Inequality, Regional Economic Development and Access to Public Services in Decentralizing Indonesia

Soenarto Rosidin, Isfandiarni

2021

Inequality, Regional Economic Development and Access to Public Services in Decentralizing Indonesia

Isfandiarni Soenarto Rosidin

Cover Design & Layout: Ashilina S.

Printing: Reprographics Vrije Universiteit Amsterdam

VRIJE UNIVERSITEIT

Inequality, Regional Economic Development and Access to Public Services in Decentralizing Indonesia

ACADEMISCH PROEFSCHRIFT

ter verkrijging van de graad Doctor aan
de Vrije Universiteit Amsterdam,
op gezag van de rector magnificus
prof.dr. C.M. van Praag,
in het openbaar te verdedigen
ten overstaan van de promotiecommissie
van de School of Business and Economics
op dinsdag 26 oktober 2021 om 13.45 uur
in een bijeenkomst van de universiteit
De Boelelaan 1105

door
Isfandiarni S. Rosidin
geboren te Jakarta, Indonesië

promotor: prof.dr. H.L.F. de Groot

copromotor: dr. P. Mulder

Table of Contents

List of Figures

List of Tables

Acknowledgements

Alhamdulillah, all praise and thank to Allah almighty for providing me with strength and blessing to finish my PhD trajectory. Pursuing a PhD is my dream since 2003, but the opportunity came in 2012 when I obtained a scholarship from the Directorate of Higher Education Ministry of Education Government of Indonesia (DIKTI) and the Scientific Program Indonesia-Netherlands (SPIN-KNAW) to enroll in the PhD program at the Department of Spatial Economics, School of Business and Economics of the Vrije Universiteit Amsterdam and started this journey in the autumn of 2013. After years and ongoing global pandemic Covid 19 in the final stage, this thesis is finally completed, and I would like to express my gratitude to many people for their continuing support during this amazing journey.

First, I would like to express my sincere gratitude to Henri de Groot, my supervisor, for his excellent guidance and valuable feedbacks to improve my works. His interesting insights helped me to understand economics better and he challenged me to perform my research beyond my own expectations. Thank you very much, Henri for your encouragement, patience and our unique collaboration. It would have been impossible to get things done without the persistent and countless discussions with Peter Mulder, my co-supervisor and my life mentor in Amsterdam. Thanks a lot, Peter, for responding to my endless questions, inspiring discussions, developing the UI-VU collaboration, helping me seeing the problems from completely different perspectives including the non-research problems and welcoming my family in Amsterdam.

Second, I would like to thank Erik Verhoef, head of the department of Spatial Economics of the Vrije Universiteit Amsterdam for his generous support and the facilitation of a very nice environment, not only to work but also to socialize among colleagues in our department. Thanks also go to the administrative supports: Elfie Bonke, Jenny Wiserma and Hedda Werkman for doing a fantastic job taking care of the PhD students even before they arrive in Amsterdam, Hadewijch van Delft as Department Manager, Esther Hartog from the International Office, Sandra Zonneveld from the human resources division and Inge Borst from the finance division. Thank you to members of the PhD committee: Wendy Janssens, Gerry van Klinken, Robert Sparrow, Budy Resosudarmo, Jan Luiten van Zanden, and Eric Koomen for the time spent reading the thesis and giving useful comments.

Third, my deep gratitude to Ari Kuncoro, Rector of Universitas Indonesia for his support and interesting insights particularly on Chapter 5. I also thank Beta Yulianti (Dean), Teguh Dartanto (Vice Dean), and Vid Adrison (Head of Department of Economics) of the Faculty of Economics and Business at Universitas Indonesia for their support in enabling me to complete this thesis. I thank Riatu Mariatul Qibthiyyah, Director of the Institute for Social and Economic Research (LPEM-FEBUI) for her generous support in accessing databases and other facilities during my study. I also thank PT PLN (Persero) to allow me accessing the electricity database that I use specifically in Chapter 6.

In doing this research, I am indebted for the help from Maureen Lankhuizen in working with the non-linear least square technique that I use in chapter 6, Ahmad Fuady for his expert in health insurance and medical science that I use in Chapter 5 and Hengky Kurniawan for his useful tips and tricks dealing with Stata since the beginning of my trajectory. Big thanks to Yusuf Sofiyandi for helping me to clean and prepare datasets. I also thank to my colleagues Gerard van der Meijden as part of the VU-UI team developing the intensive course, Jessie Bakens for her suggestions on practicalities, Gilberto Mahumane, Zhilling Wang, Xinying and Francis Ostermeijer for their contagious laughs and positive vibes, my office mates Panayiotis and Ting Zhou for our long days and nights in the office, and Sotya Fevriera, Aloysius G. Brata, and Dhaniel Ilyas as my Indonesian friends in the department of Spatial Economics. To my Indonesian fellows in the Netherlands: Lita Napitupulu, Diana Pakasi, mbak Mia, Nurmala, mbak Ira, mas Hadi, and many others, thanks so much for sharing and caring during my journey. I am grateful to Ibu Vera Basir for her understanding of my absence in her duty and Riza Riznasari for her support as my 24-hours best friend.

Finally, I am eternally grateful to my parents, R. Soenarto and Zakiah Saleh for their continuous pray to me, and especially my mother who inspired me to pursue my dream. Big thanks to my big family to take care and to ease my family business in Indonesia during my study. Last but not least, I would not have been able to come this far without a great man standing besides me, Amir Rosidin. Words cannot express

how grateful I am to have him at my side and for his constant support to continue my study since I was undergraduate student and commitment to take care of our kids. I dedicate this book to him. I feel blessed to have wonderful children Alifa, Sofi, Rashif, Ashila, son in law Amando, and grandson Alvar who always cheer my life with their big smiles and funny stories. Being apart for years while they continuously prove their best in any circumstances really makes me proud of them and helped me to finish this thesis. Thanks a lot, kids.

Jakarta, August 2021
Isfandiarni S. Rosidin

Chapter 1 Introduction

1.1 Motivation

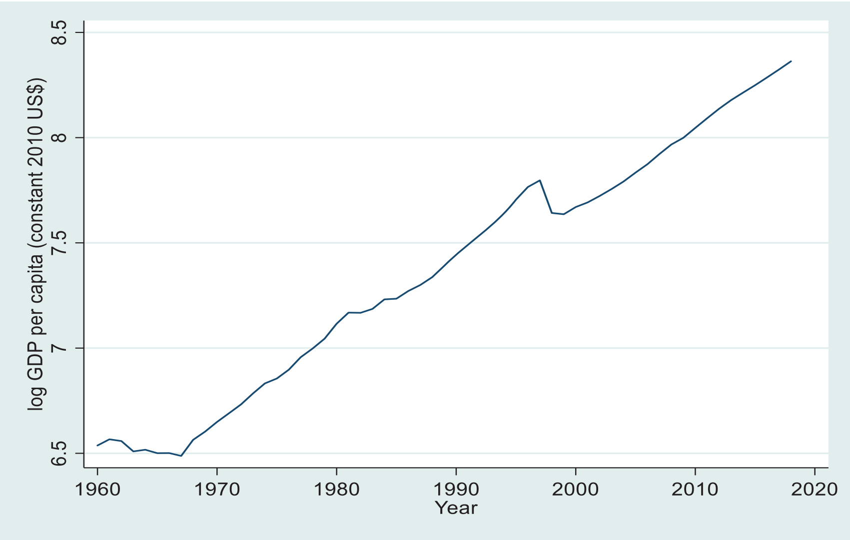
Over the past three decades, Indonesia has undergone drastic institutional changes, driven by simultaneous decentralization and democratization processes that strongly impacted the entire society. In fact, Indonesia transformed in this period from a centralized authoritarian regime into one of the largest and most decentralized democracies in the world. Amongst others, this transformation had far-reaching consequences for the level and evolution of inequality across individuals and regions as well as the organization and accessibility of public services like health and electricity.

In the aftermath of the Asian financial crisis, Indonesia followed a “big bang” decentralization policy to decentralize the responsibilities and budgets of the central government. Where the country had 316 *kabupaten*/*kota* or municipalities in 1999, this number had grown to 539 *kabupaten*/*kota* by 2014 (Ministry of Home Affairs 2019). Alongside this process, successive reforms have introduced the direct election of the president (2004), the governor and *bupati*/mayors and district heads (2005), as well as members of parliament (2009).

Since the fall of Suharto in 1998, a profound debate on inequality in Indonesia emerged that broadly resonates to society. The debate has been fueled by the observation that the distribution of income has become more skewed, with the top income groups performing better than the low income groups in terms of income growth. This raised global public concern, which is evidenced by statements that inequality is rising rapidly (World Bank 2016), the gap between the richest and the rest in Indonesia is increasing faster (Oxfam 2017) and the number of billionaires in Indonesia has grown rapidly, from just one in 2002 to 20 in 2016 (Forbes 2016). Inequality is not a question solely of income and wealth (outcomes). It also concerns unequal access to health services and reliable power supply, where the latter is characterized by clear spatial inequalities between central and more peripheral parts of the country. This thesis tries to demonstrate how the various dimensions of inequality are often closely related.

This thesis starts from the presumption that the institutional reforms have shaped socio-economic development, including patterns of economic growth and income inequality as well as quality of public service provision. Aggregate economic development of Indonesia has been positive from the late 1960s onwards (see Figure 1.1).

Figure 1.1 Per Capita GDP in Indonesia: National Level 1960–2018

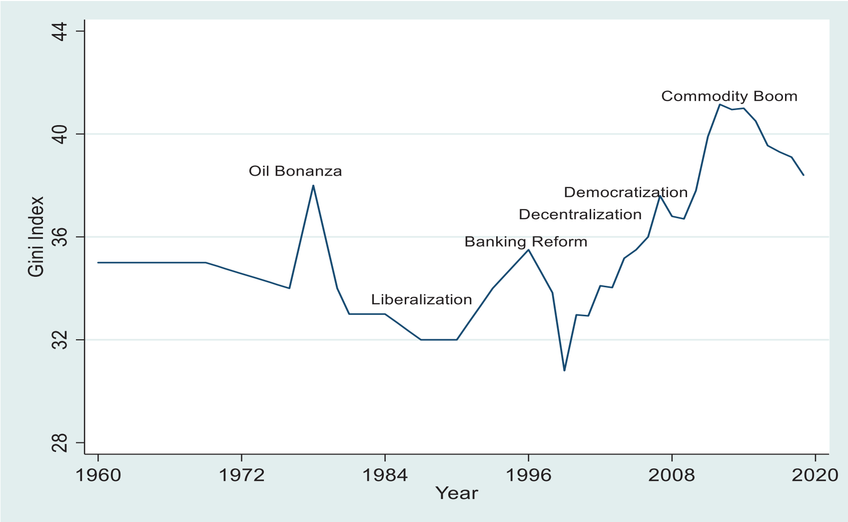


Source: World Development Indicators, World Bank.

Per capita GDP (measured in constant price USD 2011) increased with a factor six from 690 USD in 1960 to 4,284 USD in 2018. Figure 1.1 shows that per capita GDP started to increase in the early 1970s with an average annual growth rate of 8 percent until 1998. After the economic crisis in 1999, the economy has not returned to the pre-crisis growth rates, but continued to grow at an average of 5.25 percent. Recently, Indonesia's economy grew at 5.6 percent in 2019 resulting in a GDP per capita level of about USD 11,606 (World Bank 2020, measured in PPP).

Indonesia has also played a key role in the economic development of Southeast Asia, maintaining positive growth after the global financial crisis of 2008/9, but showing signs of macroeconomic weaknesses in the last few years (World Bank 2019). At the same time, income inequality, as measured by a Gini index, remains high with a Gini coefficient of 38.4 in 2019 (BPS 2020). These figures attest to the complex relationship between inequality and development.

Figure 1.2 Gini Index Indonesia: National Level 1964–2019



Source: BPS

A closer look at the evolution of inequality in Indonesia in Figure 1.2 shows that since 1960 income inequality in Indonesia – as measured by the Gini index[1] – fluctuates, with a declining tendency until the early 1990s (briefly interrupted by the Oil Bonanza in the late 1970s). Since then, however, inequality started to increase rapidly (interrupted only by the Asian crisis in 1998) to reach a peak in 2012 and decrease somewhat after the commodity boom ended in 2014. In this thesis we show that a similar pattern is also found when looking at the regional level, both for islands as well as provinces.

Several events may have caused these fluctuations. First, in the early 1970s, Indonesia benefited from windfall profits of the oil boom, followed by the

[1] We use the terms Gini Index and Gini Coefficient interchangeable as those terms refer to the same definition.

implementation of an industrialization policy in combination with substantial protection (known as an import substitution industrialization policy). However, since the mid-1980s as a result of the oil price shock in 1982 where export values dropped significantly, Indonesia moved towards an export promotion policy accompanied by liberalization and a banking reform in 1996 (Booth 2000). This new policy formed a new high-income class and thus increased the Gini index during that period. Subsequently, the Asian Financial Crisis in 1998 resulted in lower inequality and economic growth. The implementation of decentralization policies in 2000 marked the beginning of a sharp rise in inequality, especially since the recovery of the economy after 2002.

Following the big-bang decentralization, the introduction of democratization in 2005 was another major event resulting in the local district leadership to be directly elected through local elections that require huge campaign funds. On the other hand, the majority of Indonesia's billionaire wealth was generated in the industries prone to cronyism (Oxfam 2017). Thus, this finding suggests that political capture has played a role in creating many of Indonesia's modern billionaires, and has thus fueled inequality,

Besides these episodes, an increasing demand for commodities such as coal, palm oil, natural gas and rubber doubled Indonesian exports since 2004 and reached its highest level in 2011 (interrupted by the 2009 global crises). This commodity boom era seems to have instigated a rise in Gini index from 35.2 in 2004 to a peak of 41.2 in 2012. The slowdown of China's growth brought the commodity boom to an end after which the Gini coefficient tended to decline. This inequality trend bears a certain resemblance with the situation in the late 1970s after the oil boom ended and when the Gini also started declining. In an international perspective, Indonesia is catching up in terms of per capita GDP and is still characterized by a relatively low income inequality compared to many other countries, although inequality in per capita income has increased at national, island and provincial level since 2001 (see chapter 2 of this thesis for further details).

At first sight, it appears that inequality in Indonesia is mainly triggered by external factors, i.e. increasing commodity prices which benefited the high income class more, especially in commodity-exporting provinces. The spikes in the inequality trend – resulting from several events such as the oil Bonanza in the mid-1970s, the banking reform in the mid-1990s, democratization in the mid-2000s, and the commodity boom by the end of 2010 – suggest that the elite benefited disproportionately from these economic tailwinds, whereas they experienced a deterioration of their relative position after liberalization of the economy in early 1980s as well as during the Asian crisis in late 1990s and the fall of commodity prices in 2012. Thus, an increasing trend of Gini index could be perceived in two ways: as a positive sign of the effectiveness of market mechanism or as a negative sign of widening income gap across Indonesia.

1.2 Theoretical Background

Interconnections between inequality and development, particularly economic growth, can be explained by two way causal theories. The first describes how economic development affects inequality. The seminal work of Simon Kuznets (1955) provides a foundation for this relationship. He argued that as the economy grows inequality first increases and later decreases: the well-known inverted U-shape hypothesis. The second causal relationship describes how inequality affects economic development. Different strands of literature on the impact of inequality on economic development have developed, including in economic growth theory. Dominant perspectives on this relationship include: the classical approach (focusing on the impact on saving rates), the political economy approach (focusing on the impact of redistribution), the credit market imperfections channel, the rent-seeking approach, the social unrest (political instability) approach, and the unified theory of inequality and growth. A recent contribution to theories on capitalism is provided by Piketty (2014), who argues that in economies organized by capitalism over longer periods of time the rate of return on capital is persistently larger than the rate of economic growth, which will cause increases in wealth inequality.

