Essays on decentralisation and accountability in Indonesia

Ruth Nikijuluw

A thesis submitted for the degree of Doctor of Philosophy
at the Australian National University

May 2021

Declaration of originality

This thesis was written while I was studying at the Australian National University. The entire thesis is my own work, unless otherwise acknowledged in the text or Acknowledgements.

Ruth Nikijuluw

May 2021

Acknowledgements

First of all, I would like to thank my supervisor and Chair of Panel, Dr. Blane Lewis, for his invaluable guidance. Pak Blane's vast command of literature and quantitative knowledge have helped me enormously since the topic selection. He has been and will ever be my inspiration and to him I am deeply indebted. I also express my utmost gratitude to my other two panel members, Professor Budy Resosudarmo and Dr. Sarah Dong, for their resourceful inputs and insights for all three papers.

I have been fortunate to do my PhD at the Arndt-Corden Department of Economics, Crawford School of Public Policy. I have benefited greatly from the constructive interactions and encouragement from all the ACDE and Crawford family, especially during the internal seminar. To Megan Poore, Tracy McRae, and Heeok Kyung, thank you for the continuing kind support.

This PhD study would not have been achievable without tremendous inputs from Dr. Adrianus Hendrawan on local election data collection. I also thank Karin Hosking who provided professional copyediting assistance. I gratefully acknowledge financial support from the Indonesia Endowment Fund for Education (LPDP) as part of my doctorate scholarship.

Finally, I want to thank my beloved family who have been my number one supporter on this journey. To Papa and Mama, for their endless prayers which have given me strength to move forward each day even when I doubted myself. To Oma Yo, for always been such a special encouragement to me. To Amang and Inang, for understanding my struggles and supporting me relentlessly. And last but not least, to Joseph, my muse and my better half. Thank you for allowing me to talk about my research even at the dinner table. I am beyond grateful to have you beside me. One down and one more to go.

"The night has been won and I shall overcome. Yet not I, but through Christ in me.

Ad maiorem Dei gloriam."

Abstract

The first paper examines the degree of spatial interdependence in fiscal policy and service provision by local governments in Indonesia. The research forms the first attempt to test such interdependence across local areas of a developing country. A spatial Durbin model is estimated for a balanced panel of Indonesian districts for 2009–2013. The findings confirm the existence of positive spatial interdependence in both local government expenditure and access to services. Interdependence is stronger for capital-based spending and education access. The findings point to strategic competition between districts, a process that potentially boosts local government accountability and efficiency in providing service access for citizens.

The second paper contrasts two distinctive government forms at the village level in Indonesia: *desa*, which is a more citizen-based government with directly elected village heads and *kelurahan*, which is a bureaucratic-based government with appointed village heads. The paper investigates if switching to the more bureaucratic type of governance influences service access outcomes. Both village-level panel Difference in Difference (DID) for village level outcomes and pseudo-village panel DID for individual outcomes are pursued. The results show that *kelurahan* status leads to an increase in property crime incidence and reduced access to safe birth and immunisation. These findings point to the importance of horizontal accountability and village apparatus engagement, while also providing insights into the debate on whether political decentralisation should be extended to the lowest administrative tier.

The last paper investigates the consequences of electing a female leader on fiscal and service outcomes. A dataset of districts in Indonesia with close election results between male and female mayoral candidates from 2005 to 2017 is used. This study employs a randomisation-based inference in regression discontinuity design to deal with non-random assignment of female leadership. The results show that per capita expenditure, as well as budget share on social protection and infrastructure, are higher for districts that are governed by female leaders. Also, female leadership improves citizen access to both assisted birth by health professionals and safe water. Finally, the findings show that female-led districts have more prudent budget management while neither fewer nor more corruption cases. This study suggests that female leader policy preferences are different from those of their male counterparts, and that gender plays a role in determining policy choices.

Contents

List of Figures

List of Tables

1 INTRODUCTION

1.1 Decentralisation in Indonesia

In 2001, Indonesia embarked on a big bang decentralisation program, as part of the political response to a democratic transition triggered by the end of the New Order government in 1998. The initial era of decentralisation began through the implementation of Law 22/1999 on Regional Government, Law 25/1999 on the Fiscal Balance between the Central Government and the Regions, and Law 34/2000 on Regional Taxes and Levies. The decentralisation reforms grant local governments new power and responsibilities in three dimensions: political, administrative, and fiscal (Shah, 2006; Smoke and Lewis, 1996).

In terms of administration, the central government assigns some public service responsibilities to subnational governments. District government in Indonesia has the responsibility to deliver basic public services (health, infrastructure, and education), together with other tasks where districts do not have huge budgetary influence or do not play such a large role, including agriculture, communications, industry-trade, environment, land, cooperatives and labour. This shifting of functions is followed by decentralisation on the finance side to enable regions to perform their roles and responsibilities. Central government transfers some portions of its fiscal resources and expands the autonomy of regions in collecting local taxes. Assignment of services was also complemented by political decentralisation, indicated by the development of local democracy. In June 1999, the first free and fair elections were held to simultaneously elect the members of national, provincial and district parliaments. Also, the first batch of local heads of districts direct elections took place in 2005, and gradually all indirectly elected heads were replaced by directly elected ones by the end of 2010 (Lewis, 2018; Skoufias et al., 2014).

One objective of decentralisation is to bring government closer to its citizens and to improve people's access to quality public services. Subnational government in Indonesia plays a key role in delivering basic services through managing and implementing more than half of all levels' (central-province-district) government expenditure, excluding subsidy and interest payments. Sector-wise, district governments also spend more than 50 percent of infrastructure, health, and education sector national expenditure. In contrast to the expenditure side, the majority of revenue for local government budgets still comes from central government intergovernmental transfers, as major taxes are not yet decentralised to local

levels and nearly 90 percent of total government revenue is still collected by the central government.

1.2 Decentralisation and accountability

According to Bovens et al. (2014), accountability is a relational concept between government representatives or politicians who perform tasks, and their constituents who are affected by the tasks performed, and therefore make good inferences about the agents' performance. The focus of many studies on accountability is to discuss whether and how agents are or can be held to account by accountability forums, such as citizens. In the context of decentralisation, as stated by Yilmaz, Beris, and Serrano-Berthet (2010), ensuring appropriate use of such discretion in fiscal, administrative and political domains in delivering service delivery requires effective accountability systems. Their framework suggests that a good system would guarantee that citizens have the ability and opportunity to demand accountability, while local governments have means and incentives to respond to citizen demands for accountability and better service delivery. The underlying common topic of the three essays in this thesis is which aspects can influence accountability of local governments in allocating fiscal resources and delivering service access. Each paper, however, has its own theme and is different from the others.

The first paper focuses on the neighbourhood effects of local government. Given that local governments are rational agents, they will consider the policy and activities of other local governments when making their own decisions. One framework to explain this is the spillover model. A region's expenditures depend on the spending of other similarly situated regions because the provision of public goods in one region may also be enjoyed by neighbouring non-residents, in addition to the constituents (Case et al., 1993). Another potential channel of spillover is through information spillover. Information on the fiscal policy of neighbouring regions spills and citizens in one region will use the information on what the neighbouring regions do in evaluating their government's action. This creates yardstick competition amongst neighbouring regions and therefore local governments would consider other regions' policy as their response to electoral competition (Besley and Case, 1995). Identifying the fiscal and service access interdependency across districts can inform whether regional competition exists and how it affects local government accountability and effectiveness of decentralisation.

The second paper discuss two forces of accountability that local governments have in exercising their discretionary space. Local governments are accountable to higher levels of government as well as to their citizens. To contrast both types of accountability, this paper contrasts two local government forms: citizen-based governments with elected heads versus a bureaucratic-based government with appointed leaders. It draws on the proposition that citizens can hold their service providers accountable and therefore directly elected officials are more responsive to citizens' needs (Besley and Coate, 2003). However, this may come at the expense of a lack of administrative and operational skills since the elected officials will devote their time more to electoral concerns than to the norms of professional management (Carr, 2015). Contrasting these two government forms would indicate whether increased downward accountability because of direct elections may influence service access outcomes.

The last paper discusses one important feature of a leader: gender. Gender is one of the factors that shape differences in behaviour and preferences between men and women (Croson and Gneezy, 2009). As a result, female and male leaders are found to be effective in dissimilar settings and to seek different leadership styles (Eagly et al., 1995; Eagly and Carli, 2003). However, according to Anthony Downs in 1957, as summarised by Clots-Figueras (2012), preferences of politicians should not impact policy outcomes since fully accountable politicians will only care about the preferences of the median voter group. This view was later challenged by the citizen-candidate model that suggests elected candidates care about certain outcomes more than others and they will implement policies that are more closely correlated with their preferences (Besley and Coate, 1997). This paper will test these two contrasting frameworks by examining whether having female leaders in the office affects how fiscal resources are allocated and service access is delivered.

1.3 Specific research areas

Indonesia's context serves as a useful study case for all three papers. Although Indonesia's decentralisation has been widely discussed in many empirical works, there is still room to draw more empirical evidence, given that different contexts emerge from each paper compared to the existing literature.

The first paper will investigate the presence of fiscal policy interdependence among district governments in Indonesia between 2009 and 2013. By taking this time frame, the paper looks at a context post political decentralisation which began in 2005 through direct election of

district heads (Lewis, 2019a; Skoufias et al., 2014). In addition to expenditure interdependence, I will also explore the interdependence of local service access, which has not been discussed in prior literatures about Indonesia. The rationale of extending the outcome to include service access is that voters are likely to assess the level of public goods provided, rather than solely the expenditure level. I will control the plausible interaction on local public spending between neighbours when testing the service access interdependence between districts.

The second paper takes the context of village level government in Indonesia, which has two distinct forms: *desa* and *kelurahan. Desa* village is a citizen-based government with directly elected village heads and salaries for the village apparatus are funded from the village budget. The heads of *desa* villages are responsible to the village population and must submit an annual accountability report, which village council can contest (Antlöv, 2003). *Kelurahan*, on the other hand, is more bureaucratic, with appointed village heads and all village apparatus being civil servants. In this paper, I compare local service access performance between *desa* villages which switch to *kelurahan* form and *desa* villages which retain their status as *desa*, from 2001 to 2011. During this period of study, all *desa* village heads had generally been elected in free and fair elections post the Soeharto era (Olken, 2010). The period of the study does not cover the implementation period of the new Village Law in 2014 which mandates that each *desa* village receive a large influx of village funding. Before the implementation of the new Law, there were no substantial differences in responsibilities between *desa* and *kelurahan*, and the size of village budgets was considerably smaller (Antlöv et al., 2016). By looking at village level government, this paper focuses on the lowest administrative tier of government in Indonesia where the trade-off between autonomy and coordination is more apparent.

The last paper takes the context of district local elections, which elected female heads and female vice-heads, between 2005 and 2017. Direct elections offer structural opportunities for women to participate in politics since candidates can directly engage with voters without facing any institutional barriers (Dewi, 2015). In investigating the impact of female leadership on fiscal and service outcomes, I focus on district level leadership, since district government in Indonesia is responsible for the provision of basic services. I further define leaders as heads and vice-heads only, and not legislative representation, given their power in facilitating the reallocation of resources in a district. Since candidates run as pairs (head and vice-head) in

Indonesian elections, I analyse the election results in a pair setting. The paper contrasts service performance by districts that at least have one female leader in the office, either as head or vice-head, to districts that are governed by all male leaders.

1.4 Methodology and data challenges

Estimating interdependence between districts requires definition of neighbours to be previously specified. In the case for the first paper, I use geographical criteria to define neighbours as under the spillover model, the neighbourhood effects will presumably be larger for close-by jurisdictions. I implement a spatial panel model in investigating the spatial interdependency and also use the maximum-likelihood (ML) approach to account for the recursive nature of the estimation. I run the empirical analyses for the time frame of 2009 to 2013 since during this period the number of districts was fairly constant, and therefore accommodates the balanced-panel data requirement for the spatial model. Given this constraint, I also need to exclude districts that fail to report some budget and service outcomes. Therefore, I can only use 306 districts for the spending interdependence model and 436 districts for the service interdependence model. To account for the large number of districts which needed to be excluded due to missing observations, I will apply an imputation technique to the sample and compare the estimates between actual and imputed samples as robustness check.

For the second paper, I implement two specifications: village panel Difference in Difference (DID) for the village level outcomes, and pseudo-village DID for the individual level outcomes. The DID technique is suitable in this case since once a village switches its status from *desa* to *kelurahan,* the status will remain as *kelurahan* until the end of the observation period. The empirical design relies on the critical assumption that without the treatment of switching status to *kelurahan,* the time path of local service delivery of *desa* villages that switched their status would be parallel to the time path of *desa* villages that retain their status. As my observation unit, I use all *desa* villages in 2000 and then track further whether and when those villages turn to *kelurahan.* I will also present the estimated coefficients from a more restricted sample by only using villages which eventually switch into *kelurahan* during the study period. By doing this, I basically perform a ‘staggered DID’ method by utilising only ‘treated’ villages and exploit the different timings of the village status switching (Stevenson and Wolfers, 2006).

The main data used in the second paper are from five different waves of the village census, Indonesian Village Potential Statistics (PODES). The main challenge when constructing this dataset is incoherent village identifiers across waves, due to the introduction of new villages and village-splitting. Given the complexity of matching villages across waves, I only make use of villages that were consistently included in all waves, and therefore some villages were excluded from the dataset. Furthermore, for the individual level outcomes, I merged the village census data with the annual household survey data (SUSENAS) one year after the village census year.

For the last paper, estimating the effect of female leaders is not trivial since female leadership itself is not randomly distributed across districts. There could be unobserved district characteristics that are likely to affect not only electoral outcomes of female candidates, but also policy outcomes. Therefore, I identify the causal effect of a female leader by comparing districts where a pair with female candidates win the election by small margin against a pair with all male candidates, with other districts where a pair with female candidates results in second place and a pair with all male candidates win the election. I focus only on districts with such close elections, and estimate the effect using the Randomisation-based Regression Discontinuity (RD) technique (Cattaneo et al., 2016, 2015). This alternative RD procedure is more appropriate for smaller sized samples and is able to account for the limited number of districts with local elections involving women candidates (Lewis, 2019b).

The main treatment variable for the third paper is derived from the winning margin of pairs with female candidates who run local elections. This vote margin data comes from the district first-round elections data provided by the General Elections Commission (KPU) and collected by Lewis et al. (2020). However, this data does not have further information on candidates' individual characteristics and therefore to identify the gender of the candidates, I did a manual entry of that information based on desk and news reviews.

1.5 Potential contributions

There are two frameworks that make the sign and magnitude of the interaction between regions positive for several spending and service categories, but negative for others. The first is the spending spillover framework, which points at externalities that can emerge from expenditure in one region to affect people in neighbouring regions. Because of this spending externality, optimal government expenditure in each region will also depend on expenditure

policies in other regions. For instance, local governments may choose not to provide the level of service delivery demanded by its residents when they see that residents can consume and use the services delivered in other neighbouring districts. Another framework that can also explain the interaction among districts on spending and service delivery is information spillover. Information about policies in neighbouring regions spills and citizens use this information as a yardstick when evaluating the performance of their local government.

The first paper is one of the few studies that has discussed spending and service access interdependence in the case of a developing country. Many studies have identified the existence of fiscal policy interdependence in developed countries (e.g. Case, et al., 1993; Geys, 2006; Solé-Ollé, 2006; Revelli, 2005), but only a few have discussed the case of developing countries. Most discussion on developing countries has focused on the effect of spending and intergovernmental transfers on service delivery, with only limited discussion of fiscal policy interaction between regions. Also, with the exercise of looking at service access interdependence, this paper is one of the first attempts to test such interaction and it contributes to the dialogue of whether spatial interaction is one of the influential factors for local governments in providing service delivery. Lastly, the presence of interdependence can indicate some forms of regional competition which will affect local government accountability. This competition factor, furthermore, may contribute to discussion on fostering additional incentives for local government to effectively deliver service access.

On local government form, a large body of literature has contrasted various government forms, but only a small number of studies have examined this distinction at the lowest administrative tier. In Indonesia's case, some studies have discussed government forms at district levels by comparing directly and indirectly elected district heads (Lewis, 2018; Skoufias et al., 2014). The second paper will be one of the first attempts to test whether there are differences in service outcomes between two village level government forms in the case of a developing country. Village governments, as the lowest administrative tier, are crucial front-liners in service provision since they are the closest government level to citizens. The results of this second paper can contribute to discussion on which types of accountability matter more, downward or upward accountability, for government performance at village level. This paper also gives perspectives on whether service delivery should be decentralised down to the lowest possible government unit.

On analysing the effect of female leadership, the literature has noted two contrasting frameworks on this. The first assumes that political candidates have complete commitment to implement policies of their median voter and therefore policy choices will be the same for all leaders, irrespective of their gender (Anthony Downs in 1957 in Clots-Figueras, 2012). The second, in contrast, assumes that there is a divergence in policies among leaders since candidates care about certain outcomes more than others (Besley and Coate, 1997). The gender of the leaders will affect policy choices since it is one of the factors that influence candidates' preferences. Given these differing frameworks, whether female leaders will have an impact on policy outcomes remains to be tested empirically.

The third paper will contribute to the existing literature in two ways. First, it will provide a causal estimate of female leadership due to electoral outcomes rather than specific gender quota policies. Indonesia, although it adopted a quota policy for female representation in legislative nomination, does not have any specific quota policy for heads and vice-heads of districts. This paper will exploit election results to establish random assignment of female leadership among districts. By looking at elected leaders and arguing for causality, this paper differs from a previous study that observes the association between female representation in local politics and some spending and access outcomes (Suci, Yamada, & Wibowo, 2020). Second, given that candidates run as pairs in Indonesian local elections, this paper tests the impacts of female leaders using a broader definition of local leaders, including vice-heads. This unique setting in defining leaders potentially brings a new perspective to existing studies that only define leaders as someone in the highest leadership position, such as heads of districts or villages (Brollo & Troiano, 2016; Chattopadhyay & Duflo, 2004; Ferreira & Gyourko, 2014).

1.6 Structure

This thesis comprises five chapters with Chapter 2 to 4 serving as analytical chapters and presenting the core of the research. Chapter 1 provides the background of the thesis. Chapter 2 examines the degree of spatial interdependence in fiscal policy and service provision by local governments. Chapter 3 investigates if switching to a more bureaucratic form of village government influences service access outcomes at local level. Chapter 4 examines the consequences of electing a female leader on fiscal and service outcomes. The last chapter concludes by highlighting the contributions, policy implications from the findings, and possible further research directions drawn from each paper.

2 FISCAL POLICY INTERDEPENDENCE ACROSS LOCAL GOVERNMENTS

2.1 Introduction

As a rational agent, local government makes its fiscal policy decisions by considering the governmental activities of other local governments. Two broad models that can explain this strategic interaction between local governments are the spillover model and the resource-based model (Brueckner, 2003). The spillover model argues that interaction between governments arises when the consequence of any economic activity in one region affects other parties outside the designated market. Meanwhile, the resource-based model states that local governments are competing over similar mobile resources in generating their tax revenue, and hence decide their tax policy in a strategic fashion.

One classical framework for the spillover model is spending spillover that points at the beneficial or harmful externalities created by local government expenditure in one region on inhabitants in neighbouring regions. Due to this spillover, optimal government expenditure in each region also depends on expenditure policies in other regions. The interaction will potentially lead to a spatial pattern in local government expenditure policies (Case, et al., 1993;Geys, 2006). Furthermore, Solé-Ollé (2006) explains this spending spillover through two different perspectives: benefit spillover, which arises from the benefit of public good provision, and crowding spillover, which comes from congested usage of public goods due to an increase in numbers of users.

Another framework that fits into this spillover model is the information spillover that potentially creates yardstick competition amongst neighbouring regions. Information about fiscal policies in neighbouring regions spills and affects the beliefs of imperfectly informed citizens in evaluating their government's action (Besley, & Case, 1995; Revelli, 2005). As a result, citizens will use this information in choosing what they want their own local government to do and will deliberately use the performance of neighbouring regions as a basis for determining whether their government is providing an optimal amount of public goods. On the other hand, the true preference of citizens is unknown by the government and hence the best response for local government is to consider the governmental activities of other regions. These arguments bring forward the incentive for local government to mimic the fiscal policy of other regions as a rational response to voters, who will use other local governments as a yardstick to assess the competence of their own government (Caldeira, 2012; Besley, &

Case, 1995). Another voter assessment approach that is based on the information spillover framework is the reference point approach where voters use neighbouring government fiscal policies as a reference point to generate an additional utility. People experience this additional utility from their consideration of their local government policy quality relative to the other region's policies (as summarised by Geys, 2006).

Given both frameworks, the sign and magnitude of the impact of local government expenditure in neighbouring regions on the local government expenditure in one region may be positive for several expenditure categories and negative for others, leaving room for empirical studies to fill. Various studies have been conducted to identify the existence of fiscal policy interdependence in developed countries (e.g. Case, et al., 1993; Geys, 2006; Solé-Ollé, 2006; Revelli, 2005), but only a few have discussed the case of developing countries. Most of the discussions in developing countries focus on the effect of public spending or intergovernmental transfers to development outcomes per se with limited discussion on fiscal policy interaction between regions. Identifying the presence of fiscal policy interdependence is more crucial nowadays due to criticism of the lack of incentives for local government to provide better service access in a decentralised context (Lewis, 2016; Qian and Weingast, 1997). The presence of fiscal policy interdependence can indicate that some forms of regional competition will affect local government accountabilities and also the overall effectiveness of decentralisation in delivering service access.

This paper will investigate the presence of fiscal policy interdependence among district governments in Indonesia between 2009 and 2013. It will not discuss interdependence on the revenue side as most local government budget revenue in Indonesia is sourced from central government transfers and therefore will not exploit the resource-based model. Indonesia serves as a useful analytical case for developing countries, since its fiscal decentralisation scheme is mainly concentrated on the expenditure side and interjurisdictional labour mobility is limited (Gadenne and Singhal, 2014; Pepinsky and Wihardja, 2019; Smoke and Lewis, 1996). In identifying the interdependence between local fiscal policies, I will consider the multi-tiered government structure in Indonesia and therefore control for province-level expenditure policy, to confirm that the interaction is a strategic behaviour rather than a common reaction to higher-level government policy. Also, incorporation of this higher-level policy follows the suggestion that a policy might be spatially blind in design but not necessarily spatially blind in effect (Hewings, 2014).

The closest predecessor to this paper is the research conducted by Arze del Granado, Martinez-Vazquez, and Simatupang (2008), which examines the presence of jurisdictional competition in Indonesia before and after decentralisation. This paper differs from that study in several important ways. First, this paper looks at a different context, which is after political decentralisation in 2005. The local democracy process, in which district heads and local parliaments are directly elected, can trigger more competition between regions. Second, in addition to expenditure interdependence, this paper will also examine the interdependence of local services access while controlling for expenditure level and its spatial effect. In the framework of fiscal policy, government performance is not only assessed at the expenditure level but also according to intra-allocation spending and budget implementation, which eventually results in a level of service access. The reason for looking at service access variables is that rational individuals are likely to assess the cost of public good provision relative to the level of public goods provided. Hence, voters and local governments are expected to also consider service access outcomes rather than assessing the level of government expenditure alone (Geys, 2006). By also looking at service access interdependence, I will be able to disentangle the direct effect from the spending and the indirect effect coming from neighbours' service access and neighbours' spending. With this exercise, this study is one of the first attempts that tests service access interdependence in developing countries and contributes to the discussion of whether spatial interaction is one of the influential factors for service access level in a region.

The rest of the paper will be structured as follows. The next section will discuss the context of decentralisation policy in Indonesia. The third section will present a literature review of existing studies on this topic. The fourth and fifth sections will present the methodological approach in examining the existence of fiscal interdependence, including the theoretical framework and the data used. Sixth, the results will be presented in the form of summary tables focusing on the degree of interdependence variables. The last section will conclude about the presence of spatial interdependence and its implications for boosting local government accountability and efficiency.

2.2 Background

Indonesia is the world's largest archipelagic country and one of the most spatially diverse nations in its resource endowments, population settlement, location of economic activity, ecology and ethnicity (Hill et al., 2008). The era of decentralisation began in 2001 through the

implementation of Law 22/1999 on Regional Government, Law 25/1999 on the Fiscal Balance between the Central Government and the Regions, and Law 34/2000 on Regional Taxes and Levies. The former two laws were revised in 2004 with Law 32/2004 on Regional Government and Law 33/2004 on the Fiscal Balance between the Central Government and the Regions. Both laws currently are the foundation for fiscal decentralisation in Indonesia.

All those legal frameworks induce major public service responsibilities for district governments and open access to eligible resources. However, districts' geographical configurations are often arbitrary constructs and do not necessarily relate to economic zones (Hill et al., 2008). District government has the responsibility to deliver basic public services (health, infrastructure, and education), together with other tasks including agriculture, communications, industry-trade, environment, land, cooperatives and labour. Recent data from the World Bank dataset shows that subnational government plays a major role in delivering basic services through managing and implementing 54 percent of all levels' (central-province-district) government expenditure, excluding subsidy and interest payments.[1] For economic classification, 51 percent of capital expenditure is implemented by the subnational government. Sector-wise, district governments also spent more than 50 percent of infrastructure, health, and education sector expenditure. In contrast to the expenditure side, the major revenue for local government budgets still comes from central government intergovernmental transfers, as major taxes[2] are not yet decentralised to local levels and nearly 90 percent of total general government revenue is still collected by the central government. In 2015, districts' own revenues only accounted for 13 percent of total district revenue. Provinces' own revenues accounted for a higher share at 52 percent of total provinces' revenue since major user charges are being administered by the province governments. This decentralisation scheme is not only adopted by Indonesia as similar arrangements can be found in many developing countries (Gadenne and Singhal, 2014). Pursuant to Law 32/2004, the draft of the local budget (APBD) is approved by the district parliament and needs to pass the evaluation of the provincial government so it will not conflict with the national public interest or any law.

[1] Author's calculation on 2015 data from World Bank Indonesia COFIS (Consolidated Fiscal) dataset. Access to the dataset is publicly available at http://wbi.worldbank.org/boost/country/indonesia

[2] Revenue decentralisation though is part of the blue print of fiscal decentralisation. E.g. In 2013, the Government of Indonesia (GoI) gradually shifted the collection and management of property tax from central government tax to district level tax.

This assignment of services was also complemented by political decentralisation, indicated by the development of local democracy. Since the decentralisation 'big bang' in 2001, representatives of district parliaments had been elected through direct parliamentary elections in 2004, 2009 and 2014. Law 32/2004 also prescribed that the heads of local government should be directly elected. The first batch of local heads of districts direct elections took place in 2005, and gradually all indirectly elected heads were replaced by directly elected ones. Both local level direct elections, legislative and executive, were expected to increase electoral accountability and improve local governance (Sjahrir et al., 2014).

2.3 Literature review

One of the major contributions of the recent growing literature in the field of economic geography is the regional interaction model, which relies on the concept of regional boundaries (as summarised by Fujita, 2010). In this interaction model, regions are related due to linkage effects among citizens and industries, mainly because of externalities and information spillovers. One profound manifestation of this regional interaction in the public sector is fiscal policy interdependence. Ignoring this potential interaction in the policy-making process can lead to inefficiencies, since local government will neglect the externality effects of its public goods and tax decisions on the utility of non-residents (Gordon, 1983).

Case, Rosen, and Hines (1993) formalise the notion that states' expenditures depend on the spending of other similarly situated states. In this study, the authors build the theoretical framework of spending spillover using a traditional model of public choices. The spending spillover is incorporated by allowing expenditures in other states to enter the utility function of residents in one state since provision of public goods is also enjoyed by the surrounding community and not limited to constituents. The presence of this spillover eventually affects the analysis of expenditure choice, since local government may decide upon fiscal policy that will also encourage other regions to provide public goods, in which the first region can obtain benefits. One prior assumption about this interaction is that each region makes Nash assumptions[3] about other states' expenditures. Furthermore, Solé-Ollé (2006) identifies two different types of spending spillover. First, benefit spillover is a spillover that arises from the provision of public goods in which local public goods are not only enjoyed by constituents but also neighbouring non-residents. Next, in contrast, crowding spillover, in which local

[3] Ruling out the possibility of predetermined strategic interaction existence among local governments

government considers that a locality's consumption level is already influenced by the non-residents living in surrounding neighbourhoods.

Another potential channel of fiscal policy spatial interdependence is driven by the existence of information spillover. This spillover occurs when citizens, who have the principal role in the classic principal-agent setting, have asymmetric information in assessing their local government's performance (Salmon, 1987). As a result, voters' decisions in officials' elections, would not only be based on the officials' performance in office but also on the performance of neighbouring jurisdictions, which will be used as a yardstick to evaluate the incumbent. Besley and Case (1995) define this competition in revenue setting and derive a theoretical basis based on strategic games equilibrium, with citizens as the principal and incumbent local government as the agent. They find an encouraging proposition that tax-setting and vote-seeking are tied together through the nexus of yardstick competition. While the initial yardstick approach is derived from revenue setting, the possibility of competition on the expenditure side also can be explained using this argument. The main idea is that voters will use the performance of others as benchmark, and hence local governments would consider other regions' policy as a behavioural response to that electoral competition. Caldeira (2012) attempts to explain the competition from the expenditure side by deriving a modification of the yardstick model and setting federal government as the sole principal, instead of citizens, as she reflects on China's political decentralisation structure.

Another voter assessment approach that is based on the information spillover framework is the reference point approach. In this approach, voters use neighbouring government policies as a reference point to generate transaction utility, an additional utility that people experience from their consideration of the quality of the policy relative to other regions' policies (as summarised by Geys,2006). For example, Case et al. (1993) mention that citizens of one state might care about the context of welfare payments to poor residents in other states and hence derive utility from other states welfare-related expenditure. Similar to yardstick competition, the reference point approach assumes that citizens with imperfect information use the expenditures of neighbouring states as a basis for determining whether their local government is providing optimal amounts of public expenditure and public goods. This referencing behaviour affects voters' utility yet is not always necessarily reflected in voting behaviour. Both voter assessment approaches, yardstick competition and reference point, resonate

coherently with the development of political decentralisation that gives rights to citizens to directly elect and evaluate their legislators.

Studies have discussed the empirical case of fiscal policy interdependence and found mixed results based on the country context and decentralisation scheme. Using U.S. data, some analyses find that a state government's level of per capita expenditure is positively and significantly affected by the expenditure level of its neighbours (Baicker, 2005; Case et al., 1993). Some studies also observe a relevant spending spillover and fiscal interaction between local governments when using lower local government levels, for instance municipalities or district data (Ferraresi et al., 2018; Solé-Ollé, 2006). Fiscal decentralisation, which redefines the regional discretionary power of budgetary policy, is found to favour greater spatial interdependencies in the public expenditure decision process (De Siano and D'Uva, 2017). On the other hand, several studies have also observed a positive spatial correlation that can be attributed mostly to common reactions to some external shocks, rather than the actual spatial interaction form (Frère et al., 2014; Revelli, 2003). The magnitude of fiscal interaction is significantly reduced when considering fiscal policy and cooperation in a higher-level government structure. This reduction also appears when treating an increase in tax and spending as exogenous, and therefore any possible distortion generated by fiscal spillover is considerably smaller (Isen, 2014).

Despite numerous studies having posited the investigation of spatial interaction on fiscal policies among local governments, few have discussed the empirical case for developing countries. Many evaluations about decentralisation in developing countries have looked at development outcomes per se and neglected the potential fiscal policy interaction between regions. Caldeira (2012) provides empirical evidence of competition among 29 provinces in China from 1980 to 2014 in a vertical bureaucratic control and political system. Caldeira finds the existence of public spending interaction among geographically and economically close provinces. For Indonesia's case, Arze del Granado, Martinez-Vazquez, & Simatupang (2008) examine the presence of jurisdictional competition among district level governments in Indonesia between 2001 and 2004. They find evidence of yardstick and expenditure spillover effects. During that period, the relation appears to be relatively inelastic as a 1 percent increase in total discretionary spending in neighbouring districts is associated with a 0.07 percent increase in a district's own discretionary expenditure. On sectoral expenditure, they found that positive interdependence only exists for general administrative spending, which

mostly consists of payroll expenditure. One possible explanation of this relatively weak magnitude of effects is that the study's time frame did not cover the time period of the political decentralisation which started being implemented in 2005. Hence, given the current different context of the political system, among other reasons, there is still room to analyse fresh empirical evidence from Indonesia.

Local government expenditure has been acknowledged as one of the main factors that directly determines access to services. For Indonesia, Lewis (2016) finds that local government spending in Indonesia positively influences access to education, health and infrastructure services up until a certain point, then becomes negative due to political and governance factors. Still, empirical analyses of fiscal policy interdependence mainly focus on spending or revenue in isolation while neglecting the possibility of horizontal interdependence influencing service access levels. If interaction between jurisdictions is derived from the spillover of public good provision and from the relative performance assessment of voters to their local government, it is unrealistic to limit the interdependence only to fiscal outcomes such as spending or local revenue. Voters are likely to have comparative information on the level of services provided and will use the information to discipline the incumbent (Geys, 2006). Also, looking at the spending spillover framework, citizens benefit from externalities that result from the provision of public goods and services, instead of directly from local government spending (Solé-Ollé, 2006). Therefore, this paper will also examine the interdependence of local service access, after controlling for expenditure level and its spillover effect.

2.4 Empirical strategy

Theoretical framework

The main set-up for a spillover model on strategic interaction between governments is that each jurisdiction *i* chooses a level of decision variable *h*, but the decision is affected by *h* chosen elsewhere. Following Brueckner (2003), jurisdiction *i*'s objective function can be written as:

$$V(h_i, h_{-i}; X_i) \quad (1)$$

where h_{-i} is the decision variable in other jurisdictions and X_i is a vector of *i*'s characteristics that affect the preference. When maximising equation (1), $\frac{\partial V}{\partial h_i}$ is set equal to zero and hence the solution can be written as a reaction function that depends on h_{-i} and X_i:

$$h_i = R(h_{-i}; X_i) \quad (2)$$

In this paper, the level of jurisdiction will be district level government, and will be called district *i*, from this point onwards. This spillover model underlies several existing empirical studies, including the spending spillover and yardstick competition framework.

