $HHI = \sum_{p=1}^{N} S_p^2$, where $S$ denotes the share of seats won by party $p$, i.e., the fraction of the council held by each of the $N$ parties in the municipality in question. $HHI$ can range from $1/N$ to 1, where $1/N$ indicates a very heterogeneous council of "small parties" of equal size, while 1 means that all councilors belong to the same party.

largest party to gain a strong or even a majority position (e.g., an older constituency and a more secure job market both coincide with an increase in the likelihood of a single-party victory), which may in turn affect its ability to influence fiscal choices.

Two threats to identification need particular attention. The first relates to the endogeneity of $X_i$, the binary treatment ($Y_{i1}$, $Y_{i0}$), and the potential reverse causality of the outcome variable $Y_i$. For instance, the unemployment rate or the probability of a party's victory could be reversely related to public expenditures. In our context, there exists no formal way of assessing endogeneity issues (endogeneity tests are based on instrumental variable estimation methods). However, the time frame of the variables should guarantee a certain internal validity. Observed characteristics $X_i$ and the treatment are indeed measured at time $t$, where $t$ is the election year, and are thus independent of future policy choices occurring after the election. Note also that the output variable, denoted $Y\_average$ hereafter, will be calculated as a 4-year spending average over the entire term of office.

Second, if there are unobserved variables which affect both $MAJ$ and $Y$, a hidden bias might arise. Difference-in-differences cannot be used to solve this problem, since most municipalities have kept the same council structure (i.e., single-party majority or minority government) throughout the entire sample period. Regression discontinuity design (RDD) can thus be considered a more convenient identification strategy. Yet, results from the RDD analysis can only be inferred to close elections, around the 50% seat share threshold necessitated by the approach. This might pose a threat to identification in itself, since policy convergence is more likely to be observed when competition is fierce. Given our purpose, i.e., to test whether our results are stable under various intensities of political competition, we will use sample-splitting techniques (Zeileis *et al.*, 2003) in order to identify the level of political competition that is required to actually observe a significant difference in spending. Then RDD will be implemented around the optimal threshold to control for a possible unobserved selection bias.

## 2.5 Empirical results

In what follows, we will test whether the distribution of seats among parties is a key variable determining municipal expenditures. The method relies on the estimation of scores to select and pair governments with similar characteristics (Subsections 2.5.1 and 2.5.2). The strategy then proceeds with a search for endogenous thresholds in spending levels (Subsection 2.5.3).

### 2.5.1 Estimation of propensity scores

Estimates of the scores are provided in Table 2.3. Due to the binary character of the dependent variable, a probit model has been applied, using pooled (columns 1 and 2) and random effects (column 3) estimators. A fixed effects estimation would result in a considerable loss of information in this context, given the low variance of the outcome variable: the value of *MAJ* remained fixed throughout the entire period with the exception of 36 municipalities only. A pooled model with added geographical dummies, regrouping municipalities in similar areas, serves as a compromise instead, controlling for individual and potential spatial effects.[5]

According to the estimation results, single-party majorities seem more likely to be elected in poorer municipalities with a lower population density (rural areas), with a larger share of elderly inhabitants and a lower unemployment rate. Furthermore, taking into account spatial dimensions, they are more frequently found in the relatively more rural and isolated eastern and western and especially northern regions. Or, inversely, fragmented minority governments, reflecting more heterogeneous preferences, are preferred by the electorate of southern urban municipalities, richer on average and with younger

[5]The partition used is the following: *North* (Lapland, Kainuu, Northern Ostrobothnia), *East* (North Karelia, South Karelia, Northern Savonia, Southern Savonia), *West* (Central Finland, Central Ostrobothnia, Ostrobothnia, Southern Ostrobothnia, Pirkanmaa, Satakunta), *South* (Kymenlaakso, Päijänne Tavastia, Southwest Finland, Tavastia Proper) and *Capital* (Greater Helsinki).

Table 2.3. Propensity score estimation – single-party majority[a]

| | *Dependent variable:* | | |
|---|---|---|---|
| | *MAJ* | | |
| | Probit Pooled | Probit Geogr. dummies | Probit RE |
| | (1) | (2) | (3) |
| Constant | 29.986*** | 34.660*** | 57.312*** |
| | (7.945) | (8.141) | (20.739) |
| log(children) | 0.564 | 0.675 | 1.295 |
| | (0.518) | (0.581) | (1.349) |
| log(young) | 0.747 | 0.265 | 2.941* |
| | (0.614) | (0.654) | (1.714) |
| log(elderly) | 0.177 | 1.295** | 1.435 |
| | (0.676) | (0.649) | (1.571) |
| log(unemp) | −1.099*** | −1.895*** | −2.384*** |
| | (0.272) | (0.270) | (0.764) |
| log(population) | 0.029 | −0.146 | -0.140 |
| | (0.094) | (0.108) | (0.353) |
| log(density) | −0.622*** | −0.137 | −1.684*** |
| | (0.162) | (0.140) | (0.385) |
| log(income) | −3.262*** | −3.889*** | −6.645*** |
| | (0.590) | (0.660) | (1.622) |
| NORTH | | 2.649*** | |
| | | (0.449) | |
| SOUTH | | 0.272 | |
| | | (0.339) | |
| EAST | | 0.975** | |
| | | (0.384) | |
| WEST | | 0.742** | |
| | | (0.342) | |
| Observations | 736 | 736 | 736 |
| Log-Likelihood | -346.733 | -304.961 | -284.243 |

[a] Arellano's method (1987) was used to compute a robust covariance matrix, allowing a fully general structure with respect to heteroskedasticity and serial correlation in panel data. The approach was not possible for the probit RE model.

***, ** and * indicate a significance level of 1%, 5% and 10%, respectively. Standard errors in brackets.

constituencies, yet facing economic struggles like higher unemployment.

### 2.5.2 Propensity score matching

The fitted values of a pooled probit model in column 2 of Table 2.3, with the largest log-likelihood, are used as propensity scores to implement the matching, as random effects are not applicable for modeling scores. Prior to matching the municipalities, a check for common support, i.e., a sufficiently large overlap in propensity scores between the two groups, has to be performed (see Figure 2.4); observations above/below the other group's maximum/minimum propensity score are excluded.

Different matching algorithms are applied. First, all treated and non-treated units are paired based on the most similar propensity scores, using nearest neighbor matching. Next, the analysis proceeds with a more precise caliper matching algorithm, gradually decreasing the caliper width. With a caliper, we can further reduce bias by limiting the maximum permitted difference between propensity scores, and thus the underlying characteristics. The size of the matched sample may, however, be reduced as well, since any observation outside the caliper is dropped. Note that replacement, i.e., a repeated use of the same controls, is permitted. This modification should lead to a higher average quality of matching and preclude any problems with dependence on the order in which matches are made (Caliendo and Kopeinig, 2008).

Overall, the approach is successful in achieving balance (see Appendix C for a discussion on the quality of the match). Estimation results are provided in Table 2.4. The outcome of interest is the average treatment effect (ATE), which denotes the difference in public spending between treated ($MAJ = 1$) and non-treated ($MAJ = 0$) observations. Without matching, the difference between single-party majorities and their minority counterparts amounts to 251 euros, with majorities spending significantly more according to

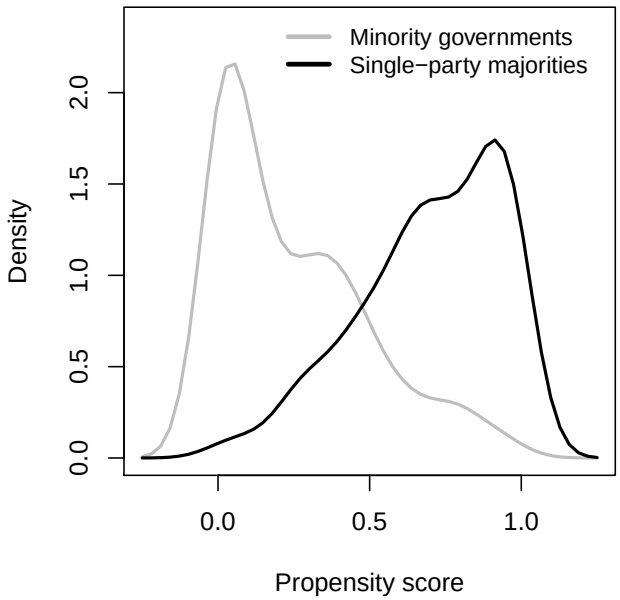



Figure 2.4. Visual analysis of propensity score density distribution in both groups.

a two-sample *t*-test. Once municipalities are matched using propensity scores, the analysis no longer reports a significant difference in *Y_average* (calculated as 4-year spending averages over each term of office). The magnitude of the difference between the groups is moreover much smaller, for instance 4 euros with the smallest reported caliper. That is, once we control for the heterogeneity of the electorate and the resulting political competition, single-party majorities are not found to raise public spending more than their more fragmented minority counterparts.

Mindful of the fact that only four outcomes of the matching process are presented in Table 2.4, i.e., nearest neighbor matching and three caliper sizes selected exogenously by us, Figure 2.5 complements the analysis by plotting the entire evolution of the average treatment effect (ATE) as a function of a decreasing caliper size. As can be observed, the results are independent of the caliper size: the *p*-values (in gray) are well above the 5% significance level. This evidence supports the idea that there is no actual difference in spending once we

Table 2.4. Results of matching algorithms – ATE for *MAJ*

| | | | | *Y_average* | |
|---|---|---|---|---|---|
| | Observations | Treated | Controls | Mean diff. | *t*-stat |
| Before matching | 736 | 334 | 402 | 250.493*** | 7.234 |
| | Matched obs. | Matched unweighted | Loss in % | ATE | *t*-stat |
| Nearest neighbor | 736 | 865 | 0 | 33.181 | 0.462 |
| Caliper = 0.06 | 713 | 842 | 3.13 | 42.071 | 0.612 |
| Caliper = 0.04 | 670 | 799 | 8.97 | 22.866 | 0.359 |
| Caliper = 0.02 | 558 | 687 | 24.18 | 4.135 | 0.076 |

***, ** and * indicate a significance level of 1%, 5% and 10%, respectively.

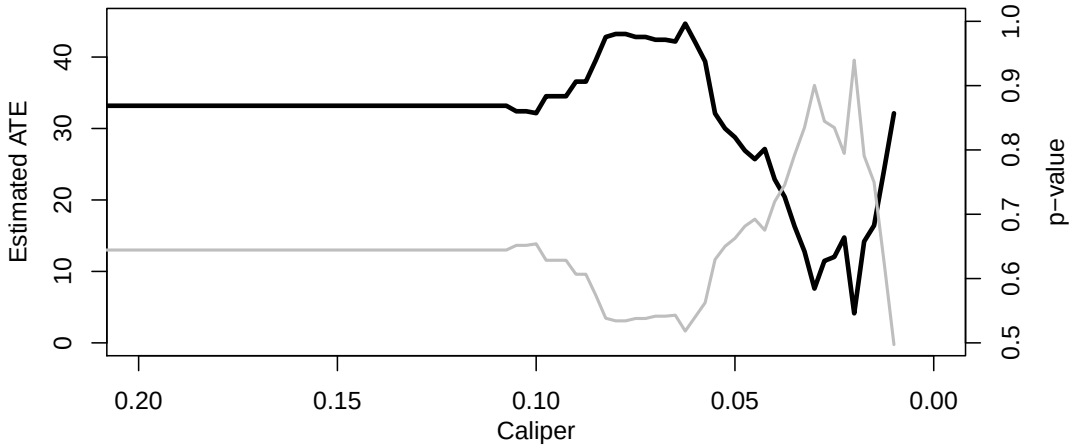



Figure 2.5. Propensity score matching: ATE (in black) and *p*-value (in gray) as a function of caliper size.

Table 2.5. Impact of political competition[a]

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | *Y_average* | | | |
| | (1) | (2) | (3) | (4) |
| Constant | 3,220.868*** | 3,222.413*** | 2,878.888*** | 2,989.158*** |
| | (31.654) | (29.465) | (130.810) | (101.506) |
| MAJ | 10.715 | | 228.183 | |
| | (87.275) | | (205.050) | |
| score | 454.966*** | 451.828*** | | |
| | (105.002) | (101.244) | | |
| MAJ × score | 47.151 | 62.040 | | |
| | (150.993) | (86.631) | | |
| largest | | | 1,159.602*** | 886.145*** |
| | | | (338.140) | (264.007) |
| MAJ × largest | | | -412.094 | 40.392 |
| | | | (416.820) | (124.575) |
| Observations | 736 | 736 | 736 | 736 |
| $R^2$ | 0.124 | 0.124 | 0.096 | 0.094 |
| Adjusted $R^2$ | 0.121 | 0.122 | 0.092 | 0.092 |

[a] Pooled OLS estimators with Arellano's (1987) robust standard errors.
***, ** and * indicate a significance level of 1%, 5% and 10%, respectively. Standard errors in brackets.

start controlling for the observed selection bias.

Last, to assess whether the intensity of political competition is of true importance, we also test the stability of the results with respect to the propensity scores by estimating the following regression model:

$$Y\_average_{i,t} = \alpha_0 + \alpha_1 MAJ_{i,t} + \alpha_2 score_{i,t} + \alpha_3 MAJ_{i,t} \times score_{i,t} + \varepsilon_{i,t}. \quad (2.3)$$

The approach is somewhat equivalent to a stratified propensity score matching (Caliendo and Kopeinig, 2008; Rosenbaum, 2002) and makes use of the whole dataset in order to estimate the interaction term $\alpha_3$. The aim is to test whether the outcome variable *Y_average* depends on *MAJ*, while controlling for the value of the estimated score. Table 2.5 provides the estimation results (columns 1 and 2). As can be seen, conclusions about the average treatment effect are stable: there is no significant difference in public spending between single-party majorities and minority governments. Only the intensity of political competition (*score*) has a significant impact: a percentage point increase in the propensity score yields on average a 4.5 euro increase in per capita spending. As a robustness check, columns 3 and 4 of Table 2.5 extend the analysis to the size of the largest party (*largest*). Again, the results for the average treatment effect are found to be stable and non-significant.[6]

### 2.5.3 Tests for endogenous thresholds

So far, our results show that (1) politics is not a matter of absolute majority of seats (the 50% seat share threshold does not show a significant effect), while (2) spending levels are in fact dependent on the intensity of political competition. Thus, the question that remains is whether there exists an optimal level of competition granting the ruling party a strategic

[6]The examination of two separate PSM analyzes based on strata arbitrarily constructed at the 50% threshold, i.e., *score* $\lessgtr$ 50%, confirms that municipalities with higher scores tend to spend more independently of their status, i.e., single-party majority or minority government.

position.

To answer this question, we test for the presence of endogenous thresholds, using the empirical approach proposed by, among others, Zeileis *et al.* (2003). The idea is to assess whether regression results are stable when examining appropriately selected subsamples, here based on the size of the largest party. The methodology is similar to a Chow test—calculated for each possible break, with the thresholds selected using the asymptotic $p$-values computed through the algorithm proposed by Hansen (1997).

Panels (a) and (b) of Figure 2.6 plot the results using first a model with a constant only, then a model including a constant and the variable *largest*, in order to locate a general breakpoint in the data. The horizontal red line represents values below which the significance of the threshold is rejected at the 1% significance level. In panel (a), the resulting split point, at approximately 59%, is shown to be statistically significant, suggesting that the sample can be split into two groups or "regimes" with different intensities of political competition. The inclusion of the variable *largest* in the test (panel (b)) obviously affects the F-statistic, but does not change this conclusion. Panel (c) corroborates these findings using recursive two-sample $t$-tests on the variable *largest*, locating again the maximum difference in spending at 59%.

To control for any potential (overt or hidden) selection bias, we now rely on regression discontinuity designs and compare the two resulting regimes in the the vicinity of the threshold—i.e., municipalities where the largest party holds a seat share just below 59% are compared with municipalities where the largest party holds a seat share just above 59%. By doing so, the two groups of municipalities are likely to be very similar in their exogenous characteristics, which eases the concerns of a selection bias (see Appendix C for differences in covariate means). If significant differences are found, the effect of government composition at the threshold is validated. As a matter of comparison, we also conduct the

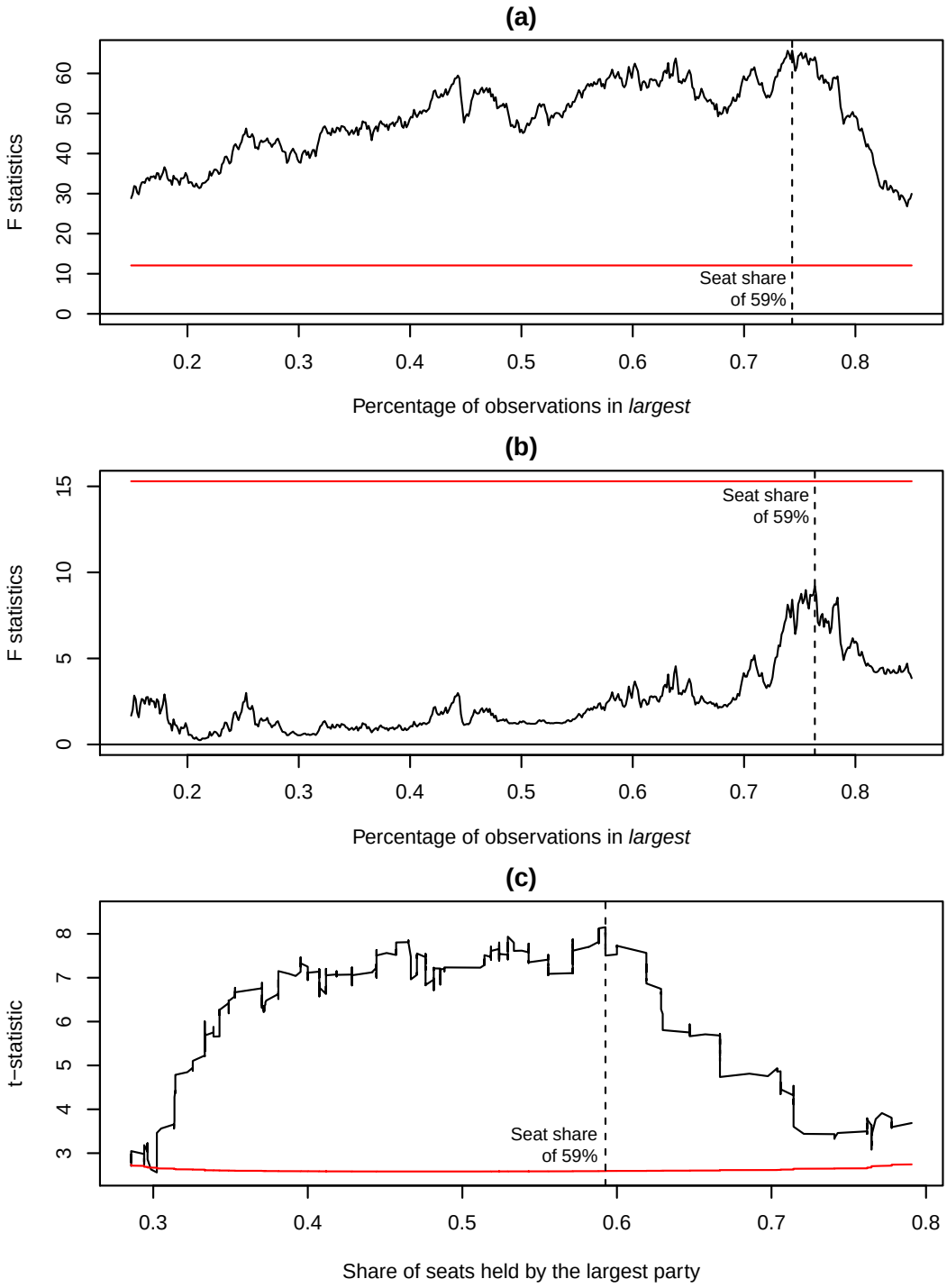



Figure 2.6. Tests for endogenous thresholds (1% critical level given by horizontal red line).

Table 2.6. RDD estimates – seat share thresholds

| | *Dependent variable:* | | | |
|---|---|---|---|---|
| | *Y_average* | | | |
| | 59% threshold | | 50% threshold | |
| | IK | CCT | IK | CCT |
| Conventional RD | 248.750*** | 238.196** | -52.477 | -57.546 |
| | (86.293) | (108.378) | (109.310) | (110.566) |
| Bias-corrected (robust) | 218.798 | 232.703* | -275.048 | -26.430 |
| | (158.913) | (120.285) | (208.396) | (139.858) |
| Observations | 507 | 260 | 397 | 387 |
| BW RD point-estimator ($h$) | 0.188 | 0.108 | 0.109 | 0.104 |
| BW bias-correction estimator ($b$) | 0.134 | 0.178 | 0.104 | 0.167 |

***, ** and * indicate a significance level of 1%, 5% and 10%, respectively. Standard errors in brackets.

tests at the 50% threshold, where no effect is expected.[7]

Estimates of the local average treatment effect (LATE) are reported in Table 2.6. The analysis employs bandwidths selected according to the method of Imbens and Kalyanaraman (2012, henceforth IK), which ensures an asymptotically optimal bandwidth under squared error loss, and of Calonico *et al.* (2014, CCT), to further guard against the use of too large bandwidths. The main bandwidths ($h$) are used to construct local linear regression point-estimators. The bias-corrected bandwidths ($b$) offer a bias-correction using local quadratic regression with a robust variance estimation.