According to the *Classical approach*, inequality is beneficial for growth. This theory suggests that the savings rate is relatively high among the wealthy (Lewis 1954; Kaldor 1956; Galor 2011). In this approach, inequality distributes resources to those individuals with higher marginal propensity to save, resulting in higher aggregate savings and more capital accumulation, subsequently increasing economic growth. Galor and Zeira (1993), Benabou (1996), and Aghion et al. (1999) find opposite relationships through various mechanisms. In addition to those criticisms, Venieris and Gupta (1986) also demonstrate that the bulk of savings is, in fact, produced by the middle-income class and not by the rich.

According to the *Political economy approach*, income inequality is harmful to growth, because it leads to policies that do not protect property rights and do not allow full private appropriation of returns from investment. High inequality will lower growth because the poor majority would vote for redistributive rather than growth-enhancing policies. The median voters choose redistribution policies (taxes and transfers), and in an unequal society, they are poorer than the mean. Taxes imposed on the margin are distortionary and slow down economic growth (Alesina and Rodrik 1994; Persson and Tabellini 1994).

The *Credit Market Imperfections Channel* proposed by Galor and Zeira (1993) demonstrates that in the presence of credit market imperfections and fixed costs associated with investments in education, occupational choices (and thus the efficient segmentation of the labour force between skilled and unskilled workers) are affected by the distribution of income. In particular, if the interest rate for borrowers is significantly higher than that for lenders, inequality may result in an under-investment in human capital.

Rent-Seeking theories explore the situation in which a widening gap between the rich and the poor results in the rich having a greater temptation to engage in rent-seeking or predatory activities at the expense of the poor (Benabou 1996). Other researchers have also proposed an institutional mechanism in which a wealthy elite will suppress democracy and equal rights before the law so as to preserve their privileged position (e.g. Bourguignon and Verdier 2000).

Social Unrest or *Political Instability* argues that the poor people's motivation to engage in crime, riots, and other disruptive activities is due to the wealth and income inequality (Barro 2000). Political institutions' stability may even be threatened by revolution so that laws and other rules have a shorter expected duration and more significant uncertainty. When the poor people participate in crime and other anti-social actions, it represents a direct waste of resources because their time and energy are not spent for productive activities. Moreover, the threats to property rights discourage investment. Through these various dimensions of sociopolitical unrest, more inequality tends to reduce the productivity of an economy, and then economic growth declines accordingly.

In the *Unified Theory of Inequality and Growth (Human Capital Mechanism),* one can reconcile the classical and the credit market imperfections approach. If imperfect capital markets exist that associates with investment in education, human capital cannot be accumulated because of the majority of poor people. On the other hand, the effect of inequality on growth depends on the relative return to physical and human capital. Physical capital is a prime engine for growth in the early stage of industrialization, but later it is substituted by human capital and the relative return to physical capital decreases. Thus, the impact of inequality on growth changes from positive to negative (Galor and Zeira 1993; Galor and Moav 2004; Galor 2011).

Finally, in his famous book *Capital in the Twenty-First Century*, Piketty (2014) developed a unified theory of the functioning of the capitalist economy by linking theories of economic growth and personal income distributions using long-run historical data series. He suggests that when the top incomes hold capital and its rate of returns keeps increasing, the rich will continue accumulating wealth, which in turn contributes to a widening income gap. In other words, rising wealth inequality and income concentration inevitably is part of the changing nature of modern capitalism. To tackle this increasing inequality, Piketty proposes redistribution through progressive taxes on wealth.

1.3 Aim and Structure of the Thesis

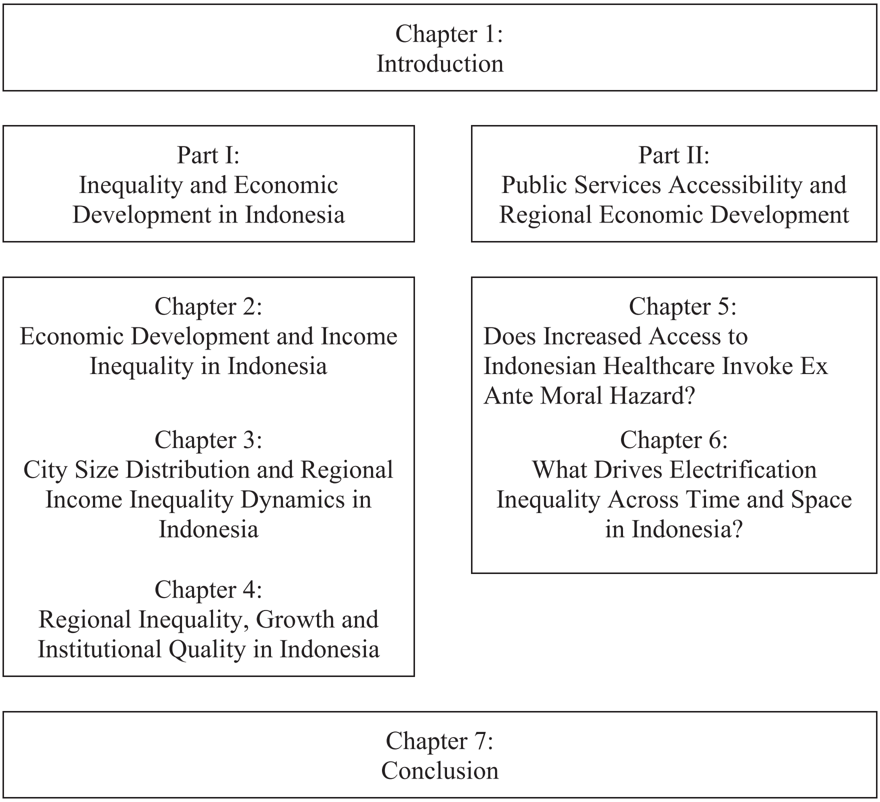
Many studies already have analyzed inequality in Indonesia. This thesis aims to contribute to the still growing body of literature by (i) focusing on empirical studies of inequality at the regional level in Indonesia; (ii) putting forward a a non-linear approach to study the relationship between inequality and economic growth; (iii) by introducing quality of institutions as potential driver of inequality patterns; (iv) by establishing a link between inequality and urban development; (v) as well as between inequality and accessibility to public services i.e. health and electricity, This thesis consists of several chapters that aim to provide insight into the nature and consequences of inequality dynamics in Indonesia during the past decades. In doing so, the thesis examines the impact of the fundamental institutional changes in Indonesia on different aspects of inclusiveness. Social inclusiveness is a central notion in the Sustainable Development Goals (SDG), especially as regards to good health, well-being, affordable and clean energy and reducing inequality. The different chapters in this thesis focus on these different dimensions of inclusiveness in the contemporary Indonesia's society.

The thesis essentially consists of two parts, as is illustrated in Figure 1.3. The first part studies various aspects of the relationship between inequality and economic development in Indonesia. A key question that is addressed in this part is how in Indonesia, as an emerging economy, the income distributions develop at various aggregation levels? Chapter 2 analyzes the implications of economic development of Indonesia on income inequality by comparing inequality in Indonesia to the rest of the world, looking at the within-country dynamics, and looking at the interpersonal dynamics. One of the particular aspects of that development is urbanization, which, amongst others, involves exclusion of the poor that cannot longer afford to live in cities. Taking an urban perspective, Chapter 3 studies the association in Indonesia between urbanization and income inequality, and identifies lessons that we can draw from the consequencs of urbanization for income inequality. Chapter 4 investigates whether inequality correlates with economic development by analyzing the relationship between changes in inequality and economic growth, and the role of institutional quality in shaping this relationship.

The second part of this thesis delves deeper into the questions regarding the accessibility of basic public services. The emphasis is on health and electricity. Given that the health care reforms in Indonesia were meant to promote inclusiveness in health access, Chapter 5 examines whether universal health care could change the behavior of people towards health i.e. ex ante moral hazard. Chapter 6 studies how the electricity diffuses at the regional level in Indonesia, against the background of the specific geography of Indonesia, implying the complexity of connecting people that live disperse across many relatively small islands, including huge disparities in terms of income and landscape. This makes Indonesia an interesting case study for analyzing electricity diffusion; given its geography, connecting everyone is not easy and the social implications are profound. Our contribution is a relevant step in the Indonesian context because the connectivity is still less than 100 percent.

Chapter 7 provides the summary of this thesis, explores its further policy relevance, and proposes future avenues of research.

Figure 1.3 Structure of Thesis



Chapter 2 Economic Development and Income Inequality in Indonesia

2.1 Introduction

In this chapter, we calculate and describe the income inequality dynamics in Indonesia over time and space, distinguishing three levels of spatial aggregation. First, we describe Indonesia's income inequality dynamics from an international perspective, in relation to the country's stage of development. Second, we document the income inequality evolution at the regional level within Indonesia, across islands and provinces. Third, we explore the interpersonal inequality dynamics by income class, based on newly constructed household expenditure data.

The aim of this chapter is to provide a comprehensive overview of income inequality dynamics in Indonesia since 1960, with a focus on the period 1987–2015. We combine the evolution of inequality over time with a spatial perspective, which allows us to assess the inequality dynamics in Indonesia in comparison with other countries and across islands and provinces as well as among individuals in relation to individuals on other islands and in other provinces. We assess inequality by exploring changes in various indicators of inequality, including the Gini index, the Theil index and the P_{90}/P_{10} ratio. Besides providing descriptive statistics, we present the results of σ- and β-convergence analyses of the variations in per capita income and key relevant indicators. Our analyses are based on consumption data provided by BPS Indonesia. As such, this chapter sets the stage for the rest of this thesis.

Over the last four decades, Indonesia has experienced, on average, high economic growth, which was triggered especially by the pro-growth policy in President Suharto's era. At the same time, however, income inequality at the national level started to rise in the early 1980s, when Indonesia started to liberalize its economy, including a turn to freer markets, deregulation and a larger role for the private sector in the economy (Booth 1992). As pointed out by many scholars, while economic growth in developing countries may reduce poverty (Dollar and Kraay

2002), its impact on income inequality is uncertain: higher economic growth may be accompanied by both lower and higher income inequality (Kanbur 2000; Banerjee and Duflo 2003). Boediono (1990) argued that the high economic growth in the 1980s in Indonesia was associated with a declining Gini during the 1980s. Yusuf et al. (2014), however, asserted that the rapid economic growth in Indonesia during the 1980s and 1990s was accompanied by a large reduction in poverty incidence, whereas its impact on inequality remains rather unclear. In this chapter, we document the relationship between economic growth and inequality in Indonesia, paying special attention to the spatial dimension of this relationship.

The chapter is structured as follows. In section 2.2, we briefly review the rapidly expanding literature on inequality in Indonesia, especially the different spatial levels at which income inequality has been studied so far. Section 2.3 presents the methods to measure inequality, and section 2.4 discusses the data sources and the construction of our own dataset. Section 2.5 discusses Indonesia's inequality from the international perspective, while section 2.6 focuses on the regional dimension of inequality in Indonesia. We then present the interpersonal inequality dynamics by income class in section 2.7. Section 2.8 presents the conclusions and options for future research.

2.2 Review of the Literature on Inequality in Indonesia

Indonesia's development and inequality patterns have been studied extensively by many researchers. For example, Yusuf et al. (2014) documented an increase in inequality in Indonesia based on estimates of expenditure inequality for 1993–2013, using several measures that draw on household expenditure data from the National Socioeconomic Survey (Susenas). In doing so, they noted that the rise in inequality reported since 2009 actually has a longer history, which used to be obscured by the fact that the Indonesian central statistics agency, *Badan Pusat Statistik* (BPS), used grouped data for its estimates of inequality until 2009. They also argued that the rise in inequality in Indonesia was largest in provinces or districts with low initial levels of inequality, implying a certain degree of convergence in spatial inequality patterns.

Yusuf et al. (2014) classified the potential drivers of inequality in Indonesia into two groups, which they labelled as exogenous and endogenous drivers. Exogenous drivers consist of shifting global trade and financial patterns and technical change, while endogenous drivers comprise macroeconomic policies, labour market policies, wealth inequality, fiscal policy (taxation and transfers) and government spending on public goods. For instance, the commodity boom in the 2000s could be perceived as a source of increasing inequality in Indonesia (Burke and Resosudarmo 2012; Yusuf et al. 2014). These commodities are grown and located outside Java, so increasing commodity prices contribute to a widening gap between Java and the rest of Indonesia. During a boom, the financial sector often grows much faster than the real sector, so the impact on income distribution can be predicted – the rich earn far more than the poor, and urban households' income grows faster than rural households' income, both exacerbating income inequality (Aziz 2015).

Yusuf et al. (2014) argued that changes in the formal labour market as an endogenous factor have increased inequality. The growth of formal employment has been rather stagnant in the last decade. Before the 1997–1998 crisis, the manufacturing sector and its employment grew at 11.2 per cent and 6.0 per cent, respectively. However, almost a decade after the crisis, the growth of the manufacturing sector was 4.7 per cent and its employment grew at only 0.9 per cent. This situation led to an excess of skilled labour in rural areas, which lowered the rural real wage, as revealed in the BPS data, which shows that the real wage in the agricultural sector has been declining over the last few years (Manning and Pratomo 2013). Thus, increasing inequality in rural areas is a logical consequence.

The Indonesian Government spends up to 25 per cent of the national budget on rice and fuel subsidies (Howes and Davies 2014). In response to an increase of 20 per cent in the rice price in the 2000s and an upsurge in international oil prices, the government increased the fuel retail prices in 2005 and distributed conditional cash transfers to the poor and near poor to compensate for the impact of inflation. This policy resulted in a slight decline of the Gini coefficient in 2006 (Yusuf and Resosudarmo 2008). However, since the international oil prices kept rising and the fuel subsidy was still high, this policy was not sufficient to maintain the lower Gini, which kept increasing. Agustina et al. (2012) also found that this fuel subsidy benefited the rich disproportionally more than the poor.

As regards the pattern of inequality across provinces and islands, Akita and Lukman (1995), applying the Williamson index, suggested a large decrease in inequality in the per capita GRP among provinces in Indonesia during the period 1975–1992. However, inequality in the non-mining per capita GRP remained stagnant during that period except in the mid-1980s, when an export-oriented policy was introduced. Across islands, it was shown by applying the same method that inequality tended to decline in Sumatera island while it tended to increase in other islands from the early 1980s until the early 1990s. Moreover, Akita et al. (2011) showed that the differences in inequality in terms of per capita GRP among Indonesia's largest regions (Java-Bali, Sumatra-Kalimantan-Papua and other regions in the eastern area) were small compared with the levels of inequality within those regions and that the levels of cross-regional inequality remained relatively constant over the years (1983–2004). They also found that an increasing level of inequality occurred not only within regions but also among districts within the provinces in those regions. Other studies (Tadjoeddin et al. 2001; Akita and Alisyahbana 2002; Hill et al. 2008) confirmed that regional inequality was relatively stable or increased slightly at the district level during the period 1993–1998. In a recent study, Vidyattama (2013) found that the inequality of the GRP per capita increased slightly at both the province and the district level in the period 1999–2009, especially as a result of the growth of Jakarta during the period 2002–2008.

Other studies have shown a widening gap between urban and rural areas and between Java and outer Java in terms of the income share after the Asian Financial Crisis and indicated that the income share of the rich is increasing while that of the poor is decreasing (Sakamoto 2007; Yusuf and Resosudarmo 2008; Mishra 2009; Chongvilaivan 2013; Miranti et al. 2013; Tadjoeddin 2013). Meanwhile, using a longer period of observation, Mishra (2009) found that the income distribution was relatively constant during the period 1963–2005, with low fluctuations between 1.9 and 2.5. Nonetheless, Yusuf et al. (2014) concluded that the gap among income groups in Java and urban areas measured by the decile dispersion ratio is worse than that measured by the Gini index, especially for urban areas and Java island. Their study, in line with Akita (2003), also demonstrated that inter-district inequality is higher than inter-province inequality, while there is no tendency for inter-regional disparity to increase in Indonesia. Furthermore, the Gini index in urban areas is much higher than that in rural areas, and, as most of the urban areas are located on Java island, this could affect the variation across islands (Sakamoto 2007; Mishra 2009; Chongvilaivan 2013; Tadjoeddin 2013).