On spending spillover, let the preference for a representative resident of district *i* be given by a quasi-concave utility function, $U(c_i, q^*_i; X_i)$, where *c* is the consumption of private goods, *X* represents district characteristics other than income, and q^* represents the level of service access being enjoyed by residents in that district. Due to the spillover, we assume that service access is a result of public goods provided by the district, q_i , and also the neighbouring district, q_{-i}. Hence, we can expand the preference equation into $U(c_i, q_i, q_{-i}; X_i)$. Using y_i as the per capita income and $c_i = c(y_i, q_i)$ the individual budget constraint yields a similar version of equation (1):

$$U(c(y_i, q_i), q_i, q_{-i}; X_i) \equiv V(q_i, q_{-i}; X_i) \quad (3)$$

The yardstick competition framework can also be explained within a similar set-up. In this framework, voters compare and look at public services in other districts to evaluate whether their government is performing well and deserves to be voted back into office. The comparison is implemented in a minimum level of public goods provision that must be delivered, observed relative to other jurisdictions, written as $q_i = \Phi(q_{-i})$. The consumer preferences are now given by:

$$U(c(y_i, q_i), q_i, \Phi(q_{-i}); X_i) \equiv V(q_i, q_{-i}; X_i) \quad (4)$$

Now, define $q_i = q(s_i, z_i)$,where s_i is the government spending of district *i*, and z_i is the cost of providing the public goods. Substituting this into the indirect objective function of (3) and (4) results in:

$$V(q_i, q_{-i}; X_i) \equiv V(s_i, s_{-i}; X_i, z_i, z_{-i}) \quad (5)$$

According to the spending spillover framework, the sign of the interaction can be positive or negative depending on the complementary degree of the decision variable. Case, Rosen, and Hines (1993), as summarised by Revelli (2005), derive the expenditure spillover model to get the slope of the reaction function formulated in (2). Looking at the expenditure side, now let preference for a representative resident of district *i* be given by $U(c_i, s_i, s_{-i}; X_i)$, subject to c_i =

$c(y_{i,}\, s_i)$. Optimising the utility function in respect to $s_{i,}$ will result in a first order condition as follows:

$$\frac{\partial U(.)}{\partial c_i}\frac{\partial c_i}{\partial s_i} + \frac{\partial U(.)}{\partial s_i} = 0 \qquad (6)$$

Equation (6) implies that the optimal level of s_i depends on $s\text{-}_i$ through raising or diminishing the marginal utility of own government expenditure ($\frac{\partial U(.)}{\partial s_i}$) or own private consumption ($\frac{\partial U(.)}{\partial c_i}$). Assuming that $\partial c_i / \partial s_i$ will be equal to -1, the sign of $\partial s_i / \partial s\text{-}_i$ can be derived from the partial derivatives of the FOC equation (6), given by:

$$\frac{\partial s_i}{\partial s_{-i}} = \frac{-\left(\frac{\partial^2 U(.)}{\partial s_i \partial s_{-i}} - \frac{\partial^2 U(.)}{\partial c_i \partial s_{-i}}\right)}{\frac{\partial^2 U(.)}{\partial s_i^2} - 2\frac{\partial^2 U(.)}{\partial s_i \partial c_i} + \frac{\partial^2 U(.)}{\partial c_i^2}} = \frac{-\left(\frac{\partial MU_S}{\partial s_{-i}} - \frac{\partial MU_C}{\partial s_{-i}}\right)}{\frac{\partial^2 U(.)}{\partial s_i^2} - 2\frac{\partial^2 U(.)}{\partial s_i \partial c_i} + \frac{\partial^2 U(.)}{\partial c_i^2}} \qquad (7)$$

The denominator is negative, as given by the second order condition of the maximisation problem. Hence, $\partial s_i / \partial s\text{-}_i$ is positive (negative) if the neighbours' expenditure is more (less) complementary with the own government expenditure compared to own private consumption. For the yardstick competition framework, as the underlying theory is the mimicking behaviour of the local government, the sign of the interaction is positive (as found by Caldeira, 2012).

Revelli (2003) points out the importance of modelling the vertical fiscal externalities in a multi-tiered government structure to better understand the interaction between lower-level government units. The study argues that an average constituent will not only enjoy the public goods provided by the local government but also benefit from those provided by the higher-level government. Taking this perspective, the preference equation can now be represented by $U(c_{i,}\, q_{i,}\, q\text{-}_i\ q_h\ ;\ X_i)$ where q_h is the public goods provided by the higher-level government. Maximising the preference subject to $c_i = c(y_{i,}\, q_{i,}\, q_h)$ leads to:

$$U(c_{i,}\, q_{i,}\, q\text{-}_i\ q_h\ ;\ X_i) \equiv V(q_{i,}\, q\text{-}_i\ q_h;\ X_i) \equiv V(s_{i,}\, s\text{-}_i\ s_h;\ X_{i\ ,}\, z_{i,}\, z\text{-}_{i,}\, z\text{-}_h) \qquad (8)$$

Expenditure interdependence model

Departing from the theoretical framework discussed in the previous section (Eq. 8), to test the existence of expenditure interdependence, this study considers a spatial lag specification in which the government expenditure of district *i* in year *t*, s_{it}, is a function of neighbouring'

government expenditure, s_{jt}. In line with earlier literature (Arze del Granado et al., 2008; Revelli, 2003), the model also allows s_{it} depending on a set of district-specific controls and higher-level (province) government expenditure, $\mathbf{X}_{it}$, and a time invariant district-specific effect, a_i. The district-specific controls include transfer (from higher-level government) per capita, GRDP per capita, population, poverty rate, and poverty gap.

$$s_{it} = \rho W s_{jt} + \mathbf{X}_{it}\boldsymbol{\beta} + \alpha_i + u_{it} \quad (9)$$

where i= 1,…n denotes a district, t = 1,…T denotes a time period and u_{it} is the random error. A set of neighbours to each district is implemented by constructing a spatial weights matrix of *W*. *W* is a (n x n) matrix that assigns neighbours to each *i,* in the sense that the element of (*i,j*) of the matrix is different from zero if district *i* and *j* are neighbours based on a predetermined condition, and zero otherwise. Matrix *W* is row standardised such that the elements of each row sum to one. This way the product of the matrix *W* and a vector of district expenditures will result in a spatially weighted average of neighbouring districts' expenditure for each of the districts.

Following Case et al.(1993), neighbours might also be subject to correlated random shocks for instance, a broader shock in regional economic growth that affects neighbouring districts. This kind of shock produces a correlation between districts' spending when there is not one present. To avoid incorrect conclusions, the model allows potential correlation among the error terms of neighbours by letting:

$$u_{it} = \varphi W u_{jt} + \varepsilon_{it} \quad (10)$$

where ε_{it} is a well-behaved idiosyncratic error that is uncorrelated between districts, and φ is a scalar parameter that indicates the degree of spatial correlation between residuals. This model is called the spatial error model. One way to express possible spatial error dependence is through Durbin representation, which nests both the spatial lag and also the spatial error (Day and Lewis, 2013; Elhorst, 2003). Formally, the spatial Durbin model contains a spatially lagged dependent variable and spatially lagged independent variables. Eq. (9) is modified as follows:

$$s_{it} = \rho W s_{jt} + \mathbf{X}_{it}\boldsymbol{\beta} + W\mathbf{X}_{jt}\boldsymbol{\gamma} + \alpha_i + \varepsilon_{it} \quad (11)$$

If $\gamma = 0$, then the Durbin simplifies to the spatial lag model (spatial autoregressive model – SAR); if $\gamma + \rho\beta = 0$, then it simplifies to the spatial error model (SEM). I will begin with the full Durbin specification, following LeSage and Pace (2009), who propose the Durbin specification which subsumes SEM and SAR as the model departure, then empirically test whether the model can be simplified to either SAR or SEM.

Service access interdependence model

For service access, based on (11), the model will take a similar form to expenditure interdependence, but separating government expenditure as one of the independent variables to get the disaggregated effects:

$$q_{it}^{*} = \rho W q_{jt}^{*} + \sigma s_{it} + \omega W s_{jt} + \mathbf{X}_{it}\boldsymbol{\beta} + W\mathbf{X}_{jt}\boldsymbol{\gamma} + \alpha_i + \varepsilon_{it} \quad (12)$$

where q_{it}^{*} is the service access enjoyed by residents in district *i* in time *t*, which is a function of neighbours' service access, q_{jt}^{*}. One main explanatory variable in this model is total local government expenditure, s_{it}. In line with earlier literature, the model also depends on a set of district-specific controls, $\mathbf{X}_{it}$, consisting of population, poverty rate, poverty gap, GRDP per capita, share of urban population, and province expenditure as a representation of higher-level government intervention. The SDM model allows all of the explanatory variables to be spatially weighted and hence we have the neighbours' spending, s_{jt}, and $\mathbf{X}_{jt}$ as a set of neighbouring district-specific controls; each is multiplied by the weight matrix.

Decomposition of direct and indirect effect

Spatial models are recursive, and hence when interpreting the coefficient of the explanatory variable, we cannot do it directly. For instance, from equation (12), if spending (s) increases then it (hypothetically) will improve service access (q^*) by σ. That increase in q^* then spills over to produce a further (hypothetical) increase in q^* and so on.

The direct effect of spending then is defined as $[(I - \rho W)^{-1}.(\sigma I + \omega W)]^{d}$, with *d* as an operator that calculates the mean diagonal element of the matrix. The first part $[(I - \rho W)^{-1}.(\sigma I)]^{d}$ can be interpreted as the effect of s_i on q_i^{*}, while the second part $[(I - \rho W)^{-1}.(\omega W)]^{d}$ can be interpreted as the effect of s_j on q_j^{*} which then also exerts changes on q_i^{*} due to its autoregressive nature. Direct effects in this case slightly differ from the coefficient estimates because in addition to the effect of each explanatory variable to the

regions, the effects also include the effect of each variable in the neighbouring regions. On the other hand, the indirect effect is $[(I - \rho W)^{-1}.(\sigma I + \omega W)]^{rsum}$, where *rsum* is an operator that calculates the mean row sum of non-diagonal elements. This effect represents the feedback effects of each variable first on the neighbouring regions. This furthermore can be interpreted as the effect of s_j on q_i^* and the effect of s_{-j} on q_i^* through q_j^* .

Specifying weight matrix

Estimation of the above model requires the definition of neighbours that must be specified a priori. Looking to the spillover model definition, as proposed by Brueckner (2003), district *j* is a neighbour of district *i*, if there is a form of spillover, either benefit or information, from district *j* that affects the fiscal choices of district *i*. In this case, the use of geographical criteria seems reasonable as under this model, close-by jurisdictions are more likely to affect each other than far away ones.

Much previous work has used either straightforward contiguity criteria – according to which neighbourliness is defined as common borders and equal weight is given to all contiguous jurisdictions – or continuous distance criteria, whereby weights are inversely related to the distance between two jurisdictions. The empirical analyses in this study are performed using weighting matrices based upon both criteria. First, the weight matrix will be developed by setting a dichotomous variable as the element of the weighting matrix, w_{ij} = 1 if districts *i* and *j* share a common border, and 0 otherwise. Second, for an alternative view, after setting d_{ij} as the distance between the centre point of district *i* and *j*, the element of the weight matrix is calculated as w_{ij} = 1/ d_{ij}. The straight contiguity matrix will be used as the main weight matrix due to the objective of checking the most parsimonious specification first. The robustness checks will employ inverse distance matrices with additional requirements that districts are in either the same province or the same region. Also checking the inverse distance matrices will shed lights on whether the results are sensitive to the fact that the dataset cannot include all districts to achieve a balanced panel set-up, due to missing data.

While acknowledging the importance of geographic proximity, some argue that other factors may be more relevant in defining neighbours. First, not all neighbours should be given equal weight and elements of the matrix can be weighted by neighbours' characteristics such as income and population size. Another argument is that it may be more sensible to build the weight matrix based on similarity in demographic or economic characteristics between

regions, regardless of the distance (Baicker, 2005; Caldeira, 2012; Case et al., 1993). However, in light of these considerations, this study will still use geographic proximity as the prior basis in defining neighbourliness since it aligns with the foundation of the theoretical framework. Also, using a geographic proximity matrix fits with the notion that implementing an elemental approach to the weight matrix is preferred to more structured approaches.

Potential econometric problem

The reaction function of the spillover model relates each jurisdiction's chosen variable, *h*, to its own characteristics and to the choices of *h* in other jurisdictions. Empirical works then used to estimate such functions that can be written in general form such as:

$$h_i = \rho W h_j + \mathbf{X}_i \boldsymbol{\beta} + \varepsilon_i \qquad (13)$$

where $\boldsymbol{\beta}$ and ρ are unknown parameters (the former is a vector), ε_i is an error term and W represents a non-negative weight matrix that has been defined a priori.

As discussed in the spatial econometrics literature (Anselin, 1988 in Brueckner, 2003), due to the strategic interaction hypothesis, the *h* values in different jurisdictions are jointly determined. Consequently, the linear combination of h_i that appears on the right-hand side of equation (13) is endogenous and correlated with the error term. The empirical strategy of this study will follow one of the methods frequently employed in the literature in addressing this issue, that is, estimating the reduced form by a non-linear maximisation technique, which is the maximum-likelihood (ML) approach (Caldeira, 2012; Case, Rosen, & Hines, 1993; Besley & Case, 1995; Day & Lewis, 2013). Since equation (13) holds for all jurisdictions and the key spatial autoregressive parameter, ρ, enters the equation non-linearly, this approach introduces a non-linear reduced form in the system and therefore obtains the key parameter via a non-linear optimisation routine. One feature of this ML estimator is that its asymptotic properties will rely on the weight matrix form. Elhorst (2003) notes that the consistency of the ML estimator will not be undermined when the spatial matrix is an inverse distance formulation or any straightforward contiguity criteria, particularly when it is a panel data estimation.

Endogeneity, albeit minor, can also come from correlation between the jurisdiction characteristics ($\mathbf{X}_i$) and the error term through natural grouping of one type of households across districts. For instance, high income households who have a specific high demand for

one type of public goods will reside in a district that has made a commitment to deliver such goods. This results in a correlation between the income level of the district (as an element of X_i) and the error term, leading to an inconsistent estimate of the income coefficient and potentially distorting the estimates of the remaining coefficients. This problem is usually addressed through finding a suitable instrument for the offending variables. An alternative approach, however, is to use panel data in which all time-invariant district characteristics can be captured by district-specific intercepts. Panel data also help to eliminate the spatial error dependence that arises through correlation of omitted variables. District-specific intercepts can capture the influence of such variables and hence the remaining error term may exhibit less spatial dependence. This study, therefore, will use a fixed-effect panel approach to capture the unobserved time-invariant district characteristics.

2.5 Data and choice of variables

This paper focuses on the second level of local government in Indonesia, below the province level, which is district government. The empirical analyses cover the period 2009–2013 for 491[4] districts, based on the Ministry of Home Affair's records for the year 2009. This number of districts excludes the districts under DKI Jakarta province that do not have budget and service delivery responsibility. To accommodate a balanced-panel data model that is required in the spatial panel model, I exclude districts that have no available data on the variable of interests, mostly due to the failure to report budget outcomes, and therefore end up with 306 districts for the spending interdependence model and 436 districts for the service interdependence model.

This time frame of 2009 to 2013 has been chosen for several reasons. First, this period ensures time and coverage consistency across categories of spending as the expenditure classification remains the same during those years. Also, during that period, the number of districts was fairly constant due to the splitting moratorium from the central government between 2009 and 2012, and therefore accommodates the balanced-panel data requirement. Second, this period includes the period after significant political decentralisation had taken place, in the form of direct voting for district' heads and district' parliaments, starting from 2009. Hence, the analyses will be able to explore different treatment effects, in comparison

[4] In 2013, excluding DKI Jakarta province, there were 505 district units. For this study, the total number of districts for 2013 has been collapsed to 491 units.

with the most recent study available on this topic, which still uses 2004 as their latest data point (Arze del Granado et al., 2008).

The dependent variable in the expenditure model is the realisation of local government expenditure per capita (2010 constant prices). To describe the allocative efficiency of the district fiscal policy, the expenditure data not only covers total government spending but also a breakdown into economic and sectoral classifications. Economic classification spending includes personnel and capital expenditure, while sectoral classification consists of health, education, infrastructure and general admin sectors. Government expenditure data is compiled from the INDO-DAPOER database[5] that comes from the Ministry of Finance.

On the service access model, this study constructs a general index of service access as the dependent variable, in addition to the sectoral services access variables. The rationale for using an index is due to the fact that one type of service access can be explained by various categories of spending, and hence using total government spending and a service access index is reasonable. Service access is defined as residents' access to public services, regardless of who provides the service, and yet fit to represent the model in equation (12). Following Skoufias et al. (2011), the index is constructed via the first principal component[6] using a correlation matrix of service access in education, health, and infrastructure, which are under the control of local governments. Service items include net enrolment rates for junior secondary (SMP) and senior secondary school (SMA), percentage of births attended by a health professional, and percentage of households with access to protected water and sanitation[7]. A correlation matrix is preferred to other measures, e.g. covariance, because this method can eliminate different variances, including among same-scaled variables. The selection of these service access items is based on data availability. Service access data are collected from the annual household survey (SUSENAS) by the Indonesian Bureau of Statistics (BPS).

[5] Indonesia Database for Policy and Economic Research (INDO-DAPOER) is accessible at http://databank.worldbank.org/data/reports.aspx?source=indo~dapoer-(indonesia-database-for-policy-and-economic-research).

[6] Principal Component Analysis (PCA) is a statistical technique used to reduce the number of variables into a smaller number of dimensions by creating uncorrelated indices or components. Each component is a linear weighted combination of the initial variables. The first component explains the largest possible amount of variation from the original variables, subject to the constraint that the sum of the squared weights is equal to one.

[7] Immunisation coverage is excluded from the service index construction due to the inconsistent definition of complete immunisation in the 2013 household survey. However, the estimation will also test interdependence in individual service access including immunisation.

As for the other variables, region characteristics data (i.e. the number of population, poverty rate, poverty gap, share of urban population and income per capita level) and district area geographical data are collected from the Indonesian Bureau of Statistics (BPS). Summary statistics of all variables used in analyses are reported in Tables 2.7 and 2.8 in the Appendix.

2.6 Results

To test whether the Durbin specification can be simplified to either SAR or SEM, I conducted several diagnostic tests. The tests revealed that the most appropriate spatial model specification to describe the data is the Spatial Durbin Model (SDM) and the model cannot be simplified to a spatial lag or a spatial error specification. Next, a panel fixed-effect model is introduced to allow the model to control for the effects of unobservable district-specific characteristics that are time invariant. The Hausman test also revealed that the fixed-effect model is preferable to a random-effect model. These specifications are similar to those of previous studies which discuss spatial interdependence in the region level across years (Arze del Granado et al., 2008; Caldeira, 2012; Day and Lewis, 2013). A summary of the diagnostic test results is reported in Table 2.9 in the Appendix.

Spending interdependence

To investigate whether there is spending interdependence between district, I estimate Eq. (11) using ML fixed-effect model with bias correction of Lee & Yu (2010)[8] and geographical lagged value of the spending variable as the variable of interest. This lagged value is constructed from the product of the weight matrix and a vector of district expenditure and therefore represents the spatially weighted average of neighbouring districts' expenditure for each district. The number of districts that are used in this spending interdependence model is 306 out of 491 official districts. Districts with missing information on the fiscal variables needed to be excluded due to the balanced-panel requirement of a spatial weight matrix.

In investigating spending interdependence, one preliminary concern that may arise is the possibility that districts' spending may always increase/decrease simultaneously over years of observation, hence creating an ambiguous perception of interaction. To account for this concern, I estimate a confidence interval for some selected outcomes based on regressions against year dummies. Figure 2.1 shows that the interval for a given year tends to overlap

[8] Lee & Yu (2010) propose an alternative estimation approach using a transformation matrix that can yield consistent parameter estimates and properly centred asymptotic distributions, when either n or T is large.

with the previous years' interval. This is true not only for total expenditure, but also for some selected expenditure categories such as capital and infrastructure spending. The overlapping intervals indicate that there is insufficient statistical evidence to assert that for all districts, spending always increases (in per capita real terms) over years. Further, Figure 2.1 suggests that temporal correlation is not problematic in the case of this study.

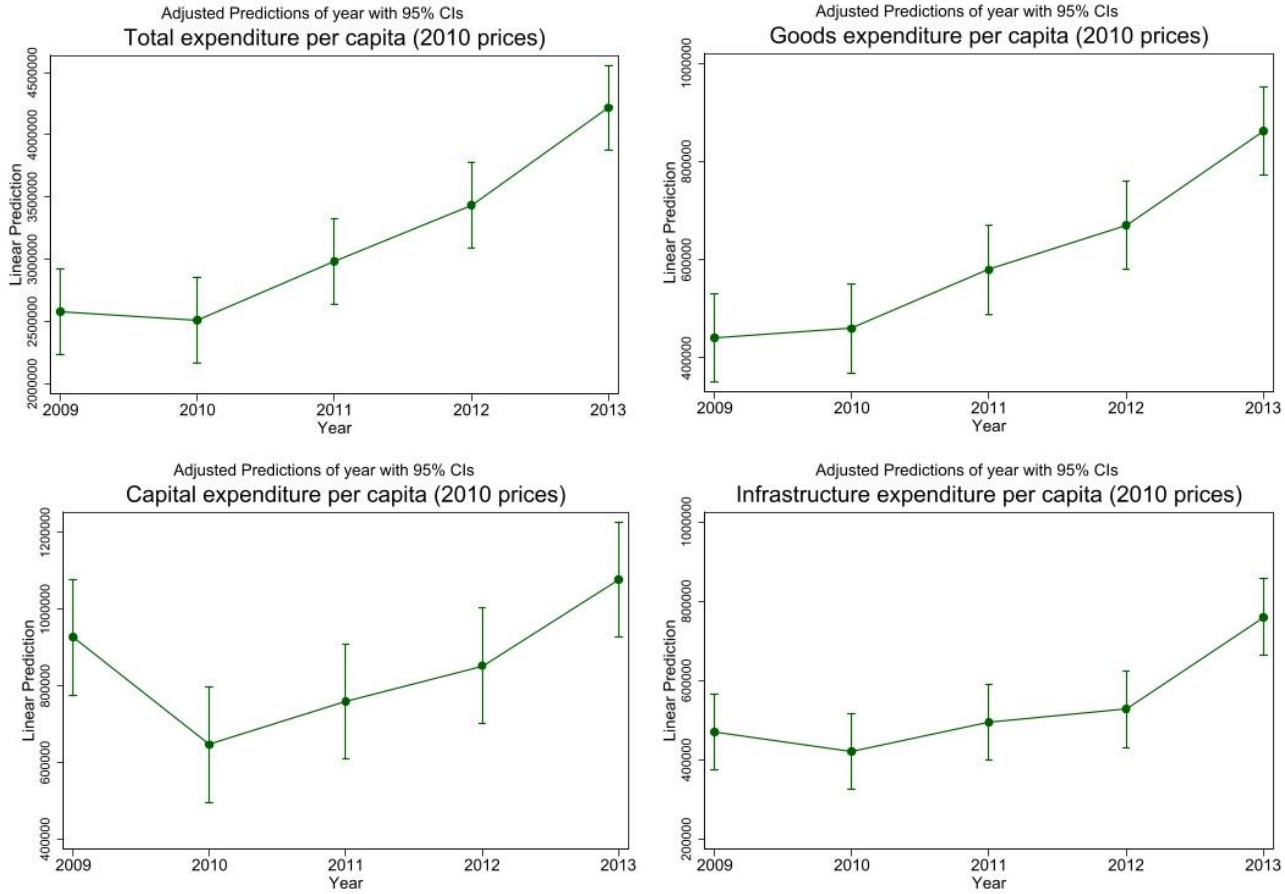


Figure 2.1 District expenditure per capita across years

Table 2.1 reports output for local government expenditure per capita, and each column represents a different category of spending. All columns provide similar results: an increase in other district's spending will lead to an increase in own district spending. On total expenditure per capita, as shown in column (1), an increase of 1% in other district's spending will induce a 0.08% increase in own district spending per capita with a significance level of $\alpha = 0.05$, conditional on transfers, province-level expenditure and socio-economic characteristics of the district. This magnitude of interdependence is similar to Arze del Granado et al.(2008), who found a neighbourhood effect of 0.07%. Positive interdependence between regions is also found in other developing countries such as China (Caldeira, 2012).

Positive neighbourhood effects also emerge when breaking the total expenditure into several categories. This is different to the results of a previous study on Indonesia which found that only general administrative spending shows positive interaction. One possible explanation is that the previous study, Arze del Granado et al. (2008), did not cover the time period of the direct election for district heads, which began gradually in 2005. Capital-based expenditure, such as capital spending in column (2) and infrastructure spending in column (6), show stronger interdependence compared to other spending categories of administration (column 5) and health spending (column 8). However, education spending has no significance coefficient and personnel spending loses its significance when the regression controls for time trend.[9] One possible explanation is most resources used for these spending categories comes from transfers. For personnel spending, the majority of resources used for this spending category comes from the block grant transfer (General Allocation Fund/Dana Alokasi Umum), which includes local civil servant salaries as one of its allocation formulas. While for education, education-based special purpose grants and School Operational Grant (*Bantuan Operasional Sekolah*, BOS*)* make a significant portion of the district education spending. As the model already controls for transfers, the absence of interaction between districts in both categories is plausible.

Another anticipated bias from this positive interaction is that the interaction is due to a common reaction to higher-level government policies, rather than strategic. While the estimation has included provincial level expenditure as a proxy for province policies, it has not covered the central government policies' factor due to data limitations. To rule out the possibility of this common reaction, I construct a placebo matrix that assigns random neighbours to each district based on geographic criteria other than distance, which are area and length of the districts. If the result shows statistical significance, it means district interaction is not based on the definition of neighbourhood distance and hence an external factor potentially exists that can explain the interaction of districts, such as nationwide policies. Using the placebo matrix, total spending and all other spending categories show insignificant results, as shown in panel (C). Overall, these results ensure that for all categories

[9] Year time trend is considered after finding that when including time fixed effects, the model with straight contiguity matrix has difficulties converging due to the complex specification, which already includes a spatial matrix with more than one hundred dimensions. I argue that, given this limitation, year time trend would suffice as the time period of the estimated model only covers five years (2009–2013) and there were no unobservable significant shocks that induced high policy shifts at district level during the period. In the robustness check, the model with inverse distance matrix has successfully converged with time fixed effects, and the results are consistent.

of spending in general, positive interdependence is not only a common reaction to certain policies, but a strategic one.

Table 2.1 Estimation results for expenditure interdependence

		By economic classification				*By sectoral classification*		
VARIABLES	Total (1)	Capital (2)	Personnel (3)	Goods (4)	Admin (5)	Infrastructure (6)	Education (7)	Health (8)
	(A) Straight contiguity matrix							
Neighbours' expenditure	0.095***	0.235***	0.096***	0.143***	0.060**	0.125***	0.032	0.067**
	[0.033]	[0.029]	[0.033]	[0.031]	[0.029]	[0.031]	[0.030]	[0.030]
R-squared	0.904	0.426	0.067	0.742	0.120	0.443	0.288	0.719
	(B) Straight contiguity matrix							
Neighbours' expenditure	0.083**	0.238***	0.050	0.133***	0.065**	0.125***	0.030	0.062**
	[0.033]	[0.029]	[0.034]	[0.031]	[0.029]	[0.031]	[0.030]	[0.029]
R-squared	0.657	0.335	0.693	0.665	0.023	0.474	0.560	0.571
	(C) Placebo matrix							
Neighbours' expenditure	0.222	-0.280	0.124	0.015	0.198	-0.328	-0.124	-0.488
	[0.266]	[0.242]	[0.325]	[0.273]	[0.280]	[0.345]	[0.304]	[0.312]
R-squared	0.410	0.000	0.574	0.556	0.194	0.049	0.309	0.106
Number of groups	306	306	306	306	306	306	306	306
Observations	1,224	1,224	1,224	1,224	1,224	1,224	1,224	1,224

Notes: Standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1. Panel B regression includes time (year) trend. Placebo matrix is constructed using principal component analysis that constructs similarity between districts based on area and length of the district. All regressions include district-specific controls and higher-level (province) government expenditure. District-specific controls are transfer per capita, GRDP per capita, population, poverty rate and poverty gap.

Service interdependence

After finding the existence of local government spending interdependence, this study also examines the interdependence of local service access, while controlling for the interaction in the spending. Similar to the spending model, I estimate Eq. (12) using ML fixed-effect model with Lee-Yu's (2010) transformation approach and geographical lagged value of service access as the variable of interest. This lagged value is constructed from the product of the weight matrix and a vector of district service access, and therefore represents the spatially weighted average of neighbouring districts' service access for each district. The number of districts that are used in this interdependence model is 436 districts out of 491 official districts. Districts with missing information on service access need to be excluded due to the balanced-panel requirement of a spatial weight matrix.

Table 2.2 indicates that positive interdependence in service access is found even after controlling for potential spillover from spending. A one unit increase in neighbours' service index is associated with a 0.2 unit increase in own district service access index, which is equal to 0.13 standard deviations, as shown in column (1). This result is conditional on local government spending, spending at provincial level, and also district-specific characteristics such as GRDP per capita, proportion of urban population, and some poverty indicators. The positive interaction persists across other single service outcome indicators, presented from column (2) to column (7). For instance, column (3) shows that a one percentage point increase in neighbours' senior high school enrolment will induce a 0.15 percentage point increase in the enrolment rate in own district. Although not many studies have discussed interaction in service access, the results of this study are consistent with the findings of Geys (2006) which uses the case of Belgium. Geys confirms the existence of neighbourhood effects in local public good policies after relating total spending to the quantity of locally provided public goods.

Table 2.2 Estimation results for service access interdependence

		Education		*Health*		*Local infrastructure*	
VARIABLES	Index service (1)	NER junior (2)	NER senior (3)	Immunisation (4)	Assisted birth (5)	Safe water (6)	Safe sanitation (7)
	(A) Straight contiguity matrix						
Neighbours' service access	0.259***	0.196***	0.171***	0.622***	0.192***	0.159***	0.096***
	[0.026]	[0.027]	[0.028]	[0.019]	[0.027]	[0.029]	[0.029]
R-squared	0.051	0.011	0.000	0.027	0.161	0.045	0.048
	(B) Straight contiguity matrix						
Neighbours' service access	0.205***	0.187***	0.148***	0.623***	0.154***	0.137***	0.080***
	[0.027]	[0.028]	[0.029]	[0.019]	[0.028]	[0.029]	[0.029]
R-squared	0.055	0.057	0.001	0.032	0.001	0.057	0.041
	(C) Placebo matrix						
Neighbours' service access	0.304	-0.315	-0.464	0.826***	0.149	-0.551**	0.395*
	[0.215]	[0.290]	[0.300]	[0.073]	[0.245]	[0.245]	[0.204]
R-squared	0.306	0.043	0.021	0.244	0.245	0.028	0.151
Number of groups	436	436	436	436	436	436	436
Observations	1744	1744	1744	1308	1744	1744	1744

Notes: Standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1. Panel B regression includes time (year) trend. Placebo matrix is constructed using principal component analysis that constructs similarity between districts based on area and length of the district. All regressions include government expenditure per capita, district-specific controls and higher-level (province) government expenditure. District-specific controls are GRDP per capita, population, poverty rate, poverty gap, and share of urban population. Index service is constructed from single indicators, except immunisation rate. Column (4) regression only covers 2009–2012 period due to the inconsistency definition of complete immunisation in 2013 household survey.

Similar to the spending interaction, one anticipated bias from this positive interaction is that the service access interaction is due to a common reaction to higher-level government policies, rather than strategic. To rule out this possibility, I estimate the same model using a placebo matrix which assigns random neighbours to each district that is not based on contiguity or distance. Using placebo matrix, as shown in panel (C), immunisation outcome shows a high significant result. This means the interaction on immunisation is potentially a common improvement on districts' service access due to the nationwide policies, rather than strategic interaction. A recent review of the health sector from the World Bank[10] finds that funding for the health centres managing immunisation, including centrally managed hospitals and community health centre (PUSKESMAS), has increased in recent years. Also,

[10] Source: http://www.anggaran.depkeu.go.id/content/Publikasi/seminar%20bidang%20kesehatan/WBGHealth%20PER%20Update%20Phase2%20MOF%20Seminar_121417.pdf.

immunisation has been one of the conditionality criteria for the conditional cash transfer program (*Program Keluarga Harapan*, PKH) which rolled out in 2007. PKH provides quarterly cash transfer to the country's poorest households, defined as households with consumption per capita less than 80 percent of the local district's poverty line. To receive the transfers, households must meet certain requirements, including sending their children to school, attending pre- and post-natal check-ups, and completing vaccinations for children. To a lesser degree, both local infrastructure outcomes, access to safe water and safe sanitation, show significant results, albeit low. These results may be attributed to the nature of the placebo matrix, which still employs other geographical criteria other than distance, which can still lead to spurious interaction between districts that are not contiguous neighbours.

Out of the remaining single indicators which have already being ruled out from potential common reaction, both education outcomes, columns (2) and (3), show stronger interdependence, compared to other indicators. This finding in education services may be attributed to the central government campaign that stresses educational improvement as their number two priority in the Medium Term Development Plan (RPJMN) 2010–2014[11]. The campaign potentially incentivises districts to search out and replicate other districts' service delivery methods.

One of the advantages of the Spatial Durbin Model (SDM) is the possibility of distinguishing direct and indirect effects of the explanatory variables. Table 2.3 presents the direct and indirect effects of the total spending (*lnexpcap*) as one of the explanatory variables of service access. The table shows that the significance level varies across categories of outcome. Particularly for senior net enrolment (column 3) and access to safe water (column 6), the indirect effect outweighs the size and significance of the direct effect, which highlights the importance of benefit spillover from neighbours' districts spending for these two service outcomes. The overall results, though, suggest that the indirect effect is not significant across all categories of service access. This points out that service interaction cannot be solely explained by the benefit spillover from neighbouring districts' spending. Furthermore, based on the other spillover framework, districts potentially mimic each other due to the information spillover, favouring the existence of yardstick competition between regions.

[11] The Medium Term Development Plan (RPJMN) 2010–2014 can be accessed through: https://www.bappenas.go.id/files/rpjmn/RPJMN%202010-2014.pdf.