Most estimators reveal a significant difference in spending between the two regimes, the effect ranging from 233 to 249 euros. On the contrary, for the 50% threshold, the estimates show no significant impact. A true difference in spending levels is thus observed if and only if the intensity of political competition is sufficiently low, independently of the 50% single-party majority threshold.

[7]For each election year (2000 and 2004), the McCrary density test rejects the null hypothesis at a 5% significance level, showing that the forcing variable *largest* is not subject to manipulation around the thresholds.

Table 2.7. Results of matching algorithms – ATE for *REG*

| | | | | *Y_average* | |
|---|---|---|---|---|---|
| | Observations | Treated | Controls | Mean diff. | *t*-stat |
| Before matching | 736 | 184 | 552 | 315.918*** | 8.125 |
| | Matched obs. | Matched unweighted | Loss in % | ATE | *t*-stat |
| Nearest neighbor | 736 | 796 | 0 | 198.070 | 1.078 |
| Caliper = 0.06 | 510 | 570 | 30.71 | 238.980*** | 3.261 |
| Caliper = 0.04 | 468 | 527 | 36.41 | 204.710*** | 3.583 |
| Caliper = 0.02 | 380 | 436 | 48.37 | 171.100*** | 4.464 |

***, ** and * indicate a significance level of 1%, 5% and 10%, respectively.

## 2.6 Robustness checks

In this section, we provide several robustness checks. First, using propensity score matching, we further compare the high competition and the low competition groups at the 59% threshold (Subsection 2.6.1). Second, we test whether the ideology of the largest party has an impact at the 50% and 59% seat share thresholds (Subsection 2.6.2). Last, a range of placebo tests is employed to assess the validity of the regression discontinuity designs (Subsection 2.6.3).

### 2.6.1 Propensity score matching at the 59% threshold

As an alternative approach to the variable *MAJ*, limiting our analysis to single-party majorities, the whole sample is now re-examined with regard to the 59% threshold. In what follows, the threshold in seat share will be used to split the sample into two regimes: the high political competition group ($REG = 0$, hereafter) and the low political competition group ($REG = 1$). In other words, municipalities where the largest party does not hold a supermajority are now compared to municipalities where the largest party holds a supermajority.

Results are provided in Table 2.7. Before matching, analogously to the comparison of single-party majorities, more powerful majorities ($REG = 1$) were observed to have higher expenditures, by about 316 euros per capita. Once matched, the remaining effect amounts to 171 euros with the smallest caliper. This provides support to the RDD results and demonstrates that the political strength of the largest party can be considered as a potential explanation for public spending differences in the Finnish municipalities.

### 2.6.2 The role of largest Finnish parties

It could be possible that the largest Finnish parties play a significant role in determining policy choices, thereby affecting our previous results. Given the distribution of election outcomes in our sample, we can test for eventual partisan effects by focusing on the two largest parties in Finland, at the local level, namely the traditionally strong *Center Party* (KESK) and the *Swedish People's Party* (RKP).

For either political party, RDD analysis estimates the impact on public spending at a 50% and a 59% threshold, where the party in question holds respectively at least a majority or a supermajority (comparison of covariate means in Appendix C). The results (Table 2.8) do not deviate from what is observed for the whole sample. In a simple comparison of means, either single-party majority is again assumed to spend significantly more on operating expenditures than fragmented minority governments: KESK by about the same amount (259 euros per capita), while for RKP majorities the positive margin is slightly smaller and less significant (199 euros at a 5% level). According to the RDD, the effect is negative instead, but no longer significant: neither KESK nor RKP majorities show a consistent statistically significant impact on public spending (right half of Table 2.8). For a supermajority, defined by the 59% general breakpoint (left half), we again observe a significant positive difference in expenditures, although the result is less clear for RKP due to sample size.

Table 2.8. RDD estimates – seat share thresholds for KESK and RKP

| KESK | *Dependent variable:* | | | |
|---|---|---|---|---|
| | *Y_average* | | | |
| | 59% threshold | | 50% threshold | |
| | IK | CCT | IK | CCT |
| Conventional RD | 263.796*** | 255.938** | -25.481 | -22.784 |
| | (102.120) | (118.329) | (116.758) | (122.803) |
| Bias-corrected (robust) | 236.466 | 249.115* | 10.604 | 29.214 |
| | (144.507) | (133.770) | (147.296) | (157.710) |
| Observations | 326 | 224 | 331 | 316 |
| BW RD point-estimator ($h$) | 0.139 | 0.091 | 0.117 | 0.095 |
| BW bias-correction estimator ($b$) | 0.124 | 0.149 | 0.174 | 0.163 |

| RKP | *Dependent variable:* | | | |
|---|---|---|---|---|
| | *Y_average* | | | |
| | 59% threshold | | 50% threshold | |
| | IK | CCT | IK | CCT |
| Conventional RD | 0.395*** | -43.774 | -177.044 | -370.256 |
| | (307.297) | (118.329) | (390.879) | (529.398) |
| Bias-corrected (robust) | 43.388 | -25.523 | -706.800 | -501.886 |
| | (403.821) | (133.770) | (772.078) | (621.800) |
| Observations | 30 | 18 | 40 | 27 |
| BW RD point-estimator ($h$) | 0.171 | 0.105 | 0.251 | 0.127 |
| BW bias-correction estimator ($b$) | 0.189 | 0.174 | 0.176 | 0.208 |

***, ** and * indicate a significance level of 1%, 5% and 10%, respectively. Standard errors in brackets.

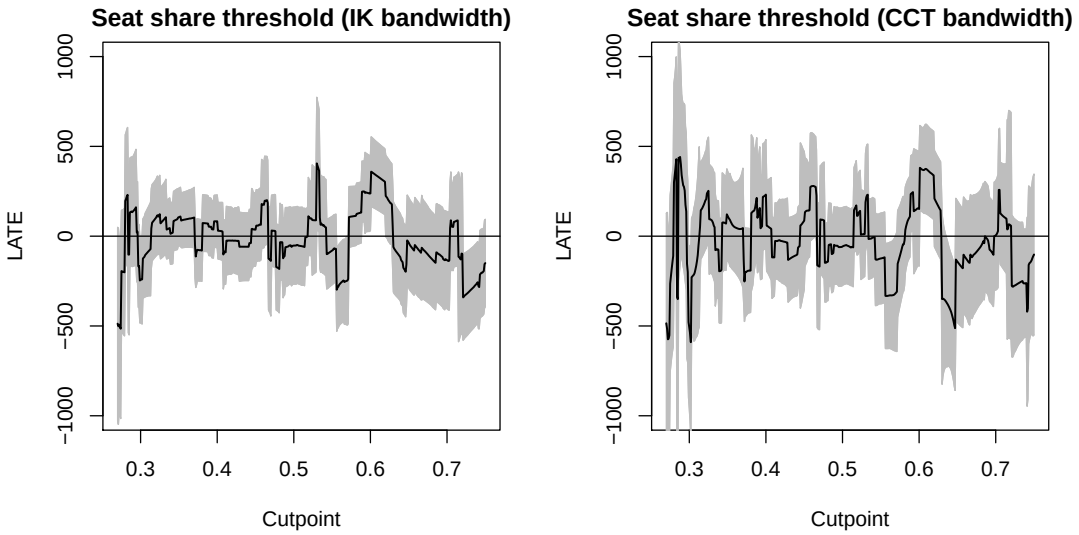



Figure 2.7. Recursive regression discontinuity design.

### 2.6.3 Placebo tests

While the RDD approach in Section 2.5 confirms the existence of a general breakpoint in the data, it does not tell us about other possible thresholds in the sample. To address this issue, we rely on placebo tests, by estimating the local average treatment effect (LATE) for all possible thresholds in our data. The approach is somewhat equivalent to the analysis of arbitrarily chosen thresholds, although we examine the full range of possible cut points.

Results are provided in Figure 2.7, with 95% confidence intervals in gray. As expected, we find a positive and significant local average treatment effect at the 59% seat share threshold for both IK and CCT bandwidths. On the contrary, no effect is apparent close to the 50% threshold. From the analysis, we can also observe that a few other cut points yield a significant effect, which is sometimes positive and sometimes negative. Yet, the estimated effects are smaller than the one observed at the 59% threshold. The findings thus corroborate the existence of two main regimes, with one group of municipalities facing a high intensity of political competition, while the other group is instead ruled by one party holding a supermajority.

## 2.7 Conclusion

Government composition is considered to be one of important determinants in public finance, expected and typically confirmed to be associated with the level of public spending (see, e.g., Bradbury and Crain, 2001; Padovano and Venturi, 2001; Perotti and Kontopoulos, 2002; Volkering and de Haan, 2001). Yet, a concern that should not be overlooked is that not just the party in power itself, but also the structure of the entire relevant political arena, i.e., the competition the party is facing, affects the outcome. Analogously to a lack of economic competition in the private sector, low political competition may lead to decreased efficiency and more pronounced rent-seeking behavior in politicians managing the public sector. Until recently, few papers have attempted to empirically verify this assumption (e.g., Ashworth *et al.*, 2014; Ferris *et al.*, 2008; Padovano and Ricciuti, 2009), furthermore hindered by varied and indirect definitions of the economic outcomes impacted. Nevertheless, the available evidence follows the intuitive conclusion that incumbents are more likely to serve their own interests when they are less at risk of losing elections.

Using data from Finnish municipalities, our analysis confirms that political power is not only a matter of absolute majority of seats, but also of party size and strength. First, we observe that governments led by a single ruling party are more likely to be elected in rural areas in the east, west and chiefly north of Finland. That is to say, fragmented minority governments are preferred by the electorate of southern urban municipalities, richer and younger on average, yet facing higher unemployment. Second, by applying quasi-experimental techniques, to control as much as possible for this electorate heterogeneity, we find that the composition of the Finnish local councils—as measured by a dummy variable indicating whether or not there is a single-party government—has little influence on their spending behavior. Last but not least, the results suggest that political competition disciplines politicians: a significant difference in spending can actually be

observed, but only when the largest party has enough influence and holds a comfortable position, e.g., at least three-fifths of the council seats.

These results raise questions about the efficiency of local governments in Finland and in other Nordic countries as well. In particular, Finnish municipalities operate in a political environment that is characterized by a strong multi-party system and, as such, by high levels of government fragmentation. Yet, according to our results, the political competition induced by this fragmentation may actually be efficient. Perhaps one explanation for this finding is that the competing parties have incentives to form alliances in order to win elections or to pass a proposal (see, e.g., Jackson and Moselle, 2002; Levy, 2004). The median voter theorem can thus hold if two main blocs are identifiable. As suggested by our empirical results, we find high spending levels only when the opposition group does not represent a serious threat.

More generally, our results point out the importance of employing careful identification strategies when examining the role of government composition in fiscal choices. First, regression discontinuity design can present a potential shortcoming if it relies solely on the examination of close contests (e.g., at the 50% threshold), where differences in spending are in theory less likely to be observed. Second, while propensity score matching allows various types of elections to be examined, it does not control for the size of the largest party *per se*. Only when the data was examined at the optimal threshold did we find a highly significant difference in spending.

# Chapter 3

# Performance of local governments: Benchmarking analysis of social welfare provision in France

---

## 3.1 Introduction

French Departments, the middle level of local government in France, are responsible for a multitude of services to the population, from maintaining roadways to running schools that provide secondary education. The singularly largest task, however, accounting for nearly two-thirds of departmental budget outlays, is the provision of social welfare services, both through direct payments and different types of in-kind aid. While local governments are compensated for their services through grant schemes, these have failed to take into account a more than 10% increase in costs to the Departments just over the past 5 years (Roucous, 2018), leading to a precarious long-term situation, where numerous Departments struggle to perform their authorized functions. One of the most affected Departments of Seine-Saint-Denis, that has seen a more than 40% rise in its social spending since 2010, was

even forced to entirely withdraw payments of unemployment benefits at the end of 2017, until reimbursed by the state (Poingt, 2017). A situation moreover not dissimilar already to the year 2015, when emergency funding had to be provided to nine additional Departments approaching a crisis (Chanut, 2016).

The competences of the Departments were originally expanded in reaction to the economic crisis of 2001, which raised the number of people depending on social welfare. After almost two decades, however, appeals are being made to again re-centralize or otherwise find a solution to lighten the burden certain welfare payments, the unemployment benefits (RSA) in particular, place on departmental budgets.

Although Departments are bound by national eligibility criteria defining beneficiaries of welfare services, the amount of individual social aid is, with few exceptions, a discretionary choice of the local governments. An understandable heterogeneity can be observed across the range of services, given the varied costs required by the different needs, yet the Departments themselves also differ wildly in the amounts they opt to allocate to the same type of welfare recipient. Taking the example of social assistance to the disabled, an average eligible individual would benefit from the highest amount of aid, equal to 41,295 euros, in Orne (Normandy), while a beneficiary in Vosges (Grand Est) would warrant a mere one-fifth of the sum.

In the literature, much formal investigation has been made into how population size impacts public good provision, starting with the works of Borcherding and Deacon (1972) and Bergstrom and Goodman (1973). In particular, emphasis is put on congestion effects, affecting the quality of a service to its users and thereby generating negative consumption externalities (e.g., Guengant *et al.*, 2002). Traditional models of public good provision consider public goods to be congestible: in a given jurisdiction, the quality of a public service, $z$ hereafter, depends on the size of the eligible population $N$. This impact is generally modeled through the use of a congestion function; for instance, a popular function adopted by Borcherding and Deacon (1972), $N^{\gamma}$, describes increasing congestion effects when $\gamma$ is negative. In general, the larger the population size $N$, the lower the quality $z$, and the lower the demand for the public good in the jurisdiction. The situation may however be further

complicated by confounding factors, such as a variable demand structure, i.e., the types of users in the jurisdiction, which may evolve jointly with population size. If so, size effects cannot be measured directly.

In this study, we examine the considerable differences observed in our data on social public expenditure in French Departments (*Départements*), focusing on population size effects and other mechanisms related to the distribution of users. We show that socio-demographics and expenditures are indeed clearly associated with population size. In order to take into account the presence of this confounding factor, we incorporate elements of benchmarking analysis methodology, particularly suitable for dealing with heterogeneous users of public services. Benchmarking presents a follow-up evaluation tool used to compare cost structures of facilities with that of a designated reference, i.e., the benchmark. In our context, the main asset of the method is that it thus allows the demand structure to be controlled for: any variation in the performance measure can thereby be attributed to size effects only.

Our analysis shows that the demand side is highly relevant in explaining differences in public spending. More specifically, both the demand structure and the demand size play a role. On the one hand, a different demand structure, i.e., different types of users, directly affects the provision costs. On the other hand, a large number of users at the same time creates congestion effects: the governments of most disadvantaged jurisdictions cannot afford to provide the same quality of service. We examine how our findings impact the position of Departments with a more rural or urban structure and discuss the financial support received through intergovernmental grant schemes.

The remainder of the paper is organized as follows. Section 3.2 describes the context of the study and provides an overview of the scale of disparities observed in social welfare provision in the French Departments. Section 3.3 focuses on the impact of population size on spending and on the demand structure. Section 3.4 presents the benchmarking methodology and provides assessment of the Departments, including further analysis of factors impacting benchmarking scores. Section 3.5 discusses the results in the context of urbanization and fiscal equalization, leading to policy implications. Last, Section 3.6concludes.

## 3.2 French Departments and social welfare provision

### 3.2.1 Context of the study

France is divided into 96 metropolitan and 5 overseas Departments (Guadeloupe, Guyane, Martinique, Mayotte and Réunion). Focusing on a more homogeneous sample of *France métropolitaine*, the five overseas and two Corsican Departments, as well as the capital of Paris, have been excluded from the sample, leaving us with 93 Departments, observed over the period of 11 years between 1998 and 2008. Table 3.1 provides an overview of our data and the key variables in our analysis.

Departments are ruled by a council elected in so-called *cantonal elections*[1], held approximately every three years: in 1998, 2001, 2004 and 2008, for our sample. Departmental social expenses, i.e., public expenditure on social welfare provision, can be divided into four main categories/services:

(1) social assistance to the mother and child – through prevention, protection, and aid to families with children, including long-term institutional placements and foster care;

(2) social assistance to pensioners and the elderly – through direct payments and subsidies to households;[2]

(3) social aid to the unemployed – through health protection, retraining, and an unemployment benefit;

[1]The departmental councilors are elected by means of direct vote by the electorate for a six-year term according to a two-ballot, uninominal majority polling procedure. An important feature of these elections is that only one-half of the councilors are renewed each election, with one councilor per constituency. A constituency is a grouping of municipalities referred to as *canton*. A candidate securing the votes of at least 25% of the canton's registered voters and more than 50% of the total number of votes is elected. If no political candidate satisfies these conditions, a second round of voting is held one week later. Entitled to present themselves in the second round are the two candidates who received the highest number of votes in the first round, plus any other candidate that received the votes of at least 10% of those registered to vote in the canton. In the second round, the candidate with the highest number of votes is elected.

[2]Note that pensions are excluded from the study because they do not belong to the competences of the French Departments.

(4) social assistance to the disabled – through subsidies to homes, direct payments, housing modifications for accessibility, etc.

In reaction to the economic crisis of 2001, which raised the number of people depending on social welfare, two competences of the Departments were further expanded: in 2002, a new welfare program targeting the elderly was created, called APA (*Allocation personnalisée d'autonomie*); then, with the *Decentralization Act* of 2004, the Departments were endowed with the responsibility for implementing social aid to the unemployed, the RMI (*Revenu minimum d'insertion*), replaced later in 2009 by the RSA (*Revenu de solidarité active*). The expenditures are mainly financed through the Departments' own tax revenues and, to a minor extent, through central grants; see, e.g., Dollery and Lorenzo (2008), or annual reports on the French local public sector by *l'Observatoire des finances et de la gestion publique locales* (OFGL).

### 3.2.2 Inter-departmental and inter-service disparities

As can be seen from Table 3.1, spending per beneficiary varies substantially from one service to another. While, on average, a Department spends 47,673 euros on one child, the cost is much lower for the other types of beneficiaries: 6,910 euros per senior, 3,384 euros per unemployed and 18,206 euros per disabled resident. This heterogeneity is unsurprising, given the varied costs required by the different needs.