As regards the dynamics of interpersonal inequality, Dick et al. (2002) asserted that the widening of the economic disparities perceived by many people in the 1980s was mainly due to excessive self-enrichment by Suharto's cronies. In line with Leigh and van der Eng (2009) and Milanovic (2016), we suggest that the top income shares in Indonesia were relatively high over the course of the twentieth century. This finding may surprise some readers as it contradicts the common "growth with equity" understanding of Indonesia's growth experience since the 1960s. Nevertheless, our results are bolstered by evidence from other sources. For example, the top wealth shares appear to be larger in Indonesia than in many other countries, whether one uses data from wealth surveys (Davies et al. 2009) or the Forbes rich lists. Piketty (2014) argued that inequality of asset and wealth ownership has driven increasing within-country income and expenditure inequality around the world. This could also have occurred in Indonesia. However, only a few studies have discussed this wealth inequality in Indonesia (Davies et al. 2009; Leigh and Van der Eng 2009, pp. 209–212).

2.3 Method of Inequality Measurement

A number of indices of inequality have been proposed to measure aspects of income distributions. A particularly convenient method of constructing such indices of inequality is to measure the extent to which an equal distribution of income deviates from the case in which all incomes are equal. Below is a list of methods and measures used for describing inequality in this chapter.

Decile Dispersion Ratio

The basic idea of this method is to measure the extent to which the top incomes differ from the bottom incomes. This presents the ratio of the average consumption (or income) of the richest 10 per cent of the population to the average consumption (or income) of the poorest 10 per cent. The interpretation is easy and simple, expressing the income of the top 10 per cent (the “rich”) as a multiple of that of those in the poorest decile (the “poor”). For general public interests, this ratio shows how big the difference is in the standard of living between the richest and the poorest members of a society.

Despite its simplicity and easiness, this method ignores information about incomes in the middle of the income distribution and does not even use information about the distribution of income within the top and bottom deciles. For further analysis, we can compare the earnings of the lowest and highest deciles relative to the median earnings to describe the dispersion of earnings. The formulas are as follows:

$$D = \frac{P_{90}}{P_{10}} \text{ and } R = \frac{P_{90}/P_{10}}{P_{90}/P_{50}}$$

where D is the dispersion of earnings, P_{10} is the lowest percentile of earnings, P_{50} is the median earnings, P_{90} is the highest percentile of earnings and R is the ratio of the richest to the poorest relative to the ratio of the richest to the median. If R is larger than 1, the gap is likely to be led by an increase in the income of the richest rather than a decrease in the income of the poorest.

Gini Coefficient

The Gini coefficient is probably the best-known indicator of income inequality. It compares every individual with every other and does not square the difference, especially for the middle class. This coefficient is basically equivalent to the size of the area between the Lorenz curve and the 45° line of equality divided by the total area under the 45-degree line of equality. The further the Lorenz curve deviates from the line of equality, the higher will be the resulting value of the Gini coefficient. Apart from its geometric counterpart in the Lorenz curve, it can also be calculated by using the following formula:

$$Gini = \frac{1}{2n^2\overline{y}}\sum_{i=1}^{n}\sum_{j=1}^{n}|y_i - y_j|$$

where y_i and y_j are the incomes of individual *i* and individual *j* with a mean of income $\overline{y}$ and where *n* is the total number of observations (Haughton and Khandker 2009). This coefficient ranges from zero to one, where zero shows perfect equality when all the income is shared equally and one is complete inequality when all the income is earned by only one individual.

The Gini coefficient has several attractive features: if all incomes are doubled, the population changes or two people swap incomes, then the coefficient does not change; a transfer from the rich to the poor will reduce inequality and lower the Gini coefficient. However, the main weakness of the Gini coefficient is its incapability of decomposing inequality by population group or income source because the total Gini coefficient of a society is not equal to the sum of the Gini coefficients of its groups. It is sensitive to the middle part of the income spectrum and not "neutral" or value free.

Theil Index

This index is part of the larger family of measures referred to as the general entropy (GE) class. It can be used to capture the variation in sub-groups of the population and regions by income source because this can be additive across different sub-groups or regions in the country. However, the Theil index does not have a straightforward representation and lacks the appealing interpretation of the Gini coefficient, which is why it is less commonly used than the Gini coefficient. The formula is as follows:

$$Theil\ Index = \frac{1}{N}\sum_{i=1}^{N}\frac{y_i}{\overline{y}}\ ln\left(\frac{y_i}{\overline{y}}\right)$$

where y_i is the income per capita of individual *i* and $\overline{y}$ is the average income per capita. The lower the index is, the lower the inequality.

Coefficient of Variation (CV)

The coefficient of variation is useful for analysing inequality as a comparison between datasets with different units because this coefficient is independent of the measurement unit. The formula of CV is simply the square root of the variance, that is, the standard deviation divided by the average income level:

$$CV = \sqrt{\frac{1}{n}\sum_{i=1}^{n}\left(\frac{y_i - \bar{y}}{\bar{y}}\right)^2} \qquad \bar{y} \equiv \frac{1}{n}\sum_{i=1}^{n} y_i$$

where y_i is the income per capita of individual i and $\bar{y}$ is the average income per capita.

Since the standard deviation of data must always be understood in the context of the mean of the data, the actual value of the CV is a dimensionless number. However, some disadvantages of the CV are as follows: the CV is not an ideal index of the certainty of a measurement when the number of replicates varies across samples; when the mean value is close to zero, the coefficient of variation will approach infinity and is therefore sensitive to small changes in the mean. Unlike the standard deviation, it cannot be used directly to construct confidence intervals for the mean.

In this chapter, we use aggregate data at the country level for the period 1961–2013, at the province level for the period 1961–2015 and at the district level for the period 1999–2014, while individual data for our consumption analysis are derived from Susenas covering the period 1987–2015.

2.4 Data Sources

We use several data sources to calculate the Gini index across countries, specifically the national bureaus of statistics of the respective countries, the inequality dataset constructed by Atkinson, the Luxembourg Income Study Database (LIS) and the Deininger and Squire dataset. To explain the relationship between the Gini index and the per capita GDP, we also use the World Development Indicators (WDI) and classify countries into four groups as follows: ASEAN countries (Malaysia, the Philippines and Thailand), emerging countries (Brazil, India, China and South Korea), African countries (Kenya, Nigeria and South Africa) and rich countries (Australia, the euro area, Japan, Singapore and the United States). However, not all countries have data for the entire period. For most countries, the earliest available data are for 1960, except for China (1969), South Africa (1964) and Singapore (1966). Furthermore, in performing panel regressions, we define country groups based on the World Bank definition and data availability as follows: (1) Latin America (18 countries): Argentina, Bolivia, Brazil, Chile, Colombia, Costa Rica, the Dominican Republic, Ecuador, El Salvador, Guatemala, Honduras, Jamaica, Mexico, Panama, Peru, Puerto Rico, Uruguay and Venezuela; (2) the Middle East (5 countries): Egypt, Iran, Israel, Jordan and Tunisia; and (3) East Asia (9 countries): China, Indonesia, Japan, South Korea, Malaysia, the Philippines, Singapore, Thailand and Vietnam.

At the regional level, we use the per capita gross regional product (GRP) and household expenditures from the microdata consumption module of the Social and Economic Survey (Susenas) provided by the central statistics agency, *Badan Pusat Statistik* (BPS). Susenas's datasets are collected every three years, and since 2011 they have been collected every single year. Our datasets are available from 1987 onwards. As a consequence of decentralization, 8 new provinces have been established since 2001, and, as of 2015, the total number of new and old provinces is 34. For consistency purposes, we keep 26 provinces as our base by regrouping new provinces with their parent provinces, namely Riau Island with Riau, Bangka Belitung with South Sumatera, Banten with West Java, North Kalimantan with East Kalimantan, Gorontalo with North Sulawesi, West Sulawesi with South Sulawesi, North Maluku with Maluku and West Papua with Papua. We also recode these provinces to construct a new panel province dataset.

In addition, we reclassify new districts into their parent districts to perform panel regressions. Our complete dataset at the district level is from 1999 onwards. Special attention is necessary when the parent districts belong to a new province and have a different name, and then all the district codes should be recoded under the new coding system provided by BPS. A detail explanation of the construction of the panel district dataset can be found in Appendix A2.1.

In calculating the inequality indicators (i.e. Gini, Theil, Ratio P_{90}/P_{10} and Decile Dispersion Ratio), we use extensively household expenditure data from Susenas, the only one nationally representative data set on socio-economic conditions in Indonesia. In this thesis, we use Susenas 1987 as the earliest available data source that involved 65,200 households. Since 1993, its sample size has grown to approximately 200,000 households, and since 2011 onwards its sample covers around 300,000 households and 1.1 million household members. The expenditure variable in Susenas consists of detailed food and non-food items where those items have been expanded since 1993, i.e. 203 food items, 28 items on housing and utilities, 37 on goods and services, 15 on clothing, 13 on durables and 5 on ceremonies and festivities. In addition, Susenas can be linked to Potensi Desa (PODES) at the district (kabupaten) level which I exploit in Chapter 4.

Despite its national representativeness, Susenas has some limitations. First, Susenas tends to exclude the very wealthy since they are hard for the enumerators to reach; if they are included, they are often considered as outliers (Mishra 2009). Edward and Sumner (2015) suggested that Susenas may be weak in capturing top incomes as only approximately 10 per cent of Indonesians consumed more than $10 per day in 2012 (in 2011 purchasing power parity (PPP) dollars). The World Bank also mentioned this selection bias problem in Susenas whereby enumerators categorize the rejection or the unwillingness of the very wealthy as selective non-response. As a result, the Gini index based on household expenditure is likely to be a lower-bound estimate for inequality (Atkinson and Brandolini 2001; Lahoti et al. 2014; Yusuf et al. 2014).

Second, the expenditure variable measures total consumption, financed by out-of-pocket money and/or by subsidies. Since total expenditure is used as a proxy of income, expenditure from out-of-pocket money in Susenas is likely to overestimate the purchasing power of the households, especially the purchasing power of low-income households who are likely to receive various kinds of economic assistance. Hence, households may appear to have high income while some parts of this income may actually be derived from subsidies (Johar et al. 2018). In addition, the variable income in Susenas is often unreliable because it is self-reported income that tends to be underreported.

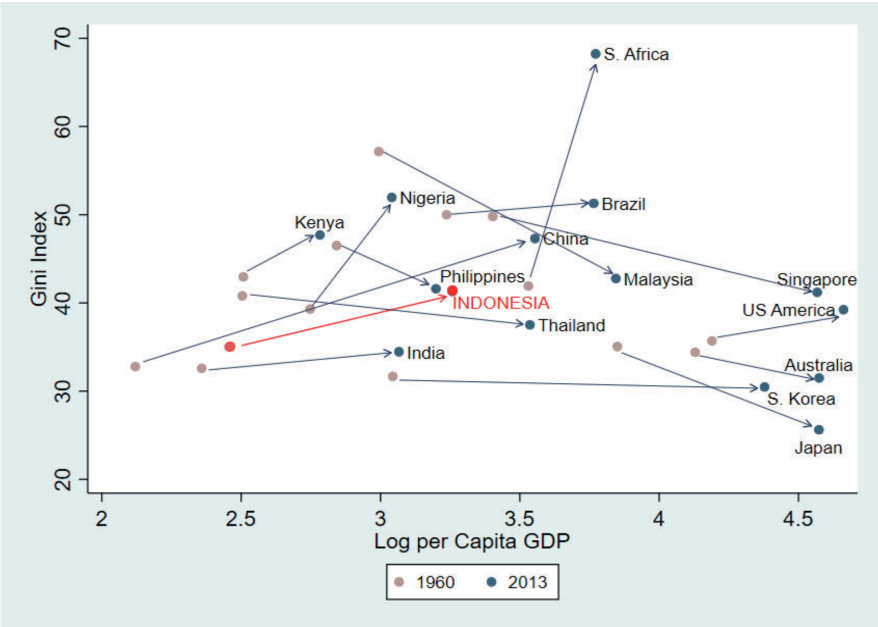
Third, there may be recall bias due to the reference period. According to Susenas Guidelines, the interviewer asks the respondent (the household representative) “How much did the household spend on [item] in the past [reference period]?" The reference period for food items is the past seven days or one month ago, while several reference periods are applied for non-food items depending on the type of non-food expenditures i.e. a month, the past three months or six months, or one year ago. These reference periods might suffer from underreporting actual expenditure by the household especially for non-food items. For instance, the respondent may not remember the exact amount spent for hospital, clinic, medicine, etc. by each household member for the last three months. Consequently, health expenditure reported in Susenas tends to be lower than the actual spending. Furthermore, health cost is financed by out-of-pocket and or health subsidy especially for the poor where they likely receive health services free of charge. As a result, health expenditure as out-of-pocket is likely underestimated (Johar et al. 2018). Further discussion concerning data limitations on health variable in Susenas can be found in Appendix A5.A1. Bearing these limitations in mind, Susenas data are still the best source available, but have to be analysed and interpreted with care.

2.5 The International Perspective

To understand Indonesia's position in relation to other countries, we compare the per capita GDP in 2005 USD prices and the Gini index from 15 countries over time. The trend of both the per capita GDP and the Gini coefficient of Indonesia has increased over time, but its position is relatively low compared with that of other countries. The per capita GDP of Indonesia has increased for the last three decades, and its figures are higher than those for India, Kenya, the Philippines and Nigeria but lower than those for Thailand, Malaysia, Brazil, South Korea, China and South Africa (2013). It appears that Indonesia is catching up with the richer countries among ASEAN and emerging countries. However, in terms of inequality measured by the Gini index, Indonesia's figure is higher than that of Thailand, India, South Korea, Australia, Japan and the USA (Figure 2.1). The same pattern emerges if we use the per capita GDP with PPP constant 2011 USD, as reported in Appendix A2.2.

In terms of the absolute value of the Gini index, Indonesia's figure is much lower and fairly stable compared with ASEAN countries, emerging countries and African countries, as shown in Figure 2.1 and Appendix A2.2. It has fluctuated between 33 and 38 for the past 20 years. The only Gini index pattern similar to Indonesia's is that of India. If we compare it with the ASEAN countries, where the Gini index ranged from 40 to 50, or with African countries with a range of 40 to 70 or even with Brazil with an index above 50 in the same period, Indonesia's figure seems unusual. In response to these findings, Booth and Sundrum (1981); Booth (1992) and Sundrum (1992) suggested that the low income inequality in Indonesia for a long time was due to the government policies, such as subsidies in the agriculture, infrastructure, education and health sectors in the early 1970s.

Figure 2.1 Relationship between the Gini Index and the per Capita GDP (Constant 2005 USD): Indonesia and Selected Countries 1960 and 2013



Source: WDI, CBS of the respective countries and various inequality datasets constructed by Atkinson, LIS, and Deininger and Squire.

In addition, the Gini index in Indonesia is calculated based on household expenditure, leading to a lower value than the index based on income.[2] Frankema and Marks (2009,

[2] BPS has conducted a three-year National Socio-Economic Survey (Susenas) since 1979 by distributing a consumption questionnaire of food and non-food items. Although BPS collects household income data, the result is inaccurate due to respondents' unwillingness to reveal

2010) proposed alternative measures for income inequality in Indonesia, specifically the ratio of unskilled wages to GDP per worker, the Theil coefficient of the inter-industry wage distribution in the manufacturing sector and the development of the relative size of the urban informal sector. Their results suggest that the major determinants of income inequality in Indonesia, that is, the wage inequality and the share of self-employed in the labour force, are similar to the findings for Brazil and Mexico, which are both known for recording among the highest levels of income inequality in the world. Moreover, adjusting the estimate of inequality using the taxation data of top incomes shows that the share of income of the richest is generally much larger in Indonesia than in other countries (see Leigh and Van der Eng 2009) and challenges the perception that Indonesia is relatively egalitarian. Thus, the claim of low income inequality in Indonesia by international standards is not based on entirely convincing evidence (Van Zanden and Mark 2012). Despite the weaknesses of Susenas in capturing the top income earners in Indonesia, the rise of inequality at all income levels occurred throughout Indonesia. We will discuss this further in section 2.5.