Table 2.3 Direct and indirect effect of spending on service access

		Education		*Health*		*Local infrastructure*	
VARIABLES	Index service (1)	NER junior (2)	NER senior (3)	Immunisation (4)	Assisted birth (5)	Safe water (6)	Safe sanitation (7)
	Straight contiguity matrix						
Neighbours' service access	0.205***	0.187***	0.148***	0.623***	0.154***	0.137***	0.080***
	[0.027]	[0.028]	[0.029]	[0.019]	[0.028]	[0.029]	[0.029]
[Main] lnexpcap	0.084	-0.528	0.580	1.874**	0.488	1.275	2.065***
	[0.056]	[0.818]	[0.828]	[0.832]	[0.728]	[0.866]	[0.589]
[Wx] lnexpcap	0.147	2.201	3.776***	0.403	-2.674**	0.284	0.741
	[0.097]	[1.407]	[1.426]	[1.309]	[1.253]	[1.495]	[1.014]
[Direct effect] lnexpcap	0.096*	-0.375	0.782	2.356**	0.401	1.559	2.110***
	[0.058]	[0.839]	[0.847]	[0.989]	[0.746]	[1.705]	[0.603]
[Indirect effect] lnexpcap	0.195*	2.477	4.277***	3.829	-2.697**	1.789**	1.004
	[0.106]	[1.513]	[1.475]	[2.912]	[1.309]	[0.859]	[0.994]
R-squared	0.055	0.057	0.001	0.032	0.001	0.057	0.041
Number of groups	436	436	436	436	436	436	436
Observations	1744	1744	1744	1308	1744	1744	1744

Notes: Standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1. All regressions include district-specific controls, local government expenditure, higher-level (province) government expenditure, and time (year) trend. District-specific controls are GRDP per capita, population, poverty rate, poverty gap, and share of urban population. Index service is constructed from single service indicators, except immunisation rate. Column (4) regression only covers 2009–2012 period due to the inconsistent definition of complete immunisation in the 2013 household survey.

2.7 Robustness check

The empirical analyses end with some robustness checks. First, for both the spending and the service interdependence model, I introduce an alternative definition of geographical neighbours which uses the inverse distance, assigning more weight for closer neighbouring districts, compared to those further away. Using inverse distance, most of the estimated coefficients are found to be higher compared to the contiguity matrix, but they are not statistically different from each other, except for administrative spending and access to safe sanitation.

Table 2.4 presents the comparison between both matrices for spending interdependence. On total expenditure interaction (column 1), using inverse distance matrix results in a 0.14 % increase in own district spending for every 1% increase of neighbouring districts' spending, with a significance level of $\alpha = 0.05$. Capital-based spending (column 2 and column 7)

maintains a higher interaction coefficient compared to other sectors. However, administration spending loses its significance. Similar to personnel spending, one possible explanation for this is because administration spending mostly covers overhead costs for the local civil servants, which serves as one of the variables that are used in the transfer allocation formula.

Table 2.4 Robustness checks for expenditure interdependence

	By economic classification				*By sectoral classification*			
VARIABLES	Total (1)	Capital (2)	Personnel (3)	Goods (4)	Admin (5)	Infrastructure (6)	Education (7)	Health (8)
	(A) Straight contiguity matrix							
Neighbours' expenditure	0.083**	0.238***	0.050	0.133***	0.065**	0.125***	0.030	0.062**
	[0.033]	[0.029]	[0.034]	[0.031]	[0.029]	[0.031]	[0.030]	[0.029]
R-squared	0.657	0.335	0.693	0.665	0.023	0.474	0.560	0.571
	(B) Inverse distance matrix for districts in one province							
Neighbours' expenditure	0.088*	0.206***	0.054	0.169***	0.090*	0.237***	0.050	0.181***
	[0.050]	[0.044]	[0.055]	[0.047]	[0.047]	[0.042]	[0.049]	[0.042]
R-squared	0.888	0.675	0.746	0.808	0.255	0.589	0.262	0.689
	(C) Inverse distance matrix for districts in one region							
Neighbours' expenditure	0.143**	0.301***	0.035	0.290***	0.099	0.344***	0.034	0.243***
	[0.071]	[0.060]	[0.082]	[0.062]	[0.072]	[0.057]	[0.075]	[0.065]
R-squared	0.822	0.612	0.731	0.779	0.131	0.565	0.258	0.694
Number of groups	306	306	306	306	306	306	306	306
Observations	1,224	1,224	1,224	1,224	1,224	1,224	1,224	1,224

Notes: Standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1. Panel A regression includes time (year) trend. Panel B and C regressions include time (year) fixed effects. All regressions include district-specific controls and higher-level (province) government expenditure. District-specific controls are transfer per capita, GRDP per capita, population, poverty rate and poverty gap.

Similarly, Table 2.5 presents comparison between both types of matrices for service interdependence. Using the inverse distance matrix, an increase of one unit in neighbours' service access (column 1) is associated with a 0.4 unit increase in the service access index of own district. Ruling out the immunisation rate due to the possible common reaction, education access (column 2 and column 3) has higher interdependence compared to other single indicators, while safe sanitation (column 7) loses its significance level. This may indicate that local infrastructure interaction is more likely to appear in districts that share the same border compared to those districts that are adjacent but not contiguous.

Overall, these results suggest that spending and service interaction are not statistically sensitive to the definition of geographical neighbourhood. The higher interdependence coefficient is anticipated since the contiguity definition of 'neighbourhood' restricts neighbours to common border-sharing regions, while data availability is limited to an incomplete sample of districts. Therefore, using the inverse distance definition, districts are assigned more neighbours compared to the straight contiguity method.

Table 2.5 Robustness checks for service access interdependence

		Education		*Health*		*Local infrastructure*	
VARIABLES	Index service (1)	NER junior (2)	NER senior (3)	Immunisation (4)	Assisted birth (5)	Safe water (6)	Safe sanitation (7)
	(A) Straight contiguity matrix						
Neighbours' service access	0.205***	0.187***	0.148***	0.623***	0.154***	0.159***	0.096***
	[0.027]	[0.028]	[0.029]	[0.019]	[0.028]	[0.029]	[0.029]
R-squared	0.055	0.057	0.001	0.032	0.001	0.045	0.048
	(B) Inverse distance matrix for districts in one province						
Neighbours' service access	0.268***	0.275***	0.250***	0.019	0.201***	0.195***	0.032
	[0.035]	[0.034]	[0.035]	[0.049]	[0.037]	[0.039]	[0.041]
R-squared	0.276	0.139	0.017	0.084	0.348	0.240	0.186
	(C) Inverse distance matrix for districts in one region						
Neighbours' service access	0.418***	0.397***	0.375***	-0.117	0.281***	0.247***	-0.020
	[0.052]	[0.051]	[0.052]	[0.085]	[0.059]	[0.063]	[0.066]
R-squared	0.201	0.133	0.008	0.001	0.317	0.214	0.131
Number of groups	436	436	436	436	436	436	436
Observations	1744	1744	1744	1308	1744	1744	1744

Notes: Standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1. Panel A regression includes time (year) trend. Panel B and C regressions include time (year) fixed effects. All regressions include district-specific controls, expenditure per capita, higher-level (province) government expenditure, and time (year) trend. District-specific controls are GRDP per capita, population, poverty rate, poverty gap, and share of urban population. Index service is constructed from single indicators, except immunisation rate. Column (4) regression only covers 2009–2012 period due to the inconsistent definition of complete immunisation in the 2013 household survey.

Further robustness checks were done for the spending interdependence model, to account for the large number of districts which needed to be excluded due to the missing observations. I constructed a complete sample of districts using the imputation method in STATA[12] and then evaluated whether it yielded findings suggestive of positive interdependence. The imputation

[12] The analysis is performed using the multiple imputation (`mi`) command package: https://www.stata.com/manuals/mi.pdf.

methods in STATA assume that missing data are MAR (missing at random). Missing data are said to be MAR if the probability that data are missing does not depend on unobserved data but may depend on observed data. Under MAR, the missing data values do not contain any additional information given observed data about the missing-data mechanism. Multivariate imputations were conducted for 1000 iterations, with the number of imputed observations varying across variables, between 0.3% and 9% of the total observations, as presented in Table 2.10 in the Appendix. Table 2.6 shows the results of different configurations of contiguity weighting matrices and the placebo matrix. All estimated coefficients of neighbours' spending maintain the same statistical significance as in the incomplete sample, while education spending still has no significance coefficients. This suggests that the results from the incomplete dataset yield consistent results as if the dataset were a complete sample.

Table 2.6 Estimation results for expenditure interdependence using imputed dataset

		By economic classification				*By sectoral classification*		
VARIABLES	Total (1)	Capital (2)	Personnel (3)	Goods (4)	Admin (5)	Infrastructure (6)	Education (7)	Health (8)
	(A) Straight contiguity matrix							
Neighbours' expenditure	0.139***	0.275***	0.146***	0.201***	0.103***	0.169***	0.043	0.097***
	[0.031]	[0.026]	[0.031]	[0.028]	[0.029]	[0.028]	[0.034]	[0.031]
	(B) Straight contiguity matrix							
Neighbours' expenditure	0.127***	0.276***	0.069**	0.189***	0.102***	0.168***	0.036	0.081***
	[0.031]	[0.026]	[0.033]	[0.028]	[0.029]	[0.028]	[0.034]	[0.031]
	(C) Placebo matrix							
Neighbours' expenditure	0.157	-0.017	0.076	0.043	0.362	0.311	-0.101	-0.114
	[0.320]	[0.240]	[0.360]	[0.282]	[0.268]	[0.295]	[0.296]	[0.337]
Number of groups	491	491	491	491	491	491	491	491
Observations	1964	1964	1964	1964	1964	1964	1964	1964

Notes: Standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1. Panel B regression includes time (year) trend. Placebo matrix is constructed using principal component analysis that constructs similarity between districts based on area and length of the district. All regressions include district-specific controls and higher-level (province) government expenditure. District-specific controls are transfer per capita, GRDP per capita, population, poverty rate and poverty gap.

2.8 Conclusion

This paper has contributed to the empirical debate on spatial interaction in fiscal policy decision making at the local level in developing countries. It had two aims: 1) to investigate the presence of spatial interdependencies in the local public expenditure decision process, and 2) to investigate whether the interdependence also persists in service access indicators. The Indonesian context, in the period of 2009–2013, was used as a useful case study, given the institutional change in the form of political decentralisation that favours greater local interaction, while on the other hand still relying mostly on the intergovernmental transfers for revenue.

Based on the existing literature, two spatial Durbin models (SDM) were constructed, one for spending and one for service access. Estimation results indicate the existence of positive horizontal interdependence in fiscal policy in Indonesia and point to spatial interdependence as one of the factors influencing public spending levels and service access at district level. Robustness checks show that positive and significant spillover effects remain when using a different weight matrix and controlling for the complete sample size and nationwide policies.

This study has some caveats that may inform future research topics. First, this paper treats spending as exogenous in the service access equation. One common alternative to overcome this is to include lagged values for service access and uses internal instruments which are the lagged value of the spending variables. However, at the time of writing, there is no established estimators in spatial econometrics, so far, that allows spatial dynamic models with one or more endogenous variables. Secondly, the weight matrices used in this study are based on geographical criteria. While previous studies mentioned the importance of geographical criteria as an initial step, it would be useful to check how the interaction differs when including socio-economic criteria in defining neighbouring districts.

Finally, this study confirms that decentralisation, both fiscal and political, can indeed provide additional incentives for local government in providing better service access through strategic competition between districts. This competition, which is reflected in the existence of positive interdependence, potentially boosts local government accountability and efficiency in delivering service access. For instance, information about positive policy outcomes in other districts that spills to the citizens of one district can help them to hold their government more accountable in meeting their needs.

Regarding policy implications, this study confirms that governments do not necessarily face a trade-off between giving regions autonomy and choosing more efficient public spending and service provision strategies. Instead, central government can nurture this strategic competition between districts by relying on knock-on effects of a good policy in one district to another district. One possible way to do this is to reward districts that perform well, which will incite other regions to do the same things. This finding serves as additional support for the development of the performance-based grants that are currently under discussion in many developing countries. The positive interaction between districts overcomes the concern that this type of grant will heavily penalise poorly performing districts, while on the other hand confirms that giving a grant to one region can also benefit other regions.

Appendix

Table 2.7 Descriptive summary for spending interdependence model

Variable		Obs	Mean	Std. Dev.	Min	Max
Total expenditure per capita	in million IDR	1,530	3.15	3.11	0.23	38.20
Capital expenditure per capita	in million IDR	1,530	0.85	1.35	0.03	18.50
Goods expenditure per capita	in million IDR	1,530	0.60	0.81	0.02	13.10
Personnel expenditure per capita	in million IDR	1,530	1.44	1.02	0.00	21.80
Admin expenditure per capita	in million IDR	1,530	1.01	1.39	0.01	20.00
Education expenditure per capita	in million IDR	1,530	0.95	0.60	0.00	11.70
Health expenditure per capita	in million IDR	1,530	0.31	0.32	0.01	6.07
Infrastructure expenditure per capita	in million IDR	1,530	0.53	0.87	0.01	13.70
Province expenditure per capita	in million IDR	1,530	0.84	0.91	0.09	7.10
Province capital expenditure per capita	in million IDR	1,530	0.19	0.28	0.01	2.98
Province goods expenditure per capita	in million IDR	1,530	0.20	0.25	0.02	1.53
Province personnel expenditure per capita	in million IDR	1,530	0.15	0.11	0.02	0.54
Province admin expenditure per capita	in million IDR	1,530	0.43	0.54	0.04	4.83
Province education expenditure per capita	in million IDR	1,530	0.06	0.06	0.01	0.31
Province health expenditure per capita	in million IDR	1,530	0.06	0.06	0.01	0.38
Province infra expenditure per capita	in million IDR	1,530	0.15	0.19	0.01	1.04
Transfer per capita	in million IDR	1,530	3.13	3.53	0.14	57.50
GRDP per capita	in million IDR	1,530	7.67	9.09	0.33	164.95
povtrate	in %	1,530	14.25	7.76	1.52	47.73
povtgap	in %	1,530	2.37	1.87	0.15	19.16
population	in person	1,530	512,416	583,152	18,365	5,202,097

Table 2.8 Descriptive summary for service interdependence model

Variable		Obs	Mean	Std. Dev.	Min	Max
Service index (pc2)		2,180	0.16	1.62	-6.84	3.66
NER_junior	in %	2,180	67.67	10.63	9.34	95.33
NER_senior	in %	2,180	49.95	12.53	3.40	86.35
Immunisation rate	in %	2,180	84.03	11.20	5.66	100.00
Birth assisted by skilled worker	in %	2,180	78.34	18.95	1.54	100.00
Safe sanitation	in %	2,180	63.06	17.10	4.55	96.75
Safe water	in %	2,180	56.47	20.64	0.93	100.00
Population number	in person	2,180	510,852	578,918	12,660	5,202,097
Poverty rate	in %	2,180	13.98	7.95	1.33	51.91
Poverty gap	in %	2,180	2.34	1.94	0.08	18.28
Share of population in urban area	in %	2,180	38.93	31.36	0.00	100.00
GRDP per capita	in million IDR	2,180	8.31	9.60	0.33	164.95
District total expenditure per capita	in million IDR	2,180	3.64	12.40	0.15	553.00
Province total expenditure per capita	in million IDR	2,180	0.89	1.00	0.09	7.10

Table 2.9 Results of diagnostic tests

	Spending interdependence		Service interdependence	
Type of test	**Coefficient**	**P-value**	**Coefficient**	**P-value**
Hausman (SDM FE vs SDM RE)	68.72	0.00	175.23	0.00
LM test (SDM FE vs SAR FE)	16.70	0.01	28.25	0.00
LM test (SDM FE vs SEM FE)	13.96	0.05	29.23	0.00
	SDM FE	SAC FE	SDM FE	SAC FE
Information criteria (AIC)	-1988.24	-1986.52	1853.47	1837.83

Notes: Regressions use straight contiguity weight matrices. Spending interdependence regression uses total expenditure per capita as outcome variable and controls for transfer per capita, GRDP per capita, population, poverty rate and poverty gap, higher-level (province) government expenditure, and time (year) trend. Service interdependence regression uses service index as outcome variable and controls for GRDP per capita, population, poverty rap, poverty gap, share of urban population, expenditure per capita, higher-level (province) government expenditure, and time (year) trend.

Table 2.10 Number of imputed observations

Variable name	Number of observations			Imputed percentage from total
	Complete	Imputed	Total	
Total expenditure per capita	2414	41	2455	1.7%
Capital expenditure per capita	2412	43	2455	1.8%
Goods expenditure per capita	2412	43	2455	1.8%
Personnel expenditure per capita	2413	42	2455	1.7%
Admin expenditure per capita	2236	219	2455	8.9%
Education expenditure per capita	2233	222	2455	9.0%
Health expenditure per capita	2233	222	2455	9.0%
Infrastructure expenditure per capita	2234	221	2455	9.0%
Transfer per capita	2402	53	2455	2.2%
Poverty rate	2443	12	2455	0.5%
Poverty gap	2443	12	2455	0.5%
GRDP per capita	2447	8	2455	0.3%

Table 2.11 Results for expenditure interdependence, straight contiguity matrix

VARIABLES	Total	By econ classification			By sectoral classification			
		Capital	Personnel	Goods	Admin	Infrastructure	Education	Health
Spatial variables:								
ln(exp)cap	0.083**	0.238***	0.050	0.133***	0.065**	0.125***	0.030	0.062**
	[0.033]	[0.029]	[0.034]	[0.031]	[0.029]	[0.031]	[0.030]	[0.029]
Main:								
lntransfercap	0.476***	0.928***	0.223***	0.516***	0.339***	0.822***	0.707***	0.610***
	[0.026]	[0.060]	[0.036]	[0.037]	[0.046]	[0.063]	[0.058]	[0.041]
lnpop	-0.433***	-0.247*	-0.593***	-0.457***	-0.481***	-0.202	-0.175	-0.367***
	[0.061]	[0.142]	[0.084]	[0.089]	[0.108]	[0.149]	[0.137]	[0.096]
povtrate	0.019***	0.029**	0.009	0.016**	0.018**	0.030**	0.004	0.016**
	[0.005]	[0.012]	[0.007]	[0.007]	[0.009]	[0.012]	[0.011]	[0.008]
povtgap	-0.023***	-0.020	-0.020**	-0.028***	-0.025**	-0.004	-0.021	-0.013
	[0.007]	[0.015]	[0.009]	[0.010]	[0.012]	[0.016]	[0.015]	[0.010]
lngrdpcap	0.024	-0.108	0.131***	0.043	0.042	-0.028	0.134*	0.063
	[0.032]	[0.073]	[0.043]	[0.046]	[0.056]	[0.077]	[0.071]	[0.050]
ln(exp)capprov	0.002	-0.036	0.056	0.079*	-0.018	-0.034	-0.160**	-0.094**
	[0.026]	[0.051]	[0.047]	[0.047]	[0.040]	[0.060]	[0.066]	[0.041]
time trend	0.070***	-0.091***	0.141***	0.096***	0.104***	0.010	0.076***	0.117***
	[0.012]	[0.026]	[0.016]	[0.019]	[0.022]	[0.028]	[0.026]	[0.019]
Wx:								
lntransfercap	-0.068	-0.320***	-0.084	-0.054	0.125*	0.055	-0.205**	-0.072
	[0.042]	[0.095]	[0.053]	[0.059]	[0.069]	[0.100]	[0.090]	[0.065]
lnpop	-0.125	-0.492***	0.083	-0.053	0.193	-0.401**	-0.002	-0.202
	[0.081]	[0.185]	[0.111]	[0.116]	[0.140]	[0.195]	[0.179]	[0.126]
povtrate	-0.014*	-0.003	-0.029***	-0.017*	0.016	-0.002	-0.031**	-0.011
	[0.007]	[0.017]	[0.010]	[0.010]	[0.013]	[0.017]	[0.016]	[0.011]
povtgap	0.010	-0.016	0.020*	0.033***	-0.034**	-0.007	0.027	0.005
	[0.009]	[0.021]	[0.012]	[0.013]	[0.016]	[0.022]	[0.020]	[0.014]
lngrdpcap	-0.022	-0.211*	0.046	-0.034	-0.052	-0.183	0.061	-0.035
	[0.048]	[0.112]	[0.066]	[0.069]	[0.085]	[0.117]	[0.108]	[0.076]
ln(exp)capprov	0.047	0.236***	-0.020	-0.040	0.115**	0.066	0.253***	0.149***
	[0.031]	[0.058]	[0.059]	[0.056]	[0.045]	[0.070]	[0.076]	[0.049]
time trend	-0.018	0.090***	-0.023	-0.020	-0.092***	-0.012	-0.015	-0.045*

	[0.015]	[0.031]	[0.020]	[0.023]	[0.026]	[0.034]	[0.031]	[0.023]
Observations	1,224	1,224	1,224	1,224	1,224	1,224	1,224	1,224
R-squared	0.657	0.335	0.693	0.665	0.023	0.474	0.560	0.571
Number of groups	306	306	306	306	306	306	306	306

Notes: Standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1. Regression includes time (year) trend.

Table 2.12 Results for expenditure interdependence, inverse distance matrix for districts in one province

VARIABLES	Total	*By econ classification*			*By sectoral classification*			
		Capital	Personnel	Goods	Admin	Infrastructure	Education	Health
Spatial variables:								
ln(exp)cap	0.088*	0.206***	0.054	0.169***	0.090*	0.237***	0.050	0.181***
	[0.050]	[0.044]	[0.055]	[0.047]	[0.047]	[0.042]	[0.049]	[0.042]
Main:								
lntransfercap	0.493***	0.948***	0.264***	0.536***	0.315***	0.775***	0.783***	0.630***
	[0.026]	[0.059]	[0.037]	[0.038]	[0.046]	[0.062]	[0.060]	[0.042]
lnpop	-0.423***	-0.162	-0.590***	-0.430***	-0.455***	-0.207	-0.119	-0.350***
	[0.058]	[0.129]	[0.081]	[0.085]	[0.102]	[0.137]	[0.132]	[0.092]
povtrate	0.017***	0.018	0.009	0.015*	0.024**	0.022*	0.008	0.013
	[0.005]	[0.012]	[0.007]	[0.008]	[0.009]	[0.013]	[0.012]	[0.008]
povtgap	-0.021***	-0.006	-0.023**	-0.027***	-0.015	-0.006	-0.021	-0.013
	[0.007]	[0.015]	[0.009]	[0.010]	[0.012]	[0.016]	[0.015]	[0.010]
lngrdpcap	0.037	-0.029	0.119***	0.062	0.035	-0.008	0.147**	0.060
	[0.031]	[0.068]	[0.043]	[0.045]	[0.054]	[0.072]	[0.070]	[0.049]
ln(exp)capprov	-0.071	-0.123	-0.010	-0.017	-0.174**	-0.140	-0.403**	-0.260***
	[0.045]	[0.115]	[0.072]	[0.088]	[0.069]	[0.095]	[0.184]	[0.072]
Wx:								
lntransfercap	-0.043	-0.345**	0.093	-0.022	-0.128	-0.142	-0.298**	-0.246**
	[0.067]	[0.139]	[0.085]	[0.092]	[0.103]	[0.147]	[0.137]	[0.098]
lnpop	-0.129	-0.236	0.152	-0.147	0.475**	-0.471	0.451	-0.219
	[0.133]	[0.284]	[0.188]	[0.191]	[0.226]	[0.305]	[0.291]	[0.204]
povtrate	-0.020**	-0.017	-0.008	-0.020	-0.033**	-0.013	-0.029	-0.017
	[0.009]	[0.019]	[0.012]	[0.013]	[0.015]	[0.021]	[0.020]	[0.014]
povtgap	0.033**	0.045	0.006	0.047**	0.002	0.033	0.056*	0.018
	[0.014]	[0.031]	[0.020]	[0.020]	[0.024]	[0.033]	[0.032]	[0.022]
lngrdpcap	0.005	-0.302*	0.131	0.062	0.117	-0.212	0.404**	0.036
	[0.070]	[0.160]	[0.099]	[0.104]	[0.123]	[0.164]	[0.160]	[0.112]
ln(exp)capprov	0.086*	0.158	0.056	0.039	0.228***	0.128	0.430**	0.279***
	[0.047]	[0.117]	[0.080]	[0.091]	[0.072]	[0.099]	[0.186]	[0.075]
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,224	1,224	1,224	1,224	1,224	1,224	1,224	1,224
R-squared	0.888	0.675	0.746	0.808	0.255	0.589	0.262	0.689

Number of groups	306	306	306	306	306	306	306	306

Notes: Standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1. Regression includes time (year) fixed effects.
Neighbouring districts are defined as districts who located at the same province.

Table 2.13 Results for expenditure interdependence, inverse distance matrix for districts in one region

VARIABLES	Total	*By econ classification*			*By sectoral classification*			
		Capital	Personnel	Goods	Admin	Infrastructure	Education	Health
Spatial variables:								
ln(exp)cap	0.143**	0.301***	0.035	0.290***	0.099	0.344***	0.034	0.243***
	[0.071]	[0.060]	[0.082]	[0.062]	[0.072]	[0.057]	[0.075]	[0.065]
Main:								
lntransfercap	0.498***	0.943***	0.274***	0.542***	0.315***	0.784***	0.763***	0.615***
	[0.023]	[0.052]	[0.032]	[0.034]	[0.041]	[0.057]	[0.053]	[0.037]
lnpop	-0.408***	-0.113	-0.599***	-0.417***	-0.491***	-0.158	-0.172	-0.366***
	[0.052]	[0.115]	[0.073]	[0.076]	[0.092]	[0.127]	[0.119]	[0.084]
povtrate	0.016***	0.020**	0.008	0.016**	0.018**	0.022**	0.012	0.009
	[0.004]	[0.010]	[0.006]	[0.006]	[0.008]	[0.011]	[0.010]	[0.007]
povtgap	-0.021***	-0.006	-0.022***	-0.029***	-0.014	-0.005	-0.020	-0.013
	[0.006]	[0.013]	[0.008]	[0.009]	[0.010]	[0.014]	[0.014]	[0.009]
lngrdpcap	0.040	-0.010	0.121***	0.059	0.015	0.016	0.132**	0.055
	[0.027]	[0.061]	[0.038]	[0.040]	[0.049]	[0.067]	[0.063]	[0.044]
ln(exp)capprov	0.008	0.034	0.079**	0.057	0.013	-0.050	-0.056	-0.078**
	[0.020]	[0.040]	[0.039]	[0.038]	[0.032]	[0.047]	[0.051]	[0.031]
Wx:								
lntransfercap	-0.166	-0.460**	0.090	-0.189	-0.096	0.015	0.033	-0.119
	[0.104]	[0.222]	[0.136]	[0.140]	[0.166]	[0.242]	[0.230]	[0.161]
lnpop	-0.359**	-0.889**	0.133	-0.139	0.853***	-1.064***	1.001***	-0.114
	[0.176]	[0.363]	[0.245]	[0.240]	[0.286]	[0.408]	[0.374]	[0.260]
povtrate	-0.033***	-0.043*	-0.013	-0.034**	-0.029	-0.040	-0.033	-0.003
	[0.011]	[0.024]	[0.016]	[0.015]	[0.019]	[0.026]	[0.024]	[0.017]
povtgap	0.048***	0.065*	0.013	0.074***	0.013	0.030	0.047	-0.005
	[0.016]	[0.035]	[0.022]	[0.022]	[0.027]	[0.038]	[0.035]	[0.025]
lngrdpcap	-0.110	-1.067***	0.266	0.066	0.324	-0.946***	0.714**	0.032
	[0.122]	[0.298]	[0.177]	[0.177]	[0.219]	[0.303]	[0.291]	[0.204]
ln(exp)capprov	-0.003	0.014	-0.103	-0.069	0.042	0.078	0.138	0.155***
	[0.037]	[0.072]	[0.076]	[0.062]	[0.051]	[0.085]	[0.087]	[0.056]
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	1,530	1,530	1,530	1,530	1,530	1,530	1,530	1,530
R-squared	0.822	0.612	0.731	0.779	0.131	0.565	0.258	0.694

Number of groups	306	306	306	306	306	306	306	306

Notes: Standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1. Regression includes time (year) fixed effects.
Neighbouring districts are defined as districts who located at the same region group. The region group are Java, Sumatera, Kalimantan, Sulawesi, Bali- Nusa Tenggara and Maluku-Papua.

Table 2.14 Results for service access interdependence, straight contiguity matrix

		Education		*Health*		*Local infrastructure*	
VARIABLES	Index service (1)	NER junior (2)	NER senior (3)	Immunisation (4)	Assisted birth (5)	Safe water (6)	Safe sanitation (7)
Spatial variables:							
Neighbours' service access	0.205***	0.187***	0.148***	0.623***	0.154***	0.137***	0.080***
	[0.027]	[0.028]	[0.029]	[0.019]	[0.028]	[0.029]	[0.029]
Main:							
lnexpcap	0.084	-0.528	0.580	1.874**	0.488	1.234	2.065***
	[0.056]	[0.818]	[0.828]	[0.832]	[0.728]	[0.846]	[0.589]
lnpop	-0.118	-0.182	-3.137	2.399	0.019	2.424	-2.105
	[0.157]	[2.280]	[2.305]	[2.025]	[2.029]	[2.357]	[1.643]
povtrate	-0.033***	0.257	-0.267	-0.416**	-0.714***	-0.291	-0.403***
	[0.013]	[0.186]	[0.188]	[0.174]	[0.165]	[0.192]	[0.134]
povtgap	0.027	0.003	0.287	-0.612***	-0.170	0.674***	0.302*
	[0.017]	[0.251]	[0.255]	[0.220]	[0.224]	[0.260]	[0.181]
urban_rate	0.013***	-0.001	0.082**	0.008	0.147***	0.169***	0.156***
	[0.003]	[0.037]	[0.037]	[0.030]	[0.033]	[0.038]	[0.026]
lngrdpcap	-0.116	-0.338	-1.524	1.127	1.566	-1.814	-2.248***
	[0.080]	[1.157]	[1.171]	[1.008]	[1.031]	[1.198]	[0.833]
lnexpcapprov	-0.136*	-0.455	1.300	2.911**	-3.009***	-2.974***	-1.179
	[0.076]	[1.101]	[1.114]	[1.344]	[0.981]	[1.139]	[0.793]
time trend	0.202***	0.259	0.543	4.616***	2.164***	3.398***	2.200***
	[0.036]	[0.515]	[0.521]	[0.564]	[0.459]	[0.533]	[0.371]
Wx:							
lnexpcap	0.147	2.201	3.776***	0.403	-2.674**	-0.022	0.741
	[0.097]	[1.407]	[1.426]	[1.309]	[1.253]	[1.457]	[1.014]
lnpop	0.253	-3.032	4.379	1.026	-3.847	5.909*	7.737***
	[0.212]	[3.070]	[3.108]	[2.688]	[2.735]	[3.179]	[2.212]
povtrate	0.061***	0.071	0.784***	0.370	0.538**	0.377	0.582***
	[0.016]	[0.238]	[0.241]	[0.228]	[0.212]	[0.246]	[0.171]
povtgap	-0.010	-0.129	-0.567	0.496*	0.600*	-0.182	0.222
	[0.024]	[0.351]	[0.354]	[0.298]	[0.313]	[0.362]	[0.253]
urban_rate	-0.002	0.133**	-0.003	-0.045	-0.139***	-0.058	-0.053
	[0.004]	[0.060]	[0.061]	[0.048]	[0.054]	[0.062]	[0.043]
lngrdpcap	0.042	0.299	0.152	-1.868	-0.481	-1.350	2.711**

	[0.104]	[1.503]	[1.521]	[1.284]	[1.340]	[1.553]	[1.082]
lnexpcapprov	0.116	0.530	-1.392	-0.163	2.589**	2.675**	0.889
	[0.086]	[1.246]	[1.260]	[1.503]	[1.110]	[1.290]	[0.897]
time trend	-0.009	0.846	0.963	-3.639***	0.012	-1.412**	-1.153***
	[0.040]	[0.577]	[0.587]	[0.626]	[0.517]	[0.599]	[0.417]
Observations	1,744	1,744	1,744	1,308	1,744	1,744	1,744
R-squared	0.055	0.057	0.001	0.032	0.001	0.057	0.041
Number of groups	436	436	436	436	436	436	436

Notes: Standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1. Regression includes time (year) trend. Index service is constructed from single indicators, except immunisation rate. Column (4) regression only covers 2009–2012 period due to the inconsistent definition of complete immunisation in 2013 household survey.

Table 2.15 Results for service access interdependence, inverse distance matrix for districts in one province

VARIABLES	Index service (1)	*Education* NER junior (2)	NER senior (3)	*Health* Immunisation (4)	Assisted birth (5)	*Local infrastructure* Safe water (6)	Safe sanitation (7)
Spatial variables:							
Neighbours' service access	0.268***	0.275***	0.250***	0.019	0.201***	0.195***	0.032
	[0.035]	[0.034]	[0.035]	[0.049]	[0.037]	[0.039]	[0.041]
Main:							
lnexpcap	0.072	-0.643	0.217	0.828	0.339	1.447*	2.187***
	[0.051]	[0.742]	[0.749]	[0.617]	[0.670]	[0.775]	[0.538]
lnpop	-0.189	-0.436	-3.674*	0.447	-0.675	1.743	-2.083
	[0.137]	[1.988]	[2.004]	[1.444]	[1.794]	[2.071]	[1.444]
povtrate	-0.048***	0.123	-0.536***	-0.494***	-0.728***	-0.456**	-0.281**
	[0.012]	[0.173]	[0.175]	[0.132]	[0.157]	[0.181]	[0.126]
povtgap	0.028*	0.109	0.382*	-0.485***	-0.175	0.486**	0.228
	[0.015]	[0.223]	[0.225]	[0.159]	[0.202]	[0.233]	[0.162]
urban_rate	0.015***	0.005	0.126***	0.013	0.132***	0.186***	0.143***
	[0.002]	[0.034]	[0.034]	[0.022]	[0.031]	[0.035]	[0.025]
lngrdpcap	-0.169**	-0.885	-1.734*	-0.240	0.907	-1.893*	-2.399***
	[0.070]	[1.020]	[1.029]	[0.717]	[0.920]	[1.064]	[0.740]
lnexpcapprov	-0.142	0.456	-0.762	-0.001	-3.042**	-2.382	-0.683
	[0.106]	[1.538]	[1.552]	[0.000]	[1.388]	[1.604]	[1.115]
Wx:							
lnexpcap	0.139	0.180	2.269	-5.257***	-0.094	2.434	1.713
	[0.125]	[1.804]	[1.824]	[1.415]	[1.629]	[1.894]	[1.319]
lnpop	0.525	-9.226*	3.798	-1.790	9.741**	7.207	19.002***
	[0.354]	[5.154]	[5.181]	[3.793]	[4.647]	[5.362]	[3.756]
povtrate	0.066***	-0.365	1.045***	0.707***	0.563**	1.133***	0.334
	[0.020]	[0.294]	[0.297]	[0.224]	[0.265]	[0.307]	[0.213]
povtgap	0.031	1.107*	-0.414	-1.871***	0.731	-0.832	0.360
	[0.041]	[0.596]	[0.601]	[0.478]	[0.537]	[0.622]	[0.431]
urban_rate	-0.002	0.090	-0.139**	0.039	0.012	-0.037	0.073*
	[0.004]	[0.061]	[0.061]	[0.040]	[0.055]	[0.064]	[0.045]
lngrdpcap	0.196	0.625	0.635	1.685	3.040	-1.594	5.182***
	[0.147]	[2.136]	[2.157]	[1.498]	[1.935]	[2.231]	[1.550]

lnexpcapprov	0.096	-1.090	0.958	0.120	2.388	1.908	0.463
	[0.114]	[1.659]	[1.673]	[0.454]	[1.497]	[1.729]	[1.202]
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,180	2,180	2,180	1,744	2,180	2,180	2,180
R-squared	0.276	0.139	0.017	0.084	0.348	0.240	0.186
Number of groups	436	436	436	436	436	436	436

Notes: Standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1. Regression includes time (year) fixed effects. Index service is constructed from single indicators, except immunisation rate. Column (4) regression only covers 2009–2012 period due to the inconsistent definition of complete immunisation in 2013 household survey. Neighbouring districts are defined as districts who located at the same province.