Yet, the Departments themselves differ wildly in the amounts they opt to distribute to the same type of welfare recipient as well. As illustrated in Figure 3.1, where Departments are ordered by increasing levels of spending per beneficiary, some councils allocated more than double the amount in the same year. The difference is most striking with regard to the aid to the unemployed, where the minimum spending per beneficiary, in the whole sample, reaches only 4% of of the largest sums. But even for families, elderly or disabled residents the lowest sums amount to no more than one-fifth of the highest levels. There is not only a large variability to the values, Departments also have different priorities with regard to where resources are allocated—e.g., the Île-de-France Departments spend relatively high sums on

Table 3.1. Summary statistics

| Variable | Description | N | Mean | St. Dev. | Min | Max |
|---|---|---|---|---|---|---|
| *socialT* | Total social expenditures (in thousands of euros). | 1,023 | 186,981 | 173,072 | 16,685 | 1,451,638 |
| *social* | Per capita social expenditures (in euros). | 1,023 | 301 | 101 | 129 | 636 |
| *socialb* | Per beneficiary social expenditures (in euros). | 1,023 | 8,376 | 1,560 | 4,543 | 14,157 |
| *familiesb* | Social assistance to families with children per beneficiary (in euros). | 1,023 | 47,673 | 12,508 | 23,199 | 119,451 |
| *elderb* | Social assistance to the elderly per beneficiary (in euros). | 1,023 | 6,910 | 2,344 | 3,102 | 21,269 |
| *unempb* | Aid to the unemployed per beneficiary (in euros). | 1,023 | 3,384 | 2,845 | 430 | 11,678 |
| *disabledb* | Social assistance to the disabled per beneficiary (in euros). | 1,023 | 18,206 | 3,638 | 8,560 | 41,295 |
| *beneficiaries* | Total number of welfare recipients. | 1,023 | 22,199 | 19,346 | 1,940 | 141,362 |
| *families* | Share of families with children in the population receiving social assistance (in %). | 1,023 | 0.188 | 0.059 | 0.060 | 0.384 |
| *elder* | Share of elderly people in the population receiving social assistance (in %). | 1,023 | 1.556 | 0.876 | 0.196 | 4.134 |
| *unemp* | Share of the unemployed in the population receiving social assistance (in %). | 1,023 | 1.519 | 0.620 | 0.552 | 3.824 |
| *disabled* | Share of disabled people in the population receiving social assistance (in %). | 1,023 | 0.388 | 0.094 | 0.184 | 0.919 |
| *WD* | Share of welfare-dependent people in the population. | 1,023 | 3.652 | 1.159 | 1.273 | 7.033 |
| *familiesS* | Share of families with children among the beneficiaries (in %). | 1,023 | 5.602 | 2.203 | 1.719 | 13.059 |
| *elderS* | Share of elderly people among the beneficiaries (in %). | 1,023 | 40.183 | 14.566 | 10.095 | 71.475 |
| *unempS* | Share of the unemployed among the beneficiaries (in %). | 1,023 | 42.895 | 13.235 | 15.023 | 79.503 |
| *disabledS* | Share of disabled people among the beneficiaries (in %). | 1,023 | 11.335 | 3.187 | 4.585 | 22.655 |
| *population* | Number of inhabitants. | 1,023 | 620,439 | 450,579 | 73,507 | 2,565,257 |
| *density* | Number of inhabitants per $km^2$. | 1,023 | 330 | 1,178 | 14 | 8,825 |
| *income* | Taxable income per capita (in euros). | 1,023 | 8,868 | 1,855 | 6,075 | 20,036 |
| *grants* | Grants per capita received by the Department (in euros). | 1,023 | 180.740 | 78.775 | 30.403 | 644.996 |

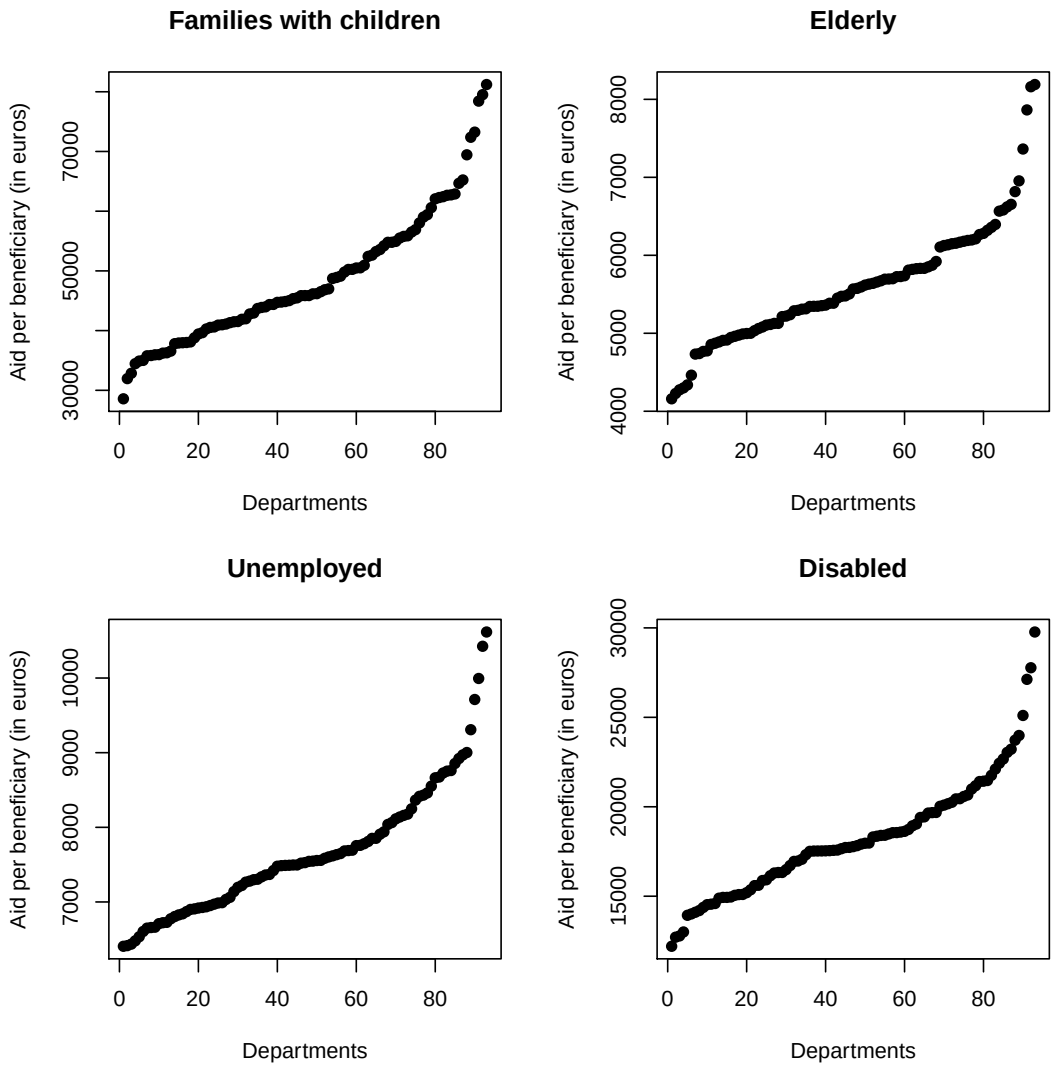



Figure 3.1. Differences in social aid per beneficiary – 2008.

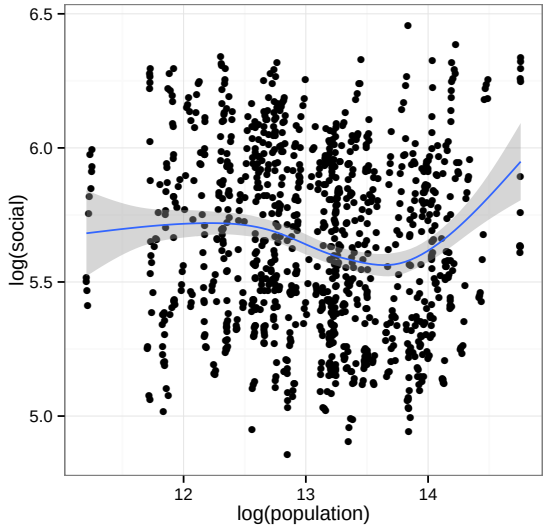
families with children and the elderly, but little on the disabled and, particularly, on the unemployed.

Inevitably, the level of resources expended towards the same goal of ensuring adequate services to the eligible population—and therefore the level of support received, i.e., the quality of the service—is far from comparable. Providing an insight into these large differences is at the heart of our analysis.

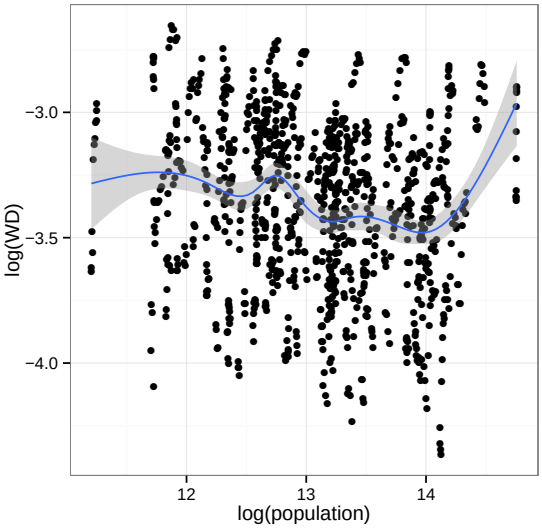
## 3.3 Population size and heterogeneity of users

Population size is linked by both theory and empirical literature to positive (economies of scale, tax-sharing advantages) or negative (congestion effects) externalities impacting public good provision. In the French local finance context, albeit on the municipal level, this relationship between population and public expenditure has been observed to be nonlinear, U-shaped (Breunig and Rocaboy, 2008). Similar findings are documented in our data as well, as illustrated in panel (a) of Figure 3.2 (variables are in logs to be consistent with the econometric analysis; see Breunig and Rocaboy, 2008, for a similar methodology).

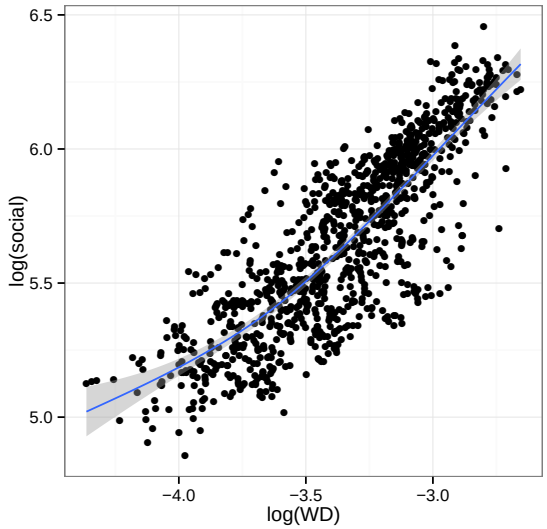
The number of welfare recipients, i.e., the constituents served as quantities of public good provided, naturally also plays a role—and furthermore serves as a link between the population of the Department on the one side, and the budget spent on the other. Indeed, the U-shaped curve in social expenditure per capita seems to reflect the share of the welfare-dependent population ($WD$) in panel (b). In theory, a greater number of welfare-dependent people in the constituency would lead to a stronger support for policies in favor of increasing the provision of public goods and services. Although, at the same time, an increasing tax burden levied on the taxpayers leads them to push for a reduction in spending instead, eventually offsetting the demand that comes from the welfare recipients (see Le Maux *et al.*, 2017, for a general model). The resulting nonlinear relationship between welfare recipients and public spending, explained by these opposite effects, emerges in panels (c) and (d), for per capita and per beneficiary social expenditure, respectively.



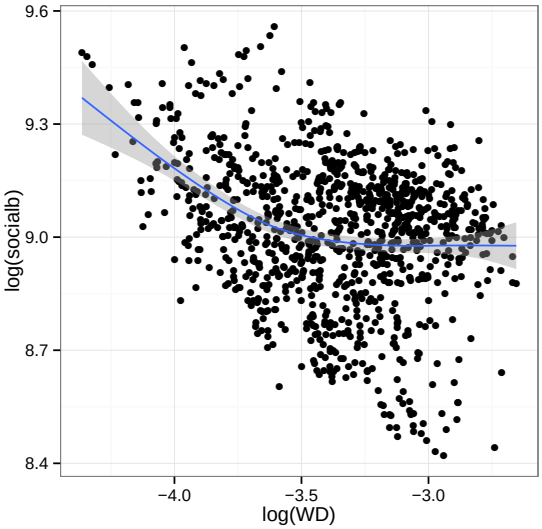

(a) Per capita social expenditures




(b) Welfare recipients (WD)

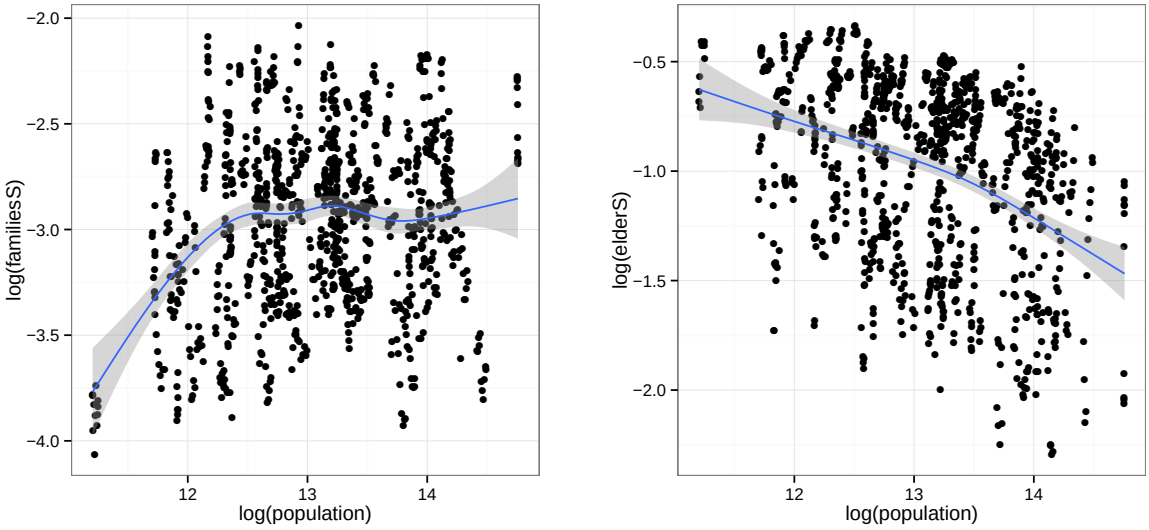



(c) WD and expenditures per capita




(d) WD and expenditures per beneficiary

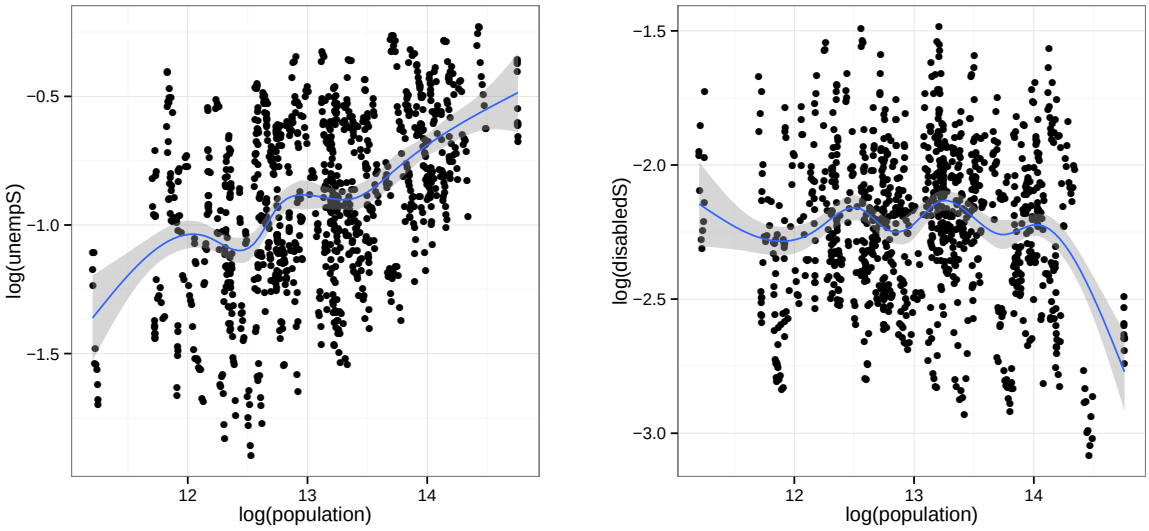
Figure 3.2. Impact of population size on spending.




(a) Families with children




(b) Elderly




(c) Unemployed




(d) Disabled

Figure 3.3. Impact of population size on demand structure.

Moreover, we can see that the distribution of users among services, i.e., the demand structure, varies with population size as well (Figure 3.3). In particular, the share of the elderly (panel (b)) in the welfare-dependent group decreases, while the share of the unemployed (panel (c)) increases. In other words, the differences in spending levels can be attributed not only to a different total number of users but, at the same time, also to which types of services are predominant in jurisdictions of different sizes. Presented in the following Section 3.4, the benchmarking methodology accounts and controls for exactly this confounding heterogeneity in the distribution of users, thereby allowing to measure size effects only.

## 3.4 Measuring the performance of French Departments

### 3.4.1 Benchmarking methodology

Benchmarking, as a follow-up performance evaluation tool, computes a series of scores comparing the unit/facility under assessment with a relevant reference point, the benchmark. The comparison is based on their ability to transform inputs (i.e., the resources used to produce and provide the good or service) into outputs (i.e., the quantity of the good or service provided, commonly defined as the number of users). A better, i.e., more "cost-efficient", performance is therefore one that maximizes the level of outputs for a given set of inputs (avoiding waste by producing as much output as input usage allows) or, alternatively, minimizes the cost of inputs required to produce a given level of outputs (avoiding waste by using as little input as output production allows).[3]

Formally, following the presentation of Josselin and Le Maux (2017), the range of services provided by unit $i$ can be classified as $s = 1, ..., S$. In our context, these represent

[3]Other methods are based on the idea of benchmarking and efficiency comparison in multiple-input/multiple-output settings, such as the efficiency-frontier techniques of Data Envelopment Analysis (DEA) or Stochastic Frontier Analysis (SFA). Our aim, however, is not to evaluate the performance of the units *per se*, but rather to take advantage of the demand structure decomposition benchmarking offers. Furthermore, our data, converting a simple monetary input into a user count, would fail to profit from the the complex differentiation the methods offer and the interpretation of the results would be largely meaningless.

the four categories of welfare programs provided by the French Departments. For each service $s$, all individual users are listed as $1,...,n_s$, where $n_s$ indicates the total number of users per service, and $N^i = n_1^i + n_2^i + ... + n_s^i$ then denotes the total number of users in unit $i$. The distribution of users among the services, i.e., the demand structure, is described in terms of relative frequencies: $f_s^i = n_s^i/N^i$, so that $f_1^i + f_2^i + ... + f_s^i = 1$. The average cost of unit $i$ can hence be defined as a weighted sum of the average costs per service:

$$\bar{c}^i = \sum_{s=1}^{S} \bar{c}_s^i \times f_s^i, \tag{3.1}$$

where $\bar{c}_s^i$ is the average cost of service $s$ in unit $i$.

Analogously, the costs of the benchmark, $B$ hereafter, are expressed as either embodying an existing unit that is relevant in the given context (e.g., for its exemplary performance, implementation of a program, location) or, quite frequently, one constructed artificially from the sample or representative units (e.g., an average).

Since the distribution of users among services is assumed to influence production costs, benchmarking then seeks to explain the observed differences in average costs by decomposing the cost ratio as:

$$\frac{\bar{c}^i}{\bar{c}^B} = \underbrace{\frac{\sum_{s=1}^{S} \bar{c}_s^i \times f_s^i}{\sum_{s=1}^{S} \bar{c}_s^B \times f_s^B}}_{\text{Total effect}} = \underbrace{\frac{\sum_{s=1}^{S} \bar{c}_s^B \times f_s^i}{\sum_{s=1}^{S} \bar{c}_s^B \times f_s^B}}_{\text{Demand effect}} \times \underbrace{\frac{\sum_{s=1}^{S} \bar{c}_s^i \times f_s^i}{\sum_{s=1}^{S} \bar{c}_s^B \times f_s^i}}_{\text{Production effect}}. \tag{3.2}$$

The first ratio, i.e., demand effect, identifies the impact of the demand structure on the average cost. It measures to which extend the distribution of users in $i$ is responsible for the cost savings, or the additional costs, compared to the benchmark. The second ratio, i.e., production effect, keeps the demand structure fixed as that of unit $i$, measuring instead the cost savings, or additional costs, generated by the production structure of $i$. That is to say, in our context, the increase (or decrease) of spending levels due to the allocative decisions of the Department.

Note that the methodology is not able to asses the value or quality of the services provided. Only the effects of the demand and production structures are judged.

### 3.4.2 Computing scores

Taking advantage of the subclass/category distinction inherent in the benchmarking analysis design, the demand structure of social welfare provision in the French Departments can be represented through the number of welfare-dependent individuals as:

$$beneficiaries = familiesN + elderN + unempN + disabledN, \tag{3.3}$$

where $disabledN$ denotes the number of people with a recognized condition that grants them, or their carers, access to financial support. Similarly, the variables $familiesN$, $elderN$ and $unempN$ represent families with children, elderly, and unemployed people, respectively, that benefit from departmental welfare programs. In order to compute relative frequencies, users of different services are taken into account as their share in the total number of welfare recipients:

$$familiesS = \frac{familiesN}{beneficiaries}, \ldots, disabledS = \frac{disabledN}{beneficiaries}, \tag{3.4}$$

with $familiesS + elderS + unempS + disabledS = 100\%$. Last, to account for possible size effects, we also define the share of welfare-dependent individuals in the population as:

$$WD = \frac{beneficiaries}{population}. \tag{3.5}$$

Note that the number of social beneficiaries in each Department depends on eligibility criteria defined by the national government. As such, our variables are not a function of the Department's public policy and can be considered exogenous. Conversely, with the exception of the APA and RMI/RSA payments, the amount of aid is within the discretion of the Departments.