Deininger and Squire (1998) pointed out that many countries that started with low levels of per capita income grew rapidly without an increase in inequality. On the other hand, other countries that failed to grow were not immune to possibly considerable swings in aggregate measures of inequality. In a few countries, where a strong relationship has emerged between growth and inequality, the Kuznets hypothesis is contradicted almost as often as it is confirmed. Following the debates about inequality and growth, two main lessons can be learned. First, Lundberg and Squire (2003) argued that, in searching for the causal connection between growth and inequality, one may not consider that both could be determined by the same set of factors over which national decision makers have some control. Still, their findings show that the determinants of growth and improved equity are not mutually exclusive and that both equity and growth can benefit significantly from increased government expenditures aimed at the redistribution of land and the provision of universal secondary education, for instance. Second, Ravallion and Chen (2003) suggested that,

---

their true income: they tend to underreport their income in the survey. Further discussion on the data limitation can be found in Section 2.4.

when considering the effect of inequality, it is equally important to ask questions about the *quality* of growth as it is to consider the *rate* of growth. Thus, countries that have high levels of initial inequality find it exceedingly difficult to reduce the incidence and depth of material deprivation through economic growth if measures are not taken to rectify the maldistribution of economic resources and opportunities (World Bank 2006).

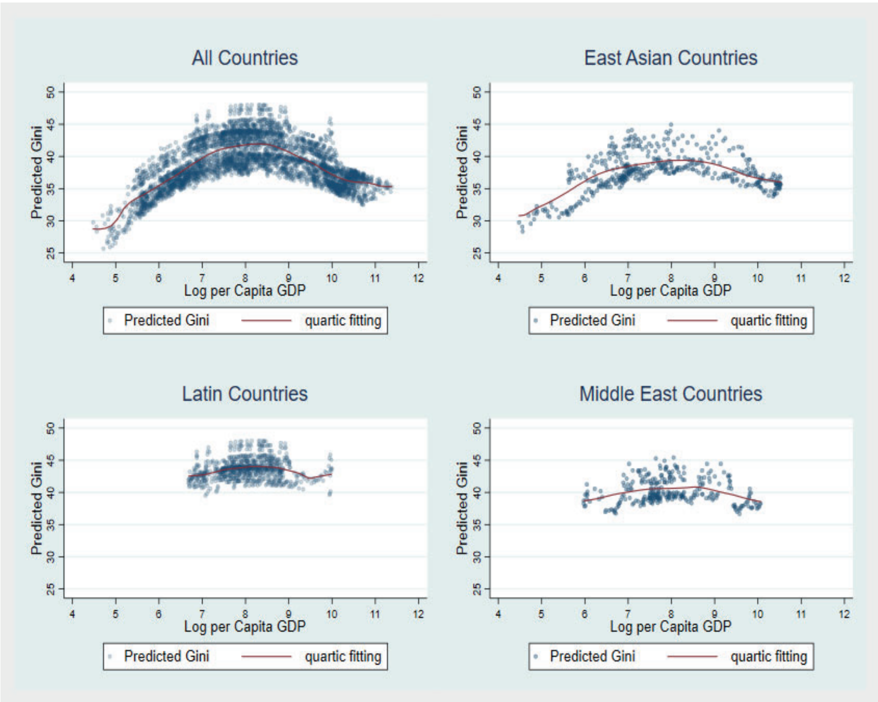
2.5.1 Kuznets at the Country Level

We continue by further examining the relationship between inequality and income in an international context, in an attempt to determine whether Indonesia is indeed characterized by exceptionally low income inequality given its stage of development. The relationship between income inequality and income per capita is assessed by performing an analysis regressing the Gini index on various forms of GDP per capita and year and dummy variables, that is, events capturing decentralization for the years after 2000, democratization for the years after 2003 and the commodity boom for the years after 2005 for country groups of Latin America, East Asia and the Middle East in ten-year periods starting in 1970, 1980, 1990, 2000 and 2010. The country group dummy variables are used to control for time-invariant omitted-variable bias, while the period of the time dummy variable is included to control for global shocks that might have an impact on inequality in any time period but are not captured by the explanatory variables in the model. We also add interaction between income and event dummies and country and period dummies. We use data for 95 countries from several sources, namely the World Bank, the LIS and the Central Bureau of Statistics of the respective countries. The unit of observation in this model is the country level, using a country-level panel dataset with annual observations for the period from 1961 to 2013.

$$\begin{aligned} Ineq_{it} = {} & U_i + T_j + \beta_0 + \beta_1 \cdot Income_{it} + \beta_2 \cdot [Income_{it}]^2 + \beta_3 \cdot [Income_{it}]^3 \\ & + \beta_4 \cdot time + \sum \beta_g \cdot d\_event_h + \sum \beta_k \cdot [d\_event_h \cdot Income_{it}] \\ & + \sum \beta_l \cdot d\_period_p + \sum \beta_m \cdot d\_country_c + \sum \beta_q \cdot [d\_period_p \cdot d\_country_c] \\ & + \sum \beta_r \cdot [d\_event_h \cdot d\_country_c] + \varepsilon_{it} \end{aligned} \tag{2.1}$$

where i represents countries and t represents time; $Ineq_{it}$ is inequality measured using the Gini coefficient; U is a country fixed effect; T_j is a year effect; *time* is a year trend; $Income_{it}$, is the per capita GDP; d_event_h represents an event dummy where h represents democratization, decentralization and the commodity boom; d_period_p represents a ten-period dummy where p is 1970, 1980, 1990, 2000 and 2010; $d\_country_c$ is a country dummy where c is Latin America, East Asia and the Middle East; and ε_{it} is the error term.

Figure 2.2 Kuznets Estimation at the Country Level



Source: Authors' estimation based on WDI, LIS and CBS.

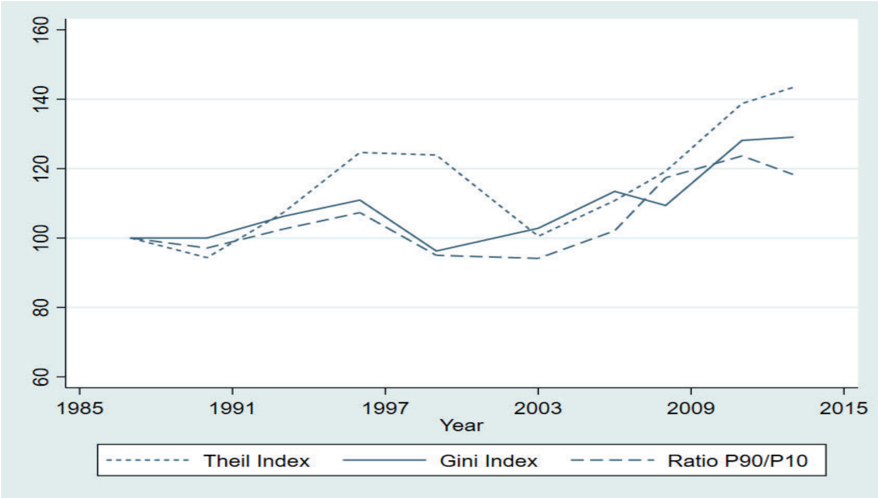
Our results show that the income per capita and its square and the interactions between the dummy Latin and the dummies 1970, 1980 and 2010 are statistically significant in affecting the Gini index (Appendix A2.3). These results suggest that the period of the oil bonanza in the 1970s, the oil shock in the 1980s and the global crisis in 2008 played an important role in determining inequality patterns across countries in Latin America. The plot of the GDP per capita and the predicted Gini index depicts an inverted U-shape (consistent with the Kuznets theory). In the beginning, income inequality is positively related to income per capita, while, after a certain point, the relationship turns negative. It appears that the relationship between the Gini index and the GDP per capita depends on the stage of development in each country. A similar pattern is presented for the country group of East Asia, while it is not very clear in the Latin American and Middle Eastern groups (Figure 2.2).

2.6 Interregional Inequality Dynamics

Indonesia's development and inequality have been studied extensively by many researchers. In the last four decades, Indonesia has experienced high economic growth, especially triggered by the pro-growth policy in President Suharto's era. However, many scholars have noted that inequality has also escalated, pointing out that, although it may reduce poverty (Dollar and Kraay 2000), high growth does not necessarily mean lowering inequality in developing countries (Kanbur 2000). Yusuf et al. (2014) also suggested that the rapid economic growth in Indonesia during the 1980s and 1990s was accompanied by a large reduction in poverty incidence but that its impact on inequality is rather unclear.

At the national level, an increasing trend in inequality consistent with the Kuznets hypothesis started in the early 1980s, when Indonesia started to liberalize its economy, promoting a freer market, a larger role for the private sector, deregulation and so on (Booth 1992). If we look at inequality figures, such as the Gini coefficient, the Theil index and the ratio of the top 10 per cent to the bottom 10 per cent, over the last 26 years, all the indicators tell the same story of an increasing trend (see Figure 2.3). The Gini coefficient increased by 34 per cent, while the Theil index increased by 43.5 per cent and the Ratio P_{90}/P_{10} increased by 18.3 per cent. The latter means that the rich spent almost five times more than the poor in 2013 as compared with 1987.

Figure 2.3 Average Inequality Indicators at the National Level, 1987–2013 (1987=100)



Source: Authors' calculation from Susenas, BPS.

Similar patterns are also shown for the ratio of the top 10 per cent to the bottom 10 per cent for both per capita food and per capita non-food expenditure. On a per capita basis, the rich spend about 3 to 3.5 times more on food than the poor and 7 to 9 times more on non-food goods. Unlike per capita food expenditure, the average P_{90}/P_{10} ratio per capita for non-food expenditure has fluctuated over time. The gap started increasing in 2003 and reached its peak in 2011. This may partly be due to a commodity boom resulting in increasing international prices of most of Indonesia's exported goods, such as coal, nickel, copper and crude palm oil. It is also reflected in a larger proportion of primary products to total exports in Indonesia during those periods (Azis 2015). Exporters, as well as those who were part of the associated value chain of those commodities, benefited from this commodity boom, which translated into a larger gap between the rich and the poor in per capita non-food expenditure and gave rise to variation in inequality across households within Indonesia.

Table 2.1 Inequality within Islands: Per Capita Household Expenditure, 1987–2013

	1987	1999	2003	2013	Changes (%)	
					1987–2013	2003–2013
Gini Coefficient						
Sumatera	25.58	25.75	28.97	36.31	42	25
Java-Bali	30.50	29.83	31.97	40.62	33	27
Kalimantan	27.25	26.78	28.95	36.90	35	27
Sulawesi	28.00	28.20	29.91	42.10	50	41
Eastern	31.33	28.55	31.10	38.20	22	23
Average	28.53	27.82	30.18	38.83	36	29
Theil Index						
Sumatera	13.96	14.56	15.63	23.02	65	47
Java-Bali	19.29	26.15	19.75	29.67	54	50
Kalimantan	15.67	14.35	17.02	22.69	45	33
Sulawesi	15.64	13.45	16.90	32.13	105	90
Eastern	21.41	20.91	15.59	24.73	16	59
Average	17.20	17.88	16.98	26.45	54	56
Ratio P_{90}/P_{10}						
Sumatera	3.31	3.30	3.24	3.80	15	17
Java-Bali	3.76	3.84	3.85	4.81	28	25
Kalimantan	3.62	3.42	3.48	4.05	12	17
Sulawesi	3.61	3.32	3.47	4.96	38	43
Eastern	4.20	4.64	3.57	4.35	3	22
Average	3.70	3.70	3.52	4.40	19	25
Decile Dispersion Ratio						
Sumatera	1.69	1.70	1.70	1.74	2	2
Java-Bali	1.72	1.73	1.72	1.84	7	7
Kalimantan	1.74	1.76	1.73	1.84	6	6
Sulawesi	1.77	1.66	1.76	1.85	5	5
Eastern	1.80	1.96	1.75	1.79	–1	2
Average	1.75	1.76	1.73	1.81	4	5

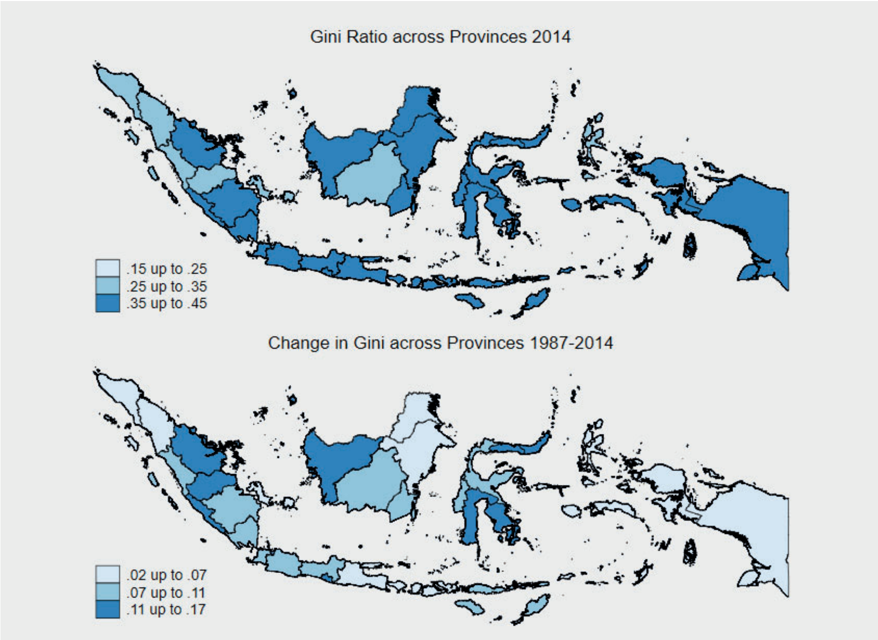
Source: Authors' calculation from Susenas, BPS.

At the regional level, we calculate four inequality indicators, viz. the Gini coefficient, the Theil index, the ratio of the top 10 per cent to the bottom 10 per cent and the decile dispersion ratio – the average ratio of the rich to the poor over the rich to the median based on per capita household expenditure from 1999 to 2014 (Table 2.1). In general, inequality within islands increased from 1987 to 2013. Eastern island was the most unequal island in 1987, but this position was taken over by Sulawesi island in 2013 based on all the indicators. Sulawesi island is the most striking case, in which the rich spent almost five times more than the poor. Sulawesi island, a centre of commodity production, benefited from the commodity boom, which led to an increase in inequality, as we previously discussed. Moreover, if we look at the decile dispersion ratio, for which all the figures are always larger than one, we can conclude that the increasing gap between the rich and the poor is due to the rich becoming richer rather than the poor becoming poorer (or, in other words, the rich benefit more than the poor). This pattern is also consistent with per capita food and non-food expenditure, for which inequality in non-food expenditure fluctuates strongly between islands compared with its food expenditure counterpart (Appendix A2.4).

Java island has an outstanding position, which may come from the role of Jakarta. Akita et al. (2011), in their study of regional income inequality, suggested that, since the share of mining has decreased, the spatial distribution of manufacturing has played a more important role in the inequality of Sumatra and Kalimantan, while the primacy of Jakarta, with strong urbanization economies facilitated by globalization and trade and financial liberalization, has determined much of the Java–Bali region's inequality and therefore the overall inequality in Indonesia. Goldberg and Pavcnik (2007) suggested that the effect of globalization on inequality depends on a trade protection pattern prior to liberalization; the particular form of liberalization and sectors it affected; the flexibility of markets in particular the degree of labor and capital mobility; and the existence of other concurrent trends (e.g., skill-biased technological change). Trade liberalization and export growth are found to be associated with lower income inequality, while increased financial openness is associated with higher inequality (Jaumotte et al. 2013).

Figure 2.4 depicts the variation in the Gini coefficient across provinces in 2014 and its change with the Gini coefficient in 1987. A high level of inequality has been experienced by most of the provinces throughout Indonesia, where more dark blue is shown in the top picture. Only a few provinces in Sumatera, Kalimantan and Sulawesi have experienced a substantial increase in the Gini index (bottom picture). The same pattern is also shown in the Gini index across districts 1999 and 2014 (Appendix A2.5).