Table 2.16 Results for service access interdependence, inverse distance matrix for districts in one region

		Education		*Health*		*Local infrastructure*	
VARIABLES	Index service (1)	NER junior (2)	NER senior (3)	Immunisation (4)	Assisted birth (5)	Safe water (6)	Safe sanitation (7)
Spatial variables:							
Neighbours' service access	0.418***	0.397***	0.375***	-0.117	0.281***	0.247***	-0.020
	[0.052]	[0.051]	[0.052]	[0.085]	[0.059]	[0.063]	[0.066]
Main:							
lnexpcap	0.053	-0.928	0.195	0.746	0.307	1.308*	2.030***
	[0.051]	[0.742]	[0.749]	[0.617]	[0.670]	[0.774]	[0.537]
lnpop	-0.179	-0.268	-3.312*	0.265	-0.565	1.522	-2.193
	[0.136]	[1.987]	[2.003]	[1.441]	[1.796]	[2.070]	[1.442]
povtrate	-0.042***	0.143	-0.433***	-0.377***	-0.686***	-0.481***	-0.300***
	[0.011]	[0.160]	[0.161]	[0.123]	[0.145]	[0.167]	[0.116]
povtgap	0.024	0.089	0.363	-0.475***	-0.200	0.432*	0.164
	[0.015]	[0.224]	[0.226]	[0.158]	[0.202]	[0.233]	[0.162]
urban_rate	0.014***	-0.001	0.123***	0.006	0.129***	0.179***	0.155***
	[0.002]	[0.033]	[0.033]	[0.022]	[0.030]	[0.035]	[0.024]
lngrdpcap	-0.148**	-0.586	-1.908*	0.235	1.198	-1.737*	-2.154***
	[0.069]	[1.012]	[1.020]	[0.709]	[0.914]	[1.055]	[0.732]
lnexpcapprov	-0.099**	-1.093	-0.779	2.428***	-1.369**	-0.082	0.246
	[0.050]	[0.733]	[0.739]	[0.819]	[0.661]	[0.764]	[0.530]
Wx:							
lnexpcap	0.349*	4.935*	1.899	-9.282***	-1.721	4.184	5.593***
	[0.202]	[2.935]	[2.960]	[2.241]	[2.652]	[3.083]	[2.150]
lnpop	0.946	-9.024	8.148	-11.037*	11.707	8.356	33.104***
	[0.577]	[8.421]	[8.478]	[6.335]	[7.643]	[8.758]	[6.169]
povtrate	0.102***	-0.238	1.445***	0.363	0.953***	1.504***	0.600**
	[0.027]	[0.397]	[0.402]	[0.294]	[0.359]	[0.415]	[0.287]
povtgap	-0.009	0.756	-1.134	-1.899***	0.549	-0.311	0.651
	[0.053]	[0.773]	[0.776]	[0.640]	[0.695]	[0.814]	[0.562]
urban_rate	0.005	0.310***	-0.160	0.026	0.037	0.056	0.035
	[0.007]	[0.102]	[0.102]	[0.066]	[0.092]	[0.108]	[0.074]
lngrdpcap	0.231	-1.205	2.878	1.591	2.484	-2.051	7.121**
	[0.274]	[3.998]	[4.036]	[2.921]	[3.625]	[4.169]	[2.888]
lnexpcapprov	0.113	1.170	1.965	-4.452***	1.377	-1.229	-0.979

	[0.095]	[1.392]	[1.406]	[1.315]	[1.255]	[1.464]	[1.007]
Time FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Observations	2,180	2,180	2,180	1,744	2,180	2,180	2,180
R-squared	0.201	0.133	0.008	0.001	0.317	0.214	0.131
Number of groups	436	436	436	436	436	436	436

Notes: Standard errors in brackets, *** p<0.01, ** p<0.05, * p<0.1. Regression includes time (year) fixed effects. Index service is constructed from single indicators, except immunisation rate. Column (4) regression only covers 2009–2012 period due to the inconsistent definition of complete immunisation in 2013 household survey. Neighbouring districts are defined as districts who located at the same region. The region group are Java, Sumatera, Kalimantan, Sulawesi, Bali-Nusa Tenggara and Maluku-Papua.

3 VILLAGE GOVERNMENT FORM AND LOCAL SERVICE ACCESS

3.1 Introduction

There is always prolonged debate about which levels of government should be decentralised, both politically and financially. The swinging pendulum on this matter often contrasts bureaucratic-based and citizen-based government forms. Bureaucratic-based government adopts more of a top-down approach with dominant roles of higher levels of government and appointed officials. Citizen-based governments reflect more autonomous government with directly elected officials. The bureaucratic form is seen as easier for coordination across government levels, while the other is expected to be more responsive to local needs while also being prone to be captured by local elites (Bardhan and Mookherjee, 2006).

In many developing countries, identifying the most effective community-level government form is a crucial issue in development, especially for rural areas where much service access is still lagging as a result of the extended centralised economy before decentralisation (Zhang et al., 2004). Therefore, empirical evaluation of these different government forms is needed. This paper contrasts those two distinctive government forms at the village level in Indonesia: *desa*, which is a more citizen-based government and *kelurahan*, which is a bureaucratic-based government.

To define the substance of its research, this study draws on propositions from two streams of related literature. The first stream of literature discusses the accountability perspective in which citizens can hold their service providers accountable. This horizontal accountability has been recognised as one of the critical elements in promoting better basic service delivery at the local level (Skoufias et al., 2014). The underlying rationale is that directly elected officials are more responsive to citizens' and communities' needs, given their increased accountability. Another stream of literature highlights the importance of administrative coordination and political alignment between government levels. Carr (2015) suggests that directly elected officials may have less administrative skills than appointed officials, who often have better management training as it is part of their career path. Also, a more democratic election system can induce considerably more fragmentation within government bodies as each leader will seek to support programs that benefit their own constituents, while the budget is funded from common pool resources (Tomsa, 2014). This fragmentation has

been perceived to have adverse effects on economic and service outcomes. However, Martinez-Bravo (2014) found that appointed officials are not neutral to electoral concerns since they also have high stakes in upper-level elections of the office, signalling vertical accountability across levels of government.

Indonesia serves as a useful empirical case, since its lowest administrative tier called village, has two types of government form: *desa* and *kelurahan*. Heads of *desa* are elected through village elections and salaries for the village apparatus are funded from the village budget. For *kelurahan*, the head is appointed by the head of the district, and all village employees are civil servants with their salaries funded from the district budget. This contradiction between *desa* and *kelurahan*, not only relating to the head or leader but also to the whole structure of the government body, gives a unique setting to examine the broader effects of government forms.

Although a large body of literature exists contrasting local government forms in developing countries, only a few studies have discussed this distinction at the lowest administrative tier. In Indonesia's case, much literature has discussed government forms at district levels, including Lewis (2017) and Skoufias et al. (2014). Their studies compare the performance of districts with directly elected district heads and indirectly elected district heads. The closest predecessor with a similar setting to this study is Martinez-Bravo (2014), who investigates the effects of different village government forms on electoral outcomes. She finds that villages with appointed heads (*kelurahan)* show stronger support for the dominant political party in district level elections. However, her study does not discuss the effects of this government form on local service access.

The purpose of this study is to examine whether there are differences in service outcomes between the two village level government forms between 2000 and 2011. In contrasting both forms, this paper investigates if switching to the more bureaucratic type of governance influences service access outcomes. Village governments, as the lowest administrative tier, are the closest government level to the people and therefore are crucial front-liners in service provision.

This study will be one of the first attempts to test this question at the lowest administrative tiers of government in the case of a developing country. Moreover, the results of this study may contribute to discussion on which types of accountability matter more for government

performance at the village level. This research can also inform perspectives on whether service access delivery should be decentralised down to the lowest government unit.

The rest of the paper will be structured as follows. The next section will discuss the context of the decentralisation policy, including the assignment of village government in Indonesia. Then, I will provide a brief review of current studies on government forms and how these are linked with the provision of service access. The fourth section will present the methodological approach including the theoretical framework and identification strategy, while the subsequent section will discuss data used in this paper and the choice of dependent variables. The sixth and seventh sections will present the results, including some tests to justify whether the identification strategy is valid. Then, I will present the results disaggregated by Java versus non-Java. The last two sections will further discuss the results and conclude which government form is more efficient in providing better service access and the implications for village level decentralisation policies.

3.2 Background

Indonesia consists of five levels of administration: central government, provinces, districts, sub-districts/*kecamatan,* and villages. In 2014, there were 34 provinces, 504 districts, around 7150 sub-districts and more than 80,000 villages in Indonesia. District government, both rural district/*kabupaten* and urban district/*kota,* has the responsibility of delivering basic public sector services (health, infrastructure, and education), together with other tasks including agriculture, communications, industry-trade, environment, land, cooperatives and labour. This assignment of services was also complemented by political decentralisation, indicated by the development of local democracy. Provincial and district parliament elections became more genuinely democratic in 1999. Also, the first batch of local heads of districts direct elections took place in 2005, and gradually all indirectly elected heads were replaced by directly elected ones by the end of 2010[13].

Each district consists of several sub-districts/*kecamatan*, which are essentially deconcentrated units of the district. Their heads are civil servants and they are selected by the district

[13] The initiation and staggered timing of direct elections of district heads has been judged as exogenously determined (Lewis, 2018; Skoufias et al., 2014). The timing of a direct election for districts which did not split depended on when the five-year term of the previous head had come to an end. For example, if a district head was last elected in January 1999 under the New Order -autocratic selection process, (s)he would have been subject to indirect election by the parliament in January 2004, and direct election in January 2009. When a district splits, the parent district retains the executive and the process of (direct) election continues as usual. For the 'child' district, the provincial government (in consultation with the Ministry of Home Affairs) appoints a new district head. This head will manage the affairs of the district until such time as an election can be arranged.

executives (Lewis, 2015). Each sub-district then consists of many villages, which is the lowest government tier in Indonesia. There are two types of villages, *desa*, which tend to be more rural and *kelurahan*, which are more urban. As such, most *kelurahan v*illages are located at the urban district/*kota.* At the end of 2014, there were 72,949 *desa*, which accounted for around 88 percent of the total villages in Indonesia, and 8,412 *kelurahan*.

The process of village classification followed Village Law No. 5/1979. Before 1979, village government varied across regions and its arrangement was largely based on local customs. This 1979 law aimed to achieve uniform administration down to the village level across Indonesia. The default classification was *desa*, and *kelurahan* was primarily formed under the directive of Ministry of Home Affairs (MOHA). Formally, villages can change their status from *desa* to *kelurahan* on the initiative of village government with approval from village legislators, by paying attention to community advice and opinion. This process is stipulated by MOHA (Ministry of Home Affairs) regulation no. 65/1999 which was then updated through MOHA regulation no. 31/2006. Later, after being approved by the district government, the head of district will issue a regulation regarding the status changes. All expenses for the transition process, including severance pay for the village head and village employees who are dismissed, are charged to the district budget. The formation of new *kelurahan*, at the minimum, shall satisfy several conditions which are mainly related to urban traits, including the population size (at least 1000 households for Java and Bali, and 1600 households for other regions), adequacy of administration buildings, economic potential, socio-cultural conditions, and improvement in quantity and quality of service access (especially transportation, communication and infrastructure). As elaborated by Jeddawi et al. (2018) in their case study, one of the reasons *desa* citizens aspire to change their government form to *kelurahan* is because of the assurance of official salaries, which hopefully will make the village office management more effective in delivering public services.

Despite that fact, certain aspects of *kelurahan* formation may generate some overlap between *desa* and *kelurahan* observable characteristics. First, although the ministerial decrees specify the requirements, none of these requirements was quantitatively or strictly enforced (Martinez-Bravo, 2014). Also as mentioned in the decrees, the proposal to switch *desa* into *kelurahan* needs to be approved first by the village legislators and citizens before being submitted to the district, giving more uncertainty on the *kelurahan* formulation process. Second, in terms of the location, in addition to *kota* districts, *kelurahan* also could be formed

in the capital of *kabupaten/*rural districts, and in the surroundings of the capital of each sub-district, even in rural districts. On the other hand, the assignment of *desa* also contributes to overlapping characteristics. In some provinces, villages are categorised as *desa*, for cultural reasons since villages want to retain their cultural values or norms, regardless of their observable characteristics. Given all these factors, inconsistencies in the implementation of the regulations in regard to *desa* and *kelurahan* formation are evident. Some *kelurahan* are quite rural in terms of their observable characteristics, and some *desa* can be quite urban as well. Table 3.1 presents cross tabulation between area status of villages, in either urban or rural areas, with their governance status, whether *desa* or *kelurahan*, during the time frame of this study.

Table 3.1 Cross tabulation between area status and administrative status, all villages

PODES wave		*Desa*	*Kelurahan*	Others	Total
2000	Urban	3,269	4,199	2	7,470
	Rural	58,759	2,129	420	61,308
		62,028	6,328	422	68,778
2003	Urban	6,706	5,269	3	11,978
	Rural	54,774	1,871	193	56,838
		61,480	7,140	196	68,816
2005	Urban	6,695	5,542	53	12,290
	Rural	54,673	1,820	1,174	57,667
		61,368	7,362	1,227	69,957
2008	Urban	6,654	5,912	43	12,609
	Rural	60,052	1,981	768	62,801
		66,706	7,893	811	75,410
2011	Urban	7,016	6,314	0	13,330
	Rural	62,726	1,769	136	64,631
		69,742	8,083	136	77,961

Notes: 'Others' category include cultural village, such as *nagari*; and new villages that are in transition status called UPT.

Desa and *kelurahan* have some differences regarding their institutional frameworks. The Law 22/1999 signified the beginning of the democratic framework of *desa* villages. Heads of *desa*, together with members of *desa* councils (*Badan Perwakilan Desa)* are directly elected by village citizens and salaries for the village apparatus are funded from the village budget. The village head is responsible to the village population and must submit an annual accountability report, which council can contest (Antlöv, 2003). During the time frame of this study, which was between 2000 and 2011, all village officials had generally been elected in free and fair village elections post the Soeharto era (Olken, 2010). For *kelurahan*, the head is appointed by the head of the district, all village apparatus are civil servants with their salaries funded from the district budget, and all benefits are stipulated by the district. District heads, furthermore, were also given rights to conduct appointments, transfers, and dismissal of *kelurahan* heads.

In 2014, the Government of Indonesia issued a new law on villages (Law no. 6/2014) which provides a framework for extending Indonesia's decentralisation and service delivery to villages (Lewis, 2015). According to this law, the explicit objectives of village decentralisation programs are to strengthen service delivery at the lowest administrative tier while reducing social inequality and poverty. Before the implementation of this law, there were no significant differences in responsibilities between *desa* and *kelurahan*. By right, *desa* government institutions have some authority over local affairs and over village budgets. However, the size of the village budget before the 2014 Village Law was considerably smaller as most of the funds available for villages were managed by central ministerial project structures (Antlöv et al., 2016).

Both *desa* and *kelurahan* are under the structure of district government, and although the funding mainly comes from the district budget[14], village government has some crucial functions. The functions can be divided into two broad categories: a governance role and a public service role. The governance role includes administrative issues such as keeping migration records, and maintaining public order and safety, including village heads who informally serve as mediators if there is a conflict in the village. The public service role includes some organisational roles in each services post that are managed by village government such as an integrated health post (POSYANDU), and village maternity post (POLINDES). POSYANDU is a village level health facility with its main activities being maternal and child health, immunisation, nutrition, and diarrhoea control. Based on instructions from the Ministry of Home Affairs, the head of village is responsible for the operation of POSYANDU. POLINDES is a specific health facility which provides maternal and child health care services. In most cases, POLINDES works together with POSYANDU, specifically implementing programs in the village. POLINDES is initiated and managed by local village officials. Most funding for POSYANDU and POLINDES comes from the district level and usually the funding amount is similar for all villages under one district. Access to both health posts will be one of the criteria in determining the individual level outcomes used in this study.

[14] This was before the implementation of the new Village Law in 2014 which mandates that each *desa* village will receive a large influx of village funding called Village Fund.

3.3 Literature review

The relationship between institutional form and government performance remains as a concern in public policy design. The main question is still about finding which government form is best to promote better basic service delivery. While modern democracy has defined elected legislators/parliaments to be one of its features, there is still plurality of forms in the case of regulators or the executive body (Besley and Coate, 2003).

The majority of literature on this topic examines ways for citizens to hold their service providers more accountable. One study on the theoretical side is Besley and Coate (2003), which explores the claim that elected regulators will yield to more pro-consumer policies. The key to their explanation is the issue of bundling. When regulators are appointed, policy issues become bundled with other issues, including the preferences of other stakeholders outside the jurisdiction, such as higher levels of government and major political parties. In contrast, rational voters will vote on the basis of candidates' policy stances and therefore for elected regulators, pro-consumer policies will gain them electoral advantage.

Although few, another stream of studies discusses this topic in the nexus of operational issues, by taking the case of council-manager government form at the municipality level in the U.S. Under this form, an elected governing body, called a council, will appoint a city/county manager to oversee daily operations and implement policies. As elaborated by Carr (2015), the appointed officers are more likely to have better administrative and operational skills, due to their training and career backgrounds. In addition, these managers will devote more of their time to manage the city/county as they are more attentive to the norms of professional management than to the demands of electoral politics. However, the assumption that appointed officials are neutral to electoral concerns is not always true. Taking the case of Indonesia, Martinez-Bravo (2014) contrasted two different forms of village government: *kelurahan* and *desa.* She found that village officials appointed by the district heads, motivated by career concerns, put more effort into influencing the outcome of national parliamentary elections in their village, compared to villages with elected officials. This reveals that appointed officials will have stakes in upper-level elections, and hence signals a need for the establishment of vertical accountability across levels of government.

Various empirical works exist on contrasting different government forms by taking the case of developing countries. Zhang et al. (2004) examine the effect of elected village leaders in

China on the structure of revenues and public expenditure. They found that elected village leaders tend to shift tax burdens from households to enterprises. Still taking the case of China, Luo et al. (2007) found that direct election of a village's leader leads to a significant amount of new investment in public goods. In the case of Indonesia, several studies have discussed the effects of different government forms on various categories of fiscal outcomes, although only at district levels. Lewis (2017) found that directly elected executives will spend less, especially on infrastructure, manage their budget more efficiently, and save more compared to parliament-appointed executives. Skoufias et al. (2014) also suggest the same results, as they found that increased electoral accountability for local governments was associated with a more prudent approach to fiscal balance. However, Kis-Katos and Sjahrir (2017) revealed that the direct election reform of district leaders did not improve their responsiveness to local public infrastructure, but reduce districts' public investment.

Despite numerous studies that have posited the investigation of different government forms in Indonesia, none of them has discussed the effects of this contrast at village level on service outcomes. The closest study that uses a similar setting at village level is Martinez-Bravo (2014), which looked at electoral outcomes. Nonetheless, some studies can inform the dynamics at village level in Indonesia. Olken (2010) contributes to the discussion of local democracy by conducting a field experiment in 49 villages. His study examined whether the plebiscite process, in which villagers can choose how to spend their village budget money, drives some results on service access. He found that the plebiscite process resulted in higher villager satisfaction and projects were more likely to be targeted to poorer hamlets. From a social capital perspective, Syukri and Mawardi (2014) observed the difference between *desa* and *kelurahan* when conducting their assessment of the Neighbourhood Upgrading and Shelter Sector Project (NUSSP)[15]. They found that *desa* tend to be more socially cohesive, with committed and highly respected leaders as their social nexus, while *kelurahan* tend to be more socially detached as in many cases *kelurahan* heads came from outside the village. Their findings are consistent with the results from a longitudinal study case of 40 Indonesian *desa* villages between 1996 and 2012, which found that village democratisation after the Soeharto era had reduced opportunities for dynastic leadership and seen more responsive *desa* leaders emerge (Antlöv et al., 2016).

[15] In 2005, with the support of the Asian Development Bank (ADB), the Government of Indonesia launched the Neighbourhood Upgrading and Shelter Sector Project (NUSSP). The aim of the NUSSP was to improve slum neighbourhoods and access to appropriate housing by low-income communities in urban areas.

3.4 Model specification

To answer the research questions, this paper examines two types of outcomes. The first type of outcome is village level outcomes and the second is individual level outcomes. All outcomes are chosen to capture the role of village heads in maintaining public order and safety and in managing village level health services such as immunisation and maternity posts. Due to data limitations, village-level outcomes for village heads' public service roles are not available, hence the estimations for immunisation and maternal health will use individual level outcomes.[16]

Village outcome specification

For the village-level estimation, this study considers outcomes of village *k* in year *t*, S_{kt} is a function of the treatment variable D_{kt}, which takes the value of 1 if the government form of the village is *kelurahan,* and 0 if the government form is *desa*[17]. Once a village switches its status to *kelurahan,* it will remain as *kelurahan* until the end of the observation period.[18]

$$S_{kt} = \rho D_{kt} + \mu U_{kt} + \beta X_{kt} + a_k + b_t + \varepsilon_{kt} \qquad (1)$$

where k = 1,..n denotes a village, t=1,..T denotes the census waves and ε_{kt} is the random error. I include the village area status, whether the village area is urban or not, U_{kt}, as control variable, hence the estimation should be interpreted as a comparison between villages in the same area, whether urban or rural. For village specific covariates, I include population number (X_{kt}) since it relates to many types of service access. The model also allows that service access at village level depends on time invariant village specific effects, a_k and time or wave specific effects, b_t.

The Difference in Difference (DID) design above relies on the key assumption that the time path of the outcomes between the control and treatment groups will be parallel in the post-treatment period in the absence of treatment. In this case, it means that without the treatment, the time path of service access of *desa* villages which switched to *kelurahan* would be

[16] Data for individual level outcomes are sourced from SUSENAS, which is not a village census and uses a sample of villages in the survey. More detailed explanation is provided in the data section. On the other hand, PODES (village census) does not include any detailed information on immunisation and maternal access.
[17] There are other village governance statuses recorded at the village census (PODES), such as *nagari* (cultural villages) and UPT, newly formed villages that are still in transition between governance statuses. Since there is no detailed breakdown of this 'other' classification and to ensure comparability, this paper will treat 'others' as missing.
[18] During the observation period, all villages that switched to *kelurahan* remain as *kelurahan.* This situation however may change after implementation of the new Village Law that gives more incentives for village to remain as *desa* or switch back to *desa.*

parallel to the time path of service access of *desa* villages which did not switch to *kelurahan.* This assumption also implies that the change in service access in the control group, which are *desa* villages that did not switch to *kelurahan,* is an unbiased estimate of the counterfactual. Therefore, in addition to estimating equation (1), I perform some procedures to test whether the parallel trends assumption in this study hold, using approaches that are summarised by Wing et al. (2018).

The first strategy to partially test the parallel trend assumption is by comparing the treatment effects with and without the group-specific linear time trends. If the coefficient of D_{kt} is not sensitive to the more restricted specification, then it indicates support for the parallel trend assumption. Hence, as an alternative specification, I extend equation (1) to include district-specific time trend, $c_d * t$.[19]

$$S_{kt} = \rho D_{kt} + \mu U_{kt} + \beta \mathrm{X}_{kt} + a_k + b_t + c_d * t + \varepsilon_{kt} \quad (2)$$

The district-specific time trend is chosen as STATA found difficulties in its routine converging when using village and or sub-district specific time trend. Also, trend variables are introduced to capture common economic shocks that also influence outcome variables, which are usually less varied between government levels lower than districts. I present both the unrestricted and more restricted specifications as results.

One risk in implementing DID is the existence of transitory shocks in pre-treatment outcomes for treated group. In this case, it means that *desa* villages which switched to *kelurahan* experience common transitory shocks that can cause difficulties in estimating the treatment effect on service access. Therefore, as the second strategy, I examine whether there are impacts of forward treatment effects by conducting a Granger type causality test. I add the lead variable of the treatment dummy and hence expect the lead variable to equal 0, $\boldsymbol{\lambda = 0}$, if there is a parallel trend (Ashenfelter, 1978).

$$S_{kt} = \rho D_{kt} + \lambda D_{kt+1} + \mu U_{kt} + \beta \mathrm{X}_{kt} + a_k + b_t + c_d * t + \varepsilon_{kt} \quad (3)$$

Third, I will check the covariate balance over the time period in order to ensure that changes in the treatment exposure are not associated with changes in the distribution of covariates. To

[19] Including district-specific time trends serve as a partial test for the parallel trend assumption. Also, district-specific time trend is chosen as the preferred specification, given the model convergence problems when employing district-time fixed effects, sub-district time trends, and village time trends.

check this, a simple covariate balancing table in the pre-treatment period cannot be very informative in validating the research design since what matters more is the distribution of covariates over the time frame. One empirical way to examine this is by estimating the covariate balance regression, which replaces the outcome variable with the covariate and fit the standard DID regression model, as shown below.

$$X_{kt} = \rho D_{kt} + \mu U_{kt} + a_k + b_t + \varepsilon_{kt} \quad (4)$$

If there is no compositional change in the covariates over time, we would expect that ρ will be equal to zero, or at least the magnitude of the coefficient would be small compared to the coefficient result when estimating against the outcome variable. In this case, I will check the balance of population variable with and without conditionality on the area status.

Individual outcomes specification

For individual level outcomes, this study implements a pseudo-village panel specification. Equation (1) is then modified to allow for individual outcomes and characteristics:

$$S_{ikt} = \rho D_{kt} + \mu U_{kt} + \mathbf{Y}_{ikt}\boldsymbol{\gamma} + a_k + d_k * b_t + \varepsilon_{ikt} \quad (5)$$

where i=1,…n denotes the number of individual surveyed in a village, k = 1,..n denotes a village, t=1,..T denotes the census waves and ε_{ikt} is the random error. Similar to the previous equation, the variable of interest is a dummy of whether the village governance status is *kelurahan* or not, $D_{\text{kt.}}$ In addition to the village specific fixed-effects, a_k, and village specific time effects, $d_k * b_t$,[20] this specification also allows controlling for relevant individual characteristics, $\mathbf{Y}_{ikt}$, which include gender, age in years, and household per capita expenditure. Also, since household observations are sampled at village level in the household surveys, the specification will cluster the standard errors at village level (Abadie et al., 2017).

The underlying identification assumption of this pseudo-village specification is that the treatment variable, the village government form status, is exogenous conditional on the various fixed effects and individual covariates. Therefore, in this specification, I introduce more restricted fixed effects, compared to the village DID specification, i.e. village specific time fixed effects, $d_k * b_t$. These village specific time fixed effects can explain most of the

[20] Village-specific fixed effects capture time-variant characteristics of the village, while the main dummy variable of *kelurahan* does not vary at the village-year level. The treatment sample for this individual model specification only includes villages that switched to *kelurahan* and remained in that status until the end of the observation period.

variations at the village level, including those which will influence the government status of the village. Also, for the dependent variables, I will construct the outcome variables based on the households' survey one year after the village census waves, which can also mitigate any possible bias due to direct reverse causality.

Difference and Difference (DID) set-up

I will present the standard DID model results for the baseline, and then present the staggered DID results as my preferred specification and main results. To implement the conventional Difference in Difference (DID) set-up, I use villages that have *desa* status in t=1 (wave/year 2000) and track further which and when villages will switch status from *desa* to *kelurahan* between 2000 and 2011. A similar set-up applies to the individual outcomes specification in which I only take sampled villages with *desa* status at t=1 and then track further when those villages switched to *kelurahan.* In addition to that, for the individual outcome specification, I only use the list of villages that are consistently available and being surveyed in all waves of the household surveys. I also apply further restrictions to the sample and perform a staggered DID approach by only using villages which eventually switch into *kelurahan* during the study period (Stevenson and Wolfers, 2006). In other words, I estimate equations (2) and (5) only for the 'treated' villages and exploit the different timings of village status switching.

In order to choose the preferred specification, I perform some procedures to test which model serves the parallel trend assumption. I found that the staggered DID model provided more supporting evidence on the parallel trend assumption, compared to the conventional DID model. By applying a staggered DID set-up, I also limit the possibility of any unobserved time-varying selection effects that differentiate between *desa* villages that switched to *kelurahan* and those that retained their status as *desa.*

3.5 Data and choice of variables

Data

The main data used in this study are from the Indonesian Village Potential Statistics (*Potensi Desa*/PODES). These data are collected by the Indonesian Central Bureau of Statistics (Badan Pusat Statistik/BPS) every two or three years. PODES is a village census that provides information about Indonesian' village characteristics, such as government administration, public goods provided in the village, socio-economic characteristics and other

information. Because each wave of PODES has a different focus, not all variables are reported in all waves of the census.

In this study, I merged information from five different waves of the village census, collected from 2000, 2003, 2005, 2008, and 2011[21]. The main problem when constructing this dataset was the inconsistent village identifiers across waves, due to the introduction of new villages and village-splitting. Given the difficulty of matching villages across waves, for the DID specification, I chose to use only villages that were consistently included in all waves, and therefore some villages were excluded from the dataset. I also excluded three provinces from the sample, which are Aceh, West Sumatera, and DKI Jakarta. Aceh and West Sumatera are excluded due to their local arrangements in which almost all villages have administrative status as *desa* or *nagari*, leaving only a few *kelurahan* in those provinces. DKI Jakarta was excluded as its districts do not have budget and service delivery responsibility.

For the pseudo-village specification, I merged the village census data with the annual household survey (SUSENAS) data. Since many questions in the SUSENAS capture conditions in the previous year, I used the SUSENAS waves one year after the village census year, which are SUSENAS 2001, 2004, 2006, 2009, and 2012. SUSENAS is neither a household panel, nor village panel data, which means not all villages that are covered in one SUSENAS wave will be covered again in the SUSENAS sampling the following year. Generally, SUSENAS uses a two-stage sampling method. First, the Probability Proportional to Size (PPS) method is used when selecting census blocks from the population census. Then, a systematic random sampling method is used to select households within each census block (approximately 16 households per census block before 2011 and 10 households per census block since 2011). Details of the number of observations, including attrition, are presented in Table 3.2.

Choice of dependent variables

This study aims to examine whether the village government forms which determine the type of government accountability have impacts on service outcomes. From both PODES and SUSENAS, I extracted several service outcomes as dependent variables, with the following

[21] Village census data for 2014 are available, however the implementation of the new Village Law in 2014 may have influenced the funding and service provision in *desa* villages. Therefore, to ensure the data is not exposed to this shock, the paper only uses village census data up to 2011.

criteria: 1) either prescribed in the regulation as duty or responsibility of the village head, or 2) main activities of service access that are administered or organised by the village.[22]

From village censuses (PODES), I use one outcome variable that can represent the role of village head to maintain safety and as mediator for any conflict, which is whether there is property crime in the village or not.[23] There is another potential question in the survey that could also fill this role, which is the question of whether there is mass conflict in the village or not. However, since this conflict question only captures massive conflicts, only a few villages have reported occurrences of such events, and therefore I prefer to use the property crime variable for several reasons. First, by only using property crime, I exclude other types of violent crimes that are usually motivated by personal drives. Second, property crime can be explained both by economic theory and also social disorganisation theory (Kelly, 2000). Indirectly, property crime has been perceived as related to social conflict measures, especially regarding economic classes, and in this case it is expected to indicate a greater spectrum of conflict at the village level. According to social disorganisation theory, informal social controls of communities over their members can create deterrent effects against committing property crime. In this study, I approach this informal social control from the role of village head in managing the network and relations inside the village.

From SUSENAS, I collected individual outcomes data, which are 1) birth assisted by medical practitioner of children 0–5 years old and 2) number of immunisations taken for BCG, DPT, measles, and Hepatitis B for individuals 0–5 years old. For immunisation outcomes, I will separate each type of immunisation, in order to differentiate the influence of immunisation that only needs to be taken once, e.g. BCG and measles, compared with immunisation that needs to be taken more than once, e.g. three times for DPT and four times for Hepatitis B.[24] Assisted birth is one of the expected outcomes of the village maternity post/POLINDES, while immunisation programs are the main activities of the village health post, POSYANDU.

[22] More detailed description of the role of village head is provided in the background section.
[23] Although more recent PODES ask an additional question regarding the qualitative frequency of the property crime, e.g. frequent, sometimes, rarely, etc., earlier PODES (village census) did not ask for this detailed information. Therefore, to improve consistency, I chose to use the incidence of property crime as the outcome variable. The use of the incidence variable also helps to address the common concern that PODES data suffers from reporting bias since it is based on self-reporting of the village apparatus. Officials are more likely underreport the frequency or magnitude of crime, rather than the incidence of the crime itself.
[24] Based on the endorsement stated by the Ministry of Health, which serves as definition of complete immunisation.