Table 3.2. Ranking 1998–2008: Total, demand and production effects

*Total effect*

| | Lowest spending | | | | Highest spending | | | |
|---|---|---|---|---|---|---|---|---|
| Rank | TE | Department | Region | Year | TE | Department | Region | Year |
| 1 | 0.559 | 66 | 76 | 1998 | 1.569 | 28 | 24 | 2001 |
| 2 | 0.567 | 13 | 93 | 2001 | 1.551 | 78 | 11 | 2000 |
| 3 | 0.570 | 66 | 76 | 1999 | 1.549 | 78 | 11 | 2001 |
| 4 | 0.580 | 13 | 93 | 1999 | 1.545 | 78 | 11 | 1999 |
| 5 | 0.581 | 34 | 76 | 1998 | 1.544 | 91 | 11 | 2002 |
| 6 | 0.590 | 34 | 76 | 1999 | 1.528 | 28 | 24 | 2000 |
| 7 | 0.593 | 9 | 76 | 2001 | 1.484 | 91 | 11 | 2004 |
| 8 | 0.596 | 9 | 76 | 1998 | 1.473 | 91 | 11 | 2007 |
| 9 | 0.599 | 13 | 93 | 1998 | 1.470 | 28 | 24 | 1999 |
| 10 | 0.601 | 13 | 93 | 2000 | 1.468 | 91 | 11 | 2008 |

*Demand effect*

| | Lowest spending | | | | Highest spending | | | |
|---|---|---|---|---|---|---|---|---|
| Rank | DE | Department | Region | Year | DE | Department | Region | Year |
| 1 | 0.473 | 13 | 93 | 2001 | 1.516 | 53 | 52 | 1998 |
| 2 | 0.487 | 13 | 93 | 2000 | 1.445 | 53 | 52 | 1999 |
| 3 | 0.511 | 13 | 93 | 1999 | 1.441 | 53 | 52 | 2000 |
| 4 | 0.546 | 13 | 93 | 1998 | 1.391 | 53 | 52 | 2001 |
| 5 | 0.579 | 83 | 93 | 1998 | 1.354 | 39 | 27 | 1999 |
| 6 | 0.582 | 83 | 93 | 2000 | 1.346 | 22 | 53 | 1998 |
| 7 | 0.583 | 13 | 93 | 2002 | 1.342 | 39 | 27 | 2001 |
| 8 | 0.583 | 83 | 93 | 2001 | 1.334 | 39 | 27 | 2000 |
| 9 | 0.584 | 13 | 93 | 2003 | 1.318 | 39 | 27 | 1998 |
| 10 | 0.590 | 66 | 76 | 2001 | 1.317 | 89 | 27 | 2001 |

*Production effect*

| | Lowest spending | | | | Highest spending | | | |
|---|---|---|---|---|---|---|---|---|
| Rank | PE | Department | Region | Year | PE | Department | Region | Year |
| 1 | 0.652 | 18 | 24 | 2002 | 1.361 | 83 | 93 | 2002 |
| 2 | 0.653 | 22 | 53 | 2000 | 1.355 | 69 | 84 | 2000 |
| 3 | 0.661 | 79 | 75 | 1998 | 1.306 | 94 | 11 | 2007 |
| 4 | 0.672 | 2 | 32 | 1998 | 1.292 | 78 | 11 | 2000 |
| 5 | 0.676 | 9 | 76 | 1998 | 1.280 | 95 | 11 | 2000 |
| 6 | 0.682 | 79 | 75 | 1999 | 1.279 | 95 | 11 | 1998 |
| 7 | 0.688 | 18 | 24 | 2000 | 1.278 | 93 | 11 | 2004 |
| 8 | 0.689 | 9 | 76 | 2001 | 1.275 | 78 | 11 | 1998 |
| 9 | 0.689 | 9 | 76 | 2000 | 1.271 | 78 | 11 | 1999 |
| 10 | 0.694 | 90 | 27 | 2000 | 1.271 | 92 | 11 | 1998 |

Scores of total (TE), demand (DE) and production (PE) effect indices are calculated according to Equation 3.2. A hypothetical average Department, computed annually from average numbers of social beneficiaries and mean costs per each category of benefit, serves as a benchmark to the assessed Departments. Table 3.2 presents ranking of the lowest and highest scores, with Departments identified by their postal code (*code postal*), the code of their Region and the year the score was measured.

The lowest total effect values, i.e., the lowest spending per beneficiary, are measured over the first years of the observed period and belong to the Departments of Pyrénées-Orientales (66), Hérault (34) and Ariège (9) in the Occitanie region (south) and and Bouches-du-Rhône (13) in Provence-Alpes-Côte d'Azur (south-east). Among these, Bouches-du-Rhône (13) clearly has the least costly structure of beneficiaries (demand effect), whereas Ariège (9) opts for the most cost-saving structure of benefit payments (production effect). Highest spending levels, on the other hand, are observed in the central Departments of Eure-et-Loir (28), Yvelines (78) and Essonne (91). In the case of Yvelines (78), together with other Île-de-France Departments, the extra cost is clearly driven by their allocation of resources.

Although benchmarking by definition cannot form conclusions about the qualitative dimension of services, and the analysis simplifies the complex welfare system through four main categories of programs, in general, lower spending levels would not lead to beneficiaries being better off—whether through direct payments, equipment provided, etc. Departments can nevertheless do very little about the socio-demographic characteristics of their population, i.e., the demand structure, they are confronted with as a given. Here in particular, an advantage seems to be presented by a larger share of the unemployed, who in our data are the less costly type of beneficiary, leading to lower demand effect scores, i.e., to savings compared to the benchmark: e.g., Bouches-du-Rhône (13) and Var (83) with 70–80% of the unemployed, compared to Jura (39) with 30–40%, or Mayenne (53) with 20–30%.

Of yet more interest to the analysis are therefore the production effect scores, i.e., Departments' spending decisions with user distribution controlled for. The lowest values

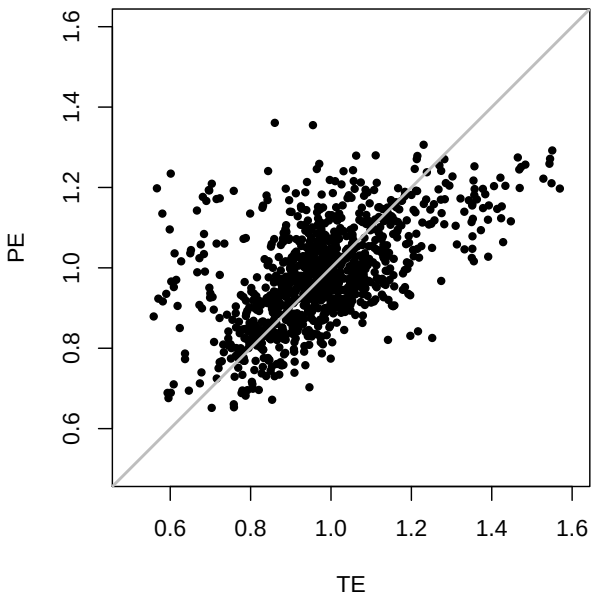



Figure 3.4. Score variation.

are linked to temporal dimension, rather than spatial: mainly observed in 1998–2002, before Departments were obliged to adopt new responsibilities for direct payments to the elderly and the unemployed (APA, RMI/RSA). Even so, allocating the least are noticeably Ariège (9), Cher (18) in Centre-Val de Loire, Côtes-d'Armor (22) in Brittany, and Deux-Sèvres (79) in Nouvelle-Aquitaine. Or, should we consider only the 2004–2008 period already affected, Haute-Marne (52) and Ardennes (8) in Grand Est, together with Allier (3) in Auvergne-Rhône-Alpes—all fairly small rural Departments. The highest sums distributed, on the other hand, clearly belong to the Region of Île-de-France, around the capital of Paris (itself excluded from our sample), together with Departments containing the large cities of Lyon (Rhône, 69) and Toulon (Var, 83).

### 3.4.3 Econometric analysis of performance scores

By construction, the total effect (TE) index simply compares the average amount each Department spends per beneficiary with the overall average values of the sample (benchmark). As such, it conveys a normalized measure of the budget councils spend to provide the bulk of the required services. Far more informative is the production effect (PE)

index, in which the demand side is already controlled for, at least partly, with regard to the composition of the group of welfare-dependent constituents. Although individual differences in scores are then reduced, a large variation still remains: many Departments end up spending substantially more in comparison (points above the 45-degree gray line in Figure 3.4), while others fall far below. To explain when and why this is the case, we next investigate the residual variance in the constituents' demand for public social expenditure.

In order to identify the factors impacting spending decisions, the PE scores are regressed on various control variables (Table 3.3), mainly socio-economic characteristics typically used in local public finance literature:

$$\begin{aligned} \log(PE_{i,t}) &= \alpha_0 + \alpha_1 \log(population_{i,t}) + \alpha_2 \log(WD_{i,t}) \\ &\quad + \alpha_3 \log(density_{i,t}) + \alpha_4 \log(income_{i,t}) + \alpha_5 \log(grants_{i,t}) \\ &\quad + \alpha_6 DEC02_t + \alpha_7 DEC04_t + \varepsilon_{i,t}, \end{aligned} \tag{3.6}$$

where $i$ and $t$ stand for Department $i$ and year $t$, respectively, and $PE_{i,t}$ are the corresponding values of the production effect index. Dummy variables $DEC02$ and $DEC04$ capture the impact of the new competences adopted in 2002 and 2004, respectively.

Model specifications in columns 3–4 take into account the nonlinear effect of population size:

$$\begin{aligned} \log(PE_{i,t}) &= \beta_0 + \beta_1 \log(population_{i,t}) + \beta_2 population_{i,t} \\ &\quad + \beta_3 \log(density_{i,t}) + \beta_4 \log(income_{i,t}) + \beta_5 \log(grants_{i,t}) \\ &\quad + \beta_6 DEC02_t + \beta_7 DEC04_t + \varepsilon_{i,t}, \end{aligned} \tag{3.7}$$

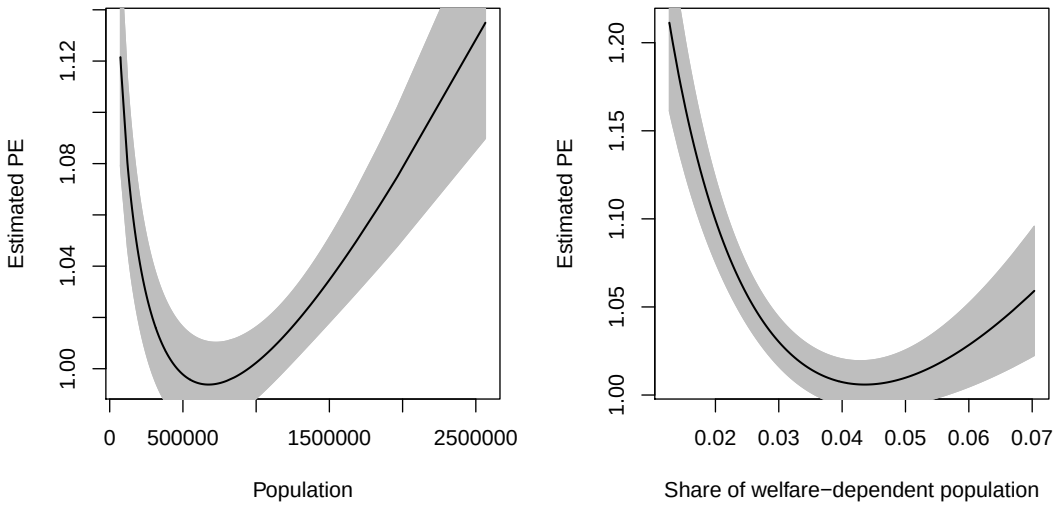



Figure 3.5. Simulating size effects in the production effect scores.

or, focusing directly on the impact of the number of welfare recipients (columns 5–6), as:

$$\begin{aligned}\log(PE_{i,t}) &= \gamma_0 + \gamma_1 \log(WD_{i,t}) + \gamma_2 WD_{i,t} \\ &\quad + \gamma_3 \log(density_{i,t}) + \gamma_4 \log(income_{i,t}) + \gamma_5 \log(grants_{i,t}) \\ &\quad + \gamma_6 DEC02_t + \gamma_7 DEC04_t + \varepsilon_{i,t}. \end{aligned} \tag{3.8}$$

Given that Departments change very little over time, all models have been estimated using pooled model, with additional random effects to avoid potential bias.

Overall (Table 3.3), lower PE scores are measured towards the beginning of our timeline, in less densely populated Departments with poorer constituencies, comprising larger numbers of welfare-dependent individuals. With a nonlinear specification, the U-shaped curve describing social spending in Section 3.3 emerges also with regard to the index scores, despite the fact that we are controlling for the demand structure. The first panel of Figure 3.5 illustrates this impact of population size on spending, according to the model in column 3. Hence, we find results similar to Breunig and Rocaboy (2008) after controlling for the confounding factors. In the second panel, the sharply declining curve of the PE measure, modeled in column 5, suggests that a larger number of users creates congestion effects: the

Table 3.3. Production effect (PE) scores – estimation results[a,b]

| | *Dependent variable:* | | | | | |
|---|---|---|---|---|---|---|
| | log(*PE*) | | | | | |
| | Pooled | RE | Pooled | RE | Pooled | RE |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Constant | $-1.238^{***}$ | $-0.904^{***}$ | $-0.602$ | $-0.061$ | $-2.768^{***}$ | $-2.136^{***}$ |
| | (0.382) | (0.219) | (0.439) | (0.253) | (0.490) | (0.384) |
| log(population) | $-0.018$ | $-0.030$ | $-0.098^{**}$ | $-0.051$ | | |
| | (0.022) | (0.023) | (0.039) | (0.038) | | |
| log(WD) | $-0.076^{**}$ | $-0.174^{***}$ | | | $-0.404^{***}$ | $-0.436^{***}$ |
| | (0.034) | (0.038) | | | (0.084) | (0.085) |
| population | | | $0.0001^{**}$ | 0.0001 | | |
| | | | (0.0001) | (0.0001) | | |
| WD | | | | | $9.286^{***}$ | $8.348^{***}$ |
| | | | | | (2.160) | (2.197) |
| log(density) | $0.046^{***}$ | $0.048^{***}$ | $0.031^{***}$ | $0.048^{***}$ | $0.039^{***}$ | $0.037^{***}$ |
| | (0.011) | (0.012) | (0.012) | (0.011) | (0.007) | (0.008) |
| log(income) | $0.096^{***}$ | $0.034^{**}$ | $0.137^{***}$ | 0.020 | $0.089^{***}$ | 0.018 |
| | (0.034) | (0.0167) | (0.039) | (0.016) | (0.034) | (0.017) |
| log(grants) | $-0.008$ | $-0.022^{*}$ | $-0.061^{***}$ | $-0.021^{*}$ | 0.006 | -0.010 |
| | (0.024) | (0.012) | (0.022) | (0.012) | (0.024) | (0.011) |
| DEC02 | $0.030^{**}$ | $0.072^{***}$ | $0.025^{**}$ | 0.014 | $0.036^{***}$ | $0.067^{***}$ |
| | (0.015) | (0.016) | (0.012) | (0.009) | (0.014) | (0.016) |
| DEC04 | 0.009 | $0.038^{***}$ | 0.002 | 0.013 | 0.001 | $0.028^{***}$ |
| | (0.013) | (0.011) | (0.012) | (0.008) | (0.012) | (0.010) |
| Observations | 1,023 | 1,023 | 1,023 | 1,023 | 1,023 | 1,023 |
| $R^2$ | 0.226 | 0.094 | 0.241 | 0.051 | 0.246 | 0.134 |
| Adjusted $R^2$ | 0.224 | 0.093 | 0.239 | 0.050 | 0.244 | 0.133 |

[a] Arellano's method (1987) was used to compute a robust covariance matrix, allowing a fully general structure with respect to heteroskedasticity and serial correlation in panel data.

[b] Population is in thousands of inhabitants, due to a large number of decimals otherwise.
***, ** and * indicate a significance level of 1%, 5% and 10%, respectively. Standard errors in brackets.

governments of more disadvantaged jurisdictions cannot afford to provide the same quality of service.

Welfare provision also exhibits significant wealth effects: the income elasticities in our sample are positive, ranging from 0.03 to 0.14. Given the low explanatory power of the estimated models, more effort nonetheless needs to be put into identifying further explanatory factors, which may present agenda for future research.

## 3.5 Discussion and policy implications

### 3.5.1 Financial difficulties of the French Departments

The question we are analyzing is rather important, since we can see the very real impact social welfare provision, as such a major part of the Departments' duties and expenditures, has on their economic situation. In the current decade, many councils find themselves encountering financial difficulties, some repeatedly bordering on insolvency (e.g., Essonne, Seine-Saint-Denis). Between 1998 and 2018, social spending has grown from around a third to over a half of annual departmental budgets. While its evolution is mainly related to the number of beneficiaries, and thus demographic and economic developments, a larger and larger role is recently being played by revaluation of benefits by the central government. Presently, such mandatory expenditures, determined at the national level (*Allocations individuelles de solidarité* (AIS), composed of the unemployment benefit (RSA), aid to the elderly (APA), and a disability compensation (PCH)), themselves form one-third of the metropolitan Departments' operating expenditures. The situation is even worse for the overseas Departments (Réunion and Guadeloupe with over 50%) (OFGL, 2018).

Consequently, despite their proclaimed autonomy, departmental councils are left with less and less room for maneuver, under pressure of the still resonating 2008 crisis (RSA has seen a decrease in the total number of beneficiaries only in 2016, and now constitutes over a half of the mandatory payments, expanding fast due to revaluations) and the process of

population ageing (the mass of APA beneficiaries is expected only to grow). Accordingly, the topic of the possible re-centralization of the RSA benefit, or partial compensation otherwise, has become one of the major political concerns advanced by the *Assemblée des Départements de France*.

### 3.5.2 Urban and rural areas

One theme that seems to emerge in our results is the role played by urbanization. In general, both with regard to the distribution of users and size effects, the results seem to contrast more rural and more urban Departments in our sample. Departments with both large populations and population densities, also correlated with higher mean income, seem to distribute more resources. Such "urban" Departments also share a distinctive demand structure, e.g., a particularly large share of the unemployed, yet few elderly among the welfare beneficiaries. Indeed, when examining the subject of social welfare provision, the overall living conditions of the constituents should be taken into consideration. Lifestyles and economic situations in traditional low-density rural areas and in more industrialized and populated urban zones differ and correspond to different demands for public goods and services.

To captures these effects, we introduce dummy variables based on the definition of INSEE (see Appendix D for details) in order to distinguish between more rural Departments and those containing large urban agglomerations. In the sample, 35 Departments, designated as *urban*, contain one or more (in three cases) of the 41 urban areas of regional and national importance (*aires urbaines*). Among them, 11 Departments (variable *metropol*) have for a capital one of the French metropoles (*aires métropolitaines*): Bordeaux, Grenoble, Lille, Lyon, Marseille, Montpellier, Nantes, Nice, Strasbourg, Toulouse and Rennes (Paris excluded from the sample). As can be seen in Figure 3.6, the urbanization dummy seems to be correlated with both the share of welfare recipients, social expenditure, and the benchmarking scores, particularly in the north and south of the country.

The INSEE classification, however, is based on the size and characteristics of municipalities. Applied to Departments, it determines the presence of a population

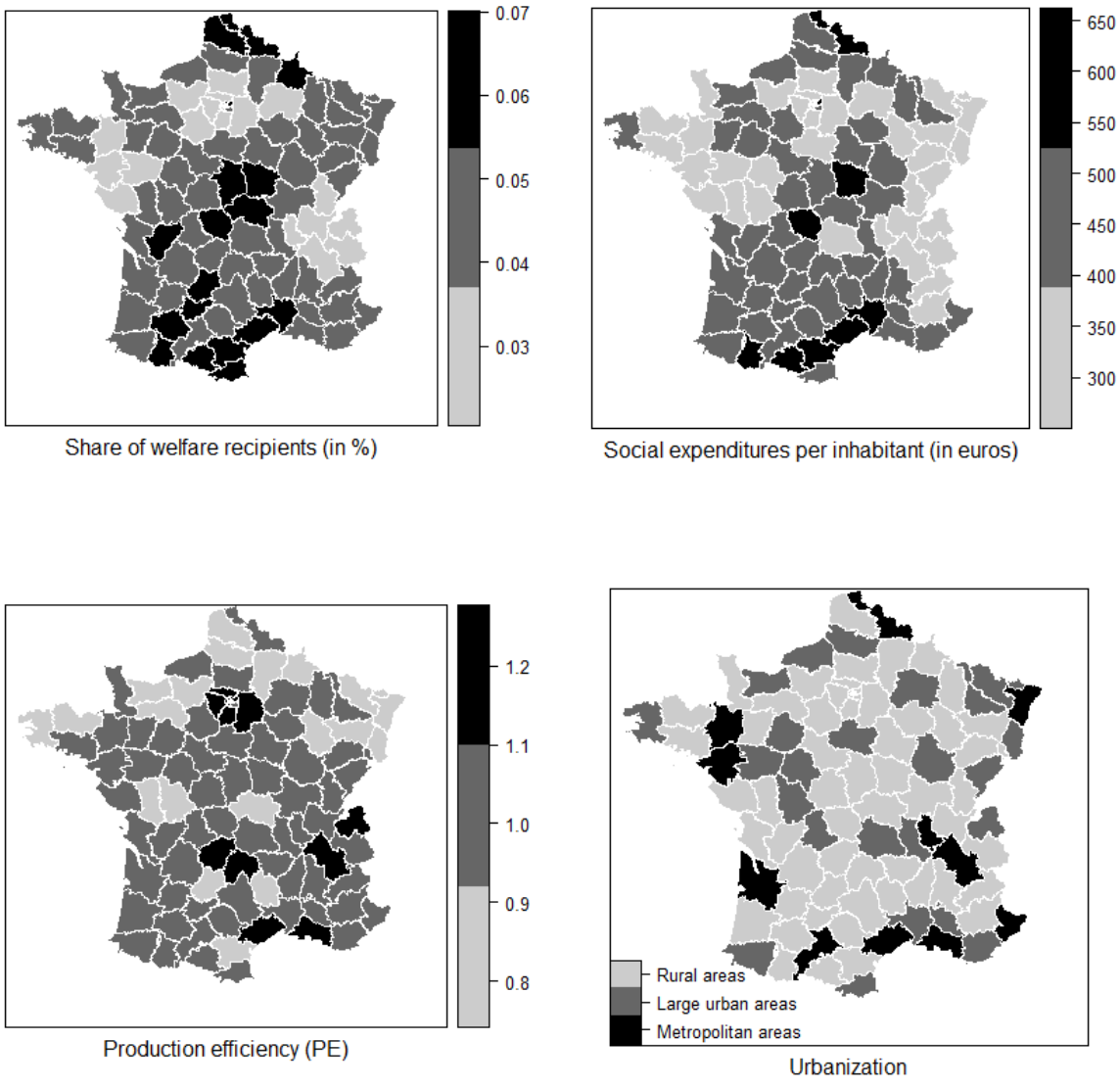



Figure 3.6. French Departments in 2008.