Figure 2.4 Gini Ratio 2014 and its Change in 1987–2014 across Provinces



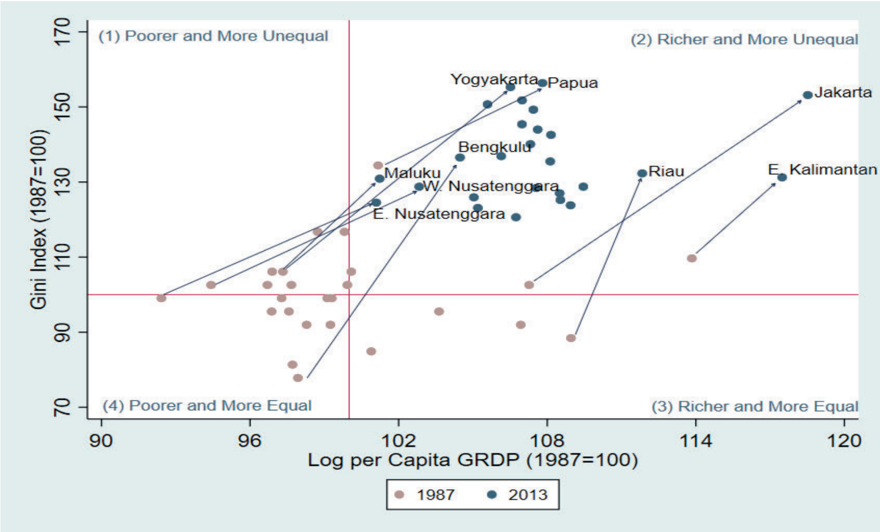
Source: Authors' calculation from Susenas, BPS.

2.6.1 Relationship between Inequality and per Capita GRDP

To understand further the relationship between inequality and per capita GRDP across provinces, we plot the change in the GRP per capita against the change in inequality between 1987 and 2013 and arrange it into four quadrants, that is, (1) "poorer" and more unequal, (2) "richer" and more unequal, (3) "richer" and more equal and (4) "poorer" and more equal. In general, more provinces are facing increasing inequality as well as better economic performance, moving to the second quadrant (Figure 2.5).[3] For instance, poor provinces that are characterized as equal or unequal provinces, such as Maluku, East Nusatenggara, West Nusatenggara, Bengkulu and Yogyakarta, respectively, are moving to the second quadrant. Jakarta has skyrocketed in terms of per capita GRDP and inequality compared with the other provinces.

[3] The same patterns are also apparent in the relationship between inequality and per capita GRDP excluding oil and gas, per capita GRDP private non-natural resources, and per capita GRDP non-agriculture in 1987, 1999 and 2013.

Figure 2.5 Inequality and Development in 1987 and 2013



Source: Statistic Indonesia and authors' calculation from Susenas, BPS.

If we look at the economic performance and inequality of Jakarta in 1987, they show that its per capita GRDP and its inequality are not the highest ones, but both sharply increase in 2013. The role of Jakarta as the capital city of Indonesia has played an important role in attracting the rich to accumulate their wealth in Jakarta rather than other provinces, especially during the commodity boom from 2004 until 2011. This phenomenon is evidenced by Papua, Riau and East Kalimantan, which are known as resource based but have different development patterns from Jakarta. We suspect that these resource-rich provinces create different institutions. For example, the license required to extract forestry or mining is different from the license required to extract oil and gas; thus, the typical rent seekers are also different. The latter would correlate with an increasing gap between the rich and the poor (Baland and Francois 2000; Wadho and Hussain 2020).

2.6.2 Kuznets at the Province Level

To assess further the Kuznets pattern at the province level, as we did at the country level, we perform a regression of the Gini index as a proxy for income inequality with the various forms of per capita GRP as a proxy for income per capita, the time trend, a dummy for islands (Sumatera, Java–Bali, Kalimantan, Sulawesi and Eastern), a dummy for fundamental policies or events such as the banking reform (1983–1984 and 1994–1995), economic reform (1986–1987), the Asian Financial Crisis (1997–1998), decentralization (2000–2001), democratization (2004–2005) and the commodity boom (2006–2011) and interaction dummies with per capita GRP and islands. We use data of 33 provinces from BPS. The unit of observation in this model is the province, and we use a province-level panel dataset with annual observations for the period from 1977 to 2013.

$$\begin{aligned} Ineq_{it} = U_i + T_j + \beta_0 + \beta_1 \cdot Income_{it} + \beta_2 \cdot [Income_{it}]^2 + \beta_3 \cdot [Income_{it}]^3 \\ + \beta_4 \cdot time + \sum \beta_g \cdot d\_event_h + \sum \beta_k \cdot [d\_event_h \cdot Income_{it}] \\ + \sum \beta_l \cdot d\_island_p + \sum \beta_q \cdot [d\_event_h \cdot d\_island_p] + \varepsilon_{it} \end{aligned} \quad (2.2)$$

where i represents provinces and t represents time; $Ineq_{it}$ is inequality measured using the Gini coefficient; U is a province effect; T_j is a year effect; t is a time trend; $Income_{it}$ is the income per capita in terms of the GRDP per capita; d_event_h represents an event dummy, where h is the oil bonanza, bank reform, economic reform, Asian Financial Crisis, decentralization, democratization and commodity boom; d_island_p represents an island dummy, where p is Sumatera, Java-Bali, Kalimantan, Sulawesi and Eastern islands; and ε_{it} is the error term.

At the province level, our results show that income per capita in all forms, bank reform, democratization and their interactions are statistically significant in affecting the Gini index. However, those events have different effects on islands. For instance, the bank reform had a greater impact on Java-Bali while democratization affected Sumatera by increasing inequality. Our results also show that the commodity boom period is statistically significant in decreasing inequality in Sumatera, Java-Bali and Kalimantan. This indicates that the commodity boom benefited not only the elites in Jakarta but also the people from those islands. Moreover, other events, such as decentralization, economic reform and the second bank reform, tend to increase inequality in Java-Bali, Sumatera and Sulawesi. A summary of the regression output can be found in Appendix A2.6.

From the regression results, at the province level, we identify Indonesia's pattern as a typical power-three function following a fraction of the Kuznets curve, and, at the district level, we find a typical Kuznets inverted U-shape.[4] A similar pattern is displayed at the island level with some variations (Appendix A2.7 and A2.8). It seems that each province has a different initial stage of development, which affects the Kuznets pattern. However, some researchers have argued that Indonesia did not follow Kuznets's prediction in its early stage of development. For three decades before the Asian Financial Crisis (AFC), Indonesia experienced sustained high growth while maintaining a stable Gini coefficient (around 0.32 to 0.36). However, the story changed after recovering from the AFC. Even though the economy was able to recover fairly quickly from the AFC and was quite robust in the face of the 2008 Global Financial Crisis (GFC), the Gini coefficient increased rapidly, reaching its highest ever peak of 0.41 in 2011 (Tadjoeddin 2013).

2.7 Interpersonal Inequality Dynamics

In this section we discuss interpersonal inequality dynamics from various aspects namely region, income class and polarization.

2.7.1 Interpersonal Inequality Dynamics by Regions

To describe the dynamics in inequality across islands and provinces, we calculate the coefficient of variation for each inequality indicator and each type of household expenditure, as presented in Table 2.2 and Table 2.3, respectively. This coefficient illustrates how one island varies from the other islands. If one island moves differently, it means that not all the provinces of that island are moving in the same direction, and then the coefficient fluctuates. Both tables describe the variation of the Gini index and Theil index as decreasing This means that income inequality tends to converge across islands and provinces based on the Theil index and the Gini index.

[4] Our regression results across districts do not show that independent variables are individually significant, as the country and province levels do, but the overall *F*-test of the district level shows that the variables are jointly significant.

However, if we observe the period of the commodity boom starting in late 2003 until 2013, the variation of all the inequality indicators decreases, implying that the income convergence between provinces was supported by a long period of relatively high commodity prices, which benefited commodity-exporting provinces.

Moreover, if we look at the pattern of non-food expenditure, the variation in the ratio of the top 10 per cent to the bottom 10 per cent tends to diverge across islands while converging across provinces. This pattern indicates that the non-food consumption inequality across provinces and within islands varies and that the gap between the poor and the rich becomes larger since the non-food expenditure of the rich is a larger proportion. As the variance fluctuates in 1999, around the Asian crisis, this suggests that the crisis associated with an increasing gap between the rich and the poor.

Table 2.2 Variance of Inequality per Capita Household Expenditure, per Capita Non-food Expenditure and per Capita Food Expenditure between Islands, 1987–2013

Coefficient of Variation	1987	1999	2003	2013	Changes (%)	
					1987–2013	2003–2013
Total Expenditure						
Gini Index	0.08	0.06	0.04	0.06	–23	44
Theil Index	0.18	0.31	0.10	0.16	–10	61
Ratio P_{90}/P_{10}	0.09	0.15	0.06	0.11	27	78
Decile Dispersion Ratio	0.02	0.07	0.01	0.03	17	97
Non-food Expenditure						
Theil Index	0.13	0.17	0.11	0.12	–12	3
Ratio P_{90}/P_{10}	0.10	0.04	0.08	0.12	19	58
Decile Dispersion Ratio	0.04	0.04	0.04	0.05	28	26
Food Expenditure						
Theil Index	0.15	0.26	0.16	0.12	–18	–26
Ratio P_{90}/P_{10}	0.08	0.20	0.07	0.07	–21	–5
Decile Dispersion Ratio	0.03	0.10	0.03	0.02	–41	–42

Source: Authors' calculation from Susenas, BPS.

Table 2.3 Variance of Inequality per Capita Household Expenditure, per Capita Non-food Expenditure and per Capita Food Expenditure between Provinces, 1987–2013

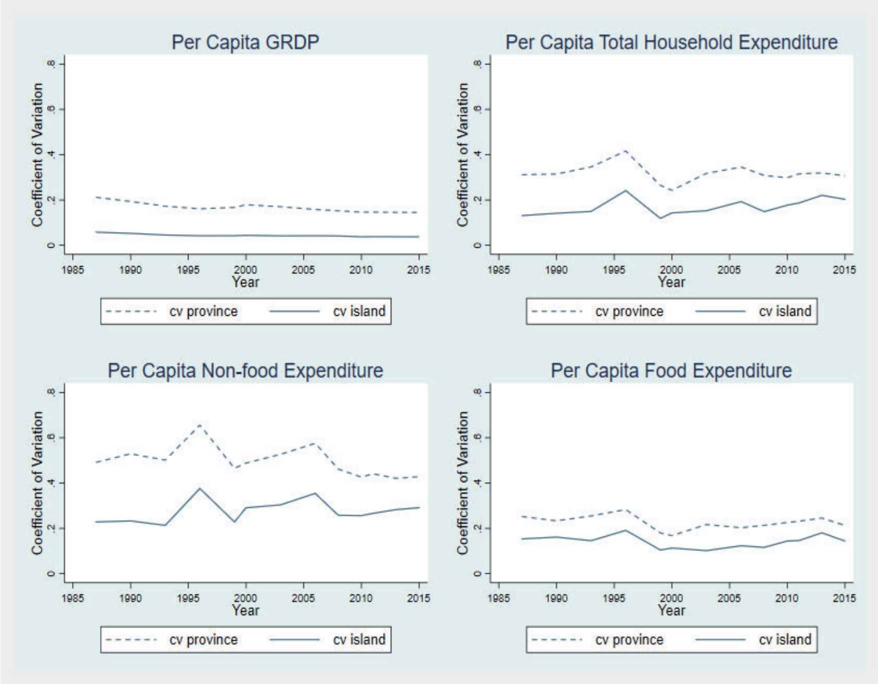
Coefficient of Variation	1987	1999	2003	2013	Changes (%)	
					1987–2013	2003–2013
Total Expenditure						
Gini Index	0.12	0.11	0.10	0.08	–30	–15
Theil Index	0.23	0.57	0.26	0.19	–19	–28
Ratio P_{90}/P_{10}	0.15	0.29	0.20	0.15	–2	–24
Decile Dispersion Ratio	0.06	0.13	0.06	0.05	–11	–16
Non-food Expenditure						
Theil Index	0.18	0.36	0.21	0.16	–12	–28
Ratio P_{90}/P_{10}	0.20	0.20	0.22	0.18	–12	–18
Decile Dispersion Ratio	0.10	0.12	0.09	0.11	10	17
Food Expenditure						
Theil Index	0.22	0.34	0.43	0.18	–18	–58
Ratio P_{90}/P_{10}	0.13	0.34	0.17	0.11	–16	–36
Decile Dispersion Ratio	0.06	0.17	0.08	0.04	–24	–43

Source: Authors' calculation from Susenas, BPS.

Regarding the variation in the decile dispersion ratio, it tends to decrease for per capita food expenditure and is likely to increase for non-food expenditure but fluctuates over time. This confirms that the gap in per capita non-food expenditure is much larger than that in per capita food expenditure. Moreover, those figures seem to be relatively stable over years and the variance correlates with inequality across provinces being likely to diverge. It appears that the gap between the rich and the average of the population has existed for a long time, and this supports the argument of the rise of new elites.

Figure 2.6 summarizes the variation in per capita GRDP and per capita household expenditure across islands and provinces. In terms of economic development, there is wide disparity across provinces but relative stability across islands. However, in terms of per capita household expenditure, the variation fluctuates more, with a tendency to increase for non-food expenditure and decrease for food expenditure as well as for total expenditure. The figure reveals a larger gap in non-food expenditure across provinces and islands than in food or total expenditure.

Figure 2.6 Variance per Capita Gross Regional Product and Household Expenditure across Provinces and Islands, 1987–2015



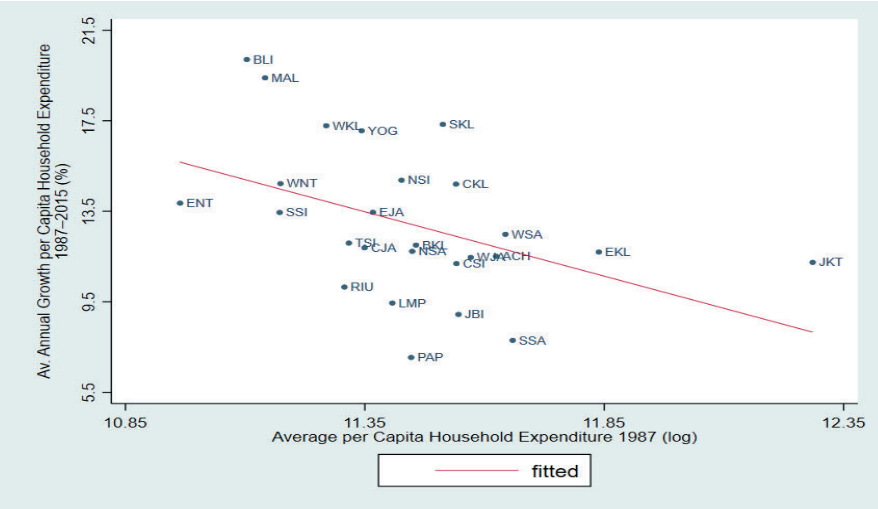
Source: Authors' calculation from Susenas, BPS.

2.7.2 Interpersonal Inequality Dynamics by Income Class

We explore further the dynamics in interpersonal inequality by income class through calculations using the income class based on the total household expenditure and then aggregating at the province level. First, we plot the average annual growth per capita expenditure and the average initial expenditure in 1987 (in log form) for all provinces, as shown in Figure 2.7. This figure suggests that β-convergence existed at the province level during the period from 1987 to 2015. The correlation between the growth in real per capita household expenditure (henceforth simply "expenditure") over time and its initial level is negative, so there is β-convergence (Barro and Sala-i-Martin 1992; Mankiw et al. 1992). If we compare this with the second period, 1999 to 2015, the catching-up process in this period was slower than that in the first period, 1987 to

2015, as presented by the flatter slope (Appendix A2.8). This might be due to the Asian crisis that occurred in 1998, which affected all of Indonesia.