Table 3.2 Number of villages and observations

	PODES 2000	PODES 2003	PODES 2005	PODES 2008	PODES 2011
Total villages	60,742	61,938	62,820	67,795	70,826
Desa status	54,799	55,220	55,273	60,203	63,077
Kelurahan status	5,553	6,534	6,846	7,326	7,613
Others	390	184	701	266	136
For Village DID – only using village that available across PODES waves and have desa status in 2000					
Total villages	27,668	27,668	27,668	27,668	27,668
Desa status	27,668	26,898	26,621	26,577	26,545
Kelurahan status	0	770	1,047	1,091	1,123
	Susenas 2001	**Susenas 2004**	**Susenas 2006**	**Susenas 2009**	**Susenas 2012***
Total individuals	940,618	1,079,885	1,156,922	1,209,564	342,458
Total villages	12,341	14,011	15,612	16,558	6,916
After matched with PODES t-1					
Matching rate	71%	82%	81%	78%	73%
Total individuals	667,842	889,424	936,504	941,676	248,995
Total villages	9,537	12,303	13,492	13,797	6,222
Desa status	7,412	9,216	10,117	10,290	4,519
Kelurahan status	2,087	3,072	3,301	3,489	1,699
Others, inc: *nagari, UPT*	38	15	74	18	4
%village coverage in SUSENAS	16%	20%	21%	20%	9%
If using villages that available across SUSENAS years and have desa status in 2000/2001					
Total individuals	11,366	12,578	11,949	11,027	6,211
Total villages	142	142	142	142	142
Desa status	142	96	76	76	74
Kelurahan status	0	46	66	66	68

Notes: ^The number of villages excludes three provinces: Aceh, West Sumatera and DKI Jakarta. In estimation, villages with 'other' status are recorded as missing. *In 2012, SUSENAS changed its sampling method, and the only data available with village identifiers are the quarterly data. This paper uses the third quarter of SUSENAS data.

3.6 Village level outcome: Property crime incidence

Results

For village level estimation, this study considers property crime incidence which represents indicators available in village census data from public services under the management or responsibility of village apparatus.

Table 3.3 presents the results on dummy variable of whether there is property crime incidence in the village or not. All columns are estimated using a Linear Probability Model (LPM) since STATA encounter problems in routine converging and maximum matrix size when using non-linear models, such as probit. The policy variable is the *kelurahan* dummy which takes a value of 1 if the village has a government form of *kelurahan.* In this set-up,

once a *desa* village is recorded as *kelurahan,* the *desa* will remain as *kelurahan* for the rest of the time period. Another thing to note is that the village government form statuses are recorded from the village census, which is not conducted every year. For all *desa* villages which have the government form recorded for the first time as *kelurahan,* the effective switching date itself may have happened anytime between the previous wave survey and the subsequent ones. For instance, village A in PODES 2000 is recorded as *desa*, then in PODES 2003, village A is recorded as *kelurahan.* This means that the effective switching of village A from *desa* to *kelurahan* may have happened anytime between 2000 and 2003. Therefore, all results of the *kelurahan* dummy should be interpreted as the lower bound coefficient, since basically this research treats all villages that switched between two waves of village census similarly and disregards how long any particular *desa* village has been a *kelurahan.*

Table 3.3 Village level outcome: property crime incidence

	DID				Staggered DID			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
kelurahan (=1)	0.073***	0.067***	0.040***	0.038**	0.069**	0.062**	0.057*	0.068**
	[0.015]	[0.015]	[0.015]	[0.015]	[0.030]	[0.030]	[0.030]	[0.031]
area (urban=1)		0.030***	0.011	0.009		0.123***	0.068**	0.059**
		[0.008]	[0.008]	[0.008]		[0.027]	[0.028]	[0.029]
Observations	138,340	138,340	138,340	138,340	5,615	5,615	5,615	5,615
R-squared	0.005	0.005	0.008	0.021	0.002	0.007	0.021	0.084
Number of iddesa	27,668	27,668	27,668	27,668	1,123	1,123	1,123	1,123
Village FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time/Wave FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
(log) population	No	No	Yes	Yes	No	No	Yes	Yes
District-specific time trend	No	No	No	Yes	No	No	No	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1, Robust standard errors in brackets. All columns are estimated using Linear Probability Model (LPM).

The results from Table 3.3 suggest that villages with *kelurahan* status experience higher probability of property crime of about 6.8 percentage points, compared with villages that retain *desa* status. When adding the population variable as covariate, as seen from columns (2) to (3), and columns (7) to (8), the standard error does not show much improvement. Furthermore, when adding district-specific time trends, as one of the ways to check for parallel trend assumptions, the coefficients change slightly from 0.04 to 0.038 in column (4), and from 0.057 to 0.068 in column (8). The results in both columns (4) and (8) show that the treatment effect is not sensitive to more restricted alternative specification and therefore indicate support for the parallel trend assumption.

By ruling out other types of crime that are usually motivated by personal drives, such as violent crime, property crime can be perceived as an indicator of socio-economic tension in the village. Property crime is viewed as an illegal way to redistribute socio-economic income due to rising social tension in a community. Also, based on Nguyen (2019), over 90 per cent of crime reported in the Indonesian household survey between 2007 and 2012 was property crime. Therefore, although not as salient as other types of conflict within a community, this higher probability of property crime in *kelurahan* villages can indicate that *desa* governments are better at performing their role of mediators in villages, compared to the *kelurahan* governments.

Testing parallel trend

In addition to the inclusion of district-specific time trends, I perform two other formal tests to examine whether the parallel trend assumption holds. Table 3.4 presents the results of the Granger type causality test, as described in equation (3), while Table 3.5 presents the results of the covariate balance regression described in equation (4). As preferred specification, the discussion on parallel trend will focus on the staggered DID estimation.

Table 3.4 presents the result of the Granger type causality test for property crime in staggered DID estimation. The decision on how many lead variables need to be used mainly depends on the total number of periods available for analysis and the timing of the treatments. In this study, I will use only one lead variable since the data is actually in intervals for every 2 to 3 years, so one lead variable (t+1) may already capture the medium-term anticipation effect. Using one lead variable will also automatically restrict the sample only up to 2008, which in this case is less problematic due to the limited switching cases between 2008 and 2011. Column (1) shows that when the model does not include the area dummy, the parallel trend assumption cannot hold, as indicated from the significant coefficient of the lead variable of the treatment dummy, *kelurahan* (t+1). The preferred specification as presented in columns (3) and (4), which includes village controls and district-specific time trends, shows insignificant results for the lead treatment variable. These results demonstrate that there is no evidence of transitory shocks for *desa* villages that switched to *kelurahan* before the switching happens, and therefore it supports the underlying parallel trend assumption.

Table 3.4 Granger type causality test: property crime

VARIABLES	(1)	(2)	(3)	(4)
kelurahan (t)	0.079**	0.074**	0.077**	0.074**
	[0.032]	[0.033]	[0.032]	[0.036]
kelurahan (t+1)	-0.065*	-0.052	-0.024	-0.031
	[0.033]	[0.033]	[0.033]	[0.040]
area (urban=1)		0.129***	0.065*	0.068*
		[0.032]	[0.033]	[0.036]
Observations	4,492	4,492	4,492	4,492
R-squared	0.005	0.011	0.028	0.113
Number of iddesa	1,123	1,123	1,123	1,123
Village FE	Yes	Yes	Yes	Yes
Time/Wave FE	Yes	Yes	Yes	Yes
(log) population	No	No	Yes	Yes
District-specific time trend	No	No	No	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1, Robust standard errors in brackets. All columns are estimated using Linear Probability Model (LPM).

Next, Table 3.5 presents the results from covariate balance tests. The table shows that the village population variable, as village covariate in the model, is eventually balanced over the period of the study, conditional on the urban area classification. The coefficient of the treatment variable (*kelurahan)* shows insignificant results in column (3) and therefore provides support for the parallel trend assumption, conditional on the area. As briefly discussed in the methodology part, the urban area status here is treated as a conditional variable rather than covariate, and therefore I can argue that the parallel trend holds for villages within the same area classification, in either urban or rural areas.

Table 3.5 Covariate balance test for village level specification

VARIABLES	(1) area	(2) (log)pop	(3) (log)pop
kelurahan (=1)	0.053***	0.056**	0.036
	[0.014]	[0.023]	[0.023]
area (urban=1)			0.364***
			[0.035]
Observations	5,615	5,615	5,615
R-squared	0.265	0.198	0.250
Number of iddesa	1,123	1,123	1,123
Village FE	Yes	Yes	Yes
Time/Wave FE	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1, Robust standard errors in brackets. All columns are estimated using Linear Probability Model (LPM).

3.7 Individual level outcomes

To complement the results from the village level specification, I also test the policy variable, *kelurahan* dummy, on several individual indicators that are parts of village roles and available in the national household survey (SUSENAS). However, the household survey data is not village level panel or individual level panel data. This condition prevents the implementation of the standard DID model, and therefore I use a pseudo-village DID specification, with treatment variables at village level and outcomes at individual level. The interpretation from this model assumes that village government status can be treated as exogenous, conditional on all fixed effects and covariates.

Testing covariate balance

The discussion on parallel trend assumptions for the individual outcome model is less extensive than for the village panel, due to the more restricted set-up of the model (by applying village specific time FE) which already absorbs many of the variations. The individual level specification is an extended version of the village level specification and uses the same treatment variable, which is the dummy of village government form. Results from the previous section show that there is support for parallel trend assumption under the treatment. The major difference between village outcomes and individual outcomes is the additional individual characteristics included as covariates. Therefore, I perform a check of whether the additional individual covariates are balanced between the *kelurahan* and *desa* villages along the period of study.

Table 3.6 shows that all individual characteristics are significant against the *kelurahan* dummy, although for gender and household expenditure per capita, the magnitudes are reduced after controlling for urban dummy. However, the next section demonstrates that the results convey consistent findings and significance levels, with and without individual characteristics. I will therefore present both models since treatment effects do not change much with and without covariates.

Table 3.6 Covariate balance regression for individual characteristics

	Male (=1)		Age in years		(log) hh exp cap	
VARIABLES	(1)	(2)	(4)	(5)	(6)	(7)
kelurahan (=1)	-0.090***	0.010***	-2.595***	9.676***	1.690***	0.935***
	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]	[0.000]
area (urban=1)		-0.100***		-12.271***		0.755***
		[0.000]		[0.000]		[0.000]
Observations	7,827	7,827	7,827	7,827	7,827	7,827
R-squared	0.058	0.058	0.108	0.108	0.678	0.678
Village FE	Yes	Yes	Yes	Yes	Yes	Yes
Village specific time FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1, Clustered standard errors at village level in brackets. All columns are estimated using Linear Probability Model (LPM). Number of observations is summation of children 0-5 years old and married woman in treated villages, which are villages that eventually switched to *kelurahan*.

Results

Table 3.7 to Table 3.9 present the results from the individual outcomes. Results from the DID model served as a baseline, and the staggered DID approach is the preferred specification. Due to the nature of the outcome variables, in addition to the linear regression, Tables 3.8 and 3.9 are also estimated using Poisson regression.[25] The results from these alternative non-linear specifications serve as robustness checks and are presented in the Appendix.

Table 3.7 presents the effect of village government form on safe birth indicator, which is run by the village maternity post (POLINDES). The table shows that villages with *kelurahan* status deliver worse results on the rate of whether a birth is assisted by a health professional or not. As shown in the table, the results are consistent across two different estimation approaches, DID and staggered DID, as well as after including relevant individual/household characteristics. Columns (4) to (6) show that switching to *kelurahan* status will decrease the probability of an individual being born through assisted birth by on average 0.8 points. Presumably, urban areas will perform better in assisted birth access, and so will *kelurahan* as there is a likelihood of strong correlation between *kelurahan* and the area status. However, this study finds a negative effect of *kelurahan* status, while the urban dummy coefficients remain positive in all cases.

[25] The safe birth indicator only can be estimated using the linear probability model (LPM) since when using probit, it encounters many dropped observations which leads to failure in prediction.

Table 3.7 Dummy of assisted birth of children 0–5 years old

	DID			Staggered DID		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
kelurahan (=1)	-0.800***	-0.793***	-0.860***	-0.800***	-0.795***	-0.852***
	[0.000]	[0.006]	[0.009]	[0.000]	[0.006]	[0.011]
area (urban=1)	0.617***	0.614***	0.473***	0.617***	0.615***	0.493***
	[0.000]	[0.007]	[0.020]	[0.000]	[0.008]	[0.025]
male (=1)		0.002	0.001		0.001	-0.001
		[0.008]	[0.008]		[0.009]	[0.009]
age in years		-0.013***	-0.012***		-0.009**	-0.009*
		[0.003]	[0.003]		[0.004]	[0.004]
(log) hh exp cap			0.130***			0.112***
			[0.016]			[0.021]
Observations	4,852	4,852	4,852	2,660	2,660	2,660
R-squared	0.465	0.466	0.481	0.372	0.373	0.388
Village FE	Yes	Yes	Yes	Yes	Yes	Yes
Village specific time FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1, Clustered standard errors at village level in brackets. All columns are estimated using Linear Probability Model (LPM).

The results on assisted birth are consistent with the results on immunisation uptake. Table 3.8 shows the effect of *kelurahan* government form on the number of measles immunisations. To achieve complete immunisation status, a child needs to have measles immunisation at least once during the first five years of childhood. Columns (4) to (6) show that although children in urban areas will have more measles immunisations, children in *kelurahan* villages, on average, will have a smaller rate of measles immunisation. These results remain consistent even after including several individual characteristics of the child, such as gender, age in years, and the income level of the household. Similar results are found for Hepatitis B, which requires multiple dosages of immunisation to be considered as complete immunisation. Table 3.9 shows that children in *kelurahan* will have smaller number of Hepatitis B immunisations compared to children in *desa* villages. I also check for BCG immunisation uptake, another one-dosage immunisation, and DPT immunisation which require multiple inoculations. The findings are consistent with the results in Tables 3.8 and 3.9 and are presented in the Appendix.

Table 3.8 Number of measles immunisations for individuals 0–5 years old

	DID			Staggered DID		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
kelurahan (=1)	-0.190***	-0.289***	-0.337***	-0.190***	-0.285***	-0.315***
	[0.000]	[0.006]	[0.041]	[0.000]	[0.009]	[0.057]
area (urban=1)	0.105***	0.143***	0.100***	0.105***	0.147***	0.120**
	[0.000]	[0.004]	[0.037]	[0.000]	[0.005]	[0.053]
male (=1)		0.021	0.021		-0.000	-0.000
		[0.017]	[0.017]		[0.024]	[0.024]
age in years		0.132***	0.132***		0.132***	0.132***
		[0.007]	[0.007]		[0.010]	[0.010]
(log) hh exp cap			0.036			0.022
			[0.031]			[0.044]
Observations	3,755	3,755	3,755	2,023	2,023	2,023
R-squared	0.333	0.425	0.426	0.333	0.420	0.420
Village FE	Yes	Yes	Yes	Yes	Yes	Yes
Village specific time FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1, Clustered standard errors at village level in brackets. All columns are estimated using linear regression.

Table 3.9 Number of Hepatitis B immunisations for individuals 0–5 years old

	DID			Staggered DID		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
kelurahan (=1)	-1.905***	-2.026***	-2.212***	-1.905***	-2.017***	-2.244***
	[0.000]	[0.012]	[0.073]	[0.000]	[0.016]	[0.106]
area (urban=1)	0.316***	0.367***	0.202***	0.316***	0.378***	0.178*
	[0.000]	[0.007]	[0.066]	[0.000]	[0.010]	[0.095]
male (=1)		0.009	0.008		-0.045	-0.045
		[0.028]	[0.028]		[0.038]	[0.038]
age in years		0.168***	0.168***		0.170***	0.171***
		[0.014]	[0.014]		[0.020]	[0.020]
(log) hh exp cap			0.137**			0.167**
			[0.055]			[0.079]
Observations	3,755	3,755	3,755	2,023	2,023	2,023
R-squared	0.547	0.578	0.580	0.536	0.568	0.571
Village FE	Yes	Yes	Yes	Yes	Yes	Yes
Village specific time FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1, Clustered standard errors at village level in brackets. All columns are estimated using linear model.

3.8 Heterogeneity of results

Indonesia is an archipelagic country. For over centuries, the main island Java has been the centre of economy and business. Therefore, in general, Java can be perceived as more developed than the other islands in Indonesia. Given this variation, I also estimate both

village level outcome and individual outcomes for the staggered DID specification, disaggregated by whether the village is in Java or outside of Java.

Table 3.10 shows that the significant impact of *kelurahan* status on property crime incidence only prevails for villages on Java island (column 4), while it is insignificant for villages outside Java (column 8). Immunisation outcomes also provide similar results with negative effects of *kelurahan* prevailing in Java and either positive or not significant outside Java.[26] On the other hand, for safe birth outcome, the magnitude of *kelurahan* status is bigger for non-Java villages compared to villages in Java (Table 3.11).

Table 3.10 Property crime incidence, by region

	Java				Non-Java			
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
kelurahan (=1)	0.116***	0.115***	0.110**	0.115**	0.035	0.024	0.016	0.034
	[0.043]	[0.043]	[0.043]	[0.046]	[0.041]	[0.041]	[0.040]	[0.043]
area (urban=1)		0.107***	0.031	0.027		0.150***	0.100***	0.094**
		[0.041]	[0.041]	[0.043]		[0.036]	[0.038]	[0.040]
Observations	2,750	2,750	2,750	2,750	2,865	2,865	2,865	2,865
R-squared	0.005	0.008	0.025	0.080	0.004	0.011	0.024	0.092
Number of iddesa	550	550	550	550	573	573	573	573
Village FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
Time/Wave FE	Yes	Yes	Yes	Yes	Yes	Yes	Yes	Yes
(log) population	No	No	Yes	Yes	No	No	Yes	Yes
District-specific time trend	No	No	No	Yes	No	No	No	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1, Robust standard errors in brackets. All columns are estimated using Linear Probability Model (LPM).

[26] Results are presented in the Appendix.

Table 3.11 Assisted birth of children 0–5 years old, by region

	Java			Non-Java		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
kelurahan (=1)	0.001	-0.017	-0.184***	-0.800***	-0.796***	-0.862***
	[0.000]	[0.017]	[0.060]	[0.000]	[0.008]	[0.015]
area (urban=1)	-0.001	0.014	0.087***	0.617***	0.616***	0.475***
	[0.000]	[0.014]	[0.030]	[0.000]	[0.010]	[0.031]
male (=1)		-0.003	-0.003		0.004	0.001
		[0.014]	[0.014]		[0.012]	[0.012]
age in years		-0.007	-0.007		-0.011*	-0.010
		[0.006]	[0.006]		[0.006]	[0.006]
(log) hh exp cap			0.093***			0.130***
			[0.030]			[0.026]
Observations	1,183	1,183	1,183	1,477	1,477	1,477
R-squared	0.455	0.456	0.469	0.314	0.316	0.333
Village FE	Yes	Yes	Yes	Yes	Yes	Yes
Village specific time FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1, Clustered standard errors at village level in brackets. All columns are estimated using Linear Probability Model (LPM).

One possible explanation is because the area of villages in Java is smaller those outside Java, as Java has a higher population per capita, compared to other main islands. People live close to their neighbours, and therefore the accountability process may work better for villages in Java. Another possible reason is that most of Java can be considered as urban areas, compared to area outside of Java. Hence, the *kelurahan* coefficient already excludes the confounding effects of *kelurahan* and *urban* areas and can better measure the government form status. A similar reason also can be applied for safe birth indicator, which stands out from the other indicators due to the lower negative effect of *kelurahan* status in Java compared to non-Java. Safe birth is an indicator that is highly related to health service provision, and for Java the number of other health facilities outside villages (hospitals, private clinics, etc.) is higher compared to non-Java. These health facilities are usually in the capitals of districts or *kecamatan*, reducing the negative effect of *kelurahan*, which are usually located closer to district capitals compared to *desa* villages.

3.9 Discussion

Overall, the results show that villages which switch to *kelurahan* perform worse on property crime, birth and immunisation service outcomes, compared to villages which stay as *desa*. The main argument for this is because the *desa* government is more accountable to its constituents since the head is directly elected and staff salaries are funded from the village budget. In other words, the results suggest that horizontal accountability influences more positive service outcomes compared to vertical (non-electoral) accountability.

Antlöv et al. (2016) support this proposition, although from a different angle. By conducting a survey of 40 *desa* villages in 1996, 2001, and 2012, their study was able to compare the *desa* villages before and after the start of decentralisation in early 2000. During the New Order era, village head elections were not independent as the elections were closely supervised by the district government and candidates had to get approval from the district government to be inaugurated, regardless of election results. Now, elections are called by village parliaments and organised by villagers themselves. Their study finds that, after decentralisation, the village heads were more likely to work in villagers' interests compared to their predecessors. Village heads who had been elected in democratic elections were also found to maintain transparency and participatory norms.

One possible way for village citizens to hold their government accountable is through participating in decision making, which takes the form of community or village meetings. These practices commonly occur in *desa* villages but not necessarily in a bureaucratic-type government form of *kelurahan.* Through conducting field experiments in 49 villages, Olken (2010) finds that engaging village citizens in decision making processes results in higher villager satisfaction and better targeted programs. I also find supporting results when comparing data from the module questionnaire of the household survey (SUSENAS) 2009, which asked whether respondents trusted their village officials in managing the village finances. Table 3.12 shows that the percentage of respondents who trust and strongly trust their village officials is higher for those who live in *desa,* compared to respondents who live in *kelurahan.* The data furthermore reveals that 10 percent of respondents who live in *kelurahan* do not even know their village level officials or headperson which indicates their low engagement with the village level community. This table is also consistent with what Syukri and Mawardi (2014) observe, that *desa* villages tend to be more socially cohesive with trusted leaders as their social nexus, while *kelurahan* in many cases even have heads who come from outside the village.

Table 3.12 Trust in village officials between *desa* and *kelurahan*, in percentages

	Q: Do you trust the village official/headperson for village financial management?						
SUSENAS 2009	Do not know	Do not care	Not trust	Less trust	Trust	Strongly Trust	Total
Desa	4.0	0.6	3.3	13.8	74.3	4.1	100
Kelurahan	10.5	0.7	4.1	15.7	66.6	2.4	100
All villages	6.0	0.6	3.5	14.6	71.8	3.5	100

Source: SUSENAS 2009, author's calculation

The elected officials' oppositions often argue that appointed officers are more likely to have better managerial skills due to their training and career backgrounds. Although education

levels cannot be fully assumed to translate into better managerial skills and hence better service access, recent data finds that the education levels of *desa* leaders are improving. Recent village survey (PODES) data showed that the percentage of *desa* village heads who had graduated from senior high school had increased from less than 40 percent in 2000 to almost 60 percent in 2011. This improvement is also acknowledged by Antlöv et al. (2016), who found in their panel survey that in 2001, some *desa* heads from the surveyed villages had only primary school education but by 2012, most *desa* heads had graduated from senior high school.

3.10 Conclusion

This study examines the effects of two different government forms at village level in Indonesia, by contrasting villages with *kelurahan* and *desa* status, between 2000 and 2011. By looking at the impact of village government form on local service access, this research contributes to the discussion on accountability and level of decentralisation. The results show that villages which switched to *kelurahan* status have more property crime compared to those which remain as *desa* villages. For individual outcomes, switching to *kelurahan* status reduces access to assisted birth by health professionals and immunisation uptake for children 0–5 years old.

Overall, these results suggest that *kelurahan,* which is a bureaucratic type of government, perform worse than *desa* villages in delivering public services. The findings challenge the common perception that people in *kelurahan* will get better access and services as all the apparatus are civil servants and the appointed head will have better access to higher levels of government. One potential rationale for this is that apparatus in *desa* villages are more accountable to their constituents, since they are directly elected. This means that horizontal accountability matters more in service access outcomes than vertical (non-electoral) accountability, even for the lowest administrative tier. Another explanation is that *desa* tend to be more socially cohesive and their heads act as the village social nexus. For instance, citizens in *desa* villages trust their leaders more compared to those who live in *kelurahan*.

However, this study has some caveats that may inform future research topics. First, due to data limitations, the treatment variable was assigned based on the first time the government form was recorded as *kelurahan,* and not the actual effective date of *kelurahan* form. The ideal robustness check would be to address this by collecting data on when the switching

from *desa* to *kelurahan* happened. Secondly, some other types of service access may have been influenced by village government forms, which were not addressed in this study due to the requirements of the panel set-up. One alternative is to only use population census years and construct some economic outcomes. However, using that approach would require another strategy for how to disentangle the influence of district government and village level government. Third, this study only covers the period before the new Village Law was implemented in 2014, which basically gives more autonomy to village governments. It would be useful to compare this distinction of government form before and after the new Village Law.

Regarding policy implications, this study provides insights into the decentralisation policy in two layers. First, this study suggests that the decentralisation process, down to the lowest administrative tier and accompanied by fair elections, will give incentives for local government to be more accountable to constituents. One concern from decentralising to the lowest level is the trade-off between accountability and bureaucracy coordination. The findings from this study indicate that the benefits of having more accountable government will exceed the additional cost of coordinating multiple tiers of government. Nevertheless, for Indonesia's case, the new Village Law in 2014 also means that each village will receive a certain amount of village funding. Therefore, the interpretation should be treated cautiously since managing a large influx of funding will require certain capacity and may induce rent-seeking behaviour that will constrain accountability. Second, this study indicates that the benefits from a directly elected government form are more notable in indicators that also include active engagement with service providers, rather than merely an access problem. For instance, immunisation uptake usually requires engagement with the health post (POSYANDU) to remind villagers of immunisation schedules. This points to the potential role of village government in providing those types of services, especially in the implementation of the village decentralisation framework.

Appendix

Table 3.13 Differences between *desa* and *kelurahan* based on PODES 2011

VARIABLES	MEAN		DIFF
	Desa (n=63077)	*Kelurahan* (n=7613)	(*Desa* – *Kelurahan*)
Whether area is urban (=1) or rural (=0)	0.103	0.768	-0.665***
Number of population	2583	7871	-5288***
Number of household	728	2046	-1318***
Whether type of widest road is asphalt (=1) or others (=0)	0.619	0.937	-0.318***
Whether village has post office (=1) or not (=0)	0.034	0.181	-0.147***
Whether village has market with permanent/semi-permanent building (=1) or not (=0)	0.163	0.335	-0.172***
Whether village has main road lighting (=1) or not (=0)	0.549	0.900	-0.350***
Proportion of household who has telephone	0.042	0.190	-0.148***
Distance to municipality (in km)	11.627	2.841	8.786***

Notes: The number above excludes three provinces: Aceh, DKI Jakarta, and West Sumatera, due to its different arrangement in village government form. The number is produced from a t-test with unequal variances between two groups: *desa* villages and *kelurahan* villages.

Table 3.14 Descriptive summary

Variable	**Observation**	**Mean**	**Std. Dev.**	**Min**	**Max**
Village DID specification – figures for all villages in the village census					
crime_prop	138,340	0.48	0.50	0	1
kelurahan (=1)	138,340	0.03	0.17	0	1
area (urban=1)	138,340	0.16	0.37	0	1
npop	138,340	3,441.32	3,047.76	13	100,884
Pseudo-village specification					
birth	4,852	0.82	0.38	0	1
fp	10,235	0.57	0.49	0	1
nimu_bcg	3,755	0.95	0.48	0	9
nimu_measles	3,755	0.80	0.56	0	9
nimu_dpt	3,755	2.06	1.16	0	9
nimu_hepb	3,755	1.86	1.22	0	9
kelurahan (=1)	53,131	0.37	0.48	0	1
area (urban=1)	53,131	0.63	0.48	0	1
ageyr	53,131	28.38	19.06	0	98
male	53,131	0.50	0.50	0	1
(log)expcap	53,131	12.49	0.73	10	17

Table 3.15 Granger type causality test: property crime, DID specification

VARIABLES	(1)	(2)	(3)	(4)
kelurahan (t)	0.103***	0.095***	0.093***	0.055**
	[0.024]	[0.024]	[0.027]	[0.027]
kelurahan (t+1)	-0.070*	-0.074*	-0.006	-0.009
	[0.038]	[0.038]	[0.047]	[0.046]
kelurahan (t+2)	-0.086	-0.087	-0.029	-0.050
	[0.098]	[0.097]	[0.094]	[0.092]
kelurahan (t+3)	-0.067	-0.071	-0.018	-0.060
	[0.112]	[0.112]	[0.114]	[0.112]
area (urban=1)		0.027**	0.028**	0.015
		[0.011]	[0.012]	[0.012]
Observations	55,336	55,336	55,336	55,336
R-squared	0.003	0.003	0.046	0.050
Number of iddesa	27,668	27,668	27,668	27,668
Village FE	Yes	Yes	Yes	Yes
Time/Wave FE	Yes	Yes	Yes	Yes
District-specific time trend	No	No	Yes	Yes
(log) population	No	No	No	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1, Robust standard errors in brackets. All columns are estimated using Linear Probability Model (LPM).

Table 3.16 Covariate balance regression, DID specification

VARIABLES	(1) area	(2) lnpop	(3) lnpop
kelurahan (=1)	0.203***	0.271***	0.237***
	[0.012]	[0.018]	[0.017]
area (urban=1)			0.168***
			[0.007]
Observations	138,340	138,340	138,340
R-squared	0.082	0.047	0.062
Number of iddesa	27,668	27,668	27,668
Village FE	Yes	Yes	Yes
Time/Wave FE	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1, Robust standard errors in brackets. All columns are estimated using Linear Probability Model (LPM).

Table 3.17 Number of BCG immunisations for individuals 0–5 years old

	DID			Staggered DID		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
kelurahan (=1)	-0.333***	-0.361***	-0.405***	-0.333***	-0.364***	-0.392***
	[0.000]	[0.005]	[0.037]	[0.000]	[0.008]	[0.058]
area (urban=1)	0.105***	0.119***	0.080**	0.105***	0.123***	0.098*
	[0.000]	[0.004]	[0.035]	[0.000]	[0.005]	[0.054]
male (=1)		-0.006	-0.007		-0.014	-0.015
		[0.015]	[0.015]		[0.023]	[0.023]
age in years		0.041***	0.041***		0.046***	0.046***
		[0.007]	[0.007]		[0.009]	[0.009]
(log) hh exp cap			0.033			0.021
			[0.029]			[0.045]
Observations	3,755	3,755	3,755	2,023	2,023	2,023
R-squared	0.434	0.446	0.446	0.387	0.401	0.402
Village FE	Yes	Yes	Yes	Yes	Yes	Yes
Village specific time FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1, Clustered standard errors at village level in brackets. All columns are estimated using linear model.

Table 3.18 Number of DPT immunisations for individuals 0–5 years old

	DID			Staggered DID		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
kelurahan (=1)	-2.000***	-2.130***	-2.343***	-2.000***	-2.124***	-2.338***
	[0.000]	[0.011]	[0.077]	[0.000]	[0.016]	[0.112]
area (urban=1)	0.211***	0.268***	0.079	0.211***	0.275***	0.086
	[0.000]	[0.007]	[0.071]	[0.000]	[0.009]	[0.102]
male (=1)		-0.002	-0.003		-0.036	-0.036
		[0.030]	[0.030]		[0.041]	[0.040]
age in years		0.183***	0.183***		0.183***	0.184***
		[0.013]	[0.013]		[0.017]	[0.017]
(log) hh exp cap			0.158***			0.158*
			[0.058]			[0.084]
Observations	3,755	3,755	3,755	2,023	2,023	2,023
R-squared	0.521	0.562	0.564	0.526	0.567	0.569
Village FE	Yes	Yes	Yes	Yes	Yes	Yes
Village specific time FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1, Clustered standard errors at village level in brackets. All columns are estimated using linear model.

Table 3.19 Number of single dosage immunisations for individuals 0–5 years old, Poisson model

	Measles			BCG		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
kelurahan (=1)	-0.251***	-0.406***	-0.445***	-0.405***	-0.440***	-0.470***
	[0.000]	[0.019]	[0.076]	[0.000]	[0.009]	[0.059]
area (urban=1)	0.111***	0.178***	0.145**	0.111***	0.131***	0.105*
	[0.000]	[0.010]	[0.063]	[0.000]	[0.006]	[0.054]
male (=1)		0.004	0.004		-0.014	-0.014
		[0.030]	[0.030]		[0.022]	[0.023]
age in years		0.179***	0.180***		0.050***	0.051***
		[0.015]	[0.015]		[0.010]	[0.009]
(log) hh exp cap			0.028			0.022
			[0.053]			[0.045]
Observations	2,023	2,023	2,023	2,023	2,023	2,023
Village FE	Yes	Yes	Yes	Yes	Yes	Yes
Village specific time FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1, Clustered standard errors at village level in brackets. All columns are estimated using Poisson model. Results are the coefficients and not the marginal effects.

Table 3.20 Number of multiple dosage immunisations for individuals 0–5 years old, Poisson model

	Hep B			DPT		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
kelurahan (=1)	-1.350***	-1.418***	-1.526***	-1.099***	-1.168***	-1.265***
	[0.000]	[0.011]	[0.055]	[0.000]	[0.010]	[0.054]
area (urban=1)	0.111***	0.149***	0.054	0.073***	0.109***	0.024
	[0.000]	[0.006]	[0.047]	[0.000]	[0.005]	[0.046]
male (=1)		-0.022	-0.022		-0.015	-0.015
		[0.019]	[0.019]		[0.019]	[0.019]
age in years		0.094***	0.094***		0.093***	0.094***
		[0.011]	[0.011]		[0.009]	[0.008]
(log) hh exp cap			0.078**			0.070*
			[0.038]			[0.038]
Observations	2,023	2,023	2,023	2,023	2,023	2,023
Village FE	Yes	Yes	Yes	Yes	Yes	Yes
Village specific time FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1, Clustered standard errors at village level in brackets. All columns are estimated using Poisson model. Results are the coefficients and not the marginal effects.

Table 3.21 Number of measles immunisations for individuals 0–5 years old, by region

	Java			Non-Java		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
kelurahan (=1)	-0.190***	-0.293***	-0.245**	-0.100***	0.090***	0.045
	[0.000]	[0.016]	[0.093]	[0.000]	[0.020]	[0.037]
area (urban=1)	-	-	-	0.105***	0.147***	0.053
				[0.000]	[0.007]	[0.060]
male (=1)		0.013	0.013		-0.011	-0.012
		[0.039]	[0.039]		[0.030]	[0.030]
age in years		0.140***	0.140***		0.125***	0.127***
		[0.018]	[0.018]		[0.012]	[0.012]
(log) hh exp cap			-0.035			0.079
			[0.073]			[0.049]
Observations	935	935	935	1,088	1,088	1,088
R-squared	0.246	0.332	0.333	0.386	0.479	0.481
Village FE	Yes	Yes	Yes	Yes	Yes	Yes
Village specific time FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1, Clustered standard errors at village level in brackets. All columns are estimated using linear model. The area dummy for Java is missing due to perfect multicollinearity between the group and variable, meaning all villages in Java in this estimation are in urban areas.