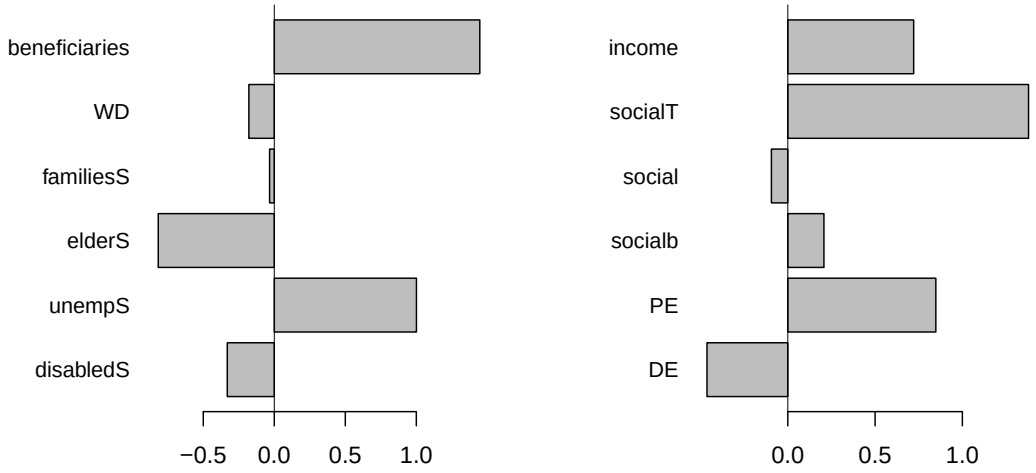



Figure 3.7. Standardized mean differences between high and low density areas.

center and may not provide an entirely accurate description: e.g., "urban" Departments Côte-d'Or (21) with the capital of Dijon, or Puy-de-Dôme (63) with Clermont-Ferrand. As a robustness check we therefore construct an additional dichotomy (variable *high_density*) following the definition of *Départements urbains* for the purpose of grant distribution (population density of over 100 inhabitants per km$^2$ and, simultaneously, urbanization rate above 65%[4]), which takes into account also the overall population density of the jurisdiction. The results are similar. As illustrated in Figure 3.7, the two groups differ significantly in their socio-economic characteristics: on average, urban Departments with a higher population density tend to be richer and have fewer welfare-dependent inhabitants, consisting largely of the unemployed; they spend less per capita, yet more per beneficiary, particularly when controlling for demand structure heterogeneity.

The effect of urbanization is further corroborated (Table 3.4) when the models in Section 3.4.3 are re-estimated as:

$$\begin{aligned}\log(PE_{i,t}) &= \delta_0 + \delta_1 \log(WD_{i,t}) + \delta_2 \log(income_{i,t}) + \delta_3 \log(grants_{i,t}) + \delta_4 URBAN_{i,t} \\ &\quad + \delta_5 DEC02_t + \delta_6 DEC04_t + \varepsilon_{i,t}, \end{aligned} \tag{3.9}$$

[4]Urbanization rate, indicated for the years 1999 and 2007, is measured as the number of municipalities (*Communes*) belonging to urban units (*unités urbaines*), i.e., contiguously built-up areas, as defined by INSEE.

Table 3.4. Production effect (PE) scores – urbanization variables[a]

| | *Dependent variable:* | | | | | |
|---|---|---|---|---|---|---|
| | log(*PE*) | | | | | |
| | Pooled | RE | Pooled | RE | Pooled | RE |
| | (1) | (2) | (3) | (4) | (5) | (6) |
| Constant | $-1.977^{***}$ | $-1.016^{***}$ | $-1.871^{***}$ | $-1.009^{***}$ | $-1.537^{***}$ | $-0.925^{***}$ |
| | (0.381) | (0.201) | (0.361) | (0.202) | (0.347) | (0.210) |
| log(WD) | $-0.071^{*}$ | $-0.186^{***}$ | $-0.074^{*}$ | $-0.187^{***}$ | $-0.073^{**}$ | $-0.177^{***}$ |
| | (0.038) | (0.036) | (0.039) | (0.036) | (0.034) | (0.037) |
| log(income) | $0.202^{***}$ | $0.046^{***}$ | $0.191^{***}$ | $0.045^{***}$ | $0.143^{***}$ | $0.038^{**}$ |
| | (0.037) | (0.017) | (0.037) | (0.017) | (0.035) | (0.017) |
| log(grants) | $-0.029$ | $-0.026^{**}$ | $-0.033$ | $-0.027^{**}$ | $-0.015$ | $-0.025^{**}$ |
| | (0.025) | (0.012) | (0.023) | (0.012) | (0.023) | (0.012) |
| urban | 0.030 | $0.039^{*}$ | | | | |
| | (0.022) | (0.023) | | | | |
| metropol | | | $0.055^{*}$ | $0.072^{**}$ | | |
| | | | (0.032) | (0.034) | | |
| high_density | | | | | $0.070^{***}$ | $0.046^{*}$ |
| | | | | | (0.023) | (0.025) |
| DEC02 | $0.033^{**}$ | $0.078^{***}$ | $0.036^{**}$ | $0.078^{***}$ | $0.031^{**}$ | $0.074^{***}$ |
| | (0.014) | (0.015) | (0.014) | (0.015) | (0.014) | (0.015) |
| DEC04 | $-0.005$ | $0.039^{***}$ | $-0.002$ | $0.040^{***}$ | 0.002 | $0.038^{***}$ |
| | (0.013) | (0.010) | (0.013) | (0.010) | (0.012) | (0.010) |
| Observations | 1,023 | 1,023 | 1,023 | 1,023 | 1,023 | 1,023 |
| $R^2$ | 0.171 | 0.085 | 0.177 | 0.086 | 0.211 | 0.088 |
| Adjusted $R^2$ | 0.170 | 0.084 | 0.176 | 0.086 | 0.210 | 0.088 |

[a] Arellano's method (1987) was used to compute a robust covariance matrix, allowing a fully general structure with respect to heteroskedasticity and serial correlation in panel data.
***, ** and * indicate a significance level of 1%, 5% and 10%, respectively. Standard errors in brackets.

where $URBAN$ designs a dummy variable *urban* (columns 1–2), *metropol* (columns 3–4), or *high_density* (columns 5–6).

Providing in-depth insight into this issue would nevertheless require more rigorous analysis and is not the aim of this paper. Rather, we point out a concern to be taken into consideration, likely linked to urbanization, which may have practical consequences for the financial situation of the Departments. According to analytical reports (Direction des études de la Banque Postale Collectivités Locales, 2015), among the most affected by the upsurge in RSA payments are urban Departments stricken with high unemployment rates,

i.e., those previously benefiting from a "cheaper" demand structure in our sample a decade ago. Similarly, APA presents a serious concern mainly to less populated rural Departments with older constituencies, to be faced in the upcoming years and decades.

### 3.5.3 Intergovernmental grants and fiscal equalization

Another important issue is raised if we consider social welfare spending in the context of fiscal equalization, i.e., how public resources are redistributed across jurisdictions in order to reduce differences in own revenues or costs. Since the constitutional revision of 2003, fiscal equalization is recognized as one of the key concerns in French local public finance (*Article 72-2*). Given that resources and expenses depend on geographical, human (e.g., income, education, age) and economic (e.g., tax bases, distribution of economic sectors) constraints, an adequate coverage of costs cannot be *a priori* guaranteed for all. In response, fiscal equalization would mitigate the disparities in resources between local authorities with regard to the burdens they face; literally, the system should "equalize" their situations.

Fiscal equalization is the essential counterpart of the increase in competences and autonomy granted to local authorities, particularly in terms of taxation. Until 2010, the principle was essentially implemented through state grants (vertical equalization), the main tools being *Dotation générale de décentralisation* (DGD) and *Dotation globale de fonctionnement* (DGF).

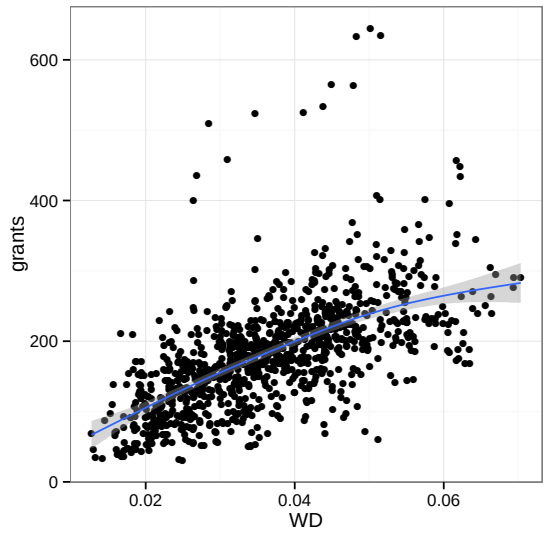
The DGD (around 5–10% of departmental revenues, before 2004) represents a long-standing grant, introduced in the early 1980s (*loi du 2 mars 1982, loi du 7 janvier 1983*), intended to compensate local authorities for the performance of additional responsibilities delegated to them by the central government. As a part of financial reform initiated in 2004, 95% of the aid have been integrated into the DGF.

The DGF (around 15–20% of revenues) labels a complex grant scheme available at different government levels: to French Regions (approximately 13% of the means), Departments (29%) and Communes (58%). For Departments, it consists of a lump

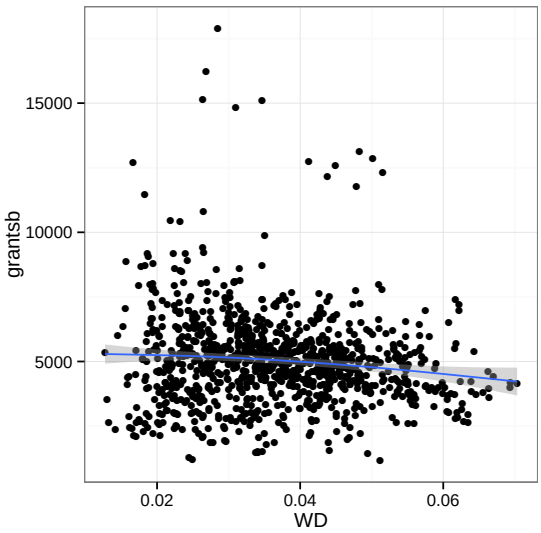
sum allocated to all and a fiscal equalization (*péréquation*) part aimed at the most disadvantaged areas, further subdivided into *Dotation de péréquation urbaine* (DPU), for urban Departments, and *Dotation de fonctionnement minimal* (DFM), for rural ones. Plagued by a long debate concerning the national costs, its undue complexity and (in)ability to accomplish the goal of equalizing the positions of the Departments, the future of the grant is uncertain. While the departmental DGF has been alternatively lightly lowered or kept constant over the past decade, leading up to marked systematic reductions since 2014, its regional equivalent has been replaced by VAT revenue transfers instead, starting from 2018 (OFGL).

Despite the criticism, our data shows that more disadvantaged jurisdictions, burdened by larger shares of welfare-dependent constituents, do in general seem to receive more support. There is a positive correlation between the share of welfare-dependent inhabitants ($WD$) and total per capita grants received by the Department, as illustrated in panel (a) of Figure 3.8. However, the DGD does not follow the trend (panel (c)), and the DGF (panel (d)) seems to disproportionally favor Departments with relatively high inequalities, such as Nord (59), Bouches-du-Rhône (13), Pas-de-Calais (62), or Rhône (69), yet not among those the most disadvantaged: Ariège (9), Aude (11), Pyrénées-Orientales (66), and others with a share of welfare-dependent population surpassing 6%. As a matter of fact, recounted per beneficiary (panel (b)), total grants received still on average decrease marginally.

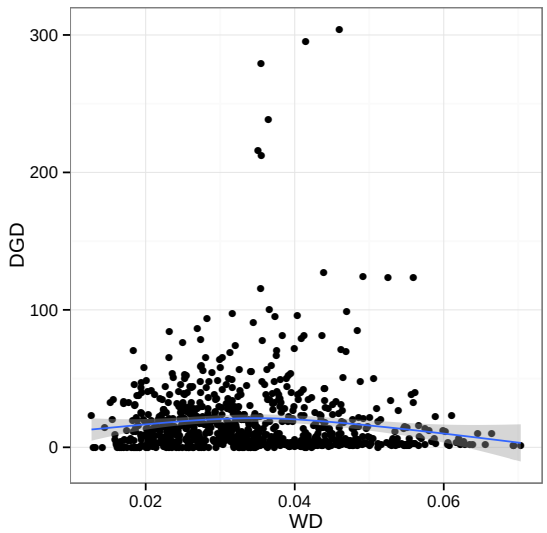
Nevertheless, the positive implication of the results is that the grant support indeed happens to consider and compensate for the burden imposed on councils' budgets by their welfare-dependent populations, and thus does contribute towards equalizing their circumstances, albeit imperfectly. Yet, while the almost horizontal trend curve in panel (b) approaches the concept of equal support per beneficiary, no matter the dependent population size, this is the case on average only. The large remaining heterogeneity (vertical distance from the trend) among Departments may help explain why we still find such differences in production effect scores, i.e., the sums allocated per beneficiary regardless of the distribution of users. It furthermore leaves room for improvement of the equalization toolkit, which should take into consideration both the demand structure and the demand size related to this



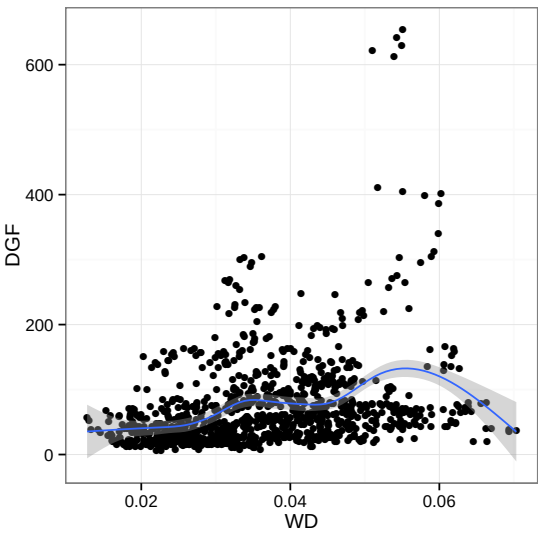

(a) Total grants per capita




(b) Total grants per beneficiary




(c) Dotation générale de décentralisation (per capita)




(d) Dotation globale de fonctionnement (per capita)

Figure 3.8. Intergovernmental grant support received by the Departments.

primary responsibility of departmental councils.

## 3.6 Conclusion

The implementation of public spending programs is often devolved to local governments, which operate independently in their geographical area to achieve the results intended by the program. How the local governments carry out their tasks is a rather important question, since there is no straightforward way of measuring their relative efficiency. They may face different constraints, various application settings, and may have chosen different organizational models. Hence, electorate's heterogeneity may not only determine the composition of the government, but also the costs of public good provision.

According to our data on social welfare provision—the primary largest task devolved to departmental councils—French Departments show considerable heterogeneity in spending, both between different welfare services and among themselves. By applying the benchmarking methodology to control for diverse demand motives, and through the subsequent nonparametric and econometric analysis of the scores, we are able to provide an insight into these large differences.

We demonstrate that social spending is characterized by a U-shaped relationship with respect to population size, associated with two opposite demand effects: a decrease in the number of elderly people on the one hand, and an increase in the number of the unemployed on the other. To account for this endogeneity problem, through benchmarking, expenditures per beneficiary (i.e., the total effect) are normalized and decomposed into the impact of the distribution of users and the residual variance in production effect scores.

The demand side of public good provision matters through two channels. First, the demand structure reflects the socio-demographics of the population, presenting Departments with more or less costly distribution of users. The spending levels depend considerably on this distribution of users (families with children, elderly, unemployed or disabled beneficiaries), as evidenced by the reduced individual differences in production effect scores.

Furthermore, the demand structure evolves with population size, leading to relatively large shares of the unemployed and fewer elderly beneficiaries in more populated Departments.

Second, although individual differences in spending are partly explained by the structure of services, a large variation remains, presenting the impact of the size of the welfare-dependent population. Large numbers of beneficiaries naturally constitute a constraint on the Departments' budgets and their ability to provide social services, demand structure notwithstanding. For instance, the Department of Var (83), facing virtually identical demand structure as Bouches-du-Rhône (13), yet providing for one-third the number of users, allocated in 1998 on average 680 euros more per beneficiary. Particularly affected by these congestion effects seem to be smaller, less densely populated Departments with relatively poorer constituencies.

The performance of French Departments in social welfare provision thus depends both on the total number of users, i.e., the demand size, and on the distribution of users, i.e., the demand structure. Further policy implications of our finding were discussed in the context of urbanization (related both to particular demand structure and size effects) and fiscal equalization (attempting to but not quite succeeding at mitigating the disparities between departmental councils).

Funding of welfare expenses is at the heart of the debates between the Departments and the state, since the transferred resources are proving insufficient. At the same time, assessing how the Departments perform in their task can be difficult, particularly in this case of complex heterogeneous services that cannot be reduced to the elementary logic of minimizing spending to maximize efficiency. Further thorough analysis of the overall situation and how it developed, taking into account findings concerning the structure and size of the demand, may bring important insights that could be used also by individual departmental councils, who presently often strive desperately to plan and meet their budgets. Moreover, it might aid the stalemate plaguing the debate about the *Dotation globale de fonctionnement* (DGF) and other tools aimed at improving and equalizing the fiscal situation of the French Departments, which has reached the point where some councils consider rather deliberately defaulting on the RSA payments instead (Vives, 2018).

# General conclusion

Analogously to the private resources and private individuals of the private sector, the public sector of the economy operates with public resources in order to provide goods and services to its constituents. Accordingly, how the resources, typically obtained through taxation, are used, i.e., how the limited budgets are distributed and spent on public policies, is the product of interaction between supply and demand. In representative democracies, the demand side, i.e., the consumers of goods and services, is represented by the citizens—the voters casting ballots in elections. The supply side, i.e., the producers/providers of the goods and services, is embodied by the politicians—the elected representatives.

Demand driven models, such as the median voter theorem or the Meltzer and Richard hypothesis, state that the characteristics and preferences of the electorate are the main forces driving public decisions. Models focusing on the supply side argue instead that the composition of governing bodies (e.g., number of parties, ideology of elected officials) is the key explanatory factor. The aim of the thesis has been to shed light on the connection between the two categories of explanations, through a series of empirical essays, using a quasi-experimental approach to study the local public sector in France and Finland.

In the first chapter, we have examined whether and how partisan effects may play a role in the spending decisions of French Departments. The political economy literature concurs that left-wing governments tax and spend more than right-wing ones. We first theoretically demonstrate that this result is more complex than it seems: what may seem a partisan effect, due to the direct impact of parties' ideology on public spending, may actually be a selection bias, because changes in the distribution of voters' preferences determine changes in the ideology of the government in office. We have attempted to overcome this problem of observational equivalence by applying two identification strategies, regression discontinuity design and propensity score matching. Using data from 93 departmental councils, over a period of eleven years, between 1998 and 2008, we show that left-wing governments facing the same economic situation as right-wing ones do not spend more, particularly in the case of social expenditures.

In the second chapter, we have focused on the local public sector in Finland, formed by a single level of municipal councils. In the political economy literature, there is a continuous debate about whether government composition truly affects public spending levels. One reason is that several confounding factors may blur our understanding of the phenomenon: on the one hand, the degree of government fragmentation and the intensity of political competition can be interrelated; on the other hand, the distribution of voters' preferences may determine the number of parties that compete in a given polity. To address these issues, our study examined a panel of 368 Finnish municipalities over two election years, employing propensity score matching and regression discontinuity design methods in order to control for both electorate heterogeneity and the intensity of political competition. We demonstrate that there is no significant difference in public spending between single-party majorities and minority governments. Instead, using sample-splitting techniques, we have located a significant breakpoint in our data, where the largest party holds a supermajority of approximately three-fifths of the seats. At this threshold, spending levels are found to increase significantly, by nearly 250 euros per capita.

Finally, in the third chapter, we have returned to the French Departments, focusing instead on the demand side of public good provision, i.e., the electorate and, in particular,

the beneficiaries of the social welfare services provided by the departmental councils—the welfare-dependent population. In most attempts to test the impact of population size on public spending, socio-demographic concerns are excluded from the analysis. We have departed from the literature by proposing a new approach, in the form of benchmarking analysis, particularly suitable for dealing with heterogeneous users of public services. We show that social spending is characterized by a U-shaped relationship with respect to the population size, associated with two opposite demand effects: a decrease in the number of elderly people on the one hand, and an increase in the number of the unemployed on the other. Social welfare expenditures depend both on the demand size and on the demand structure, proving the demand side of public good provision to be highly relevant in explaining policy choices. Policy implications of our results were discussed in the context of urbanization and fiscal equalization.