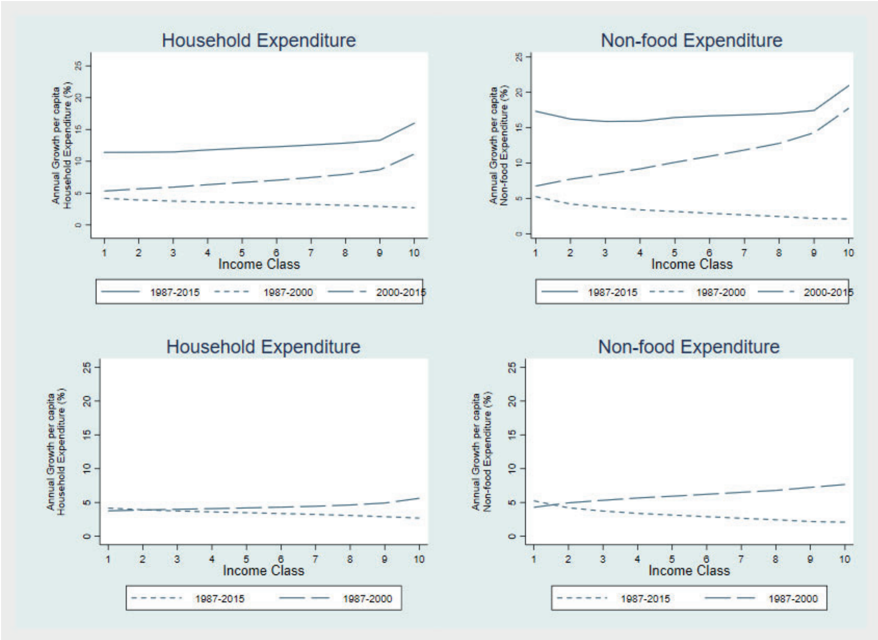
Figure 2.7 Beta Convergence in Household Expenditure at the Province Level in 1987–2015



Source: Authors' calculation from Susenas, BPS.
Note: ACH: Aceh; NSA: North Sumatera; WSA: West Sumatera; RIU: Riau; JBI: Jambi; SSA: South Sumatera; BKL: Bengkulu; LMP: Lampung; BBL: Bangka Belitung; RIS: Riau Islands; JKT: Jakarta; WJA: West Java; CJA: Central Java; YOG: Yogyakarta; EJA: East Java; BTN: Banten; BLI: Bali; WNT: West Nusa Tenggara; ENT: East Nusa Tenggara; WKL: West Kalimantan; CKL: Central Kalimantan; SKL: South Kalimantan; EKL: East Kalimantan; NSI: North Sulawesi; CSI: Central Sulawesi; SSI: South Sulawesi; TSI: Southeast Sulawesi; GOR: Gorontalo; WSI: West Sulawesi; MAL: Maluku; NML: North Maluku; WPA: West Papua; PAP: Papua.

Second, we plot the growth of household expenditure by income class for several periods. The higher income class grew faster than the lower income group during the periods 1987–2015 and 2000–2015, but the poor grew faster than the rich from 1987 to 2000 (Figure 2.8). It seems that the time break occurred in 2000. If we look further, the growth of non-food expenditure per income class is higher than its counterpart of total expenditure. This implies that the rich, instead of the middle class, gained most from the globalization in Indonesia, while Milanovic (2016), regarding global income inequality, claimed that the winners of globalization are the middle class from emerging countries, including Indonesia, and the global top 1 per cent. Our data, using household expenditures from Susenas, show that the expenditure of the Indonesian middle class grew faster than that of the other income groups only during the period from 2005 to 2015.[5]

Figure 2.8 Growth Incidence Curve for Indonesia, 1987–2015



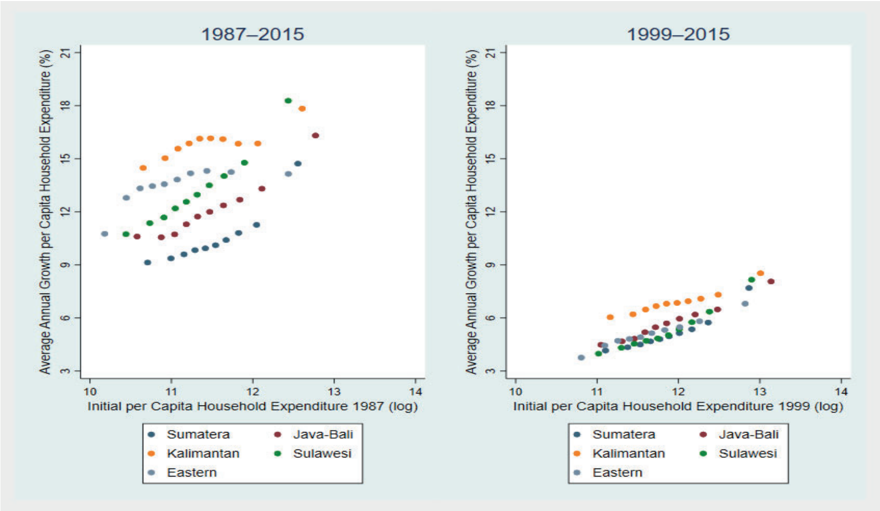
Source: Authors' calculation from Susenas, BPS.

In line with Milanovic, Leigh and Van der Eng (2009) found that the top income shares in Indonesia were relatively large over the course of the twentieth century. This finding may surprise some readers as it contradicts the common "growth with equity" understanding of Indonesia's growth experience since the 1960s. Nevertheless, our results are bolstered by evidence from other sources. For example, the top wealth shares appear to be larger in Indonesia than in many other countries, whether one uses data from wealth surveys (Davies et al. 2009) or the *Forbes* rich lists. Piketty (2014) argued that the inequality of assets and wealth ownership has driven increasing

[5] Unlike the household expenditure surveyed by Susenas, the information on wealth, tax payments and income at the individual level in Indonesia is very limited. As a consequence, inequality in wealth could not be calculated and analysed, especially at the province or district level.

within-country income and expenditure inequality around the world. This could have occurred in Indonesia. However, only a few studies have discussed this wealth inequality in Indonesia (Davies et al. 2009; Leigh and Van der Eng 2009).

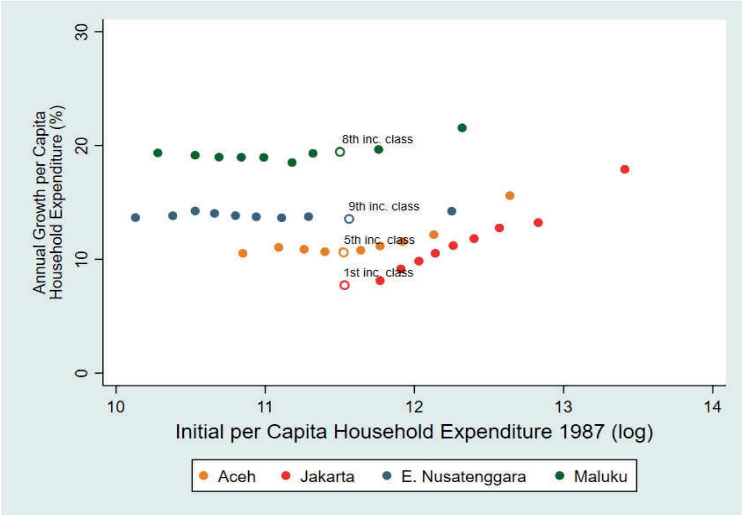
Figure 2.9 Interpersonal Inequality Dynamics at the Island Level by Income Class



Source: Authors' calculation from Susenas, BPS.

If we look at interpersonal inequality at the island level, not all member provinces grew at the same rate within islands. Similar to the province pattern, the low income class grew more slowly than the higher income group (Figure 2.9). The initial level of per capita expenditure of the lowest income class in Sumatera is equal to the initial level of the fourth income class on Eastern island in the full period from 1987 to 2015. Sulawesi is an interesting island, which started in 1987 with a low level of expenditure and has grown very fast during the past decades, converging with the rich islands. The second period of 1999 to 2015 or post crisis era shows that all the islands moved in the same direction at a lower growth, with the better-off classes growing faster than any other income classes, as depicted by their steeper slopes.

Figure 2.10 Interpersonal Inequality Dynamics in Four Provinces by Income Class, 1987–2015



Source: Authors' calculation from Susenas, BPS.

To understand further the dynamics of changing inequality at the province level, we observe how the income classes differ in each province. For each income class, we plot the annual growth of per capita household expenditure and the initial level of per capita consumption. The striking figure shows that the initial level of expenditure of the lowest income class in the richest province is much higher than the initial level of expenditure of the highest income class in the poorest province or the middle class in the particular provinces. For instance, the initial level of expenditure of the lowest income class in Jakarta, as the richest province, is more or less the same as the initial level of expenditure of the second-highest income class in East Nusatenggara, as the poorest province, or almost equal to the initial level of the middle class in Aceh (Figure 2.10). Provinces with low initial expenditure grew faster than those with a high initial level of expenditure, catching up the rich provinces; this is typical β-convergence. These patterns clearly show that people's individual experiences of growth and convergence differ vastly depending on their position on the income ladder. For each income class, the initial level of per capita household expenditure in Jakarta is the highest, and this is nominated as the richest province. Maluku province is known as a relatively poor province. However, all of its income classes grew faster than any income class in Jakarta. This triggered the question of which factors drive income class growth in particular provinces. It might be the quality of human capital, quality institutions or development advancement. We will discuss this topic further in Chapter 4.

Figure 2.11 also shows that the initial expenditure of the highest income class is higher than that of the other income classes and that it grew much faster. It implies that the higher initial household expenditure led to the higher income class growing faster and thus the gap between the low and the high income class becoming wider in each province. On the other hand, the lowest initial level of per capita expenditure of the highest income class is found in South Sulawesi (light green circle, SSI), which is equal to the initial level of per capita expenditure of the middle class in Jakarta (Panel A). A different pattern is shown in Papua (cranberry circle, PAP), where the income classes all grew at almost the same rate. The lowest growth of the highest income class in Papua is equal to the growth of the middle class in South Sumatera. People from the low income class in Papua seem unlikely to catch up the higher income class as the gap between the rich and the poor remains wide. Papua is known as a resource-based province dominated by mining, but its stage of development is indeed lagging behind the other provinces. On the other hand, the lowest initial level of per capita expenditure of the highest income class in 1987 is found in South Sulawesi; it is equal to the initial level of per capita expenditure of the middle class in Jakarta. A similar pattern is displayed for the second period of 1999 to 2015 but with lower growth of per capita household expenditure for all income classes in all provinces due to the economic crisis. This suggests that the income classes in each province kept growing during the crisis (Panel B).

Figure 2.11 Inequality Dynamics at the Province Level by Income Class

Source: Authors' calculation from Susenas, BPS.
Note: ACH: Aceh; NSA: North Sumatera; WSA: West Sumatera; RIU: Riau; JBI: Jambi; SSA: South Sumatera; BKL: Bengkulu; LMP: Lampung; BBL: Bangka Belitung; RIS: Riau Islands; JKT: Jakarta; WJA: West Java; CJA: Central Java; YOG: Yogyakarta; EJA: East Java; BTN: Banten; BLI: Bali; WNT: West Nusa Tenggara; ENT: East Nusa Tenggara; WKL: West Kalimantan; CKL: Central Kalimantan; SKL: South Kalimantan; EKL: East Kalimantan; NSI: North Sulawesi; CSI: Central Sulawesi; SSI: South Sulawesi; TSI: Southeast Sulawesi; GOR: Gorontalo; WSI: West Sulawesi; MAL: Maluku; NML: North Maluku; WPA: West Papua; PAP: Papua.

In a recent study on interprovincial income inequality, Kataoka (2018) suggested that overall technical inefficiency largely contributed to the reduction in income inequality. He also found that the convergence of inequality in resource utilization inefficiency had a greater impact on inequality convergence in overall technical inefficiency than in resource allocation inefficiency. For instance, in 2010, pure labour productivity became a substantial new factor in determining income inequality. Since Kataoka used this measurement, which is affected by per capita physical and human capital and technology, the spatial allocation imbalance of these factors has become a new concern in Indonesia.

2.7.3 Income Polarization

To extend our discussion on the dynamics of interpersonal inequality in Indonesia, we observe how income polarized across islands as well as provinces by classifying income into ten deciles; specifically, D_1 is the poorest and D_{10} is the richest. A greater degree of heterogeneity at the very top and bottom of the income distribution could explain the disparities within countries (Palma 2011). We calculate the ratio of the richest (D_{10}) to various income shares to understand the disparity within Indonesia, that is, the share of the richest to the poorest (D_{10}/D_1), the share of income of the top 10 per cent to the share of income of the bottom 40 per cent ($D_{10}/(D_4–D_1)$) to show heterogeneity in the tails, the share of the top 10 per cent to the middle ($D_{10}/(D_9–D_5)$) and upper middle ($D_{10}/(D_9–D_7)$) to show the homogeneity in the middle and various ratios of the share of the top 10 per cent to other income shares.

Table 2.4 Island Mean Values for Different Income Ratios, 1987–2015

Island	D_{10}/D_1		D_{10}/D_2		D_9/D_2		$D_{10}/(D_4–D_1)$	
	1987	2015	1987	2015	1987	2015	1987	2015
Sumatera	6.34	9.83	4.76	7.22	2.87	3.39	1.11	1.66
Kalimantan	7.05	8.58	5.39	6.33	3.12	3.28	1.26	1.45
Sulawesi	7.34	12.05	5.47	8.54	3.20	4.08	1.27	1.95
Java-Bali	8.95	13.36	6.65	9.96	3.44	4.26	1.55	2.27
Eastern	9.54	12.29	7.32	8.04	3.62	4.01	1.69	1.86

Source: Authors' calculation from Susenas, BPS.

Java-Bali is the island with the highest levels of inequality and polarization, but it moves rapidly closer to the middle of the distribution compared with other islands (Table 2.4). The greater inequality in Yogyakarta, Jakarta and Papua decreases rapidly, and it moves closer to the middle of the distribution. For instance, Jakarta's multiples of "D_{10}/D_1", "D_{10}/D_2", "D_9/D_2" and "$D_{10}/(D_4–D_1)$" in 2015 are about twice those of the other provinces (Appendix A2.10).

Of all the statistics in Table 2.5, the coefficient of variation best shows the distributional contrast between the homogeneous middles and the heterogeneous tails – the figures of the ratio of the richest to the poorest (D_{10}/D_1) is the highest while the ratio of the richest to the middle ($D_{10}/(D_9–D_5)$) and to the upper class ($D_{10}/(D_9–D_7)$) is around 0.1. This table also indicates that greater inequality rapidly decreases close to the middle of the distribution across provinces in Indonesia.

Table 2.5 Measures of Spread for Income Groups (26 Provinces), 2015

Income Ratio	Median	Mean	Standard Deviation	Coefficient of Variation
D_{10}/D_1	9.44	10.23	2.09	0.20
$D_{10}/(D_4–D_1)$	1.67	1.75	0.30	0.17
$D_{10}/(D_9–D_5)$	0.62	0.62	0.06	0.10
$D_{10}/(D_9–D_7)$	0.88	0.88	0.08	0.09
D_{10}/D_9	2.10	2.09	0.15	0.07
D_{10}/D_2	7.25	7.65	1.42	0.19
D_9/D_2	3.46	3.65	0.58	0.16

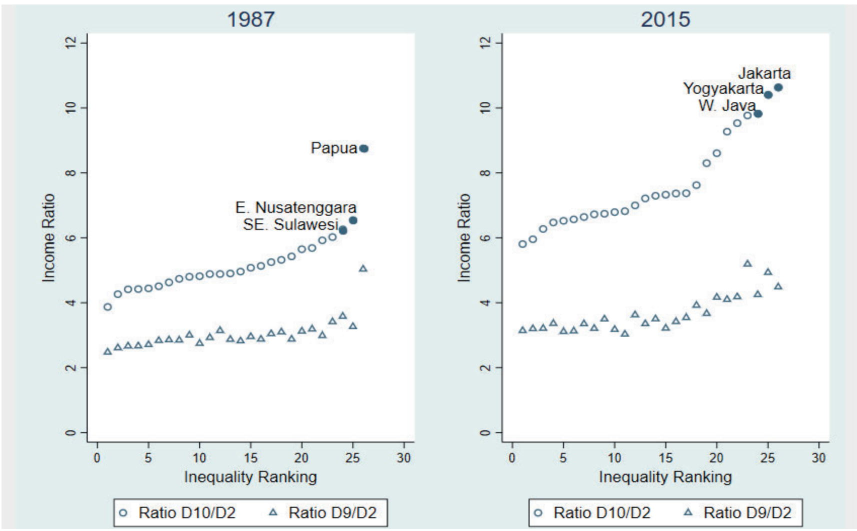
Source: Authors' calculation from Susenas, BPS.