Table 3.22 Number of Hepatitis B immunisations for individuals 0–5 years old, by region

	Java			Non-Java		
VARIABLES	(1)	(2)	(3)	(4)	(5)	(6)
kelurahan (=1)	-1.905***	-2.009***	-2.016***	0.150***	0.428***	0.239**
	[0.000]	[0.020]	[0.110]	[0.000]	[0.048]	[0.089]
area (urban=1)	-	-	-	0.316***	0.381***	-0.012
				[0.000]	[0.014]	[0.149]
male (=1)		-0.042	-0.042		-0.046	-0.051
		[0.061]	[0.061]		[0.049]	[0.049]
age in years		0.157***	0.157***		0.179***	0.184***
		[0.023]	[0.023]		[0.031]	[0.031]
(log) hh exp cap			0.005			0.329**
			[0.086]			[0.123]
Observations	935	935	935	1,088	1,088	1,088
R-squared	0.557	0.588	0.588	0.494	0.529	0.538
Village FE	Yes	Yes	Yes	Yes	Yes	Yes
Village specific time FE	Yes	Yes	Yes	Yes	Yes	Yes

Notes: *** p<0.01, ** p<0.05, * p<0.1, Clustered standard errors at village level in brackets. All columns are estimated using linear model. The area dummy for Java is missing due to perfect multicollinearity between the group and variable, meaning all villages in Java in this estimation are in urban areas.

4 THE IMPACT OF FEMALE LEADERSHIP ON PUBLIC SPENDING AND SERVICE ACCESS

4.1 Introduction

In terms of both economic and political positions, women remain underrepresented in many crucial areas of society. Parliaments and leadership roles throughout the world are still dominated by men, making the under-representation of women a challenge for governance that allegedly represents all groups in society (Hillman, 2017). However, in the past decade, there has been an increase in women taking up leadership positions in both the public and private sectors of many countries and the United Nations' Sustainable Development Goals (SDGs) has incorporated the effective participation of women in political leadership as one of its long-term action goals. These recent changes have attracted the interest of social science studies that seek to understand whether there are implications of female leadership, or a lack thereof, for public policy outcomes.

Previous studies on the impacts of female leadership concentrate on the notion of politicians' preferences. According to Anthony Downs' classic work of 1957, any politician, regardless of their gender, adjusts their policy platforms to accommodate the median voters' preferences (Clots-Figueras, 2012; Ferreira and Gyourko, 2014). Downs argues further that the preferences of a politician should not influence policy outcomes since whoever is elected implements policies that reflect electoral preferences. This view, however, has been contested by another hypothesis that shows a divergence in policies by elected officials in the absence of complete policy commitment from the politicians to accommodate voters' preferences. Based on this framework, a politician's preferences may differ according to their gender and therefore female leaders may implement policies that are more highly correlated with their preferences for public goods (Besley and Coate, 1997; Clots-Figueras, 2012, 2011).

This study investigates whether the gender of district leaders has influence on fiscal and service outcomes in Indonesia. The sample used in this study is drawn from districts that had local direct elections for heads and vice-heads between 2005 and 2017. District government in Indonesia is responsible for the provision of basic public services and is accountable to the electorate. Following several prior studies (Brollo and Troiano, 2016; Ferreira and Gyourko, 2014), I focus my analysis on executive leadership, such as head and vice-head level leadership, and not legislative representation. In Indonesia, although budget laws need to be

passed by local parliaments, heads and vice-heads have more power in facilitating the reallocation of resources in a district. On the other hand, legislators, in passing legislation, may need to negotiate with other parliament members, so the impact of female parliament members on policy outcomes may only be noticeable when there is a significant increase in female members in parliament.

The key challenge in estimating the causal effect of female leaders is that female leadership is not distributed randomly across districts. There could be unobserved district characteristics that influence female candidacy. These characteristics are likely to affect not only electoral outcomes, but also policy outcomes. To account for this challenge, I identify the causal effect of a female leader by comparing districts where a female candidate won the election by a small margin against a male candidate with other districts where a female candidate results in second place and a male candidate is the winner. This design will allow me to assume that female leadership is distributed as good as randomly assigned. By focusing on sufficiently close elections, the correlation between policy outcomes and unobserved district characteristics can be assumed to be close to zero. In other words, female leaders who are elected through close elections against male candidates are presumably elected under similar socio-economic and political factors as male leaders who win by only a small margin against female candidates.

Several studies have empirically tested the impact of female leadership against certain policy issues that are traditionally perceived as falling under women's responsibilities in society. Duflo (2012) discusses that women and men have different policy preferences and women should be in favour of policies that better reflect their priorities and their traditional role in the society. Previous literature has found that women leaders improve policy outcomes in areas of society that women generally take charge of, such as health, education, and social welfare. Female leaders in office and cabinet are associated with higher government spending for social welfare and public health (Brollo and Troiano, 2016; Chattopadhyay and Duflo, 2004; Chen, 2010; Mavisakalyan, 2014). In terms of service access, female political representation increases the probability that individuals in society will attain a primary level of education (Clots-Figueras, 2011). In the case of Indonesia, female parliament members and district heads are associated with higher health spending and better access to sanitation (Suci et al., 2020).

The closest predecessor to this paper is the unpublished work conducted by Suci et al. (2020), which observes the impacts of female representation on local politics by using Indonesian district data between 2010 and 2017. This paper differs from Suci's in several important ways. First, this paper provides a causal estimate of female leadership impacts in Indonesia. In analysing these impacts, I analyse the effect of variation in gender of district leaders due to electoral outcomes rather than gender quota policies. The identification strategy used in this study will utilise the election results to establish random assignment of female leadership. Second, given that candidates run as pairs (head and vice-head) in Indonesian elections, this paper will be able to test such impacts not only for the highest structural positions such as head of district, but also for the broader definition of leadership positions, which includes vice-heads.

The remainder of the paper is organised as follows. Section 2 collates the current literature on how and why female political leaders affect fiscal and service outcomes. Section 3 describes the institutional context of female political leadership in Indonesia. Section 4 explains the identification strategy and describes the data used and the choice of outcome variables. Section 5 illustrates and explains the results obtained. The sixth and seventh section present the results on sensitivity analyses and discussion on potential mechanism. The final section concludes and provides a policy discussion on female leadership.

4.2 Female leaders and policy outcomes

Literature has documented fundamental differences on preferences between men and women, which includes risk, social, and competitive preferences (Croson and Gneezy, 2009). In particular, for social preference, women are observed to be more sensitive to social cues than men. As a result, female and male leaders are found to be differentially effective in many settings (Eagly et al., 1995). One mechanism to explain this different preference is that the gender role expectation factor further spills over into leadership roles and therefore produces differences between male and female leaders. Another factor that shapes differences in behaviour between men and women is the constrained environment. Given the difficulty for women to gain recognition and leadership positions, successful female leaders generally seek leadership styles that are more egalitarian and supportive of others (Eagly and Carli, 2003).

According to the seminal work of Anthony Downs in 1957 as summarised by Clots-Figueras (2012), the preferences of a politician should not impact policy outcomes. This framework

assumes that political candidates have a firm commitment to implement specific policies and only care about election results. If this is the case, the gender of political leaders should not matter for policy outcomes because the equilibrium of the policies will reflect the preferences of the electorate, which is the median voter group. Women make up part of the citizens who vote, and hence their preferences will be represented by the elected leaders, irrespective of their gender.

This view of the political process was later challenged by studies that show divergence in policies among political leaders. Besley and Coate (1997) developed a citizen-candidate model that suggests the motivation of citizens to run in an election or not is a trade-off between the cost of running and the probability of being elected and getting to implement their preferred policies. In this setting, elected candidates care about certain outcomes more than others, and once in power they cannot completely commit to the moderate or preferred policies of the electorate. Hence, the gender of the politician matters for policy purposes, as female leaders will implement policies that are more correlated with their preferences.

Given these two contrasting frameworks, whether female politicians will have an impact on policy outcomes remains to be tested empirically. Cross country analyses have found that an increase in the share of women in cabinet is associated with an increase in public health spending (Mavisakalyan, 2014). A similar finding has been reported for the U.S. case where health care spending, particularly for poorer children, disabled people, and the elderly increased when the number of women representatives was substantial, regardless of party platforms (Courtemanche and Green, 2017). For the case of developing countries, many studies have discussed female representatives in the Indian parliament. Increasing female legislators points to more investment in health and early education, and also redistributive policies such as land reform (Clots-Figueras, 2011). This increase in the number of female members of parliament is also found to expand the attainment of primary level education in urban areas (Clots-Figueras, 2012).

The correlation between female representation in politics and policies leaves another question of whether this outcome reflects the gender of the politicians or the preferences of voters who elect female leaders in the first place. Given this concern, several studies have implemented more careful research methods to observe plausible exogenous variation in female leader assignment and argue for causality. Chattopadhyay and Duflo (2004) is one of the early studies that uses the random setting of reservation policies in India, in which one-third of

village head positions have been randomly reserved for women since the mid-1990s. They found that elected female village leaders invest more in drinking water access. Several studies also exploit close election results between male and female candidates, to implement the regression discontinuity design in identifying the causal effect of electing a female mayor. In Brazil, female mayors spend more on education, health care, and social assistance (Brollo and Troiano, 2016). This influence of female mayors is not conditional on the share of women members in local councils. In contrast to most research, Ferreira and Gyourko (2014) reveal that the gender of the mayor does not seem to have effects on short- or long-run policy choices of U.S. cities. They argue that the Tiebout sorting mechanism from the citizens and intense competition among local governments may prevent divergent policies among local governments.

Overall, most of the studies mentioned above report consistent evidence that female politicians influence health, social protection, and education outcomes. However, the literature also notes some findings that corroborate the median voter view of the political process, in which the gender of leaders may not matter. These differing results point to the importance of country-specific contexts in which female leadership is undertaken, since gender effects may differ by market institution and socio-economic settings (Ferreira and Gyourko, 2014).

4.3 Female local leaders in Indonesia

The era of decentralisation in Indonesia began through the effective implementation of several laws that contain the spirit of local autonomy: 1) Law 22/1999 on Regional Government which began being implemented in 2001, 2) Law 25/1999 on Fiscal Balance between the Central Government and the Regions which also began being implemented in 2001, and 3) Law 34/2000 on Regional Taxes and Levies.[27] In addition to the fiscal side, political decentralisation has become an important concern in recent decades (Lewis and Hendrawan, 2019).

At the beginning of the decentralisation era in 1999, local leaders were appointed by members of local parliaments, whereas before this, central authorities appointed mayors. Indonesia commenced the direct election of district heads in 2005. Since then, direct elections

[27] After the implementation of these laws, the government replaced the former two laws with Law 32/2004 on Regional Government and Law 33/2004 on Fiscal Balance between the Central Government and the Regions. Later, in 2008, Law 32/2004 was further revised with Law 12/2008.

have been implemented in a gradual fashion, as the appointed heads' terms ended. Heads and vice-heads run as pairs and can be nominated by a party, a coalition of parties, or run as independent candidates. The standard office term for each pair is five years and a two-term limit is applied for district heads. This initiation of direct election has encouraged women's participation in local politics to a certain extent. Dewi (2015) elaborates an argument in which direct election provides structural opportunities for women to be recruited into politics. Female politicians can now directly engage with voters and gain votes from constituents directly without facing the institutional barriers of oligarchic and male dominated political systems, such as the local parliament.

However, cultural and structural factors may still play some role in the obstacles faced by women in the political arena. In the 2018 World Values Survey (WVS) conducted in Indonesia, 72 per cent of respondents either agreed or strongly agreed that men made better political leaders than women.[28] Taking the case of India, Gangadharan, Jain, Maitra, & Vecci (2016) found that backlash behaviour towards female leaders is not a result of perceived ineffectiveness or performance of women leaders but rather linked to the norms of gender identity. Therefore, social norms are important drivers of perceptions and behaviour towards female leaders.

Although Indonesia does not have a gender quota policy for local leader positions, it adopted a quota policy for parliament members via Law 12/2003, which was revised by Law 10/2008. Law 12/2003 stipulates that each participating political party should have a minimum of 30 percent female candidates in legislative nominations. Furthermore, Law 10/2008 strengthens the requirement by making the minimum 30 percent a compulsory precondition for a party in order to be eligible to contest elections. The revised law also demands that parties have at least one female candidate among the top three ranked candidates on each party nomination list.[29] Despite the lack of penalties for non-compliance, the law contributed to a higher proportion of women candidates in 2009[30] compared to previous elections (Hillman, 2017). Women accounted for almost 35 percent of candidates contesting national, provincial, and

[28] Source: https://www.worldvaluessurvey.org/WVSDocumentationWV7.jsp. The most recent survey data for Indonesia is from 2018, while the previous wave is from 2006. In 2006, 59 per cent of respondents supported this statement.
[29] According to Law 10/2008, local parliament seats are allocated to a party in proportion to its vote share. Then, for candidates within a party to be elected, they need to obtain at least 30% of the electoral divisor number (number of valid votes divided by the number of seats in the electoral area). If the number of candidates within a party who pass the minimum votes is greater than the number of seats obtained by that political party, the seat will be given to the candidate who is ranked higher in the party-nomination list.
[30] Since the beginning of decentralisation, representatives of sub-national parliaments have been elected through parliamentary elections in 1999, 2004, 2009, 2014 and 2019.

district parliaments. However, the proportion of women who made it to the assembly remains limited. In 2014, women only won 14 percent of district parliament seats. This is because the legal candidate quota does not guarantee that women candidates will be elected. According to Hillman (2017), whether quotas translate into more female representation depends on the degree of secularisation and whether females are actively involved in political activities such as voting in the election and being ranked members in political parties.

During the early waves of local direct elections, there was significant variation in the number of women joining political contests across regions. Local mayoral elections that involve female candidates (either as heads or vice-heads) are more evident in Java. From 2005 to 2008, Indonesia held 466 direct elections consisting of 355 elections outside Java and 111 elections in Java. Throughout that period, the percentage of elected pairs including at least one female in Java was almost 10 percent (11 elected pairs out of 111 elections), compared to outside Java which was about 4 percent (15 elected pairs out of 355 elections). Under the previous assignment mechanism before 2005, within districts in Java, only five female leaders were ever elected to office by members of local parliaments (Dewi, 2015). This means that the number of female leaders in local politics in Java has doubled under the direct election era.

Figure 4.1 shows the frequency distribution of districts that elected females as either their heads or vice-heads between 2005 and 2017. The blue shade represents the number of elections in which districts have elected female candidates in office, either as head or vice-head, with darker blue meaning more periods. Between 2005 and 2017, 93 districts had elected pairs that included at least one female. Almost half of those, 43 districts, are located in Java. The central and eastern parts of Java have the highest numbers of districts that have ever elected female leaders. There were 17 and 12 districts with female elected leaders during that period in Central and East Java, respectively. Over the years, 30 districts elected female leaders in the office more than once, in which six of those districts elected female leaders for three electoral periods.

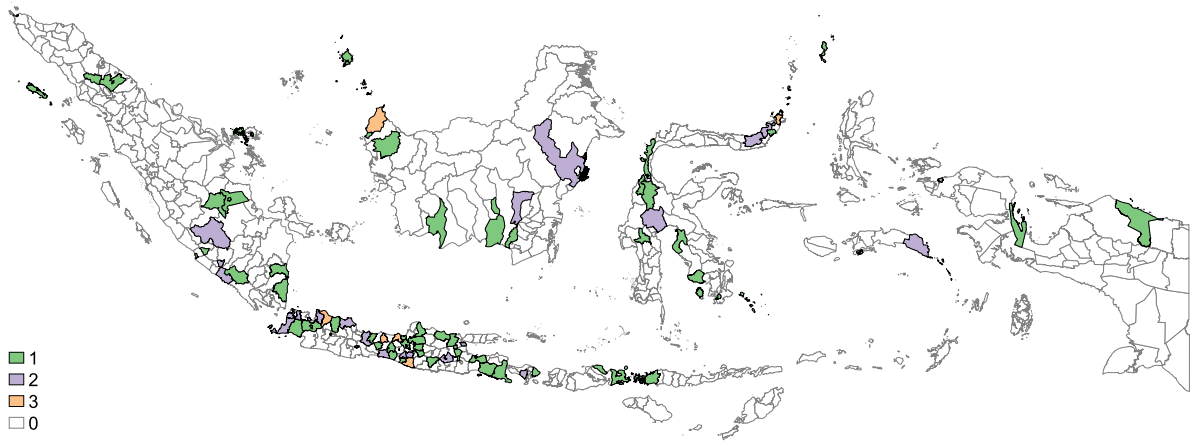


Figure 4.1 Number of elections that elected female leaders between 2005 and 2017

Over the years, it has been widely expected that women's involvement in political contests would increase, given the rising awareness and media coverage of women candidates (Hillman, 2017). Table 4.1 displays the number of elections that involved pair(s) with at least one female between 2005 and 2017. In total, 5538 pairs contested district elections during that period, and almost 10 percent of those (543 pairs) included at least one female candidate either as head and or vice-head. As shown in Table 1, out of 543 pairs, more than 40 percent had female candidates for the head of the district. Looking at the election results, almost half of the 543 pairs ended up either as the election's winner or runner-up.

Table 4.1 Number of pairs with female candidates in each election

Election year	Total	Female as candidates for			Election results	
		Head	Vice-head	Both	Win	Runner-up
2005	60	15	45		12	13
2006	21	9	12		8	2
2007	15	4	11		1	3
2008	42	14	28		3	7
2010	109	44	64	1	23	26
2011	36	18	18		5	14
2012	34	19	15		8	6
2013	72	34	37	1	13	15
2015	117	52	63	2	44	37
2017	37	21	16		14	10
Total	543	230	309	4	131	133

4.4 Identification and data

Identification strategy

This study examines the impact of female leadership on fiscal and service outcomes. However, identifying causal effects of the policymaker's gender on policy outcomes is not straightforward. Female leadership is not randomly assigned; hence, comparison between districts with a female leader and those with a male leader may generate biased estimates due to endogeneity issues. Policy decisions might be correlated with district-specific characteristics that could also influence female candidacy and gender of the district leaders. For example, citizens in one district may have specific attitudes toward women that make them favour female executive leadership and higher levels of maternal health interventions. If so, an estimated effect of female leaders on maternal health service access could be partially due to an unobserved factor, namely the preference of the citizens towards public goods. Exclusion of this factor would result in omitted variable bias.

This paper adopts a randomisation-based regression discontinuity (RD) design to deal with these endogeneity issues (Cattaneo et al., 2017, 2015). The intuition of this strategy is that districts in which a male candidate won against a female candidate by a narrow margin can be a good counterfactual for those districts in which the opposite election result occurred: a female candidate won the election and a male candidate was in second place. The identification basically derives from the uncertainty elements in the final outcome of elections with small margins of victory between the winner and runner-up. Following similar studies that exploit close elections as their identification strategy (Brollo and Troiano, 2016; Ferreira and Gyourko, 2014), I assume that for mixed-gender elections with small wining margins between winner and runner-up, the probability of winning is the same for both female and male candidates.

The limited number of districts with local elections involving women candidates creates some practical concerns in estimating an RD design. The small number of observations reduces the statistical power associated with the derived estimation results under the standard continuity-based RD approach (Lewis et al., 2020). Therefore, this study applies alternative RD procedures that are more appropriate for smaller sized samples, named randomisation-based RDD (Cattaneo et al., 2017, 2015). This framework assumes there is a window around the cut-off that is determined by the winning vote margin, where estimation and inference echo

the randomised experiment. The first step of the procedure is to find the widest feasible bandwidth which guarantee that all covariates are balanced. Covariate balance can be achieved if its mean value is not statistically significantly different between each side of the cut-off. Once a suitable window has been selected, the treatment effect is the difference in means adjusted by the winning margins and the interaction between the treatment dummy and winning margin on either side of the cut-off.

In the context of this study, the winning margin (X) of the female candidate is the running variable and the cut-off is 0. The right of the cut-off indicates a positive winning margin of pairs with female candidates in which female leaders are elected. The left of the cut-off indicates a negative winning margin, meaning that a pair with female candidates ends up as runner-up and male leaders are elected. As this framework is an extension of the standard RD design, the basic empirical model of this study is similar to the one proposed by Imbens and Lemieux (2008). I define Y as potential public spending and service access outcomes of interest, where $Y_{it}(0)$ is the outcome for districts with male elected leaders at year t and $Y_{it}(1)$ is the outcome for districts with female elected leaders at district i. Therefore, the impact of female leaders is given by $Y_{it}(1) - Y_{it}(0)$. However, for each elected pair, I can only observe one of the outcomes, and not simultaneously for both outcomes, so attention turns to to the average effects over subgroups of population. Let $D_{it} = 1$ if a district was exposed to a female elected leader and $D_{it} = 0$ if district was not exposed to a female elected leader. Observed policy outcomes, Y_{it}, are therefore equal to $Y_{it}(0)$ if $D_{it} = 0$ and $Y_{it}(1)$ if $D_{it} = 1$. The average causal effect of the female elected leaders, τ, given the cut-off, c=0, is:

$$\tau = \mathrm{E}\left[Y_{it}(1) - Y_{it}(0) \mid X_{it} = c\right] = \mathrm{E}\left[(Y_{it}(1) \mid X_{it} = c\right] - \mathrm{E}\left[Y_{it}(0) \mid X_{it} = c\right] \qquad (1)$$

Equation (1) implies that outcomes just to the left of the cut-off can be employed as valid counterfactual for those just to the right of the cut-off, within the selected window or bandwidth. If there is a positive influence of having female leaders on policy outcomes, then τ is expected to be positive and statistically significant.

Furthermore, I argue that winning vote margins can be directly related to public spending and service outcomes. For instance, elected candidates with larger winning vote margins have a greater mandate to implement their desired policies and are therefore better at their role in providing service access. Based on this argument, I assume that winning margin as the running variable does not meet the necessary exclusion restriction that requires it to has no

direct impact on outcomes, and therefore I employ a polynomial of degree one of the running variable. This is the standard assumption that usually holds in continuity-based RD and I adopt the same assumption to be applied in this randomisation-based RD model.

By setting the polynomial degree at one, I perform a linear transformation, letting the slopes be subtracted from the outcomes and hence leaving a residualised version of the outcome that differs only in the intercept between sides of the cut-off (Cattaneo et al., 2016). In other words, for each side of the cut-off, a regression of the policy outcomes on the treatment dummy (D), the winning margin (X), and the treatment dummy interacted with a polynomial degree one of the winning margins is estimated, and then the difference in the intercepts between both sides is used as the test statistic. I also use a triangular kernel density that employs higher weights for observations closer to the cut-off.

$$Y_{it} = a + \rho D_{it} + g(X_{it}) + D_{it} * g(X_{it}) + \varepsilon_{kt} \quad (2)$$

In this paper, I define female leaders as females who have positions as district head or district vice-head. Given that candidates run as pairs (head and vice-head) in Indonesian local elections, I therefore analyse the election results in a pair setting. The treatment group comprises districts that elected pairs with at least one female candidate to run the office, whereas the control group includes districts that elected all-male pairs for head and vice-head positions. Throughout the paper, the implicit assumption is that vice-heads have significant influence over policy-making and therefore can be treated as local leaders.

Data and choice of variables

I use data from 1294 district first-round elections between 2005 and 2017 provided by the General Elections Commission (KPU), collected by Lewis et al. (2020)[31]. Out of these elections, I only utilise districts with local elections involving close winning margins between women and men candidates (mixed-gender elections). As a treatment group, I use districts that elected pairs with at least one female candidate as election winner and pairs with all male candidates as runner-up.[32] For the control group, I use districts that elected pairs with all male candidates as election winners and pairs with at least one female candidate as runners-up. I

[31] These data are not yet publicly available. They were compiled by Adrianus Hendrawan who kindly provided the data.

[32] Initially I also considered estimating another set of samples where I would define a treatment group as districts with a female head, regardless of the vice-head's gender. Using this definition, I only have 123 observations with 59 winning cases and 64 cases of runner-up. The number of observations apparently is not sufficient to draw the randomisation inference RD and the window selection process is problematic. Therefore, I decide to exclude this analysis from the study and consider this set as potential future work as more data becomes available.

exclude districts that have pairs with female candidates both as winner and runner-up of the local elections and districts with single candidate elections.

Before 2015, to win an election, a pair of head and vice-head needed to secure a certain portion of the votes. If no candidate secured this minimum share, then the top two pairs faced off in a second round.[33] According to Lewis (2019), only 7 percent of total district elections were not concluded during the first round. In the cases included in this study, only two local elections involving female candidates as winner or runner-up did not conclude in the first round. Given so few cases, I decided to exclude these two elections from the analysis. The final relevant sample of the analysis is 214 observations consisting of 106 cases of female candidates as election winners and 108 cases with female candidates as runner-up.

On the outcome variables, deciding which policy outcomes will be affected by the gender of the leaders is not straightforward. Some outcomes, especially public spending, are not targeted specifically according to the gender of the beneficiaries (Clots-Figueras, 2012). Therefore, in this paper, I will further assume that the overall improvement on outcomes benefits women and reflects better representation of women's preferences. For fiscal outcomes, the dependent variable is the average of per capita expenditure (using 2010 constant prices) as well as share of total expenditure, from one year after election (t+1) to four years after election (t+4), or to the most recent year available. On per capita expenditure, I look at total expenditure as well as the sectoral classification, such as health and social protection expenditure. Government expenditure data is compiled from Ministry of Finance via the INDO-DAPOER database[34]. For service outcomes, the measure of local service delivery that I employ in this paper is service access. This service access data represents both demand and supply side interaction. I use the average annual delta of service access from one year after election (t+1) to four years after election (t+4), or in some cases, to the most recent year available. The service outcomes include access to sectoral service that are the responsibility of district government: education, health, and local infrastructure. Outcomes for education access are net enrolment rate for primary, junior secondary and senior secondary level. For health, I use the percentage of births attended by health professionals. I also employ three outcomes that represent whether citizens have access to such infrastructure:

[33] Between 2005 and 2008, the vote threshold was 25%. Between 2008 and 2013, the threshold was 30% of the votes. Since 2015, the elections have only been administered for one round and the pair with the largest vote is elected.
[34] Indonesia Database for Policy and Economic Research (INDO-DAPOER) is accessible at http://databank.worldbank.org/data/reports.aspx?source=indo~dapoer-(indonesia-database-for-policy-and-economic-research)

percentage of households with access to safe water, percentage of households with access to safe sanitation, and percentage of villages with asphalt road. The selection of these services also considers data availability. Service access data are collected from the annual household survey (SUSENAS) by the Indonesian Bureau of Statistics (BPS). As for other covariates which gauge district characteristics, such as poverty rate, population, area, and GRDP per capita, the data are compiled from BPS. A table of summary statistics of the variables is presented in the Appendix.

4.5 Results

Window selection

The window selection procedure is a data-driven process of conducting balance tests on pre-treatment covariates (Cattaneo et al., 2016). I select the window based on predetermined covariates that would influence both election outcomes as well as policy outcomes in a district. As recommended by Cattaneo et al. (2015), I set the minimum sample size to roughly include at least 10 observations on either side of the threshold in order to have enough observations to conduct the balance tests in the smallest possible window. Therefore, the minimum window that I consider is [-0.05, 0.05], because within this window there are 14 observations to the left and 16 observations to the right of the cut-off.

In each window, I perform randomisation-based tests of the sharp null hypothesis of no treatment effect of each of the predetermined covariates. Following Lewis (2019), the list of covariates that I use to determine the appropriate bandwidth are: poverty rate, log of population, log of area (in km square), log of intergovernmental transfer per capita, log of GRDP per capita, effective number of political parties, ethnic fractionalisation index, average service access, dummy of Java island, and dummy of urban districts.[35] All covariates are measured in the year just prior to the election year. I also add two additional election covariates: a dummy variable of whether candidates are incumbent or not and the number of candidates running in the election.[36] I use a standard significance level of 0.15 to test the

[35] Districts are identified based on the official GoI classification on *kabupaten* (rural districts) and *kota* (urban districts).

[36] Studies have also acknowledged that political dynasties are quite apparent in the case of female leaders and may play a role in winning elections. For instance, Wardani and Subekti (2021) provide evidence suggesting that 44 percent of female candidates elected to Indonesia's national parliament in 2019 were members of political dynasties. However, their study did not investigate the case of district heads elections and, after all, male candidates may also be parts of dynasties. From a technical point of view, the unavailability of political dynasty information, together with more specific data on the socio-economic background of each candidate running, makes these factors plausible but unobservable, which is also assumed to be balanced between the specified windows.

randomisation assumption within the window, as suggested by Cattaneo et al. (2015). I also specify the number of observations to be added at each side of the cut-off at each step as a minimum as 1 and use 1000 replications for the randomisation test. The largest window in an ordered sequence of trials where the minimum p-value exceeds 0.15 is the valid window that I use for estimating the treatment effects.

The window selection procedure suggests that a bandwidth of 0.066 on either side of the cut-off is appropriate. Table 4.2 provides the minimum p-values for the different consecutive windows considered for this analysis. The minimum p-value in the chosen window is 0.1922 and the minimum p-value for the next largest window, [-0.070, 0.070], is 0.118, which is smaller than significance level of 0.15. Figure 4.2 then illustrates that the p-values vary for very short windows but once the window is large enough it stabilises approaching zero. The vertical reference line denotes the suggested window. It shows that after the suggested bandwidth, the minimum p-value has reached below 0.15. The figure supports the decision to use the selected window [-0.066, 0.066] with 21 observations to the left of the cut-off and 19 observations to the right of the cut-off [37]. This selected bandwidth is around a third from the average winning margin, both for the winning cases (0.17) and losing cases (-0.18) of pairs with female candidates, and therefore can also be considered as a suitable approximation in defining close election margins. Detailed statistics and the distribution of winning margin are presented in the Appendix.

Table 4.2 Randomisation-based p-values from balance tests for different windows

Window	Minimum p-value	Covariate with minimum p-value	Obs<cut-off	Obs>cut-off
[-0.050,0.050]	0.176	Incumbent dummy	14	16
[-0.053,0.053]	0.158	Incumbent dummy	15	17
[-0.054,0.054]	0.224	Average service access	16	18
[-0.063,0.063]	0.274	Incumbent dummy	18	18
[-0.066,0.066]	0.192	Area	21	19
[-0.070,0.070]	0.118	Incumbent dummy	22	20
[-0.072,0.072]	0.108	Incumbent dummy	22	20
[-0.074,0.074]	0.092	Incumbent dummy	23	20
[-0.076,0.076]	0.048	Incumbent dummy	24	21
[-0.079,0.079]	0.072	Incumbent dummy	24	22
[-0.083,0.083]	0.024	Incumbent dummy	25	23
[-0.097,0.097]	0.034	Incumbent dummy	25	26

Notes: All covariates, except for the election variables, are measured in the year just prior to the election year (t-1, with t is the election year)

[37] When running the estimation, the number of observations varies across outcomes due to some missing data.

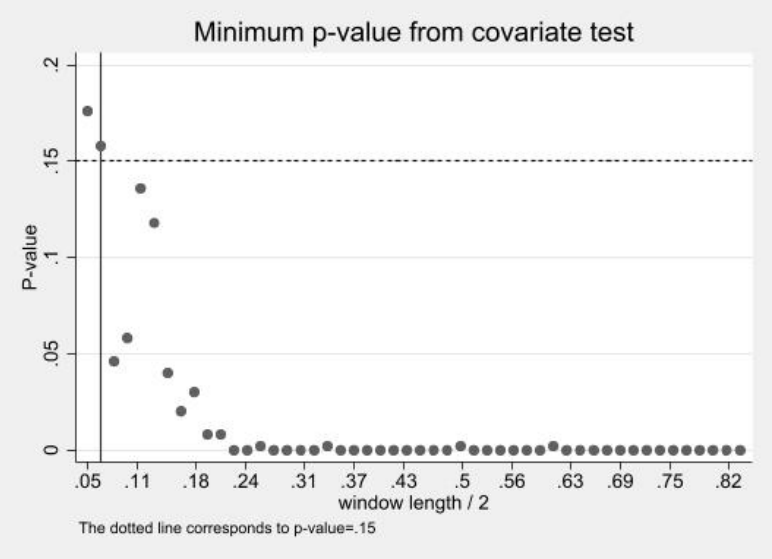


Figure 4.2 Plot of p-values

Fiscal and service delivery outcomes

For policy outcomes, I firstly investigate the fiscal performance of female leaders compared to male leaders, and then check whether the effect persists in more concrete service delivery performance, which is observable at the district level.

Table 4.3 provides the estimation results of female leadership effect on fiscal outcomes. The table presents that overall, districts governed by at least one female leader have higher expenditure per capita compared to districts governed by male leaders. Looking in detail into sectoral spending, districts with female leaders in the office either as head and or vice-head, have higher spending per capita on all major basic services spending: health, education, social protection and infrastructure. Social protection spending has the highest coefficient of estimates followed by infrastructure then health spending. Districts with female leaders had more than doubled infrastructure spending per capita and spent more than twice as much per capita on health, compared to districts without female leaders. However, the interpretation of these coefficients' magnitudes should be treated with caution, given the small number of districts in the sample. The table further shows that there is no statistically significant difference in the generation of own-source revenue per capita across districts governed by female and male leaders. Hence, given that transfer per capita is balanced among treatment and control groups, this increase in total spending may be attributed to a smaller surplus than otherwise, or a budget deficit[38].

[38] Unfortunately, the data does not allow for detailed investigation of surpluses or reserves at district level.

Table 4.3 Per capita fiscal outcomes under local randomisation

Dependent variable	Difference in means (finite sample)	
Average from (t+1) election to (t+4) election	**τ**	**p**
(log)total exp per capita	0.959	0.000
(log)education exp per capita	0.587	0.000
(log)health exp per capita	0.936	0.000
(log)social protection exp per capita	1.321	0.000
(log)infra exp per capita	0.939	0.000
(log)own-source revenue per capita	-0.214	0.294

Notes: Selected bandwidth is [-0.066,0.066]. The dependent variables are listed in column one. Budget figures are in 2010 constant prices. τ is the estimated treatment effect and p is the p-value associated with the estimated treatment effect. The outcome variable is the average across the office period from (t+1) to (t+4) election year, or until the most recent data available. Other sectors outside basic service sectors are general administration, housing, agriculture, economy, environment, and agriculture. Results for these other sectors are available in the Appendix.

To further examine the effect of female leaders on budget policy, I then estimate the effect of having female leaders on share of total expenditure for each of the major sectors. Table 4.4 shows that districts with female leaders spend a bigger portion of their budget on social protection and community infrastructure[39]. For instance, female leaders induced 5.5 percentage points more infrastructure spending share out of the total budget. These findings corroborate the per capita increase in the previous table, that involvement of female leaders in the office leads to more investment in both sectors. However, the education budget share shows a negative coefficient and therefore indicates that female leadership may shift spending from education to other sectors. One plausible explanation for this is that most districts in Indonesia have already met the mandatory spending requirement of 20 percent of the total budget on education. The law on this mandatory spending was introduced in 2003 but did not become officially effective until 2009 after the constitutional court issued its ruling on the interpretation of the mandated rule. Higher education public spending has also been found to have negligible effect on education improvement for districts with significant levels of corruption (Suryadarma, 2012).