While the results cannot be easily generalized to other types of government levels and/or countries, our findings show how closely the demand (voters) and the supply (politicians) sides of public policies are connected. In particular, we demonstrate that empirical results in favor of supply side explanations have to be interpreted with caution: what traditional econometric methods may deem supply side effects, might actually constitute a selection bias, since the characteristics of the electorate jointly affect the configuration of the political arena and policy choices. Without controlling for this essential link, the results of empirical studies are inevitably biased.

Surprisingly, while much research has been done into the topics discussed in the thesis, there is at times little consistency in the empirical results and many empirical studies are in fact not causal (see, e.g., Potrafke, 2018). The main underlying concern, while hardly rare and isolated to this field of study, seems to be that each discipline (statistics, econometrics, political economy, political science) evolves separately from the others. For one, the political economy and political science literature often examine technically corresponding subject matters. In reality, however, much more focus is given by one side to certain issues (e.g., the endogeneity of government composition in political science). Unfortunately, this is done in a manner that is not easily accessible to the other side without eventual aid and

clarification by multidisciplinary experts or teams. Both fields would likely greatly benefit from an even closer collaboration, yet this is made difficult by the routine clear division in textbooks, journals, general overviews, etc., as well as different preferences in methodology and terminology. Another example is that of the quasi-experimental approach, which became increasingly popular after 2000, and yet only a few papers have incorporated the techniques in addressing the impact of government composition.

Second, for practical convenience, many theories in political economy are examined and tested separately. The question of a combination of effects, i.e., several processes working at the same time, is rarely addressed. While pragmatic and understandable, new insights would be gained from the integration of related subjects that may have an impact—such as the close relationship between political fragmentation and political competition. Furthermore, key notions are not always clearly and uniformly defined across the literature (e.g., partisan effects). This can be an issue, as empirical results are often commented on as being extremely sensitive to definitions. For instance, Ashworth *et al.* (2005) obtain different effects depending on whether political fragmentation is defined as the number of parties or the *effective* number of parties; later, the same author (Ashworth *et al.*, 2014) already considers the impact of political fragmentation through the number of parties competing in elections (electoral competition) and the volatility of election outcomes over time (intertemporal competition) besides.

Last but not least, the choice of the identification strategy can also be subject to methodology quarrels. For instance, the "traditional" regression discontinuity design tends to be promoted by some researchers, who oppose the application of matching techniques to the point of excluding them even as a robustness check. On the contrary, propensity score matching is preferred and frequently used in health and cognitive sciences in order to estimate average treatment effects. While all methods present advantages and limitations, the idea should be to be aware of both and benefit from methodological combinations that allow to offset the potential shortcomings in a given context.

In the meantime, these obstacles offer a formidable agenda for future research. The political economy literature only stands to benefit from the recent and future improvements

in the statistical sciences, and the increasing availability of larger databases. Just as it will continue to learn from interactions with the political and other social sciences. Hence, new theories encompassing several types of behavior could be developed and tested in a more robust way. The PhD thesis can thus end on a rather positive note—while, at this time, much remains to be done in the literature, we can see that the tools and the path to forge ahead are being prepared.

# Appendices

# Appendix A

## Computation of partial derivatives

The first order-condition is given by $\frac{d\Omega}{dz} = 0$:

$$\Omega' = \theta_1^j \frac{\partial U_1(y_1, z)}{\partial z} - \theta_2^j p_2 \frac{\partial U_2(y_2 - p_2 z, z)}{\partial x} + \theta_2^j \frac{\partial U_2(y_2 - p_2 z, z)}{\partial z} = 0.$$

We can figure out the derivatives of $z^*$ with respect to any other exogenous variable, say $k$, using the implicit function theorem: $\frac{dz^*}{dk} = -\frac{\partial \Omega'/\partial k}{\partial \Omega'/\partial z}$. We have:

$$\frac{\partial \Omega'}{\partial z} = \theta_1^j \frac{\partial^2 U_1}{\partial z^2} - \theta_2^j p_2 \frac{\partial^2 U_2}{\partial x \partial z} + \theta_2^j \frac{\partial^2 U_2}{\partial z^2},$$

which is negative since $U_1$ and $U_2$ are concave and their cross partials are assumed to be non-negative.

The partial derivative of $\Omega'$ with respect to $y_1$ is given by:

$$\frac{\partial \Omega'}{\partial y_1} = \theta_1^j \frac{\partial^2 U_1}{\partial z \partial x},$$

which is positive. Hence, we get $dz^*/dy_1 > 0$.

The derivative of $\Omega'$ with respect to $y_2$ is:

$$\frac{\partial \Omega'}{\partial y_2} = -\theta_2^j p_2 \frac{\partial^2 U_2}{\partial x^2} + \theta_2^j \frac{\partial^2 U_2}{\partial z \partial x},$$

which is positive. We have $dz^*/dy > 0$.

The derivative of $\Omega'$ with respect to $p_2$ is:

$$\frac{\partial \Omega'}{\partial p_2} = -\theta_2^j \frac{\partial U_2}{\partial x} + \theta_2^j (p_2)^2 \frac{\partial^2 U_2}{\partial x^2} - \theta_2^j p_2 \frac{\partial^2 U_2}{\partial z \partial x},$$

which is negative. We get $dz^*/dp_2 > 0$.

The derivative of $\Omega'$ with respect to $\theta_1$ is:

$$\frac{\partial \Omega'}{\partial \theta_1^j} = \frac{\partial U_1(y_1, z)}{\partial z},$$

which is always positive. We have $dz^*/d\theta_1 > 0$.

The derivative of $\Omega'$ with respect to $\theta_2^j$ is:

$$\frac{\partial \Omega'}{\partial \theta_2^j} = -p_2 \frac{\partial U_2}{\partial x} + \frac{\partial U_2}{\partial z},$$

i.e., we have $dz^*/d\theta_2 < 0$ iff $\frac{\partial U_2}{\partial z} < p_2 \frac{\partial U_2}{\partial x}$, which is true given the first-order condition.

Last, since $p_2 = \frac{cN^\alpha}{n_2}$, the sign of the derivative with respect to $\alpha$ and $c$ is directly given by the sign of $dz^*/dp_2$.

# Appendix B

## Principal component analysis of the French political spectrum

Principal component analysis (PCA) has been implemented to identify the link between the share of welfare recipients, the share of seats held by extreme parties, and those held by mainstream center-left, center-right and centrist coalitions (see Table B1). The advantage of implementing PCA is that it reduces the multidimensionality of the problem to a two-axis dimension.

Figure B1 shows the relationship between the shares of seats and the demand variables ($families$, $elder$, $unemp$, $disabled$). The first two dimensions sum up almost 50% of the total inertia. Consequently, the PCA illustration can be readily interpreted.

The first component (horizontal axis) represents the most information and is in line with our analysis. It opposes the far-right, the center-right and centrist parties to both the center-left and far-left parties. The construction of the variable $L$ is based on this observation. As can be seen from the first row of Table B2, the opposition is linked to the share of welfare recipients, as illustrated by the correlation of each variable with the first dimension.

The second component (vertical axis) reveals that the political spectrum is even more complex (second row of Table B2). There is an opposition between (1) Departments with a large number of unemployed recipients (south quadrant), urban jurisdictions for the most part, and (2) Departments with a high share of elderly and disabled people who benefit from social assistance (north quadrant), mainly rural.

Table B1. Political parties on the left, center, and right[a]

| | |
|---|---|
| $L = 1$ | *Far.left* – share of seats in the council held by far-left parties<br>PRG, CAP, Les Alternatifs, PC/PCF, Parti de gauche, PT |
| | *Center.left* – share of seats in the council held by center-left parties<br>ADS, MDC, MRC, Majorité Présidentielle, PS, ADD, MDR, GE, MEI, Les verts, Ecologie, Europe-écologie les verts, CAP21 |
| $L = 0$ | *Center* – share of seats in the council held by centrist parties<br>UDF - PSD, UDF radical, PRV, UDF-CDS, UDF-PR, UDFP et R, UDF, MODEM, Alliance centriste, NC |
| | *Center.right* – share of seats in the council held by center-right parties<br>RPR, UDI, MPF, DL, RPF, UMP, PCD |
| | *Far.right* – share of seats in the council held by far-right parties<br>CNPT, CNI, CNIP, FN, Alsace d'abord, DLR, LDS, Unser Land |

[a] Some of the candidates were independent, i.e., did not belong to a political party. However, we knew the ideology of these independent candidates, i.e., far left-wing, left-wing, right-wing or far right-wing.

Table B2. Correlation with the first and second axes

| | far.left | center.left | center | center.right | far.right | families | elder | unemp | disabled |
|---|---|---|---|---|---|---|---|---|---|
| First axis | 0.409 | 0.776 | -0.618 | -0.658 | -0.064 | 0.339 | 0.576 | 0.590 | 0.446 |
| Second axis | -0.302 | -0.197 | -0.355 | 0.582 | -0.320 | 0.069 | 0.654 | -0.463 | 0.659 |

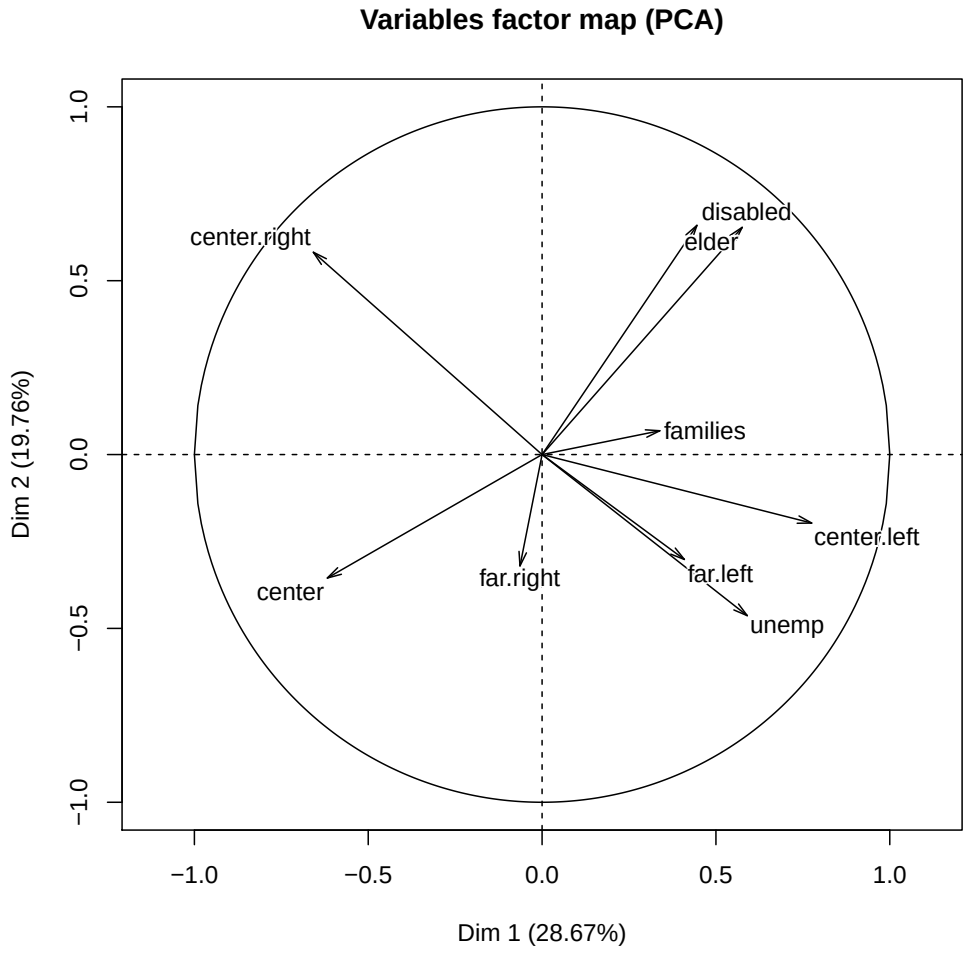



Figure B1. Correlation between the share of welfare-dependent individuals and the share of seats held by centrist and extreme parties

# Appendix C

## Quality of PSM and RDD comparisons

In order to ascertain that the PSM method was successful in achieving balance, i.e., that the treated and non-treated observations are near-identical, various techniques can be used to formally asses the quality of the match. The basis of most is to examine and verify the similarity of the matched units with regard to various covariates representing their key characteristics, i.e., to perform a comparison between group means or medians. No great or significant difference should be found particularly for the variables identified as significant by the propensity score estimation, since it is indirectly on these attributes that the pairs are matched.

Two of the more popular customary methods, a calculation of average standardized bias (SB) and $t$-tests of exogenous variables, are presented below. The standardized bias or "standardized difference" was defined by Rosenbaum and Rubin (1985) as:

$$SB = 100 \times \frac{\bar{x}_T - \bar{x}_C}{\sqrt{0.5 \times (s_T^2 + s_C^2)}}. \tag{C.1}$$

It is computed as the difference of sample means for each covariate in the treated group $\bar{x}_T$ and the matched comparison group $\bar{x}_C$, divided by the square root of the average of sample variances $s_T^2$ and $s_C^2$ in both groups. Although there are no formal rules, SB between 3 and 5% is seen as sufficient (Caliendo and Kopeinig, 2008), while SB of 20% after matching is considered large (Rosenbaum and Rubin, 1985).

A similar approach using a simple two-sample $t$-test moreover supplies the statistical significance of differences in covariate means between the treated and non-treated groups. Generally, the tighter the caliper, the higher the quality of the match, as demonstrated by fewer significant differences in control variables between the groups and by reduced SB.

Tables C1 and C2 present matching quality indicators for each PSM iteration in Chapter 2: observations matched on *MAJ* and *REG*, respectively. As can be seen, a tighter caliper indeed corresponds to a higher quality of the match. For SB, the calculated values are below 20% and often well below or around 10%. SB is particularly useful in its ability to illustrate the marked reduction of bias observed before matching: e.g., from 115% to 2% for mean income in the *MAJ* match.

According to Table C2, some variables (e.g., *population* or *density*) still show a significant difference after matching, pointing out the importance of the selection bias, as well as the inherent difficulty of finding balance on some of the observed covariates. Yet, differences between the group means are strongly reduced as the caliper decreases. For instance, before matching, the difference in population between minority governments ($MAJ = 0$) and single-party majorities ($MAJ = 1$) was $19,355 - 4,566 = 14,789$ inhabitants. After matching, with the smallest caliper (0.02), the difference amounts to only $13,671 - 8,577 = 5,094$. This convergence process is observed for virtually all covariates. What should thus be retained here is that the impact of government composition on public spending becomes insignificant when we attempt to reduce the selection bias. The few differences that persist between the treatment and comparison groups are of less importance, relatively speaking, as illustrated by the SB measure.

Based on the results of two-sample *t*-tests, note that propensity score matching is able to better reduce the observed differences in the exogenous variables than the RDD approach (see Table C3), and thus in theory the selection bias in our data.

Table C1. Matching quality indicators – standardized bias (SB) in percentage

| Matched on *MAJ* | | *Matching algorithm:* | | | |
|---|---|---|---|---|---|
| Variable | Before matching | Nearest neighbor | Caliper 0.06 | Caliper 0.04 | Caliper 0.02 |
| children | -23.42 | -5.67 | -6.50 | -7.15 | -5.57 |
| young | 7.35 | 8.71 | 7.28 | -7.15 | 7.20 |
| elderly | 60.92 | 6.89 | 6.13 | 5.59 | 2.52 |
| unemp | 22.77 | 7.47 | 6.32 | 5.15 | 9.00 |
| population | -42.28 | -16.06 | -16.36 | -16.37 | -18.26 |
| density | -44.92 | -14.96 | -15.16 | -15.65 | -16.62 |
| income | -114.57 | -0.32 | 1.18 | 1.99 | 1.95 |

| Matched on *REG* | | *Matching algorithm:* | | | |
|---|---|---|---|---|---|
| Variable | Before matching | Nearest neighbor | Caliper 0.06 | Caliper 0.04 | Caliper 0.02 |
| children | -6.85 | 1.57 | -3.97 | -2.76 | -2.14 |
| young | 32.99 | 8.94 | -6.21 | -2.76 | -6.35 |
| elderly | 41.99 | 4.68 | -3.34 | -3.50 | -5.01 |
| unemp | 7.84 | -53.55 | -26.76 | -20.93 | -17.10 |
| population | -37.67 | -28.40 | -3.22 | 3.45 | 11.32 |
| density | -39.84 | -30.61 | -26.04 | -24.03 | -19.40 |
| income | -102.66 | 33.37 | 39.20 | 32.83 | 30.12 |

Table C2. Matching quality indicators – difference in exogenous variables: two-sample $t$-tests

| | | | *Matching algorithm:* | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Before matching | | Nearest neighbor | | Caliper 0.06 | | Caliper 0.04 | | Caliper 0.02 | |
| *MAJ* | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| children | 7.565 | 7.153*** | 7.267 | 7.174 | 7.279 | 7.172 | 7.249 | 7.133 | 7.213 | 7.123 |
| young | 10.209 | 10.337** | 10.086 | 10.226* | 10.098 | 10.214 | 7.249 | 7.133 | 10.003 | 10.110 |
| elderly | 18.135 | 20.822*** | 19.568 | 19.867 | 19.587 | 19.854 | 19.618 | 19.858 | 19.737 | 19.845 |
| unemp | 10.862 | 11.875*** | 12.371 | 12.698 | 12.394 | 12.669 | 12.453 | 12.677 | 12.434 | 12.821* |
| population | 19,355 | 4,566*** | 11,999 | 7,971*** | 12,240 | 8,087*** | 12,545 | 8,286*** | 13,671 | 8,577*** |
| density | 95.937 | 9.111*** | 56.043 | 33.916*** | 57.332 | 34.622*** | 59.698 | 35.682*** | 66.072 | 38.729*** |
| income | 10,354 | 8,313*** | 9,065 | 9,059 | 9,059 | 9,083 | 9,096 | 9,135 | 9,138 | 9,177 |

| | | | *Matching algorithm:* | | | | | | | |
|---|---|---|---|---|---|---|---|---|---|---|
| | Before matching | | Nearest neighbor | | Caliper 0.06 | | Caliper 0.04 | | Caliper 0.02 | |
| *REG* | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 | 0 | 1 |
| children | 7.532 | 7.413 | 7.308 | 7.332 | 7.110 | 7.050 | 7.101 | 7.059 | 7.024 | 6.993 |
| young | 10.236 | 10.811*** | 10.250 | 10.385* | 10.188 | 10.091 | 7.101 | 7.059 | 10.180 | 10.084 |
| elderly | 18.639 | 20.420*** | 19.301 | 19.490 | 20.653 | 20.525 | 20.698 | 20.563 | 20.908 | 20.722 |
| unemp | 12.305 | 12.678 | 13.118 | 10.433*** | 12.954 | 11.621*** | 12.955 | 11.901*** | 13.129 | 12.258** |
| population | 15,400 | 4,119*** | 11,767 | 4,589*** | 4,614 | 4,498 | 4,488 | 4,612 | 4,350 | 4,753* |
| density | 75.312 | 6.638*** | 54.415 | 9.949*** | 10.160 | 8.505*** | 9.924 | 8.399*** | 9.270 | 8.163*** |
| income | 9,470 | 7,733*** | 8,936 | 9,645*** | 8,235 | 8,869*** | 8,177 | 8,687*** | 8,059 | 8,496*** |

***, ** and * indicate a significant difference at a 1%, 5% and 10% level, respectively.

Table C3. RDD quality indicators – difference in exogenous variables for *MAJ* and *REG*: two-sample *t*-tests

| | | | *Bandwidth:* | | | |
|---|---|---|---|---|---|---|
| | Whole sample | | 0.109 | | 0.104 | |
| *MAJ* | 0 | 1 | 0 | 1 | 0 | 1 |
| children | 7.666 | 7.306*** | 7.493 | 7.119** | 7.443 | 7.119* |
| young | 10.300 | 10.476 | 10.338 | 10.114 | 10.318 | 10.114 |
| elderly | 17.877 | 20.538*** | 19.130 | 20.633*** | 19.310 | 20.633*** |
| unemp | 11.956 | 12.930*** | 12.261 | 13.599*** | 12.301 | 13.599*** |
| population | 19,198 | 4,615*** | 7,226 | 5,361*** | 6,554 | 5,361*** |
| density | 98.619 | 9.427*** | 39.467 | 11.930*** | 36.775 | 11.930** |
| income | 9,947 | 7,939*** | 9,394 | 8,143*** | 9,321 | 8,143*** |

| | | | *Bandwidth:* | | | |
|---|---|---|---|---|---|---|
| | Whole sample | | 0.188 | | 0.108 | |
| *REG* | 0 | 1 | 0 | 1 | 0 | 1 |
| children | 7.532 | 7.413 | 7.317 | 7.455 | 7.286 | 7.245 |
| young | 10.236 | 10.811*** | 10.209 | 10.847*** | 10.154 | 10.659** |
| elderly | 18.639 | 20.420*** | 19.946 | 20.275 | 20.267 | 20.422 |
| unemp | 12.305 | 12.678 | 12.715 | 13.451 | 12.976 | 14.755*** |
| population | 15,400 | 4,119*** | 5,901 | 4,424*** | 6,047 | 4,657*** |
| density | 75.312 | 6.638*** | 26.707 | 6.508*** | 33.430 | 6.236** |
| income | 9,470 | 7,733*** | 8,835 | 7,647*** | 8,570 | 7,767*** |