If we look at the income ratios across provinces, as presented in Appendix A2.11, Jakarta is in the top three of all the income ratios. This suggests that Jakarta has the highest inequality measured by any ratio. The second province that lies in the top three income ratios is West Java, the province with the second-highest inequality in Indonesia. Other provinces, such as Yogyakarta, Papua, South Sulawesi and Lampung, are also members of the top three in inequality for particular income ratios. For instance, Yogyakarta has a high ratio of the rich to the poor, suggesting that the inequality in Yogyakarta is led by the growing number of the poor rather than the rich.

The most interesting pattern is displayed by Lampung and South Sulawesi. Lampung occupies the third position of the high income ratio measured by the ratio of the richest to the second-highest income (D_{10}/D_9). This ratio suggests that the cumulative wealth is concentrated in the highest income class and thus inequality is worse in Lampung together with Jakarta and West Java. On the other hand, South Sulawesi lies in the third position of the middle and upper-middle income ratios. South Sulawesi, with Makassar port as a hub for eastern Indonesia and the main port for export-import activities, benefited from the commodity boom era, when many commodity-exporting provinces were located on Sulawesi island. This increase in the middle and upper-middle classes led to higher inequality, as measured by $D_{10}/(D_9–D_5)$ and $D_{10}/(D_9–D_7)$. These ratios also exhibit the most homogeneous provinces, namely Jambi, Central Kalimantan and West Sumatera. These three provinces are known as agricultural and non-resource-based regions.

Moreover, Figure 2.12 shows the uniqueness of the income polarization in Indonesia. The top three provinces in 1987 (Papua, East Nusatenggara and South East Sulawesi) are from different islands, while their counterparts in 2015 are from Java island (West Java, Yogyakarta and Jakarta). This indicates that inequality across provinces varies by islands. The ranges for the rankings are also different. For instance, D_{10}/D_2 ranges from 5.8 to 10.6 while D_9/D_2 only ranges from 3.1 to 4.5 in 2015.

Figure 2.12 Inequality Ranking in 26 Provinces, 1987 and 2015



Source: Authors' calculation from Susenas, BPS.
Note: We use the Ratio D_{10}/D_2 as a base to rank provinces in 1987 and 2015.

These patterns show that Indonesia is following the global trend of rising inequality in which the highest and lowest income classes are the most volatile while the middle income class is homogeneous, as explained by Palma (2011). He introduced two terms regarding this global phenomenon: “centrifugal” forces, which increase the diversity in the shares of the top 10 and the bottom 40 per cent, and “centripetal” forces, which lead to growing uniformity in the income share appropriated by deciles 5 to 9. The greatest volatility in the income distribution lies at the extreme ends (rich and poor), with markedly higher stability in the middle and upper-middle classes. This pattern is known as the homogeneous middle, as shown in Appendix A2.12. A similar pattern is also apparent in per capita non-food expenditure. Thus, the within-Indonesia disparity of inequality basically relates to the distributional fact of the income share of the rich. Palma (2011) suggested that political institutional factors and the nature of the political settlements in the real world are likely to have a far greater influence on the determination of income distribution than purely economic factors.

2.8 Conclusion

This chapter studied the income inequality dynamics in Indonesia since 1960, with a focus on the period 1987–2015. It used different levels of spatial aggregation: (i) the national level (from an international perspective); (ii) the regional level (islands and provinces); and (iii) the individual level (income classes). Our analysis at the cross-country level and at the province level covered the period 1961–2015, while our analysis for income classes was based on micro data (from Susenas) for the period 1987–2015. This approach allowed us to present new insights into the inequality dynamics in Indonesia.

The main findings of this chapter are as follows. Indonesia shows a rising trend in inequality measured by the Gini index since 2001 but at a much lower level than other countries for many years prior to 2001 and still much lower than leading countries in the ASEAN and emerging and African countries. However, this finding should be interpreted with care because comparisons across countries face measurement problems due to data availability, such as income versus expenditure as discussed by Leigh and Van der Eng (2009) and Van Zanden and Marks (2012).

Instead, our analysis at the regional and individual level mainly used expenditure data from the Susenas household survey. We found increasing inequality at all the aggregated levels, that is, national, island and province, suggesting the existence of increasing spatial disparities along a geographical dimension. Our findings confirm the results of Akita and Lukman (1995) and Yusuf et al. (2014). In contrast, the variation in inequality across islands and provinces has decreased persistently over time, suggesting that the differences in inequality between islands as well as between provinces are becoming smaller.

From the individual perspective, we found that the rapidly increasing interpersonal inequality in Indonesia has mainly been driven by strong income growth at the top of the income distribution – and this is true for all the provinces and islands during the observed period from 1987 to 2015. We also discovered that the highest income classes in the relatively poor provinces experienced substantially higher income growth than the lowest income classes in the relatively rich provinces. This finding is in line with the study conducted by Yusuf et al. (2014), which covered a shorter period, 1993 to 2013, and regional dimensions including urban–rural areas and Java–non Java regions. Our findings have two implications. First, we can confirm that interpersonal inequality kept increasing regardless of the data and methodology chosen, as shown by our calculation based on expenditure from the household consumption dataset versus top income from the taxation dataset as calculated by Davies et al. (2009) and Leigh and Van der Eng (2009). Agustina et al. (2012) also found that fuel subsidies benefit the rich disproportionally more than the poor. Second, the middle class is not the winner, as concluded by Milanovic (2016), who argued that especially the middle class has benefited from globalization. Our calculation shows that the middle class grew only relatively fast during the period 2005–2015. The rich grew quickly over the entire period, while the poor were persistently left behind in terms of income growth. Hence, our results are consistent with Leigh and Van der Eng (2009), who also concluded that the top income shares in Indonesia have been relatively high for a long time, and Van der Weide and Milanovic (2014) who concluded that inequality is bad for growth of the poor but not for that of the rich.

Together, these trends lead to the conclusion that *on average* the relatively poor provinces are catching up with the relatively rich provinces while at the same time interpersonal inequality in Indonesia is rising at all levels of spatial aggregation – with individual income growth being strongly positively correlated with the individual income level.

Income polarization exists uniquely in Indonesia. We found that the top three inequality rankings in 1987 are for Papua, East Nusatenggara and South East Sulawesi, and those provinces are from different islands. Meanwhile, the top three inequality rankings in 2015 are all from Java islands (West Java, Yogyakarta and Jakarta). This indicates that inequality across provinces varies by island as income is highly polarized within islands.

These findings suggest that the increase in inequality may come at the expense of the accessibility of certain public services, which we will discuss in Chapter 5 and Chapter 6. Those chapters will delve deeper into the accessibility of public services. After this descriptive study on inequality dynamics, in the next two chapters, we will examine the urban and governance drivers of inequality dynamics.

Appendices

Appendix A2.1 Constructing True Panel Districts

We make intensive use of the dataset from BPS, that is, Podes, the Susenas consumption module and the Susenas core module. The first two datasets are collected every three years, while the latter is collected on a yearly basis. Since 2011, the Susenas consumption module has also been collected every year. Our complete datasets are available from 1999 onwards. Hence, in total, we have six Podes datasets (1999, 2002, 2005, 2008, 2011 and 2014); eight Susenas consumption module datasets (1999, 2002, 2005, 2008, 2011, 2012, 2013 and 2014); and sixteen Susenas core module datasets from 1999 to 2014.

As a consequence of decentralization, since 2001, many new districts and provinces have been established. As of 2014, the total number of new and old districts is 511. If we only count the number of old districts by recoding new districts back to their parents in 2000, the total number of districts becomes 341. However, not all districts are surveyed in Podes, the Susenas consumption module and the Susenas core module. Thus, we should select districts that are consistently recorded in those 3 main datasets to construct a true panel, and the total number of selected districts is 242. In terms of the size of the population, these 242 districts cover about 80 per cent of the total population in Indonesia.

Our main challenge in constructing the true district panel is the coding system provided by BPS, that is, not only creating an old code for old districts and a new code for new districts but also applying the old code to new districts and the new code to old districts. Therefore, to obtain the panel of 242 districts, we undertake the following procedures. First, we recode new districts back to their parent districts using the identification code base year of 2000 issued by BPS, so the recoding process starts in 2001. New districts that were established prior to 2000 are considered as parent districts. We find some inconsistencies in the coding system. Each dataset has its own code reference. For instance, Podes 1999 has two IDs, namely ID 1998 and ID 1999; Susenas Core 2000 refers to ID 1999 but Susenas Consumption 2000 refers to ID 1998; and Podes 2000 uses the ID reference of 1998, while Podes 2005 refers to IDs

2002 and 2003. Moreover, the names of districts are not written in the same style. If this becomes too confusing, we check the official name written on the decree of district establishment at the Ministry of Home Affairs.

Second, we carefully check the new codes due to the splitting up of issues. BPS has changed provinces' ID, especially Papua, several times. Papua's original province name was Irian Jaya, and its ID was 82 from 1993 until 1998. Irian Jaya province then split into two provinces: Papua and West Papua. The ID for both new provinces varies according to the year's reference (Table A2.1.1). Facing these typical inconsistencies, we have to recode districts manually by matching the names of districts as well as the reference ID. Below is a summary of the process of changing IDs for new and parent provinces.

Table A2.1.1 New Code and Its Parent Provinces

Name of Provinces	Province Code	Year of Reference
Riau Islands	14	1993–2001
	20	2002–2005
	21	2006–2014
Bangka Belitung	16	1993–1998
	19	1999–2014
Banten	32	1993–1998
	36	1999–2014
East Kalimantan	64	1993–2009
	65	2010–2014
Gorontalo	71	1993–1998
	75	1999–2014
Maluku	81	1993–2014
	81	1993–1998
North Maluku	82	1999–2014
Irian Jaya	82	1993–1998
Papua	82	1993–1998
	92	1999–2000*
	93	1999–2000**
	94	2001–2014
West Papua	82	1993–1998
	91	1999–2000
	94	2001–2002
	91	2003–2014

Note: * Podes, ** Susenas

Third, we recode districts from step 1 to match the ID in each dataset. Podes and Susenas have different ID references for particular years. For instance, Podes 2011 refers to IDs 2008 and 2009 while Susenas 2011 refers to ID 2009. After finishing this step, we have a number of districts that were always surveyed in each dataset, specifically 293 districts in Podes, 268 districts in the Susenas consumption module and 253 districts in the Susenas core module.

Last, we select districts that are consistently available in all 3 datasets. Some provinces, like Aceh, Papua and Maluku, are excluded from Susenas 2000, 2001, 2002 and 2005 due to conflict areas and tsunami. Hence, we finally have 242 districts as the full panel from 1999 to 2014 (Table A2.1.2). Moreover, there are 165 districts that have not been split up since 1993, but our data are only available from 1999 onwards. We perform this recoding process for all the datasets.

Table A2.1.2 List of 242 Panel Districts

Name of Districts	District Code	Name of Districts	District Code
Kab. Tapanuli Selatan	1203	Kab. Lampung Barat	1801
Kab. Tapanuli Tengah	1204	Kab. Lampung Selatan	1803
Kab. Tapanuli Utara	1205	Kab. Lampung Tengah	1805
Kab. Labuhan Batu	1207	Kab. Lampung Utara	1806
Kab. Asahan	1208	Kota Bandar Lampung	1871
Kab. Simalungun	1209	Kab. Bangka	1901
Kab. Dairi	1210	Kab. Belitung	1902
Kab. Karo	1211	Kota Pangkal Pinang	1971
Kab. Deli Serdang	1212	Kota Jakarta Selatan	3171
Kab. Langkat	1213	Kota Jakarta Timur	3172
Kota Sibolga	1271	Kota Jakarta Pusat	3173
Kota Tanjung Balai	1272	Kota Jakarta Barat	3174
Kota Pematang Siantar	1273	Kota Jakarta Utara	3175
Kota Tebing Tinggi	1274	Kab. Bogor	3201
Kota Medan	1275	Kab. Sukabumi	3202
Kota Binjai	1276	Kab. Cianjur	3203
Kab. Sawahlunto/Sijunjung	1304	Kab. Bandung	3204
Kab. Padang Pariaman	1306	Kab. Garut	3205
Kota Padang	1371	Kab. Tasikmalaya	3206
Kota Padang Panjang	1374	Kab. Ciamis	3207
Kota Bukittinggi	1375	Kab. Kuningan	3208
Kota Payakumbuh	1376	Kab. Cirebon	3209
Kab. Indragiri Hulu	1402	Kab. Majalengka	3210
Kab. Indragiri Hilir	1403	Kab. Sumedang	3211
Kab. Kampar	1406	Kab. Indramayu	3212
Kab. Bengkalis	1408	Kab. Subang	3213
Kab. Kepulauan Riau	1410	Kab. Purwakarta	3214
Kota Pekan Baru	1471	Kab. Karawang	3215
Kota Batam	1472	Kota Bogor	3271
Kab. Kerinci	1501	Kota Sukabumi	3272
Kab. Sarolangun	1503	Kota Bandung	3273
Kab. Tanjung Jabung Barat	1507	Kota Cirebon	3274
Kab. Tebo	1508	Kota Bekasi	3275
Kota Jambi	1571	Kab. Cilacap	3301
Kab. Ogan Komering Ulu	1601	Kab. Banyumas	3302
Kab. Ogan Komering Ilir	1602	Kab. Purbalingga	3303

Continued

Kab. Muara Enim	1603	Kab. Banjarnegara	3304
Kab. Lahat	1604	Kab. Kebumen	3305
Kab. Musi Rawas	1605	Kab. Purworejo	3306
Kab. Musi Banyuasin	1606	Kab. Wonosobo	3307
Kota Palembang	1671	Kab. Magelang	3308
Kab. Bengkulu Selatan	1701	Kab. Boyolali	3309
Kab. Rejang Lebong	1702	Kab. Klaten	3310
Kab. Bengkulu Utara	1703	Kab. Jombang	3517
Kota Bengkulu	1771	Kab. Nganjuk	3518
Kab. Sukoharjo	3311	Kab. Madiun	3519
Kab. Wonogiri	3312	Kab. Magetan	3520
Kab. Karanganyar	3313	Kab. Ngawi	3521
Kab. Sragen	3314	Kab. Bojonegoro	3522
Kab. Grobogan	3315	Kab. Sampang	3527
Kab. Blora	3316	Kab. Pamekasan	3528
Kab. Rembang	3317	Kab. Sumenep	3529
Kab. Pati	3318	Kota Kediri	3571
Kab. Kudus	3319	Kota Blitar	3572
Kab. Jepara	3320	Kota Malang	3573
Kab. Demak	3321	Kota Probolinggo	3574
Kab. Semarang	3322	Kota Pasuruan	3575
Kab. Temanggung	3323	Kota Mojokerto	3576
Kab. Kendal	3324	Kota Madiun	3577
Kab. Batang	3325	Kota Surabaya	3578
Kab. Pekalongan	3326	Kab. Pandeglang	3601
Kab. Pemalang	3327	Kab. Lebak	3602
Kab. Tegal	3328	Kota Tangerang	3671
Kab. Brebes	3329	Kab. Jembrana	5101
Kota Magelang	3371	Kab. Tabanan	5102
Kota Surakarta	3372	Kab. Badung	5103
Kota Salatiga	3373	Kab. Gianyar	5104
Kota Semarang	3374	Kab. Klungkung	5105
Kota Pekalongan	3375	Kab. Bangli	5106
Kota Tegal	3376	Kab. Karang Asem	5107
Kab. Kulon Progo	3401	Kab. Buleleng	5108
Kab. Gunung Kidul	3403	Kota Denpasar	5171
Kab. Sleman	3404	Kab. Lombok Barat	5201
Kota Yogyakarta	3471	Kab. Lombok Tengah	5202
Kab. Pacitan	3501	Kab. Lombok Timur	5203
Kab. Ponorogo	3502	Kab. Sumbawa	5204
Kab. Trenggalek	3503	Kab. Dompu	5205
Kab. Lumajang	3508	Kab. Bima	5206
Kab. Jember	3509	Kota Mataram	5271
Kab. Banyuwangi	3510	Kab. Sumba Barat	5301
Kab. Bondowoso	3511	Kab. Sumba Timur	5302
Kab. Situbondo	3512	Kab. Kupang	5303
Kab. Probolinggo	3513	Kab. Timor Tengah Selatan	5304
Kab. Pasuruan	3514	Kab. Timor Tengah Utara	5305
Kab. Sidoarjo	3515	Kab. Belu	5306
Kab. Mojokerto	3516	Kab. Alor	5307

Source: BPS, authors' code.