[39] Infrastructure spending in local government includes construction of waste management facilities, provision of drinking water supply systems and irrigation construction, in addition to some betterment of road facilities. Most road construction is funded through line ministry spending (central government spending). More detailed description of sectoral spending is available in the Appendix.

Table 4.4 Sectoral share from total budget under local randomisation

Dependent variable	Difference in means (finite sample)	
Average from (t+1) election to (t+4) election	**τ**	**p**
Education budget share	-0.108	0.000
Health budget share	-0.010	0.312
Social protection budget share	0.005	0.006
Infrastructure budget share	0.055	0.008

Notes: Selected bandwidth is [-0.066,0.066]. The dependent variables are listed in column one. τ is the estimated treatment effect and p is the p-value associated with the estimated treatment effect. The outcome variable is the average across the office period from (t+1) to (t+4) election year, or until the most recent data available. Other sectors outside basic service sectors are general administration, housing, agriculture, economy, environment, and agriculture. Results for these other sectors are available in the Appendix.

The next analysis is the effect of having female leaders on service delivery performance. I investigate the service delivery of three basic services sectors, which are education, health, and infrastructure. I use enrolment rates as an outcome for education, and access to water, sanitation, and roads for infrastructure. However, due to constraint of data availability, I can only use one outcome for health which is the percentage of births attended by health professionals.

Table 4.5 presents the effect of female leadership on service access and three key findings emerge. First, districts with female elected leaders on average have higher rates of assisted birth compared to districts governed by male elected leaders. Districts with female leaders on average have a 2.6 percentage point higher annual increment and this difference is significant at confidence level 99%. Second, infrastructure service access shows mixed results, with a statistically significant positive effect of female leadership on safe water, but insignificant effects on safe sanitation and village road quality. Districts that are governed by female leaders experience a 1.6 percentage point higher annual increment in access to safe water. One plausible explanation is that infrastructure projects require multi-year implementation and therefore there might be longer lags between the beginning of the project and the realised service access. Third, there is no significant difference on education access across districts governed by female and male leaders, as judged by overall enrolment rates.

Table 4.5 Service outcomes under local randomisation

Dependent variable	Difference in means (finite sample)	
Average annual delta from (t+1) to (t+4) election	**τ**	**p**
NER primary	-0.356	0.464
NER junior	0.649	0.386
NER senior	-0.086	0.882
Attended births	2.568	0.002
Safe water	1.563	0.026
Safe sanitation	0.563	0.370
Villages with asphalt road	-0.112	0.926

Notes: Selected bandwidth is [-0.066,0.066]. The dependent variables are listed in column one. The outcome variable is the average annual delta during the office period from (t+1) to (t+4) election year, or until the most recent data available. τ is the estimated treatment effect and p is the p-value associated with the estimated treatment effect.

Lastly, I test the robustness of these results to the underlying assumption that there must be no manipulation of the running variable, that is, the winning margin of female candidates. Given the existence of corruption cases associated with local election (Lewis and Hendrawan, 2019), candidates might be able to increase their chance of winning through methods of poll fraud. However, although all candidates are likely to engage in such fraud, it is unlikely that any single candidate can precisely manipulate the winning margins.

Figure 4.3 shows the estimated density distribution of the winning margin for pairs with female candidates, along with the 95 percent confidence interval. The figure shows support for the no-manipulation assumption. Following Cattaneo, Jansson, & Ma (2018), I perform a formal test on the null hypothesis which states that there is no discontinuity at the cut-off. The robust bias-corrected test statistic with polynomial of degree two, a triangular kernel, and jack-knifed standard errors is -0.354 and the p-value is 0.723. This test statistic indicates that the null hypothesis cannot be rejected. The evidence suggests that there is no manipulation of the running variable and therefore the underlying assumption of the RD approach is valid.

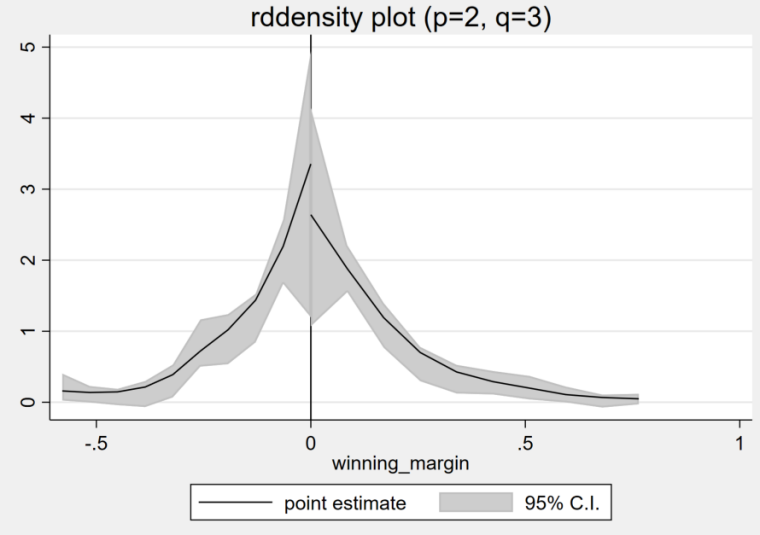


Figure 4.3 Density of female candidates' winning margin

4.6 Sensitivity analysis

Since the window selection process is an initial crucial step in examining the inference, I conduct sensitivity analyses to assess whether the point estimate would change across different window lengths. Following procedures introduced by Cattaneo et al. (2015), the null hypothesis for this sensitivity test is that the treatment effect is equal to a stated value. Therefore, for a 95% confidence level, in order not to reject the null hypothesis, a p-value higher than 0.05 is expected. For each of the considered outcomes, I generate a matrix of p-values over a range of possible treatment effects in the row and various window lengths in the column. In other words, the matrix will show how the interval for the point estimate will change over a range of different consecutive windows.

For fiscal outcomes, I run the sensitivity analysis for two types of sectoral spending per capita, which also shows a higher budget share out of total expenditure: social protection and infrastructure. I increase the window length by 0.01 and hence I use eight other windows in addition to the [-0.066,0.066] to check for sensitivity. I use a range of treatment effects from -1 to 1.5 to accommodate the point estimate of each category spending. In this section, I present the results in plots, while the complete matrix of the sensitivity test is available in the Appendix.

Figure 4.4 presents the graphical display of the sensitivity analysis matrix, for health and social protection spending. The plots depict a range of window lengths on the horizontal axis and the grid of (stated) treatment effects under the null hypothesis. The shade corresponds to

the p-value associated with the null hypothesis, that the treatment effect equals the stated value on the vertical axis, for each pair of window lengths. The lightest shade represents zero and the darkest represents a p-value of one. Therefore, darker shades mean higher p-values, which point to more support for not rejecting the null hypothesis.

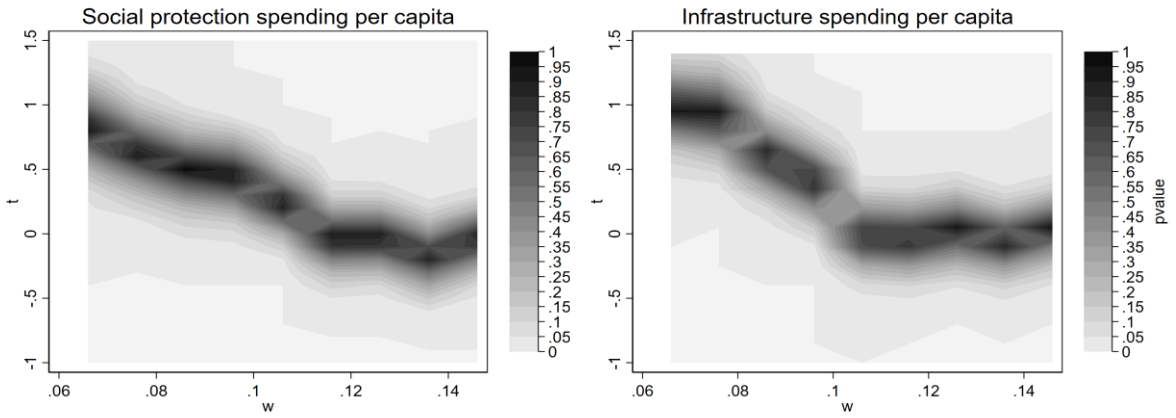


Figure 4.4 Sensitivity analysis for social protection and infrastructure spending

The results reported in Table 4.3 indicate that the point estimate for social protection spending per capita is 1.32, and for infrastructure it is 0.93. Both plots in Figure 4.4 show that these point estimates seem to be relatively stable for windows before [-0.1,0.1] as indicated by the dark shade around both values. As larger windows are considered, the darker shade also covers the value of zero on the vertical axis. This means that within these larger windows, the null hypothesis that the treatment effects are zero cannot be rejected. However, the initial window selection process revealed that for windows bigger than [-0.1,0.1], the covariates are not well-balanced. These sensitivity analyses suggest that the point estimates for health and social protection are robust but only up to narrowly selected windows.

I also run sensitivity analyses for both statistically significant service outcomes: assisted birth rate and percentage of households with access to safe water. Similar to the fiscal outcomes, I use the same increase in window length of 0.01, but then the treatment effects are between -0.5 and 3.5 to account for the results of the point estimates from Table 4.4. The estimation results in the previous section find that the point estimate for assisted birth rate is 2.6 while for access to safe water it is 1.6.

Figure 4.5 presents the plots of the matrix both for assisted birth and safe water. For assisted birth, the point estimate is relatively stable across different windows, as indicated by the dark shade that is consistently evolving around values bigger than 1. Even for the largest tested

window, [-0.146,0.146], the p-value against the null hypothesis of zero treatment effect is small, as shown from the lighter shade for value of zero on the vertical axis. This means there is enough evidence to reject the hypothesis of zero treatment effect across different windows. On the other hand, the test results for safe water are not as stable as those for the birth rate. As the windows are getting larger, the plot shows a darker shade for zero value on the vertical axis, which means that the probability of zero treatment effect is approaching one. These plots suggest that the point estimate for assisted birth rate is relatively more robust to any window specification compared to the point estimate for safe water access.

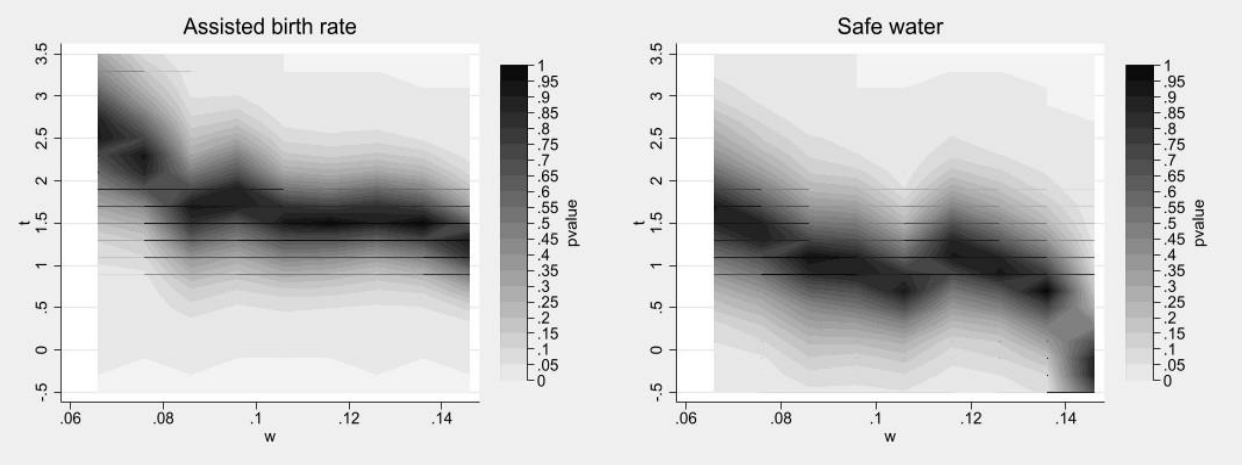


Figure 4.5 Sensitivity analysis for assisted birth rate and safe water

4.7 Possible mechanism

Overall, the results show that districts with female leaders tend to spend more budget, both in total and in sectoral spending on health, social protection and infrastructure. Female leadership also induces higher assisted birth rates and access to safe water. These results may be driven by differences in policy preferences between female leaders and their male counterparts. Studies have found that there appear to be gender differences in preferences for public policies and services (Mavisakalyan, 2014). For instance, most women prioritise health more than men due to their traditional role as primary caregivers in the family. Hence, inadequate health care would directly impact them and how they manage household budgets. The same notion also can be applied in explaining assisted birth rate services. Women overall will support better assistance at births due to their role as child-bearers. Therefore, female leaders must have a superior understanding of women's preferences, in addition to their own, which then influences their policy decisions upon taking office. These policy preferences may operate through multiple mechanisms. In this section, I discuss two plausible mechanisms: fiscal resources and governance.

For fiscal outcomes, the findings reveal that districts governed by female leaders spend higher per capita than districts governed by male leaders. For districts to afford more spending, there should be additional fiscal resources. However, the results also reveal that districts with female leaders do not systematically generate more of their own-source revenue per capita. Another possible source of budgetary resources is transfers from higher levels of government, which in the setting of this study is one of the covariates that is balanced among the treatment and control groups. Therefore, the only remaining untested source of additional resource is from the financing side of the budget, which due to data limitations cannot be formally tested[40]. Ruling out own-source revenue and transfers, the additional fiscal resources to fund increases in total spending may be attributed to smaller surpluses or budget deficits, which may require drawing down on fiscal reserves.

In addition to fiscal resources, another factor that may affect the capacity of districts in spending and providing service access is governance. Corruption and audit results are two variables that are usually used to represent the quality of public management of a government. Table 4.6 presents the results of estimating a similar RD model as in equation (1) with proxies of governance as the outcomes. The table shows a positive significant effect for audit score, with female leadership inducing a higher probability of recording unqualified opinions of audit scores (audit score > 3). Across the relevant cases, more than half of the districts recorded audit scores equal to or less than 3, therefore having a female leader in the office increases the probability of districts recording better audit scores than their counterparts. This higher median audit score also implies that female-led districts potentially have better spending efficiency which means more prudent budget management with the same amount of resources. For corruption, the preponderance of evidence from the table suggests no significant influence of female leaders on corruption, both in terms of incidence and the number of cases.

[40] Official data on surpluses and subnational fiscal reserves is not available. One way to proxy the surplus is by taking the difference between total revenue and total expenditure. I estimated the difference as outcome, but the result is not significant. On fiscal reserves, bank deposit balances of subnational governments are sometimes used as a proxy. Deposit data can be accessed at https://www.bi.go.id/id/statistik/ekonomi-keuangan/sekda/Default.aspx, however only at province level.

Table 4.6 Governance outcomes under local randomisation

	Difference in means (finite sample)	
Dependent variable	**τ**	**p**
Dummy median audit >3 (=1)	0.261	0.040
Dummy of whether there is corruption case or not	0.088	0.362
Number of corruption cases	0.168	0.102
Number of corruption convicts	0.194	0.122

Notes: Selected bandwidth is [-0.066,0.066]. The dependent variables are listed in column one. The number of cases and convicts are the sum of corruption cases done by government officials across office period from (t+1) to (t+4) election year or until most recent year available. Dummy of corruption measures whether there is corruption case done by government officials or not across the office period. Dummy of median audit measures whether the median of annual audit score across the office period bigger than 3 or not. Corruption variables are based on the corruption years and not the court years. Audit score category: 1- Disclaimer, 2- Adverse Opinion, 3- Qualified Opinion, 4- Unqualified Opinion. τ is the estimated treatment effect and p is the p-value associated with the estimated treatment effect.

The modest results in Table 4.6 suggest the presence of other transmission mechanism, particularly for service access outcomes, that has not yet captured in audit and corruption variables. One possible explanation may be that female leaders adopt a different way of delivering services within the same amount of available budget. For instance, female leaders might employ some local regulatory changes that incite improvement in health and water outcomes. Also, female leaders might introduce better program management when they take office, such as monitoring and evaluation activities, within a particular sector, which eventually contribute to service access improvement.

4.8 Conclusion

This study has examined whether female leadership influences a district's policy outcomes. Using districts with close election results between male and female candidates, this research has estimated the causal impact of having at least one female leader, either as head or vice-head, on fiscal and service delivery outcomes. The results found that districts with female elected leaders have higher per capita spending as well as budget share on social protection and infrastructure. Likewise, female leadership results in a higher assisted birth rate and greater access to safe water. Among these service outcomes, assisted birth rate was found to be the most robust across different specifications, while there was less support from the data for infrastructure outcomes. These findings suggest that female leaders influence local policy outcomes in sectors that continually concern women, such as birth and water-sanitation improvement. Therefore, this study provides some evidence in favour of the citizen-based model which highlights the influence of leaders' identity in determining their policy choices.

There are some caveats on this study that may inform future research topics. First, some policy outcomes used in this paper are gender-neutral outcomes that affect all constituents.

Although there is no significant influence on overall service access, female leaders may have influenced the gap between women and men accessing services. For instance, instead of using overall enrolment rate, using female enrolment rate over a longer time period may suggest different findings. Assessing gender gaps is crucial in examining the potential long-term impact of female leaders and their role in influencing the aspirations of future generations and opinions on women's ability (Beaman et al., 2012). Secondly, due to the limited number of observations, this study has not been able to thoroughly explore other political factors and set-ups, such as whether there is any difference between districts that have female leaders as head and districts that have female vice-heads. Although I implicitly assume that vice-heads have significant influence over policy-making in this paper, different roles may still translate to different levels of exposure to policy formulation. It will be useful to discuss this political set-up once more election data that involves female candidates becomes available. Lastly, due to the small number of districts in the sample, the findings from this study would not be suitable to represent a coherent interpretation on the magnitude of impact. As more data become available, further study that investigates the magnitude will become more appropriate and will be able to complement the general direction of outcomes that have been discussed in this study.

Regarding policy discussion, this study provides two insights. First, female leaders are found to pay more attention to sectors that disproportionately affect women such as safe birth and safe access to water, which are among the main problems in developing countries such as Indonesia. Improved assisted birth benefits women in their role as child-bearers and better access to safe water is crucial for household activities such as cooking and laundry. These findings are similar to what prior studies have found in other developing countries such as India and Brazil. Nevertheless, these findings give the novel perspective that female leaders, who are elected in a democratic election without a quota policy, thrive well and generate positive outcomes, even in a society where a male-dominated political environment prevails. The positive impacts answer concerns that female leaders may perform worse due to the difficulty of navigating social norms.

Second, this research points to the importance of having equal opportunities for women and men to participate in politics. One common policy response that has been adopted in many countries to achieve this is through gender quota policies. While there are various views about this quota issue, quota policy alone may not necessarily guarantee and lead to better

female representation outcomes. Therefore, policies are required to address more structural problems, such as the origin of women's low levels of representation in politics. This will then result in discussion of deeply entrenched social norms regarding women's appropriate roles. For instance, do women simply have a lack of interest in politics or do the prevailing norms make women think that it would be mostly futile to run for office? Increased exposure to women in leadership positions can mitigate this backlash attitude towards women leaders and may reduce behavioural barriers to women's leadership.

Reform in the political finance system could also be an option. Although candidates running for mayoral election are typically not political party officials or party members, candidates do seek out alignment with political parties to run in elections. Through forming coalitions with political parties, candidates expect to secure nomination and receive votes from voters who identify with particular parties (Lewis, 2019a). Under the current system, political parties receive limited public funding while facing increasing electoral costs. Parties tend to raise money by selling nominations on a party list and nominate candidates who are in its networks. This process disadvantages non-elite candidates but is particularly disadvantageous to women (Perdana and Hillman, 2020). An increase in public funding to political parties would enable them to proceed with a more objective nomination process and nominate more capable candidates, which may include female candidates.

Appendix

Table 4.7 Descriptive statistics

Variable	Obs	Mean	Std. Dev.	Min	Max
Covariates					
Poverty rate	207	13.96	7.96	1.67	54.94
(log) Population	212	12.86	1.14	9.76	15.44
(log) Area	196	7.10	1.43	2.77	10.07
Ethnic fractionalisation index	137	0.42	0.32	0.00	0.99
Average service access	201	72.27	9.89	46.62	92.12
(log) Transfer per capita	202	14.35	0.80	12.99	17.39
(log) GRDP per capita	212	16.93	0.66	15.78	19.83
Dummy urban districts (=1)	214	0.23	0.42	0	1
Dummy districts are located in Java (=1)	214	0.40	0.49	0	1
Effective number of political parties (ENP)	200	7.50	2.36	2.02	15.24
Dummy candidates are incumbent (=1)	214	0.33	0.47	0	1
Number of candidates running in the election	214	4.05	1.79	2	12
Fiscal outcomes (average over office periods)					
(log) Total expenditure per capita	192	14.61	0.71	13.41	17.22
(log) Education expenditure per capita	214	13.38	0.54	11.39	14.88
(log) Health expenditure per capita	214	12.47	0.75	10.88	14.92
(log) Social protection expenditure per capita	214	9.70	1.14	7.02	13.41
(log) Infrastructure expenditure per capita	214	11.92	1.29	8.64	16.07
(log) Own-source revenue per capita	192	12.11	0.75	10.33	14.19
Education budget share	192	0.31	0.10	0.03	0.56
Health budget share	192	0.11	0.04	0.02	0.31
Social protection budget share	192	0.01	0.01	0.00	0.04
Infrastructure budget share	192	0.13	0.06	0.00	0.56
Governance outcomes					
Dummy median audit score >3 (=1)	192	0.48	0.50	0	1
Dummy of whether there is corruption case or not	192	0.19	0.40	0	1
Number of corruption cases	192	0.24	0.53	0	2
Number of corruption convicts	192	0.31	0.72	0	4
Service outcomes (average annual delta over office periods)					
NER primary	191	0.42	1.61	-5.45	7.14
NER junior	214	1.06	2.32	-8.48	8.95
NER senior	214	1.56	2.73	-6.84	11.22
Attended births	214	1.16	2.79	-17.78	8.69
Safe water	214	1.67	3.61	-18.67	21.35
Safe sanitation	214	1.71	2.50	-12.46	12.42
Villages with asphalt road	213	2.29	4.60	-4.06	32.82
Running variable					
Winning margin of pairs with female candidates	214	-0.006	0.246	-0.826	0.854
Winning margin of pairs with female candidates > 0	106	0.168	0.165	0.004	0.854
Winning margin of pairs with female candidates < 0	108	-0.177	0.186	-0.826	0.000

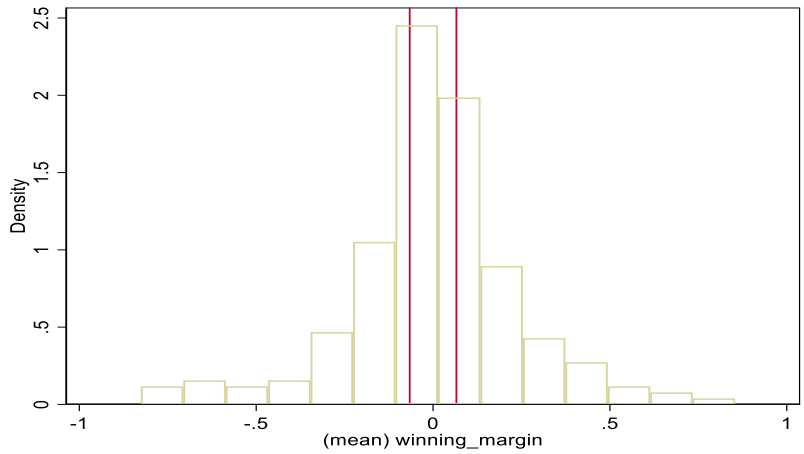
Table 4.8 Pre-treatment covariates balance test within specified window [-0.066,0.066]

Pre-treatment covariates balance test within selected bandwidth [-0.066,0.066]	Treatment	Control	Difference (Treatment – Control)	p-value
	(1)	(2)	(3)	(4)
Poverty rate	15.326	18.598	-3.272	0.136
(log) Population	12.773	12.638	0.135	0.618
(log) Area	7.022	7.249	-0.227	0.540
Ethnic fractionalisation index	0.337	0.418	-0.082	0.383
Average service access	69.977	71.651	-1.674	0.465
(log) Transfer per capita	14.277	14.489	-0.212	0.325
(log) GRDP per capita	16.845	16.828	0.017	0.907
Dummy urban districts (=1)	0.258	0.237	0.021	0.842
Dummy districts are in Java (=1)	0.419	0.316	0.114	0.279
Effective number of political parties (ENP)	7.871	7.489	0.382	0.547
Dummy candidates are incumbent (=1)	0.387	0.263	0.124	0.279
Number of candidates running in the election	4.903	4.421	0.482	0.359
(log) Total expenditure per capita	14.377	14.559	-0.182	0.389

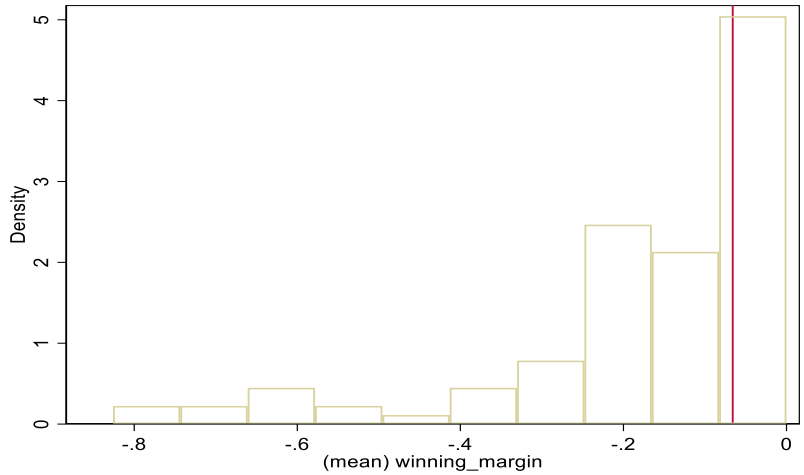
Figure 4.6 Winning margin distribution for pair with female candidates

Notes: Red line marks the selected bandwidth [-0.066,0.066]

(i). Winning margin distribution across sample



(ii). Winning margin distribution for losing cases (winning margin < 0)



(iii). Winning margin distribution for winning cases (winning margin > 0)

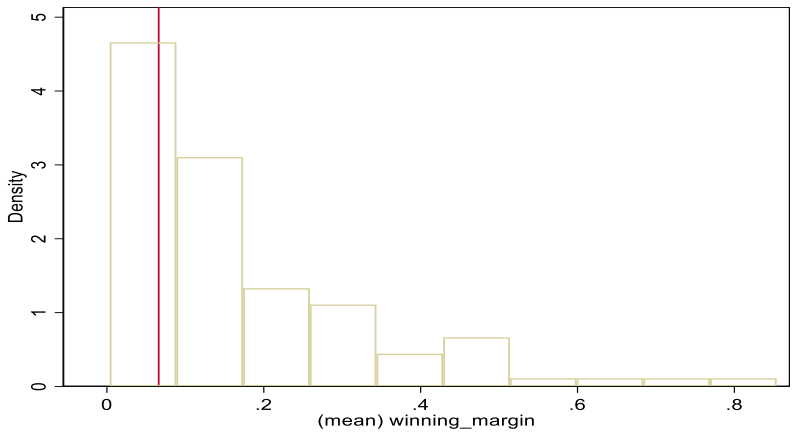
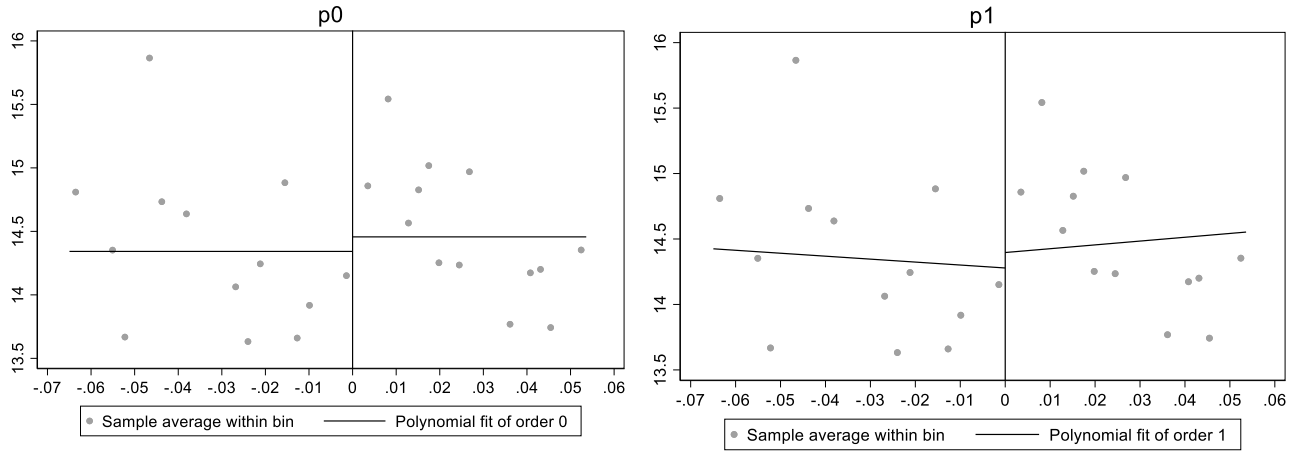


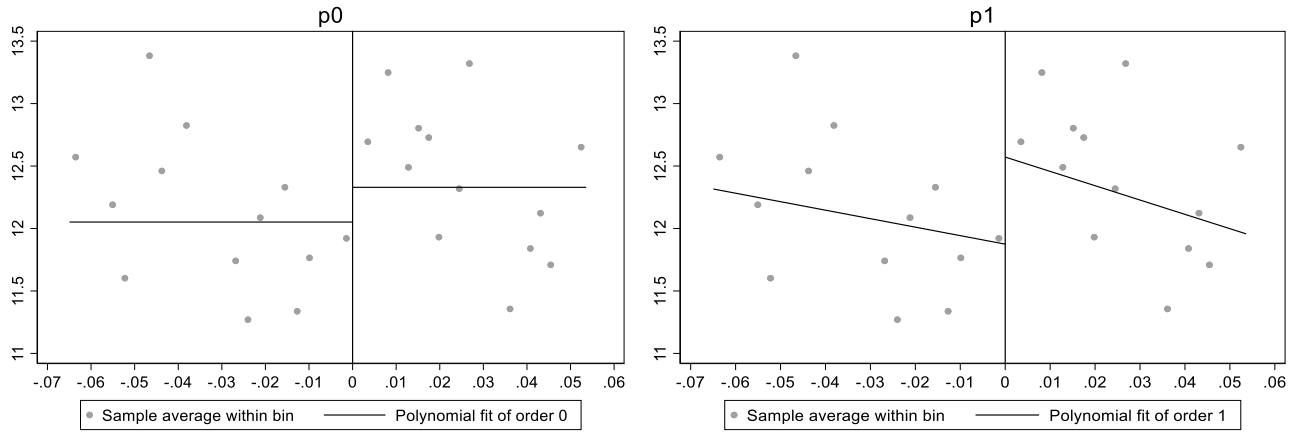
Figure 4.7 Plot between female leadership and fiscal outcomes

Notes: Plots are produced using `rdplot` STATA commands within selected bandwidth [-0.066,0.066]

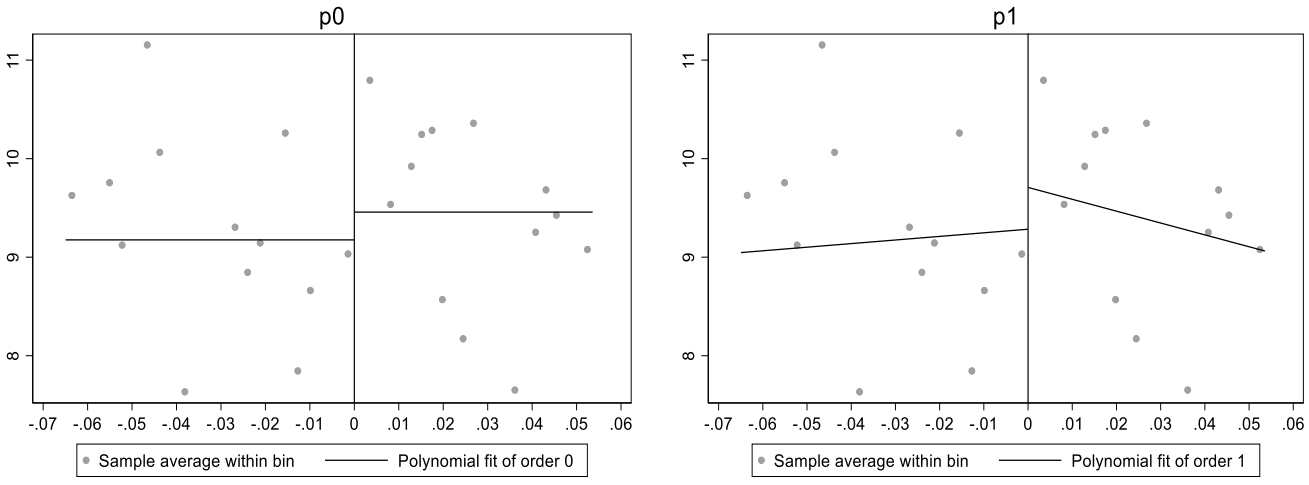
(i). (log) Total expenditure per capita



(ii). (log) Health expenditure per capita



(iii) (log) Social expenditure per capita



(iv) (log) Infrastructure expenditure per capita

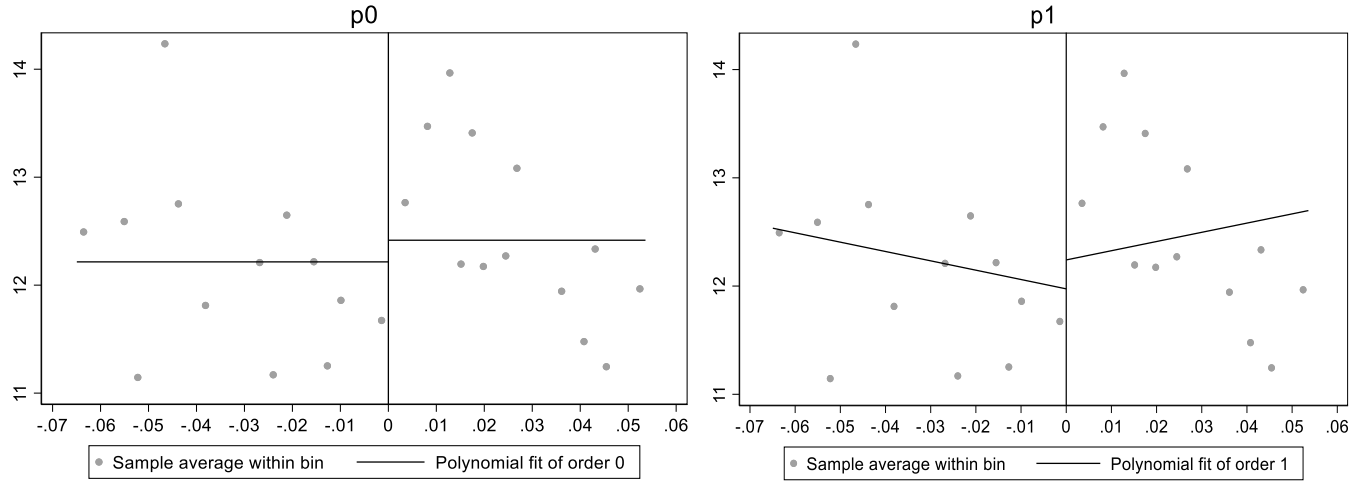


Table 4.9 Fiscal outcomes under local randomisation, all sectors

Dependent variable	**Difference in means (finite sample)**	
Average from (t+1) election to (t+4)^ election	**T**	**P>|T|**
Per capita value		
(log)total exp per capita	0.959	0.000
(log)education exp per capita	0.587	0.000
(log)health exp per capita	0.936	0.000
(log)social protection exp per capita	1.321	0.000
(log) infra exp per capita	0.939	0.000
(log) general admin exp per capita	1.228	0.000
(log) housing exp per capita	-0.092	0.796
(log) agriculture exp per capita	1.554	0.000
(log) economy exp per capita	1.396	0.000
(log) environment exp per capita	1.137	0.012
(log) tourism exp per capita	1.077	0.000
Share from total expenditure		
education	-0.108	0.000
health	-0.010	0.312
social protection	0.005	0.006
infrastructure	0.055	0.008
general admin	0.066	0.002
housing	-0.011	0.002
agriculture	0.020	0.004
economy	0.014	0.000
environment	-0.001	0.990
tourism	0.001	0.910

Notes: Selected bandwidth is [-0.066,0.066]. The dependent variables are listed in column one. Budget figures are in 2010 constant prices. τ is the estimated treatment effect and p is the p-value associated with the estimated treatment effect. The outcome variable is the average across the office period from (t+1) to (t+4) election year, or until the most recent data available.