***, ** and * indicate a significant difference at a 1%, 5% and 10% level, respectively.

Table C4. RDD quality indicators – difference in exogenous variables for KESK and RKP: two-sample *t*-tests

| | | | *Bandwidth:* | | | |
|---|---|---|---|---|---|---|
| | Whole sample | | 0.117 | | 0.095 | |
| KESK majority | 0 | 1 | 0 | 1 | 0 | 1 |
| children | 7.478 | 7.341 | 7.469 | 7.150* | 7.377 | 7.143 |
| young | 10.352 | 10.542 | 10.372 | 10.124 | 10.315 | 10.110 |
| elderly | 18.989 | 20.476*** | 19.374 | 20.635** | 19.636 | 20.710** |
| unemp | 12.363 | 13.740*** | 12.553 | 13.931*** | 12.580 | 13.808** |
| population | 8,342 | 4,405*** | 6,287 | 5,059*** | 5,801 | 4,894** |
| density | 20.590 | 8.167*** | 14.527 | 9.922*** | 13.923 | 10.015*** |
| income | 9,009 | 7,749*** | 8,821 | 7,985*** | 8,800 | 7,995*** |

| | | | *Bandwidth:* | | | |
|---|---|---|---|---|---|---|
| | Whole sample | | 0.139 | | 0.091 | |
| KESK supermajority | 0 | 1 | 0 | 1 | 0 | 1 |
| children | 7.380 | 7.466 | 7.269 | 7.304 | 7.205 | 7.203 |
| young | 10.253 | 10.971*** | 10.179 | 10.760*** | 10.076 | 10.656*** |
| elderly | 19.597 | 20.291* | 20.411 | 20.476 | 20.677 | 20.520 |
| unemp | 12.774 | 13.957*** | 13.306 | 14.514** | 13.505 | 15.196** |
| population | 7,094 | 3,971*** | 5,147 | 4,427* | 4,875 | 4,546 |
| density | 16.992 | 5.934*** | 11.677 | 6.042*** | 10.593 | 5.773*** |
| income | 8,663 | 7,474*** | 8,208 | 7,557*** | 8,048 | 7,653*** |

| | | | *Bandwidth:* | | | |
|---|---|---|---|---|---|---|
| | Whole sample | | 0.251 | | 0.127 | |
| RKP majority | 0 | 1 | 0 | 1 | 0 | 1 |
| children | 8.649 | 7.119*** | 8.647 | 7.496* | 8.865 | 6.985** |
| young | 11.041 | 10.075 | 11.089 | 10.719 | 11.544 | 10.310 |
| elderly | 16.224 | 21.033*** | 16.314 | 20.067*** | 16.824 | 20.511** |
| unemp | 8.635 | 6.884** | 8.278 | 7.377 | 7.472 | 7.909 |
| population | 17,332 | 5,461*** | 16,540 | 6,908** | 8,898 | 7,352 |
| density | 207.565 | 11.584** | 211.872 | 14.428** | 231.177 | 16.310* |
| income | 12,165 | 9,135*** | 12,279 | 9,453*** | 12,455 | 10,002* |

| | | | *Bandwidth:* | | | |
|---|---|---|---|---|---|---|
| | Whole sample | | 0.171 | | 0.105 | |
| RKP supermajority | 0 | 1 | 0 | 1 | 0 | 1 |
| children | 8.223 | 7.162** | 8.425 | 7.984 | 7.612 | 7.305 |
| young | 10.823 | 10.051 | 11.323 | 11.209 | 10.770 | 10.402 |
| elderly | 17.427 | 21.033*** | 18.135 | 19.207 | 18.944 | 19.779 |
| unemp | 8.476 | 6.605*** | 7.312 | 6.823 | 7.291 | 7.773 |
| population | 15,001 | 4,824*** | 8,078 | 5,930 | 9,195 | 5,228** |
| density | 160.169 | 9.984** | 160.518 | 11.270 | 243.264 | 11.094 |
| income | 11,577 | 8,965*** | 11,441 | 9,131** | 12,455 | 9,800 |

***, ** and * indicate a significant difference at a 1%, 5% and 10% level, respectively.

# Appendix D

## Urban areas in France

According to the French *National Institute of Statistics and Economic Studies* (INSEE), an urban area, or a "large urban area" (354 in metropolitan France), is *"...a group of touching municipalities, without pockets of clear land, encompassing an urban center (urban unit) providing at least 10,000 jobs, and rural districts or urban units (urban periphery) among which at least 40% of employed resident population works in the center or in the municipalities attracted by this center."*

Among these, 41 are recognized to be of regional and national importance, regrouping areas with strong economic, cultural and political potential (see Table D1). The 12 "metropolitan areas" (*aires métropolitaines*), each with over 500,000 inhabitants and a corresponding workforce structure (research and development, intellectual industries, business-to-business, management, culture and leisure), represent the largest French metropoles: Bordeaux, Grenoble, Lille, Lyon, Marseille, Montpellier, Nantes, Nice, Paris, Strasbourg, Toulouse and Rennes. The remaining 29 "large urban areas" (*grandes aires urbaines*) each account for over 200,000 inhabitants.

The metropolitan areas are located mostly along the borders, while the "large urban areas" are found in the Paris Basin or near the metropoles (Figure D1).

Table D1. Urban areas in France according to INSEE

| | *Aires métropolitaines* (12) | Department | Region |
|---|---|---|---|
| 1 | Paris | Paris (excluded) | Île-de-France (11) |
| 2 | Lyon | Rhône (69) | Auvergne-Rhône-Alpes (84) |
| 3 | Marseille | Bouches-du-Rhône (13) | Provence-Alpes-Côte d'Azur (93) |
| 4 | Toulouse | Haute-Garonne (31) | Occitanie (76) |
| 5 | Lille | Nord (59) | Hauts-de-France (32) |
| 6 | Bordeaux | Gironde (33) | Nouvelle-Aquitaine (75) |
| 7 | Nice | Alpes-Maritimes (6) | Provence-Alpes-Côte d'Azur (93) |
| 8 | Nantes | Loire-Atlantique (44) | Pays de la Loire (52) |
| 9 | Strasbourg | Bas-Rhin (67) | Grand Est (44) |
| 10 | Grenoble | Isère (38) | Auvergne-Rhône-Alpes (84) |
| 11 | Rennes | Ille-et-Vilaine (35) | Bretagne (53) |
| 12 | Montpellier | Hérault (34) | Occitanie (76) |
| | *Grands aires urbaines* (29) | Department | Region |
| 13 | Toulon | Var (83) | Provence-Alpes-Côte d'Azur (93) |
| 14 | Douai | Nord (59) | Hauts-de-France (32) |
| 15 | Rouen | Seine-Maritime (76) | Normandie (28) |
| 16 | Metz | Moselle (57) | Grand Est (44) |
| 17 | Clermont-Ferrand | Puy-de-Dôme (63) | Auvergne-Rhône-Alpes (84) |
| 18 | Nancy | Meurthe-et-Moselle (54) | Grand Est (44) |
| 19 | Valenciennes | Nord (59) | Hauts-de-France (32) |
| 20 | Tours | Indre-et-Loire (37) | Centre-Val de Loire (24) |
| 21 | Caen | Calvados (14) | Normandie (28) |
| 22 | Orléans | Loiret (45) | Centre-Val de Loire (24) |
| 23 | Angers | Maine-et-Loire (49) | Pays de la Loire (52) |
| 24 | Dijon | Côte-d'Or (21) | Bourgogne-Franche-Comté (27) |
| 25 | Saint-Étienne | Loire (42) | Auvergne-Rhône-Alpes (84) |
| 26 | Avignon | Vaucluse (84) | Provence-Alpes-Côte d'Azur (93) |
| 27 | Brest | Finistère (29) | Bretagne (53) |
| 28 | Le Mans | Sarthe (72) | Pays de la Loire (52) |
| 29 | Reims | Marne (51) | Grand Est (44) |
| 30 | Le Havre | Seine-Maritime (76) | Normandie (28) |
| 31 | Mulhouse | Haut-Rhin (68) | Grand Est (44) |
| 32 | Perpignan | Pyrénées-Orientales (66) | Occitanie (76) |
| 33 | Amiens | Somme (80) | Hauts-de-France (32) |
| 34 | Dunkerque | Nord (59) | Hauts-de-France (32) |
| 35 | Limoges | Haute-Vienne (87) | Nouvelle-Aquitaine (75) |
| 36 | Nîmes | Gard (30) | Occitanie (76) |
| 37 | Besançon | Doubs (25) | Bourgogne-Franche-Comté (27) |
| 38 | Pau | Pyrénées-Atlantiques (64) | Nouvelle-Aquitaine (75) |
| 39 | Bayonne | Pyrénées-Atlantiques (64) | Nouvelle-Aquitaine (75) |
| 40 | Poitiers | Vienne (86) | Nouvelle-Aquitaine (75) |
| 41 | Annecy | Haute-Savoie (74) | Auvergne-Rhône-Alpes (84) |

Source: INSEE, 2011.

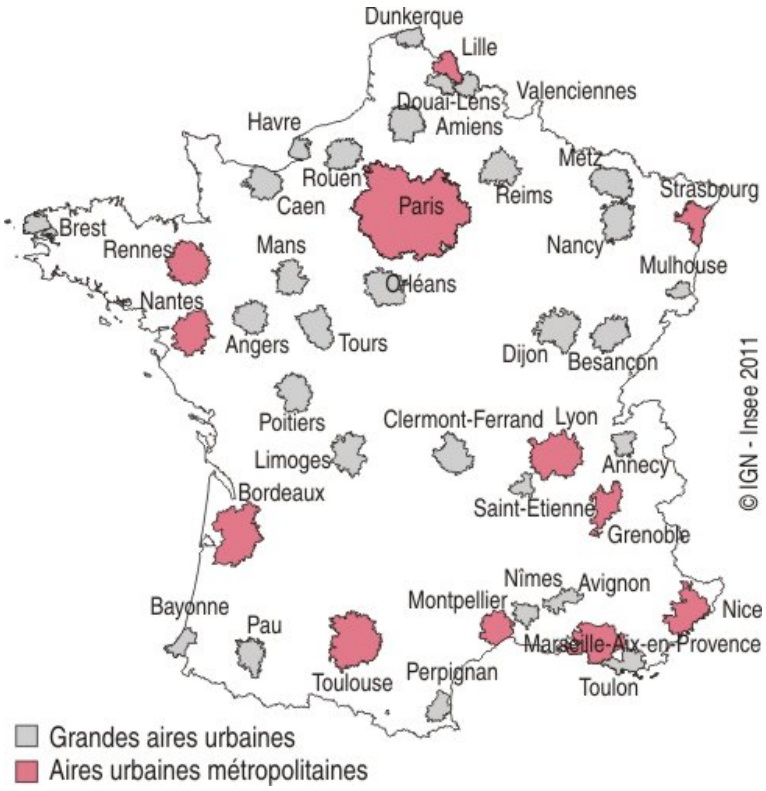



Figure D1. Metropolitan and large urban areas.

# References


Abadie, A. and Imbens, G. W. (2006). Large sample properties of matching estimators for average treatment effects. *Econometrica*, 74(1):235–267.

Abadie, A. and Imbens, G. W. (2016). Matching on the estimated propensity score. *Econometrica*, 84(2):781–807.

Abramowitz, A. I. and Saunders, K. L. (2006). Exploring the bases of partisanship in the American electorate: Social identity vs. ideology. *Political Research Quarterly*, 59(2):175–187.

Ahmed, S. and Greene, K. V. (2000). Is the median voter a clear-cut winner?: Comparing the median voter theory and competing theories in explaining local government spending. *Public Choice*, 105(3–4):207–230.

Alesina, A., Baqir, R., and Easterly, W. (2000). Redistributive public employment. *Journal of Urban Economics*, 48(2):219–241.

Alesina, A. and Drazen, A. (1991). Why are stabilizations delayed? *American Economic Review*, 81(5):1170–1188.

Alesina, A. and Perotti, R. (1995). The political economy of budget deficits. *Staff Papers*, 42(1):1–31.

Alesina, A. and Tabellini, G. (1990). A positive theory of fiscal deficits and government debt. *Review of Economic Studies*, 57(3):403–414.

Andreoni, J. (2007). Giving gifts to groups: How altruism depends on the number of recipients. *Journal of Public Economics*, 91(9):1731–1749.

Arellano, M. (1987). Computing robust standard errors for within group estimators. *Oxford Bulletin of Economics and Statistics*, 49(4):431–434.

Aronsson, T., Lundberg, J., and Wikstrom, M. (2000). The impact of regional public expenditure on the local decision to spend. *Regional Science and Urban Economics*, 30(2):182–202.

Arpaia, A. and Turrini, A. (2008). *Government expenditure and economic growth in the EU: long-run tendencies and short-term adjustment* (No. 300). Directorate General Economic and Financial Affairs (DG ECFIN), European Commission.

Ashworth, J., Geys, B., and Heyndels, B. (2005). Government weakness and local public debt development in Flemish municipalities. *International Tax and Public Finance*, 12(4):395–422.

Ashworth, J., Geys, B., Heyndels, B., and Wille, F. (2014). Competition in the political arena and local government performance. *Applied Economics*, 46(19):2264–2276.

Ashworth, J. and Heyndels, B. (2005). Government fragmentation and budgetary policy in "good" and "bad" times in Flemish municipalities. *Economics and Politics*, 17(2):245–263.

Bäck, H. (2008). Explaining and predicting coalition outcomes: Conclusions from studying data on local coalitions. *European Journal of Political Research*, 42(4):441–472.

Barr, J. and Davis, O. (1966). An elementary political and economic theory of the expenditures of local governments. *Southern Economic Journal*, 33(2):149–165.

Beck, N. (2001). Time-series–cross-section data: What have we learned in the past few years? *Annual Review of Political Science*, 4(1):271–293.

Bénabou, R. (2000). Unequal societies: Income distribution and the social contract. *American Economic Review*, 90(1):96–129.

Bender, B. and Lott, J. R. (1996). Legislator voting and shirking: A critical review of the literature. *Public Choice*, 87(1–2):67–100.

Bergstrom, T. C. and Goodman, R. P. (1973). Private demand for public goods. *American Economic Review*, 63(3):280–296.

Besley, T. and Case, A. (2003). Political institutions and policy choices: evidence from the United States. *Journal of Economic Literature*, 41(1):7–73.

Besley, T. and Coate, S. (1997). An economic model of representative democracy. *Quarterly Journal of Economics*, 112(1):85–114.

Bjørnskov, C. and Potrafke, N. (2012). Political ideology and economic freedom across Canadian provinces. *Eastern Economic Journal*, 38(2):143–166.

Borcherding, T. E. (1997). Why does government's share of national income grow? An assessment of the recent literature on the US experience. In Mueller, D. C., editor, *Perspectives on Public Choice: A Handbook*, pages 562–590. Cambridge: Cambridge University Press.

Borcherding, T. E. and Deacon, R. T. (1972). The demand of services of non-federal governments. *American Economic Review*, 62(5):891–901.

Borge, L.-E. and Rattsø, J. (2002). Spending growth with vertical fiscal imbalance: decentralized government spending in Norway, 1880–1990. *Economics and Politics*, 14(3):351–373.

Borge, L.-E. and Rattsø, J. (2004). Income distribution and tax structure: Empirical test of the Meltzer-Richard hypothesis. *European Economic Review*, 48(4):805–826.

Borrelli, S. A. and Royed, T. J. (1995). Government "strength" and budget deficits in advanced democracies. *European Journal of Political Research*, 28(2):225–260.

Bradbury, J. C. and Crain, W. M. (2001). Legislative organization and government spending: cross-country evidence. *Journal of Public Economics*, 82(3):309–325.

Brennan, G. and Buchanan, J. (1980). *The Power to Tax: Analytic Foundations of a Fiscal Constitution*. Cambridge: Cambridge University Press.

Breunig, R. and Rocaboy, Y. (2008). Per-capita public expenditures and population size: a non-parametric analysis using French data. *Public Choice*, 136(3):429–445.

Brutel, C. (2011). *Un maillage du territoire français* (No. 1333). Insee Première, Institut national de la statistique et des études économiques (INSEE). Retrieved from https://www.insee.fr/fr/statistiques/1280952.

Buchanan, J. M. and Tullock, G. (1962). *The Calculus of Consent: Logical Foundations of Constitutional Democracy*. Ann Arbor: University of Michigan Press.

Caliendo, M. and Kopeinig, S. (2008). Some practical guidance for the implementation of propensity score matching. *Journal of Economic Surveys*, 22(1):31–72.

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

Calvert, R. (1985). Robustness of the multidimensional voting model: Candidate motivations, uncertainty, and convergence. *American Journal of Political Science*, 29(1):69–95.

Campbell, D. E., Green, J. C., and Layman, G. C. (2011). The party faithful: Partisan images, candidate religion, and the electoral impact of party identification. *American Journal of Political Science*, 55(1):42–58.

Caplan, B. (2001). Has Leviathan been bound? A theory of imperfectly constrained government with evidence from the states. *Southern Economic Journal*, 67(4):825–847.

Chanut, J.-C. (2016, January 13). Versement des aides sociales : les départements au bord du gouffre. *La Tribune*. Retrieved from http://www.latribune.fr.

Deacon, R. T. (1979). The expenditure effects of alternative public supply institutions. *Public Choice*, 34(3–4):381–397.

Debus, M. and Gross, M. (2016). Coalition formation at the local level: Institutional constraints, party policy conflict, and office-seeking political parties. *Party Politics*, 22(6):835–846.

Direction des études de la Banque Postale Collectivités Locales (2015). *Regard financier sur les départements : 33 ans depuis les premières lois de décentralisation*. Thematic report, Assemblée des Départements de France et La Banque Postale Collectivités Locales. Retrieved from http://www.departements.fr/wp-content/uploads/2016/10/20150428_regard-financier-departements.pdf.

Dollery, B. and Robotti, L., editors (2008). *The Theory and Practice of Local Government Reform*. Cheltenham: Edward Elgar Publishing.

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

Dubois, E., Leprince, M., and Paty, S. (2007). The effects of politics on local tax setting: evidence from France. *Urban Studies*, 44(8):1603–1618.

Duit, A., Feindt, P. H., and Meadowcroft, J. (2016). Greening Leviathan: the rise of the environmental state? *Environmental Politics*, 25(1):1–23.

Duncombe, W., Miner, J., and Ruggiero, J. (1997). Empirical evaluation of bureaucratic models of inefficiency. *Public Choice*, 93(1–2):1–18.

Duverger, M. (1954). *Political Parties: Their Organisation and Activity in the Modern State*. New York, NY: Wiley.

Ferris, J., Park, S.-B., and Winer, S. (2008). Studying the role of political competition in the evolution of government size over long horizons. *Public Choice*, 137(1):369–401.

Folke, O. (2014). Shades of brown and green: party effects in proportional election systems. *Journal of the European Economic Association*, 12(5):1361–1395.

Foucault, M., Madies, T., and Paty, S. (2008). Public spending interactions and local politics. Empirical evidence from French municipalities. *Public Choice*, 137(1–2):57–80.

Freier, R. and Odendahl, C. (2015). Do parties matter? Estimating the effect of political power in multi-party systems. *European Economic Review*, 80:310–328.

Freitag, M. and Sciarini, P. (2001). The political economy of budget deficits in the European Union: the role of international constraints and domestic structure. *European Union Politics*, 2(2):163–189.

Glasgow, G., Golder, M., and Golder, S. N. (2012). New empirical strategies for the study of parliamentary government formation. *Political Analysis*, 20(2):248–270.

Glasgow, G. and Golder, S. N. (2015). A new approach to the study of parties entering government. *British Journal of Political Science*, 45(4):739–754.

Green, D., Palmquist, B., and Shickler, E. (2002). *Partisan Hearts and Minds: Political Parties and the Social Identities of Voters*. New Haven, CT: Yale University Press.

Guengant, A., Josselin, J.-M., and Rocaboy, Y. (2002). Effects of club size in the provision of public goods. Network and congestion effects in the case of the French municipalities. *Papers in Regional Science*, 81(4):443–460.

Hahm, S. D., Kamlet, M. S., and Mowery, D. C. (1996). The political economy of deficit spending in nine industrialized parliamentary democracies: The role of fiscal institutions. *Comparative Political Studies*, 29(1):52–77.

Hansen, B. E. (1997). Approximate asymptotic P values for structural-change tests. *Journal of Business & Economic Statistics*, 15(1):60–67.

Hansson, Å. and Stuart, C. (2003). Peaking of fiscal sizes of government. *European Journal of Political Economy*, 19(4):669–684.

Hansson, I. and Stuart, C. (1984). Voting competitions with interested politicians: Platforms do not converge to the preference of the median voter. *Public Choice*, 44(3):431–442.

Hayes, K. J., Razzolini, L., and Ross, L. B. (1998). Bureaucratic choice and nonoptimal provision of public goods: theory and evidence. *Public Choice*, 94(1–2):1–20.

Hinich, M. J. (1977). Equilibrium in spatial voting: The median voter result is an artifact. *Journal of Economic Theory*, 16(2):208–219.

Hinich, M. J. (1978). Some evidence on non-voting models in the spatial theory of electoral competition. *Public Choice*, 33(2):83–102.

Hirano, K. and Imbens, G. W. (2004). The propensity score with continuous treatments. In Gelman, A. and Meng, X., editors, *Applied Bayesian Modeling and Causal Inference from Incomplete-Data Perspectives: An Essential Journey with Donald Rubin's Statistical Family*, pages 73–84. Hoboken, NJ: John Wiley & Sons.

Horstmann, I. J. and Scharf, K. A. (2000). *The new federalism: Distributional conflict, voluntarism, and segregation* (No. 2000-11). Research report, Department of Economics, University of Western Ontario.

Hyytinen, A., Meriläinen, J., Saarimaa, T., Toivanen, O., and Tukiainen, J. (2018). Public employees as politicians: evidence from close elections. *American Political Science Review*, 112(1):68–81.

Imbeau, L. M., Pétry, F., and Lamari, M. (2001). Left-right party ideology and government policies. *European Journal of Political Research*, 40(1):1–29.

Imbens, G. W. and Kalyanaraman, K. (2012). Optimal bandwidth choice for the regression discontinuity estimator. *Review of Economic Studies*, 79(3):933–959.

Imbens, G. W. and Wooldridge, J. M. (2009). Recent developments in the econometrics of program evaluation. *Journal of Economic Literature*, 47(1):5–86.

Jackson, J. E. and Kingdon, J. W. (1992). Ideology, interest group scores, and legislative votes. *American Journal of Political Science*, 36(3):805–823.

Jackson, M. O. and Moselle, B. (2002). Coalition and party formation in a legislative voting game. *Journal of Economic Theory*, 103(1):49–87.

Josselin, J.-M. and Le Maux, B. (2017). *Statistical Tools for Program Evaluation: Methods and Applications to Economic Policy, Public Health, and Education*. Springer International Publishing AG 2017.

Kinder, D. R. and Kiewiet, D. R. (1979). Economic discontent and political behavior: The role of personal grievances and collective economic judgments in congressional voting. *American Journal of Political Science*, 23(3):495–527.

Kolluri, B., Panik, M., and Wahab, M. (2000). Government expenditure and economic growth: evidence from G7 countries. *Applied Economics*, 32(8):1059–1068.

Kreps, D. M. and Wilson, R. (1982). Reputation and imperfect information. *Journal of Economic Theory*, 27(2):253–279.

Laver, M. and Benoit, K. (2015). The basic arithmetic of legislative decisions. *American Journal of Political Science*, 59(2):275–291.

Le Maux, B., Dostálová, K., and Padovano, F. (2017). *Ideology and public policies: A quasi-experimental test of the hypothesis that left-wing governments spend more* (No. 2017-01-ccr). Economics Working Papers, Condorcet Center for Political Economy.

Le Maux, B. and Rocaboy, Y. (2016). Competition in fragmentation among political coalitions: theory and evidence. *Public Choice*, 167(1–2):67–94.

Le Maux, B., Rocaboy, Y., and Goodspeed, T. (2011). Political fragmentation, party ideology and public expenditures. *Public Choice*, 147(1–2):43–67.

Le Maux, B. and Zhang, W. (2013). Does incrementalism stem from political fragmentation? An empirical test. *Papers in Regional Science*, 92(3):535–553.

Levy, G. (2004). A model of political parties. *Journal of Economic Theory*, 115(2):250–277.

Lewis-Beck, M. S. and Stegmaier, M. (2000). Economic determinants of political outcomes. *Review of Political Science*, 3(1):183–219.

Lijphart, A. (1990). The political consequences of electoral laws, 1945–85. *American Journal of Political Science*, 84(2):481–496.

Lindbeck, A. and Weibull, J. (1987). Balanced-budget redistribution as the outcome of political competition. *Public Choice*, 52(3):273–297.

Lindbeck, A. and Weibull, J. (1993). A model of political equilibrium in a representative democracy. *Journal of Public Economics*, 51(2):195–209.

Lundqvist, L. J. (2001). A green fist in a velvet glove: The ecological state and sustainable development. *Environmental Values*, 10(4):455–472.

Magazzino, C. (2012). The nexus between disaggregated public spending and GDP in the Euro area. *Economics Bulletin*, 32(3):2560–2579.

Marlow, M. L. (1988). Fiscal decentralization and government size. *Public Choice*, 56(3):259–269.

Martin, L. W. and Stevenson, R. T. (2001). Government formation in parliamentary democracies. *American Journal of Political Science*, 45(1):33–50.

Martin, L. W. and Stevenson, R. T. (2010). The conditional impact of incumbency on government formation. *American Political Science Review*, 104(3):503–518.

Matsusaka, J. G. (2010). Popular control of public policy: A quantitative approach. *Quarterly Journal of Political Science*, 5(2):133–167.

Mattos, E. and Rocha, F. (2008). Inequality and size of government: evidence from Brazilian states. *Journal of Economic Studies*, 35(4):333–351.

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

Mehay, S. L. (1984). The effect of governmental structure on special district expenditures. *Public Choice*, 44(2):339–348.

Mehay, S. L. and Gonzales, R. A. (1985). Economic incentives under contract supply of local government services. *Public Choice*, 46(1):79–86.

Meltzer, A. H. and Richard, S. F. (1981). A rational theory of the size of government. *Journal of Political Economy*, 89(5):914–927.

Meriläinen, J. (2018). Single-party rule, public spending and political rents: Evidence from Finnish municipalities. *Scandinavian Journal of Economics*. Advance online publication. doi.org/10.1111/sjoe.12288.

Milanovic, B. (2000). The median voter hypothesis, income inequality, and income redistribution: an empirical test with the required data. *European Journal of Political Economy*, 16(3):367–410.

Milesi-Ferretti, G. M., Perotti, R., and Rostagno, M. (2002). Electoral systems and public spending. *Quarterly Journal of Economics*, 117(2):609–657.

Mohl, P. and Pamp, O. (2009). Income inequality & redistributional spending: An empirical investigation of competing theories. *Public Finance & Management*, 9(2):1–56.

Moisio, A., Loikkanen, H., and Oulasvirta, L. (2010). *Public services at the local level–The Finnish way* (No. 2). VATT Policy Reports, VATT Institute for Economic Research.

Mueller, D. C. (2003). *Public Choice III*. Cambridge: Cambridge University Press.

Nelson, M. A. (1986). An empirical analysis of state and local tax structure in the context of the Leviathan model of government. *Public Choice*, 49(3):283–294.

Niskanen, W. A. (1971). *Bureaucracy and Representative Government*. New Brunswick, NJ: AldineTransaction.

Oates, W. E. (1985). Searching for Leviathan: An empirical study. *American Economic Review*, 75(4):748–757.

OFGL (1996–). Rapports de l'Observatoire des finances et de la gestion publique locales (OFGL). In *Collectivités locales | Le portail de l'Etat au service des collectivité*. Retrieved from https://www.collectivites-locales.gouv.fr/rapports-lobservatoire-des-finances-et-gestion-publique-locales-ofgl.

Olson, M. (1993). Dictatorship, democracy, and development. *American Political Science Review*, 87(3):567–576.

Ordeshook, P. and Shvetsova, O. (1994). Ethnic heterogeneity, district magnitude, and the number of parties. *American Journal of Political Science*, 38(1):100–123.

Osborne, M. J. and Slivinski, A. (1996). A model of political competition with citizen candidates. *Quarterly Journal of Economics*, 111(1):65–96.

Osterloh, S. and Debus, M. (2012). Partisan politics in corporate taxation. *European Journal of Political Economy*, 28(2):192–207.

Padovano, F. (2013). Are we witnessing a paradigm shift in the analysis of political competition? *Public Choice*, 156(3–4):631–651.

Padovano, F. and Petrarca, I. (2014). Are the responsibility and yardstick competition hypotheses mutually consistent? *European Journal of Political Economy*, 34:459–477.

Padovano, F. and Ricciuti, R. (2009). Political competition and economic performance: evidence from the Italian regions. *Public Choice*, 138(3–4):263–277.

Padovano, F. and Venturi, L. (2001). Wars of attrition in Italian government coalitions and fiscal performance: 1984–1994. *Public Choice*, 109(3–4):15–54.

Perotti, R. and Kontopoulos, Y. (2002). Fragmented fiscal policy. *Journal of Public Economics*, 86(2):191–222.

Persson, M. (1995). Why are taxes so high in egalitarian societies? *Scandinavian Journal of Economics*, 97(4):569–580.

Persson, T. and Svensson, L. (1989). Why a stubborn conservative would run a deficit: Policy with time-inconsistent preferences. *Quarterly Journal of Economics*, 104(2):325–345.

Persson, T. and Tabellini, G. (2000). *Political Economics: Explaining Economic Policy*. Cambridge, MA: MIT Press.

Pettersson-Lidbom, P. (2008). Do parties matter for economic outcomes? A regression-discontinuity approach. *Journal of the European Economic Association*, 6(5):1037–1056.

Pickering, A. and Rockey, J. (2011). Ideology and the growth of government. *Review of Economics and Statistics*, 93(3):907–919.

Pickering, A. and Rockey, J. (2013). Ideology and the size of US state government. *Public Choice*, 156(3–4):443–465.

Poingt, G. (2017, November 10). Pourquoi la Seine-Saint-Denis refuse de verser le RSA de décembre. *Le Figaro*. Retrieved from http://www.lefigaro.fr.

Potrafke, N. (2018). Government ideology and economic policy-making in the United States—a survey. *Public Choice*, 174(1–2):145–207.

Pradhan, R. P. and Bagchi, T. P. (2012). The expenditure-GDP nexus: evidence from a panel of SAARC 7-countries. *International Journal of Public Policy*, 8(4–6):295–307.

Rattsø, J. and Tovmo, P. (2002). Fiscal discipline and asymmetric adjustment of revenues and expenditures: Local government responses to shocks in Denmark. *Public Finance Review*, 30(3):208–234.

Roberts, K. (1977). Voting over income tax schedules. *Journal of Public Economics*, 8(3):329–340.

Romer, T. (1975). Individual welfare, majority voting, and the properties of a linear income tax. *Journal of Public Economics*, 4(2):163–185.

Rosenbaum, P. R. (2002). *Observational Studies*. New York, NY: Springer, 2nd edition.

Rosenbaum, P. R. and Rubin, D. B. (1983). The central role of the propensity score in observational studies for causal effects. *Biometrika*, 70(1):41–55.

Rosenbaum, P. R. and Rubin, D. B. (1985). Constructing a control group using multivariate matched sampling methods that incorporate the propensity score. *The American Statistician*, 39(1):33–38.

Roubini, N. and Sachs, J. D. (1989a). Government spending and budget deficits in the industrial countries. *Economic Policy*, 4(8):99–132.

Roubini, N. and Sachs, J. D. (1989b). Political and economic determinants of budget deficits in the industrial democracies. *European Economic Review*, 33(5):903–933.

Roucous, D. (2018, February 5). Aide sociale départementale, c'est 64% du budget des départements. *L'Humanité*. Retrieved from http://www.humanite.fr.

Saarimaa, T. and Tukiainen, J. (2015). Common pool problems in voluntary municipal mergers. *European Journal of Political Economy*, 38:140–152.

Santolini, R. (2008). A spatial cross-sectional analysis of political trends in Italian municipalities. *Papers in Regional Science*, 87(3):431–451.

Schneider, M. (1986). Fragmentation and the growth of local government. *Public Choice*, 48(3):255–263.

Shapiro, P. and Sonstelie, J. (1982). Did proposition 13 slay Leviathan? *American Economic Review*, 72(2):184–190.

Sigelman, L. and Buell, E. H. (2004). Avoidance or engagement? Issue convergence in presidential campaigns. *American Journal of Political Science*, 48(4):650–661.

Skjæveland, A., Serritzlew, S., and Blom-hansen, J. (2007). Theories of coalition formation: An empirical test using data from Danish local government. *European Journal of Political Research*, 46(5):721–745.

Solé-Ollé, A. (2003). Electoral accountability and tax mimicking: the effects of electoral margins, coalition government, and ideology. *European Journal of Political Economy*, 19(4):685–713.

Solé-Ollé, A. (2006). The effects of party competition on budget outcomes: Empirical evidence from local governments in Spain. *Public Choice*, 126(1):145–176.

Stadelmann, D., Marco, P., and Eichenberger, R. (2013). Quantifying parliamentary representation of constituents' preferences with quasi-experimental data. *Journal of Comparative Economics*, 41(1):170–180.

Stigler, G. (1972). Economic competition and political competition. *Public Choice*, 13(1):91–106.

Tabellini, G. and Alesina, A. (1990). Voting on budget deficit. *American Economic Review*, 80(1):37–49.

Tridimas, G. and Winer, S. L. (2005). The political economy of government size. *European Journal of Political Economy*, 21(3):643–666.

Tullock, G. (1959). Some problems of majority voting. *Journal of Political Economy*, 67(6):571–579.

Velasco, A. (2000). Debts and deficits with fragmented fiscal policymaking. *Journal of Public Economics*, 76(1):105–125.

Vives, A. (2018, October 15). Val-de-Marne : toujours plus d'allocataires du RSA. *Le Parisien*. Retrieved from http://www.leparisien.fr.

Volkerink, B. and de Haan, J. (2001). Fragmented government effects on fiscal policy: New evidence. *Public Choice*, 109(3–4):221–242.

Wahab, M. (2004). Economic growth and government expenditure: evidence from a new test specification. *Applied Economics*, 36(19):2125–2135.

Weatherford, M. S. (1978). Economic conditions and electoral outcomes: Class differences in the political response to recession. *American Journal of Political Science*, 22(4):917–938.

Weingast, B. R., Shepsle, K. A., and Johnsen, C. (1981). The political economy of benefits and costs: A neoclassical approach to distributive politics. *Journal of Political Economy*, 89(4):642–664.

Wildavsky, A. D. (1986). *Budgeting : A Comparative Theory of the Budgetary Process*. New Brunswick, NJ: Transaction Publishers.

Wittman, D. (1973). Parties as utility maximizers. *American Political Science Review*, 67(2):490–498.

Wright, R. (1986). The redistributive roles of unemployment insurance and the dynamics of voting. *Journal of Public Economics*, 31(3):377–399.

Zeileis, A., Kleiber, C., Krämer, W., and Hornik, K. (2003). Testing and dating of structural changes in practice. *Computational Statistics & Data Analysis*, 44(1–2):109–123.

# List of Figures


1.1 Distribution of voters and public good provision. . . . . 22

1.2 Cross-sectional comparison – 2008. . . . . 27

1.3 Vote function: simulating the probability of being on the left. . . . . 31

1.4 Spending and inequalities: simulations of an average Department (for graphical convenience, confidence intervals are not drawn). . . . . 34

1.5 Propensity score matching: ATE (in black) and *p*-value (in gray) as a function of caliper size. . . . . 43

1.6 Comparison of observations taken into account by RDD and PSM. . . . . 45

2.1 Changes in the structure of Finnish municipalities. . . . . 58

2.2 Distribution of seats after the 2000 and 2004 elections. . . . . 61

2.3 Comparison of electoral results and estimated propensity scores. . . . . 64

2.4 Visual analysis of propensity score density distribution in both groups. . . . 69

2.5 Propensity score matching: ATE (in black) and *p*-value (in gray) as a function of caliper size. . . . . 70

2.6 Tests for endogenous thresholds (1% critical level given by horizontal red line). 74

2.7 Recursive regression discontinuity design. . . . . 79

3.1 Differences in social aid per beneficiary – 2008. . . . . 89

3.2 Impact of population size on spending. . . . . 91

3.3 Impact of population size on demand structure. . . . . 92

3.4 Score variation. . . . . 98

3.5 Simulating size effects in the production effect scores. . . . . 100

3.6 French Departments in 2008. . . . . 104

3.7 Standardized mean differences between high and low density areas. . . . . 105

3.8 Intergovernmental grant support received by the Departments. . . . . 109

B1 Correlation between the share of welfare-dependent individuals and the share of seats held by centrist and extreme parties . . . . . 128

D1 Metropolitan and large urban areas. . . . . 140

# List of Tables


1.1 Summary statistics . . . . . 26

1.2 Geographical dummies . . . . . 29

1.3 Vote function – estimation results . . . . . 30

1.4 Spending equation – estimation results . . . . . 33

1.5 Regression discontinuity design – estimation results: LATE . . . . . 37

1.6 Propensity score matching – results of matching algorithms: ATE . . . . . 39

1.7 RDD quality indicators – difference in exogenous variables: two-sample $t$-tests 41

1.8 Matching quality indicators – standardized bias (SB) in percentage . . . . . 41

1.9 Matching quality indicators – difference in exogenous variables: two-sample $t$-tests . . . . . 42

1.10 PSM – cross-sectional analysis and robustness of the results with respect to election years . . . . . 44

2.1 Summary statistics . . . . . 60

2.2 Evidence of a selections bias – two-sample $t$-tests . . . . . 62

2.3 Propensity score estimation – single-party majority . . . . . 67

2.4 Results of matching algorithms – ATE for *MAJ* . . . . . 70

2.5 Impact of political competition . . . . . 71

2.6 RDD estimates – seat share thresholds . . . . . 75

2.7 Results of matching algorithms – ATE for *REG* . . . . . 76

2.8 RDD estimates – seat share thresholds for KESK and RKP . . . . . 78

3.1 Summary statistics . . . . . 88

3.2 Ranking 1998–2008: Total, demand and production effects . . . . . 96

3.3 Production effect (PE) scores – estimation results . . . . . 101

3.4 Production effect (PE) scores – urbanization variables . . . . . 106

B1 Political parties on the left, center, and right . . . . . 127

B2 Correlation with the first and second axes . . . . . 127

C1 Matching quality indicators – standardized bias (SB) in percentage . . . . . 132

C2 Matching quality indicators – difference in exogenous variables: two-sample $t$-tests . . . . . 133

C3 RDD quality indicators – difference in exogenous variables for *MAJ* and *REG*: two-sample $t$-tests . . . . . 134

C4 RDD quality indicators – difference in exogenous variables for KESK and RKP: two-sample $t$-tests . . . . . 135

D1 Urban areas in France according to INSEE . . . . . 139

UNIVERSITE BRETAGNE ECONOMIE LOIRE ET GESTION

UNIVERSITÉ DE RENNES 1

Titre : Etude de la demande et de l'offre de biens publics locaux : une approche quasi-expérimentale

**Mots clés :** biens publics ; gouvernements locaux ; effets partisans ; fragmentation politique ; quasi-expérimentation

**Résumé :** L'objectif de cette thèse de doctorat est d'étudier l'adéquation entre l'offre de biens publics locaux et la demande des citoyens. D'un point de vue théorique, les modèles analysant la demande de service public, tel le modèle de l'électeur médian ou l'hypothèse de Meltzer et Richard, démontrent que les caractéristiques de l'électorat ainsi que ses préférences sont les principaux déterminants des choix publics. Les modèles analysant l'offre de service public concluent au contraire que c'est la composition du gouvernement (nombre de partis, idéologie des élus) qui joue un rôle décisif. Cette recherche a pour objet de faire le lien entre ces deux catégories d'explications, à travers une série d'essais empiriques utilisant une approche quasi-expérimentale, et se basant sur une analyse du secteur public local français et finlandais. L'analyse des dépenses de fonctionnement des départements français montre que les gouvernements ayant une majorité de sièges à gauche ne dépensent pas plus en aide sociale que leurs homologues à droite. Les niveaux de dépense dépendent au contraire du nombre d'usagers et de leur répartition entre les quatre risques que sont le chômage, l'aide à la famille, l'aide aux personnes handicapées et aux personnes âgées. Sur données finlandaises, l'approche quasi-expérimentale montre en moyenne qu'il n'y a pas de différence significative entre les dépenses des gouvernements majoritaires et celles des gouvernements minoritaires. La demande des citoyens semble donc jouer un rôle prédominant dans l'explication des niveaux de dépenses publiques locales pour ces deux pays.

Title: On the demand and supply of local public goods: a quasi-experimental approach

**Keywords:** public goods; local governments; partisan effects; political fragmentation; quasi-experiments

**Abstract:** The aim of the PhD thesis is to study the demand and supply sides of local public good provision. Demand driven models, such as the median voter theorem or the Meltzer and Richard hypothesis, state that the characteristics and preferences of the electorate are the main forces driving public decisions. Models focusing on the supply side argue instead that the composition of governing bodies (e.g., number of parties, ideology of elected officials) is the key explanatory factor. The thesis seeks to shed light on the connection between the two categories of explanations, through a series of empirical essays, using a quasi-experimental approach to study the local public sector in France and Finland. The analysis of operating expenditures of the French Departments shows that left-wing governments do not spend more on social assistance than their right-wing counterparts. Spending levels depend instead on the number of welfare recipients and their type: children, unemployed, disabled and elderly people. For Finnish municipalities, the quasi-experimental approach indicates that there is no significant difference between the expenditures of majority governments and those of minority governments. The demand side, i.e., the electorate, thus seems to play an important role in explaining policy choices in the two countries.