Table 4.10 Sensitivity analysis matrix for (log) social protection spending per capita

	Different bandwidths								
Stated value on null hypothesis	**0.066**	**0.076**	**0.086**	**0.096**	**0.106**	**0.116**	**0.126**	**0.136**	**0.146**
-1	0	0	0	0	0	0	0	0	0
-0.9	0	0	0	0	0	0	0	0	0
-0.8	0	0	0	0	0	0	0	0.004	0.002
-0.7	0	0	0	0	0	0.004	0.006	0.016	0.004
-0.6	0	0	0	0	0	0.016	0.008	0.05	0.012
-0.5	0	0	0	0	0	0.046	0.032	0.158	0.032
-0.4	0	0	0	0	0	0.12	0.094	0.314	0.096
-0.3	0.002	0	0.002	0.004	0.016	0.242	0.236	0.544	0.232
-0.2	0.002	0.002	0.008	0.016	0.042	0.474	0.45	0.866	0.46
-0.1	0.002	0.004	0.026	0.034	0.128	0.794	0.784	0.7	0.768
0	0.008	0.018	0.06	0.078	0.288	0.89	0.886	0.41	0.854
0.1	0.022	0.038	0.13	0.166	0.57	0.552	0.56	0.216	0.534
0.2	0.04	0.112	0.23	0.332	0.876	0.298	0.292	0.094	0.246
0.3	0.062	0.206	0.43	0.564	0.8	0.124	0.126	0.028	0.092
0.4	0.14	0.36	0.674	0.852	0.48	0.056	0.052	0.008	0.026
0.5	0.248	0.59	0.98	0.866	0.262	0.016	0.02	0.008	0.01
0.6	0.414	0.902	0.712	0.558	0.134	0.004	0.01	0.002	0.004
0.7	0.63	0.792	0.438	0.332	0.05	0	0.004	0	0.002
0.8	0.926	0.534	0.238	0.168	0.014	0	0	0	0.002
0.9	0.8	0.316	0.132	0.05	0.004	0	0	0	0
1	0.54	0.17	0.048	0.024	0	0	0	0	0
1.1	0.346	0.08	0.02	0.004	0	0	0	0	0
1.2	0.206	0.034	0.01	0.002	0	0	0	0	0
1.3	0.092	0.02	0.004	0	0	0	0	0	0
1.4	0.042	0.008	0.002	0	0	0	0	0	0
1.5	0.022	0.004	0.002	0	0	0	0	0	0

Table 4.11 Sensitivity analysis matrix for (log) infrastructure spending per capita

	Different bandwidths								
Stated value on null hypothesis	**0.066**	**0.076**	**0.086**	**0.096**	**0.106**	**0.116**	**0.126**	**0.136**	**0.146**
-1	0	0	0	0	0	0	0	0	0
-0.9	0	0	0	0	0	0	0	0	0
-0.8	0	0	0	0	0.002	0	0	0.002	0
-0.7	0	0	0	0	0.002	0.004	0	0.002	0
-0.6	0	0	0	0	0.012	0.018	0.008	0.02	0.006
-0.5	0	0	0	0	0.028	0.042	0.014	0.032	0.018
-0.4	0	0	0	0	0.078	0.106	0.054	0.12	0.046
-0.3	0	0	0	0.004	0.194	0.236	0.166	0.26	0.116
-0.2	0	0	0.006	0.01	0.426	0.492	0.326	0.536	0.244
-0.1	0	0	0.012	0.034	0.692	0.778	0.542	0.852	0.472
0	0	0	0.018	0.092	0.928	0.874	0.89	0.788	0.82
0.1	0.004	0	0.038	0.224	0.574	0.568	0.746	0.47	0.774
0.2	0.006	0.004	0.094	0.408	0.368	0.322	0.456	0.226	0.454
0.3	0.02	0.02	0.202	0.67	0.172	0.15	0.242	0.082	0.186
0.4	0.038	0.052	0.392	0.98	0.058	0.06	0.102	0.038	0.096
0.5	0.068	0.118	0.67	0.724	0.02	0.026	0.038	0.012	0.022
0.6	0.122	0.208	0.992	0.422	0.008	0.002	0.012	0.008	0.008
0.7	0.246	0.376	0.696	0.236	0.002	0	0.002	0	0.006
0.8	0.484	0.634	0.414	0.106	0	0	0	0	0.004
0.9	0.772	0.916	0.224	0.048	0	0	0	0	0.002
1	0.906	0.768	0.106	0.018	0	0	0	0	0
1.1	0.634	0.528	0.052	0.004	0	0	0	0	0
1.2	0.386	0.326	0.018	0	0	0	0	0	0
1.3	0.206	0.192	0.002	0	0	0	0	0	0
1.4	0.1	0.094	0.002	0	0	0	0	0	0
1.5	0.026	0.044	0	0	0	0	0	0	0

Table 4.12 Sensitivity analysis matrix for assisted birth rate

	Different bandwidths								
Stated value on null hypothesis	**0.066**	**0.076**	**0.086**	**0.096**	**0.106**	**0.116**	**0.126**	**0.136**	**0.146**
-0.5	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-0.3	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000	0.000
-0.1	0.002	0.000	0.004	0.000	0.000	0.000	0.002	0.000	0.004
0.1	0.002	0.002	0.010	0.006	0.004	0.006	0.004	0.004	0.018
0.3	0.002	0.004	0.028	0.016	0.022	0.012	0.012	0.016	0.036
0.5	0.002	0.008	0.082	0.038	0.060	0.050	0.042	0.048	0.112
0.7	0.008	0.018	0.140	0.094	0.128	0.126	0.094	0.110	0.270
0.9	0.028	0.038	0.252	0.172	0.282	0.242	0.182	0.248	0.494
1.1	0.066	0.094	0.382	0.294	0.452	0.430	0.344	0.458	0.794
1.3	0.102	0.160	0.604	0.454	0.666	0.666	0.576	0.718	0.920
1.5	0.200	0.278	0.816	0.640	0.924	0.988	0.892	0.996	0.622
1.7	0.316	0.426	0.924	0.898	0.802	0.750	0.818	0.712	0.340
1.9	0.434	0.654	0.708	0.900	0.538	0.470	0.566	0.456	0.180
2.1	0.568	0.860	0.484	0.660	0.306	0.268	0.310	0.250	0.082
2.3	0.706	0.924	0.306	0.476	0.168	0.134	0.166	0.126	0.026
2.5	0.914	0.730	0.196	0.294	0.082	0.066	0.076	0.040	0.012
2.7	0.898	0.562	0.120	0.166	0.036	0.020	0.028	0.014	0.004
2.9	0.740	0.374	0.062	0.078	0.012	0.010	0.010	0.006	0.002
3.1	0.554	0.244	0.030	0.024	0.004	0.006	0.004	0.000	0.000
3.3	0.410	0.158	0.012	0.006	0.000	0.000	0.000	0.000	0.000
3.5	0.302	0.076	0.008	0.002	0.000	0.000	0.000	0.000	0.000

Table 4.13 Sensitivity analysis matrix for safe water access

	Different bandwidths								
Stated value on null hypothesis	**0.066**	**0.076**	**0.086**	**0.096**	**0.106**	**0.116**	**0.126**	**0.136**	**0.146**
-0.5	0.006	0.004	0.038	0.036	0.046	0.020	0.022	0.060	0.674
-0.3	0.012	0.020	0.082	0.070	0.090	0.034	0.048	0.098	0.890
-0.1	0.016	0.040	0.144	0.122	0.170	0.062	0.098	0.198	0.878
0.1	0.052	0.094	0.214	0.218	0.270	0.116	0.156	0.346	0.678
0.3	0.082	0.152	0.318	0.334	0.444	0.216	0.266	0.500	0.456
0.5	0.162	0.246	0.430	0.504	0.682	0.354	0.456	0.722	0.286
0.7	0.248	0.364	0.618	0.678	0.916	0.506	0.646	0.998	0.172
0.9	0.374	0.522	0.832	0.890	0.814	0.702	0.920	0.770	0.104
1.1	0.520	0.678	0.956	0.884	0.586	0.912	0.868	0.566	0.050
1.3	0.708	0.882	0.730	0.666	0.414	0.866	0.640	0.368	0.024
1.5	0.864	0.880	0.522	0.490	0.236	0.668	0.452	0.232	0.008
1.7	0.922	0.670	0.364	0.334	0.134	0.450	0.294	0.142	0.004
1.9	0.732	0.492	0.238	0.212	0.054	0.288	0.188	0.074	0.002
2.1	0.546	0.330	0.148	0.116	0.020	0.168	0.108	0.044	0.002
2.3	0.388	0.228	0.094	0.052	0.006	0.094	0.052	0.018	0.000
2.5	0.290	0.148	0.056	0.034	0.004	0.054	0.018	0.010	0.000
2.7	0.214	0.096	0.020	0.026	0.002	0.026	0.006	0.002	0.000
2.9	0.122	0.048	0.008	0.010	0.002	0.010	0.004	0.000	0.000
3.1	0.072	0.022	0.002	0.000	0.000	0.004	0.002	0.000	0.000
3.3	0.038	0.012	0.002	0.000	0.000	0.000	0.000	0.000	0.000
3.5	0.026	0.006	0.002	0.000	0.000	0.000	0.000	0.000	0.000

Table 4.14 Sectoral expenditure classification for government spending

Sector	Description
Education	*includes* provision of education services and facilities for all levels of education under the domain of local government *excludes* youth and sport funding
Health	*includes* medical supplies, hospital services and facilities; community health services; family planning programs; research and development in health sector
Social protection	*includes* all social assistance programs (sickness, disability, old age, family, housing assistance, etc) *excludes* family planning programs and efforts to eliminate unemployment; social assistance programs that take the form of subsidies such as rice subsidy (RASKIN)
Infrastructure	*Includes* construction of waste management facilities; provision of drinking water supply systems; irrigation construction; improving facilities and infrastructure to access energy (electricity); betterment and construction of roads and transportation ports, telecommunication facilities
General admin	*includes* administration and operation for all executive institutions; costs of ensuring local autonomy research and development of technology unrelated to specific function; salary spending which could not be related to particular sector i.e: salary of finance office and planning agency
Housing and public facilities	*includes* all activities related to housing development, community empowerment *excludes* building drinking water supply systems
Agriculture	*includes* activities in agriculture and extractive sector (forestry and marine/fisheries)
Economy	*includes* all activities undertaken to develop, promote and improve economic activities within country *excludes* construction of infrastructure and facilities as well as any activities in agriculture and extractive sector
Environment	*includes* activities related to maintaining natural resources and environment as well as limiting pollution *excludes* waste system construction and betterment
Tourism and culture	*includes* culture promotion; organisation of festivals / tourist attractions; tourism development support; youth and sport funding; improving sport facilities

5 CONCLUSION

This thesis provides empirical evidence on which aspects can improve the accountability of local governments in allocating fiscal resources and delivering service access to citizens. Each paper focuses on one aspect and is presented in Chapters 2 to 4. Chapter 2 examines the neighbourhood aspect and looks at how policies in other districts influence fiscal policy and service delivery in one district. Chapter 3 investigates different forms of village government and investigates which type is better in providing village level service access. Chapter 4 explores whether having female leaders in the office influences spending policy and service access provision.

Findings from Chapter 2 confirm the existence of positive horizontal interdependence and that spatial interdependence influences public spending levels and service access at district level. This means that an increase in public spending and service access in neighbouring districts is associated with an increase in those outcomes in own district. This positive interaction indicates that strategic competition between neighbouring districts may exist, given that the paper looks at the period after substantial reform through political decentralisation in 2005. The local democracy process, in which district heads and local parliaments are directly elected, can trigger more competition between regions. This competition can serve as an additional incentive for local government and potentially boost accountability in spending the budget and delivering service access. One plausible channel for this is through information spillover, in which information about positive policy outcomes in neighbouring districts informs citizens, who then hold their government more accountable for meeting their needs.

Although further examination is required, the findings from Chapter 2 might be relevant to other developing countries, especially those countries that have not decentralised major taxes to local levels. The results suggest that interdependence between spending and service access can be evident even when there is no major tax competition between regions. Regarding policy discussion, Chapter 2 provides insights that central government can expect a knock-on effect of a good policy in one district to another district. One way to do this is to reward districts that perform well in the form of performance-based grants, which are currently under discussion in many developing countries. The positive interdependence found in this chapter confirms that giving an additional grant to one region can also benefit other regions. On

future research directions, this paper defines neighbours based on geographical criteria. While geographical contiguity is important as an initial step, it would be useful to check whether district interaction would be different when considering socio-economic criteria in defining neighbours.

Chapter 3 suggests that citizen-based government performs better in delivering local service access compared to the bureaucratic type. This indicates that downward accountability to citizens matters more in service delivery than non-electoral upward accountability to the higher levels of government, even for the lowest administrative tier. Heads and apparatus in *desa* villages are more accountable to their constituents as the heads are directly elected and their salaries are funded from the village budget. One reflection of this accountability is that heads of *desa* act as the village social nexus and citizens in *desa* villages are found to trust their leaders more compared to citizens at *kelurahan.*

The unique setting with different types of village level government might not occur in other countries, hence this study is more of a country-specific case study. Nevertheless, the findings can still inform broader discussion about how decentralisation, accompanied by fair elections, will give incentives for local government to be more accountable to its constituents. In the context of Indonesia, the benefits of having more accountable government are more notable for service provision that includes active engagement with service providers, such as immunisation uptake. This points to the potential role of village government in administering these types of services. This paper, however, only discusses the period before the new Village Law was implemented in 2014. This new law basically gives more autonomy to village level government and each *desa* village will receive a certain amount of village funding. Therefore, for future research, it would be useful to revisit the comparison between these two government forms after the implementation of the new law.

Chapter 4 is the only chapter that discusses personal attributes of leaders, instead of collective aspects such as government form and neighbourhood effects. The results found that districts with female elected leaders in the office have higher per capita spending as well as higher budget shares spent on social protection and community infrastructure, such as water and sanitation. Likewise, districts with female leaders also have higher assisted birth rates and greater access to safe water. Therefore, this study provides some evidence in favour of the citizen-based model, which highlights the influence of leaders' identity in determining their policy choices.

Findings from Chapter 4 also suggest that female leaders pay more attention to areas that continually concern women, such as safe birth and water-sanitation improvement, which are among the remaining development problems in developing countries. While more studies are required for specific country cases, the results from this chapter answer concerns that female leaders may find difficulty in implementing such changes in a developing country with a male-dominated political environment and social norms. This study also points to the importance of having equal opportunities for women and men to participate in politics. Policies are required to address more structural problems, such as how to change social norms related to women's roles in the community.

Regarding future research directions, it would be useful to test the female leadership effect on gender gaps in accessing services, in addition to overall service access. For instance, in addition to the overall enrolment rate, using the gap between male and female enrolment rates may suggest new insights. Gender gaps in accessing services are usually long-term outcomes of changed aspirations and social norms due to female role-models as leaders, and therefore offer potential for future research once a longer time period of study is possible. Another research direction is to differentiate between the roles of female head and female vice-head. Although vice-heads have significant influence over policy-making, each role may translate to different exposure to policy formulation. Once more election data that involves female candidates becomes available, it would be useful to explore whether there are differences between districts that have female leaders as heads and vice-heads.

REFERENCES

Abadie, A., Athey, S., Imbens, G., Wooldridge, J., 2017. When Should You Adjust Standard Errors for Clustering? https://doi.org/10.3386/w24003

Antlöv, H., 2003. Village government and rural development in Indonesia: The new democratic framework. Bull. Indones. Econ. Stud. 39, 193–214. https://doi.org/10.1080/00074910302013

Antlöv, H., Wetterberg, A., Dharmawan, L., 2016. Village Governance, Community Life, and the 2014 Village Law in Indonesia. Bull. Indones. Econ. Stud. 52, 161–183. https://doi.org/10.1080/00074918.2015.1129047

Arze del Granado, J., Martinez-Vazquez, J., Simatupang, R., 2008. Local government fiscal competition in developing countries: The case of Indonesia. Urban Public Econ. Rev. 1, 13–45.

Ashenfelter, O., 1978. Estimating the Effect of Training Programs on Earnings. Rev. Econ. Stat. 60, 47–57.

Baicker, K., 2005. The spillover effects of state spending. J. Public Econ. 89, 529–544. https://doi.org/10.1016/j.jpubeco.2003.11.003

Bardhan, P., Mookherjee, D., 2006. Decentralisation and Accountability in Infrastructure Delivery in Developing Countries. Econ. J. 116, 101–127.

Beaman, L., Duflo, E., Pande, R., Topalova, P., 2012. A Policy Experiment in India. Science (80-.). 335, 582–586.

Besley, T., Case, A., 1995. Incumbent Behavior : Vote-Seeking , Tax-Setting , and Yardstick Competition 85, 25–45.

Besley, T., Coate, S., 2003. Elected Versus Appointed Regulators : Theory and Evidence. J. Eur. Econ. Assoc. 1, 1176–1206.

Besley, T., Coate, S., 1997. An economic model of representative democracy. Q. J. Econ. 112, 85–114. https://doi.org/10.1162/003355397555136

Bovens, M., Goodin, R.E., Schillemans, T., Bovens, M., Schillemans, T., Goodin, R.E., 2014. Public Accountability. Oxford Handb. Public Account. 1–22. https://doi.org/10.1093/oxfordhb/9780199641253.013.0012

Brollo, F., Troiano, U., 2016. What happens when a woman wins an election? Evidence from close races in Brazil. J. Dev. Econ. 122, 28–45. https://doi.org/10.1016/j.jdeveco.2016.04.003

Brueckner, J.K., 2003. Strategic Interaction Among Governments: An Overview of Empirical Studies. Int. Reg. Sci. Rev. 26, 175–188. https://doi.org/10.1177/0160017602250974

Caldeira, E., 2012. Yardstick competition in a federation: Theory and evidence from China. China Econ. Rev. 23, 878–897. https://doi.org/10.1016/j.chieco.2012.04.011

Carr, J.B., 2015. What have we learned about the performance of Council-Manager government? A review and synthesis of the research. Public Adm. Rev. 75, 673–689. https://doi.org/10.1111/puar.12415.What

Case, A.C., Rosen, S., Hines, J.R., 1993. Budget Spillovers Interdependence: Evidence and Fiscal Policy from the States. J. Public Econ. 52, 285–307. https://doi.org/10.1016/0047-2727(93)90036-S

Cattaneo, M.D., Frandsen, B.R., Titiunik, R., 2015. Randomization Inference in the Regression Discontinuity Design: An Application to Party Advantages in the U.S. Senate. J. Causal Inference 3, 1–24. https://doi.org/10.1515/jci-2013-0010

Cattaneo, M.D., Jansson, M., Ma, X., 2018. Manipulation testing based on density discontinuity. Stata J. 18, 234–261. https://doi.org/10.1177/1536867x1801800115

Cattaneo, M.D., Titiunik, R., Vazquez-Bare, G., 2017. Comparing Inference Approaches for RD Designs: A Reexamination of the Effect of Head Start on Child Mortality. J. Policy Anal. Manag. 36, 643–681. https://doi.org/10.1002/pam.21985

Cattaneo, M.D., Titiunik, R., Vazquez-Bare, G., 2016. Inference in Regression Discontinuity Designs under Local Randomization. Stata J. 16, 331–367. https://doi.org/10.1177/1536867x1601600205

Chattopadhyay, R., Duflo, E., 2004. Women as Policy Makers: Evidence from a India-Wide Randomized Policy Experiment. Econometrica 72, 1409–1443. https://doi.org/10.2139/ssrn.288974

Chen, L.-J., 2010. Do Female Politicians Influence Public Spending ? Evidence from Taiwan. Mimeo (Academia Sin. Taiwan) 10, 1–19.

Clots-Figueras, I., 2012. Are Female Leaders Good for Education? Am. Econ. J. Appl. Econ. 4, 212–244.

Clots-Figueras, I., 2011. Women in politics. Evidence from the Indian States. J. Public Econ. 95, 664–690. https://doi.org/10.1016/j.jpubeco.2010.11.017

Courtemanche, M., Green, J., 2017. The Influence of Women Legislators on State Health Care Spending for the Poor. Soc. Sci. 6, 40. https://doi.org/10.3390/socsci6020040

Croson, R., Gneezy, U., 2009. Gender differences in preferences. J. Econ. Lit. 47, 448–474. https://doi.org/10.1257/jel.47.2.448

Day, J., Lewis, B., 2013. Beyond univariate measurement of spatial autocorrelation: disaggregated spillover effects for Indonesia. Ann. GIS 19, 169–185. https://doi.org/10.1080/19475683.2013.806353

De Siano, R., D'Uva, M., 2017. Fiscal decentralization and spillover effects of local government public spending: the case of Italy. Reg. Stud. 51, 1507–1517. https://doi.org/10.1080/00343404.2016.1208814

Dewi, K.H., 2015. Indonesian Women and Local Politics. NUS Press Singapore, Kyoto University Press, Singapore.

Duflo, E., 2012. Women empowerment and economic development. J. Econ. Lit. 50, 1051–1079. https://doi.org/10.1257/jel.50.4.1051

Eagly, A.H., Carli, L.L., 2003. The female leadership advantage: An evaluation of the evidence. Leadersh. Q. 14, 807–834. https://doi.org/10.1016/j.leaqua.2003.09.004

Eagly, A.H., Karau, S.J., Makhijani, M.G., 1995. Gender and the Effectiveness of Leaders: A Meta-Analysis. Psychol. Bull. 117, 125–145. https://doi.org/10.1037/0033-

2909.117.1.125

Elhorst, J.P., 2003. Specification and Estimation of Spatial Panel Data Models. Int. Reg. Sci. Rev. 26, 244–268. https://doi.org/10.1177/0160017603253791

Ferraresi, M., Migali, G., Rizzo, L., 2018. Spillover effects in local public spending. Reg. Stud. 0, 1–15. https://doi.org/10.1080/00343404.2017.1415429

Ferreira, F., Gyourko, J., 2014. Does gender matter for political leadership? The case of U.S. mayors. J. Public Econ. 112, 24–39. https://doi.org/10.1016/j.jpubeco.2014.01.006

Frère, Q., Leprince, M., Paty, S., 2014. The Impact of Intermunicipal Cooperation on Local Public Spending. Urban Stud. 51. https://doi.org/10.1177/0042098013499080

Fujita, M., 2010. The evolution of spatial economics: From th??nen to the new economic geography. Japanese Econ. Rev. 61, 1–32. https://doi.org/10.1111/j.1468-5876.2009.00504.x

Gadenne, L., Singhal, M., 2014. Decentralization in Developing Economies. Annu. Rev. Econom. https://doi.org/10.1146/annurev-economics-080213-040833

Gangadharan, L., Jain, T., Maitra, P., Vecci, J., 2016. Social identity and governance: The behavioral response to female leaders. Eur. Econ. Rev. 90, 302–325. https://doi.org/10.1016/j.euroecorev.2016.01.003

Geys, B., 2006. Looking across borders: A test of spatial policy interdependence using local government efficiency ratings. J. Urban Econ. https://doi.org/10.1016/j.jue.2006.04.002

Gordon, R.H.., 1983. An Optimal Taxation Approach to Fiscal Federalism Published by : Oxford University Press Stable URL : http://www.jstor.org/stable/1881778 Accessed : 11-07-2016 01 : 08 UTC. Q. J. Econ. 98, 567–586. https://doi.org/https://doi.org/10.2307/1881778

Hewings, G.J.D., 2014. Spatially blind trade and fiscal impact policies and their impact on regional economies. Q. Rev. Econ. Financ. 54, 590–602. https://doi.org/10.1016/j.qref.2014.04.007

Hill, H., Resosudarmo, B., Vidyattama, Y., 2008. Indonesia's changing economic geography. Bull. Indones. Econ. Stud. 44, 407–435. https://doi.org/10.1080/00074910802395344

Hillman, B., 2017. Increasing Women's Parliamentary Representation in Asia and the Pacific: The Indonesian Experience. Asia Pacific Policy Stud. 4, 38–49. https://doi.org/10.1002/app5.160

Imbens, G.W., Lemieux, T., 2008. Regression discontinuity designs: A guide to practice. J. Econom. 142, 615–635. https://doi.org/10.1016/j.jeconom.2007.05.001

Isen, A., 2014. Do local government fiscal spillovers exist? Evidence from counties, municipalities, and school districts. J. Public Econ. 110, 57–73. https://doi.org/10.1016/j.jpubeco.2013.12.005

Jeddawi, M., Rahman, A., Nawawi, M., Bakri, R., Sagrim, B.G., 2018. Studi Kemungkinan Perubahan Status Desa Teluk Kapuas Menjadi Kelurahan di Kabupaten Kubu Raya. J. Ilmu Pemerintah. Suara Khatulistiwa III, 31–50.

Kelly, M., 2000. Inequality and crime. Rev. Econ. Stat. 82, 530–539. https://doi.org/10.1162/003465300559028

Kis-Katos, K., Sjahrir, B.S., 2017. The impact of fiscal and political decentralization on local public investment in Indonesia. J. Comp. Econ. 45, 344–365. https://doi.org/10.1016/j.jce.2017.03.003

Lee, L. fei, Yu, J., 2010. Estimation of spatial autoregressive panel data models with fixed effects. J. Econom. 154, 165–185. https://doi.org/10.1016/j.jeconom.2009.08.001

Lewis, B.D., 2019a. How do mayors get elected? The causal effects of pre-electoral coalitions on mayoral election outcomes in Indonesia. Local Gov. Stud. 00, 1–20. https://doi.org/10.1080/03003930.2019.1627334

Lewis, B.D., 2019b. Legislature Size, Local Government Expenditure and Taxation, and Public Service Access in Indonesia. Stud. Comp. Int. Dev. 54, 274–298. https://doi.org/10.1007/s12116-019-09278-1

Lewis, B.D., 2018. Local Government Form in Indonesia: Tax, Expenditure, and Efficiency Effects. Stud. Comp. Int. Dev. 53, 25–46. https://doi.org/10.1007/s12116-017-9236-z

Lewis, B.D., 2016. Local government spending and service delivery in Indonesia: the perverse effects of substantial fiscal resources. Reg. Stud. 0, 1–13. https://doi.org/10.1080/00343404.2016.1216957

Lewis, B.D., 2015. Decentralising to villages in Indonesia: Money (and other) mistakes. Public Adm. Dev. 35, 347–359. https://doi.org/10.1002/pad.1741

Lewis, B.D., Hendrawan, A., 2019. The impact of majority coalitions on local government spending, service delivery, and corruption in Indonesia. Eur. J. Polit. Econ. 58, 178–191. https://doi.org/10.1016/j.ejpoleco.2018.11.002

Lewis, B.D., Nguyen, H.T.M., Hendrawan, A., 2020. Political accountability and public service delivery in decentralized Indonesia: Incumbency advantage and the performance of second term mayors. Eur. J. Polit. Econ. 64, 101910. https://doi.org/10.1016/j.ejpoleco.2020.101910

Luo, R., Zhang, L., Huang, J., Rozelle, S., 2007. Elections, fiscal reform and public goods provision in rural China. J. Comp. Econ. 35, 583–611. https://doi.org/10.1016/j.jce.2007.03.008

Martinez-Bravo, M., 2014. American Economic Association The Role of Local Officials in New Democracies: Evidence from Indonesia. Am. Econ. Rev. 104, 1244–1287. https://doi.org/10.1257/aer.104.4.1244

Mavisakalyan, A., 2014. Women in cabinet and public health spending: Evidence across countries. Econ. Gov. 15, 281–304. https://doi.org/10.1007/s10101-014-0141-x

Nguyen, H.T.M., 2019. Do more educated neighbourhoods experience less property crime? Evidence from Indonesia. Int. J. Educ. Dev. 64, 27–37. https://doi.org/10.1016/j.ijedudev.2018.12.005

Olken, B.A., 2010. Direct democracy and local public goods: Evidence from a field experiment in Indonesia. Am. Polit. Sci. Rev. 104, 243–267. https://doi.org/10.1017/S0003055410000079

Pepinsky, T.B., Wihardja, M.M., 2019. Decentralization and Economic Performance in Indonesia. J. East Asian Stud. 11, 337–371.

Perdana, A., Hillman, B., 2020. Quotas and ballots: The impact of positive action policies on women's representation in Indonesia. Asia Pacific Policy Stud. 1–13. https://doi.org/10.1002/app5.299

Qian, Y., Weingast, B.R., 1997. Federalism as a Commitment to Preserving Market Incentives. J. Econ. Perspect. 11, 83–92. https://doi.org/10.1257/jep.11.4.83

Revelli, F., 2005. On spatial public finance empirics. Int. Tax Public Financ. 12, 475–492. https://doi.org/10.1007/s10797-005-4199-9

Revelli, F., 2003. Reaction or interaction? Spatial process identification in multi-tiered government structures. J. Urban Econ. 53, 29–53. https://doi.org/10.1016/S0094-1190(02)00512-0

Salmon, P., 1987. Decentralization as an incentive scheme. Oxford Rev. Econ. Policy 3, 24.

Shah, A., 2006. Local Governance in Developing Countries. The World Bank. https://doi.org/10.1007/978-94-017-3457-8_13

Sjahrir, B.S., Kis-Katos, K., Schulze, G.G., 2014. Administrative Overspending in Indonesian Districts: The Role of Local Politics. World Dev. 59, 166–183. https://doi.org/10.1016/j.worlddev.2014.01.008

Skoufias, E., Narayan, A., Dasgupta, B., Kaiser, K., 2014. Electoral accountability and local government spending in Indonesia 42. https://doi.org/http://dx.doi.org/10.1596/1813-9450-6782

Skoufias, E., Narayan, A., Dasgupta, B., Kaiser, K., 2011. Electoral Accountability, Fiscal Decentralization and Service Delivery in Indonesia. World 5614, 56.

Smoke, P., Lewis, B.D., 1996. Fiscal decentralization in Indonesia: A new approach to an old idea. World Dev. 24, 1281–1299. https://doi.org/10.1016/0305-750X(96)00042-3

Solé-Ollé, A., 2006. Expenditure spillovers and fiscal interactions: Empirical evidence from local governments in Spain. J. Urban Econ. 59, 32–53. https://doi.org/10.1016/j.jue.2005.08.007

Stevenson, B., Wolfers, J., 2006. Bargaining in the Shadow of the Law: Divorce Laws and Family Distress. Q. J. Econ. 121, 267–288. https://doi.org/10.1093/qje/121.1.267

Suci, B.K., Yamada, K., Wibowo, K., 2020. Does Female Politician Affect Policy Outcomes in Indonesia? Bandung.

Suryadarma, D., 2012. How corruption diminishes the effectiveness of public spending on education in Indonesia. Bull. Indones. Econ. Stud. 48, 85–100. https://doi.org/10.1080/00074918.2012.654485

Syukri, M., Mawardi, M.S., n.d. Sharing Knowledge on Community- Driven Development in Indonesia : An Assessment of the Neighborhood Upgrading and Shelter Sector Project.

Tomsa, D., 2014. Party System Fragmentation in Indonesia: The Subnational Dimension. J. East Asian Stud. 14, 249-278,316.

Wardani, S.B.E., Subekti, V.S., 2021. Political Dynasties and Women Candidates in Indonesia's 2019 Election. J. Curr. Southeast Asian Aff. 40, 28–49. https://doi.org/10.1177/1868103421991144

Wing, C., Simon, K., Bello-gomez, R.A., 2018. Designing Difference in Difference Studies : Best Practices for Public Health Policy Research.

Yilmaz, S., Beris, Y., Serrano-Berthet, R., 2010. Linking local government discretion and accountability in decentralisation. Dev. Policy Rev. 28, 259–293. https://doi.org/10.1002/pad.579

Zhang, X., Fan, S., Zhang, L., Huang, J., 2004. Local governance and public goods provision in rural China. J. Public Econ. 88, 2857–2871. https://doi.org/10.1016/j.jpubeco.2003.07.004