Appendix A2.2 Gini Index and per Capita GDP (PPP Constant 2011 USD): Indonesia and Selected Countries, 1990–2013

Country	1990		2000		2010		2013		Changes 1990–2013 (%)	
	GDP/ capita	Gini Index	GDP/ capita	Gini Index	GDP/ capita	Gini Index	GDP/ capita	Gini Index	GDP/ capita	Gini Index
ASEAN										
Indonesia	4,295	32.0	5,552	32.9	8,027	38.0	9,254	41.3	115	29
Malaysia	10,155	44.6	15,688	44.9	20,390	43.8	22,556	42.8	122	–4
Philippines	4,010	42.8	4,243	46.1	5,614	42.6	6,325	41.6	58	–3
Thailand	6,369	45.2	8,939	43.2	12,822	39.4	13,932	37.5	119	–17
EMERGING										
Brazil	9,997	60.4	11,015	58.6	14,043	53.1	14,555	51.3	46	–15
China	1,554	34.6	3,609	45.8	9,230	48.1	11,525	47.3	642	37
India	1,812	29.7	2,600	32.7	4,638	33.9	5,238	34.5	189	16
South Korea	12,087	29.5	20,757	31.7	30,440	31.0	32,708	30.5	171	3
AFRICAN										
Kenya	2,009	57.5*	1,812	49.0	2,080	45.1	2,193	47.7	9	–17
Nigeria	3,050	42.9	2,855	49.6	5,148	48.8	5,676	51.9	86	21
South Africa	9,935	67.0	9,519	68.0	11,651	69.5	12,106	68.2	22	2
RICH										
Australia	28,546	27.9	35,247	31.3	41,328	31.0	42,810	31.5	50	13
Japan	29,550	23.9	32,186	24.6	34,571	25.5	35,481	25.6	20	7
Singapore	34,202	43.6	51,491	41.4	71,816	42.7	76,237	41.2	123	–6
USA	36,982	34.9	45,956	35.6	49,307	38.0	51,451	39.2	39	12

Note: * The GINI index for Kenya pertains to 1992.
Source: WDI, UNDP, OECD, CBS from respective countries.

Appendix A2.3 Regression Results for Kuznets Countries

log income	31.66***
	(10.69)
log income²	–3.056**
	(1.37)
log income³	0.0899
	(0.06)
year	0.222
	(1.41)
decentralization	–2.886
	(4.77)
democratization	–1.289
	(4.02)
commodity boom	–4.051
	(6.26)
period 1970s	–8.190
	(26.85)
period 1980s	–2.893
	(14.15)
period 1990s	0.164
	(14.14)
period 2000s	–0.258
	(1.63)
period 2010s	–0.0255
	(4.92)
<u>Interaction Income and Event Dummies</u>	
income and decentralization	0.281
	(0.23)
income and democratization	0.0917
	(0.31)
income and commodity boom	0.281
	(0.28)
<u>Interaction Event Dummies and Country Dummies</u>	
<u>Latin</u>	
decentralization and Latin	0.519
	(1.39)
democratization and Latin	–1.335
	(1.15)
commodity boom and Latin	–0.404
	(1.12)
<u>East Asia</u>	
decentralization and East Asia	–0.753
	(1.76)
democratization and East Asia	0.378
	(1.43)
commodity boom and East Asia	0.696
	(1.36)
<u>Middle East</u>	
decentralization and Middle East	1.463
	(3.24)
democratization and Middle East	–1.510
	(2.28)
commodity boom and Middle East	0.867
	(2.18)

	Continued
Interaction Period Dummies and Country Dummies	
Latin	
period 1970s and Latin	3.114***
	(1.05)
period 1980s and Latin	2.813***
	(0.91)
period 1990s and Latin	–0.785
	(0.78)
period 2000s and Latin	1.203
	(1.28)
period 2010s and Latin	–3.576**
	(1.46)
East Asia	
period 1970s and East Asia	1.206
	(0.94)
period 1980s and East Asia	–1.076
	(0.82)
period 1990s and East Asia	–0.927
	(0.77)
period 2000s and East Asia	–0.0160
	(1.60)
period 2010s and East Asia	–1.558
	(1.16)
Middle East	
period 1970s and Middle East	3.325
	(2.18)
period 1980s and Middle East	–2.050
	(1.99)
period 1990s and Middle East	–0.135
	(1.65)
period 2000s and Middle East	–1.404
	(2.66)
period 2010s and Middle East	–0.356
	(2.13)
Constant	–495.5
	(2,766)
Number of Observations	2,084
R-squared	0.163
Number of countries	95

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Appendix A2.4 Inequality within Islands:
Per Capita Food and Non-food Expenditure, 1987–2013

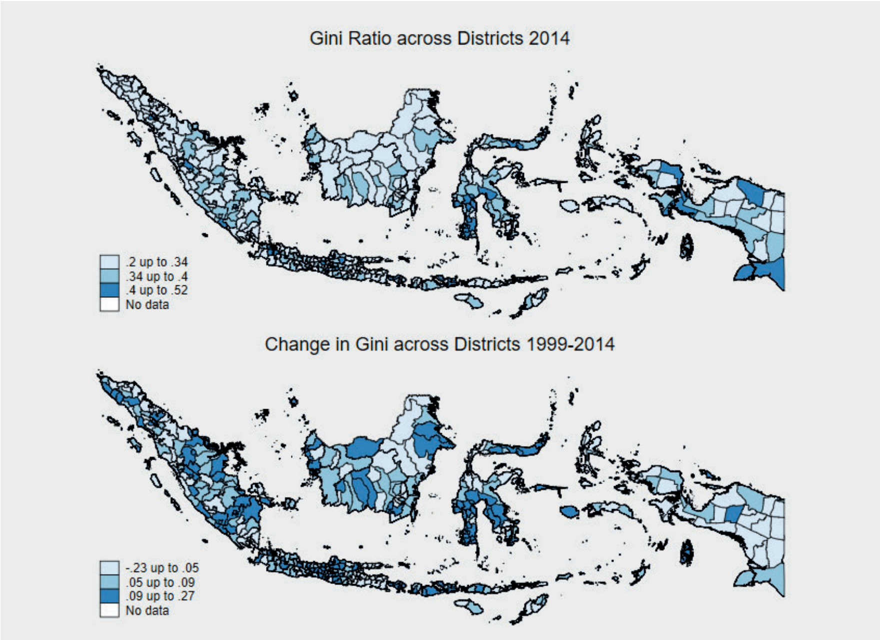
	1987	1999	2003	2013	Changes (%)	
					1987–2013	2003–2013
Per Capita Food Expenditure						
Theil Index						
Sumatera	9.82	10.46	11.55	10.48	6.7	–9.3
Java-Bali	11.08	12.24	16.46	12.62	13.9	–23.3
Kalimantan	11.23	10.97	12.84	11.62	3.4	–9.5
Sulawesi	11.17	9.81	11.01	14.33	28.2	30.2
Eastern	14.47	17.56	13.83	13.56	–6.3	–2.0
Average	11.55	12.21	13.14	12.52	8.4	–4.7
Ratio P_{90}/P_{10}						
Sumatera	2.84	2.92	2.91	2.98	4.8	2.3
Java-Bali	3.03	3.13	3.25	3.32	9.7	2.2
Kalimantan	3.11	3.00	3.14	3.13	0.7	–0.3
Sulawesi	3.10	2.88	3.01	3.53	13.7	17.0
Eastern	3.56	4.41	3.48	3.38	–5.0	–2.8
Average	3.13	3.27	3.16	3.27	4.4	3.4
Decile Dispersion Ratio						
Sumatera	1.61	1.63	1.66	1.62	0.3	–2.3
Java-Bali	1.63	1.65	1.67	1.66	1.8	–0.8
Kalimantan	1.67	1.69	1.72	1.66	–0.5	–3.1
Sulawesi	1.71	1.60	1.71	1.68	–1.9	–1.9
Eastern	1.74	2.01	1.80	1.70	–2.2	–5.3
Average	1.67	1.72	1.71	1.66	–0.5	–2.7
Per Capita Non-food Expenditure						
Theil Index						
Sumatera	37.52	40.92	38.19	52.66	40.3	37.9
Java-Bali	40.39	56.43	36.59	58.01	43.6	58.5
Kalimantan	39.94	39.51	40.37	47.77	19.6	18.3
Sulawesi	36.97	37.45	43.53	64.47	74.4	48.1
Eastern	50.31	41.90	32.08	52.45	4.3	63.5
Average	41.03	43.24	38.15	55.07	34.2	44.3
Ratio P_{90}/P_{10}						
Sumatera	6.65	6.79	5.96	6.92	4.1	16.1
Java-Bali	6.47	7.39	6.57	8.50	31.4	29.2
Kalimantan	7.73	7.57	6.48	7.34	–5.0	13.4
Sulawesi	7.37	7.28	6.97	9.32	26.5	33.6
Eastern	8.21	7.26	5.80	8.43	2.7	45.3
Average	7.28	7.26	6.36	8.10	11.2	27.4

Continued

<u>*Decile Dispersion Ratio*</u>						
Sumatera	2.27	2.34	2.19	2.22	–2.5	1.3
Java-Bali	2.17	2.29	2.19	2.29	5.3	4.7
Kalimantan	2.41	2.52	2.28	2.53	5.0	11.3
Sulawesi	2.39	2.30	2.38	2.46	3.0	3.1
Eastern	2.36	2.33	2.14	2.35	–0.4	9.7
Average	2.32	2.36	2.24	2.37	2.1	6.0

Source: Susenas, authors' calculation.

Appendix A2.5 Gini Ratio 2014 and Its Change 1999–2014 across Districts



Source: Authors' calculation from Susenas, BPS.

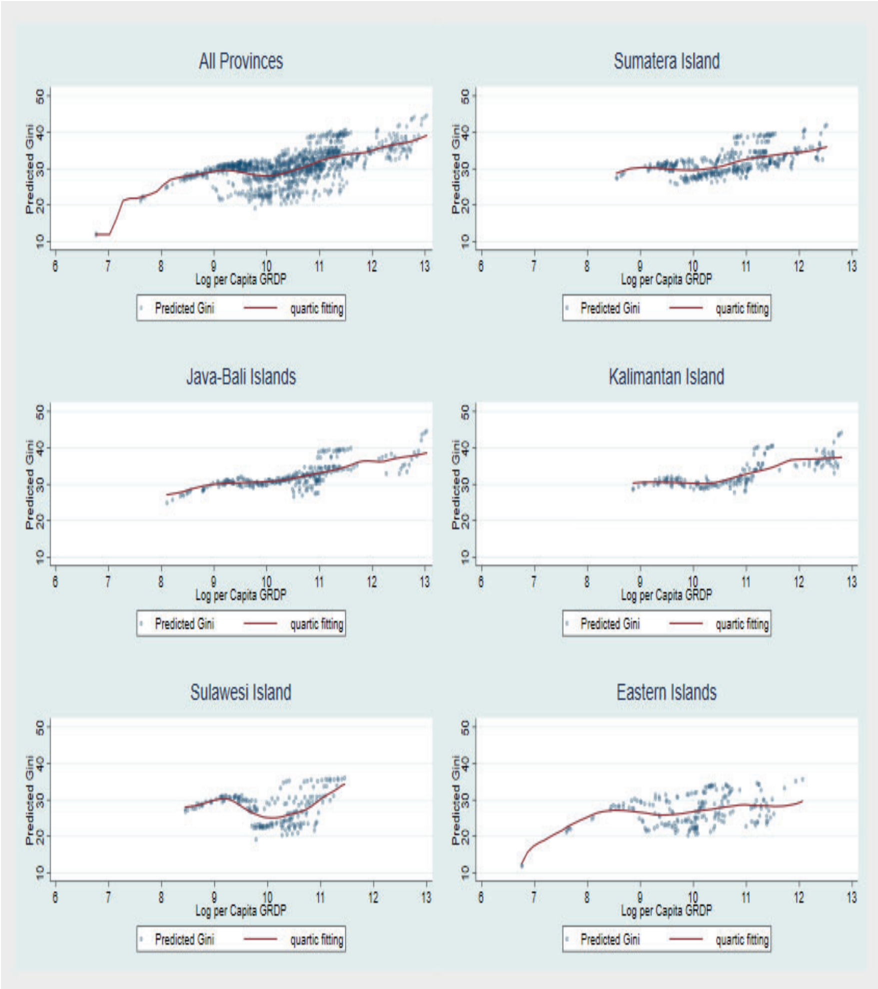
Appendix A2.6 Regression Results for Kuznets Provinces

log income	97.40**
	(42.06)
log income^2	–9.106**
	(3.92)
log income^3	0.287**
	(0.12)
year	0.222
	(0.59)
bank reform1	–24.94***
	(7.54)
economic reform	–0.823
	(7.59)
bank reform2	–7.521
	(6.64)
Asian crisis	5.218
	(7.56)
decentralization	10.58
	(7.52)
democratization	26.53***
	(7.81)
commodity boom	–11.08
	(8.12)
<u>InteractionIncome and Event Dummies</u>	
income and bank reform1	1.791***
	(0.55)
income and economic reform	–0.376
	(0.54)
income and bank reform2	0.439
	(0.61)
income and asian crisis	–0.796
	(0.70)
income and decemtralization	–0.947
	(0.69)
income and democratization	–2.244***
	(0.73)
income and commodity boom	1.622**
	(0.64)
<u>Interaction Event Dummies and Island Dummies</u>	
<u>Sumatera Island</u>	
bank reform1 and Sumatera	–0.963
	(1.21)
economic reform and Sumatera	3.951***
	(1.13)
bank reform2 and Sumatera	1.390
	(1.03)
asian crisis and Sumatera	0.571
	(1.26)
decentralization and Sumatera	1.476
	(1.27)
democratization and Sumatera	2.503*
	(1.37)
commodity boom and Sumatera	–4.239***
	(1.16)
<u>Java-Bali Island</u>	
bank reform1 and Java-Bali	2.214*
	(1.22)

	Continued
economic reform and Java-Bali	3.436***
	(1.17)
bank reform2 and Java-Bali	0.342
	(1.09)
Asian crisis and Java-Bali	–1.080
	(1.34)
decentralization and Java-Bali	2.611*
	(1.33)
democratization and Java-Bali	1.046
	(1.43)
commodity boom and Java-Bali	–3.670***
	(1.23)
Kalimantan Island	
bank reform1 and Kalimantan	2.236
	(1.43)
economic reform and Kalimantan	1.792
	(1.34)
bank reform2 and Kalimantan	1.040
	(1.25)
Asian crisis and Kalimantan	0.450
	(1.56)
decentralization and Kalimantan	2.270
	(1.55)
democratization and Kalimantan	0.383
	(1.65)
commodity boom and Kalimantan	–2.480*
	(1.43)
Sulawesi Island	
bank reform1 and Sulawesi	0.106
	(1.24)
economic reform and Sulawesi	–1.713
	(1.17)
bank reform2 and Sulawesi	2.772**
	(1.14)
Asian crisis and Sulawesi	–2.006
	(1.39)
decentralization and Sulawesi	1.201
	(1.35)
democratization and Sulawesi	1.932
	(1.42)
commodity boom and Sulawesi	–1.072
	(1.22)
Constant	–756.2
	(1,186)
Number of Observations	984
R-squared	0.698
Number of provinces	33

Notes: standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

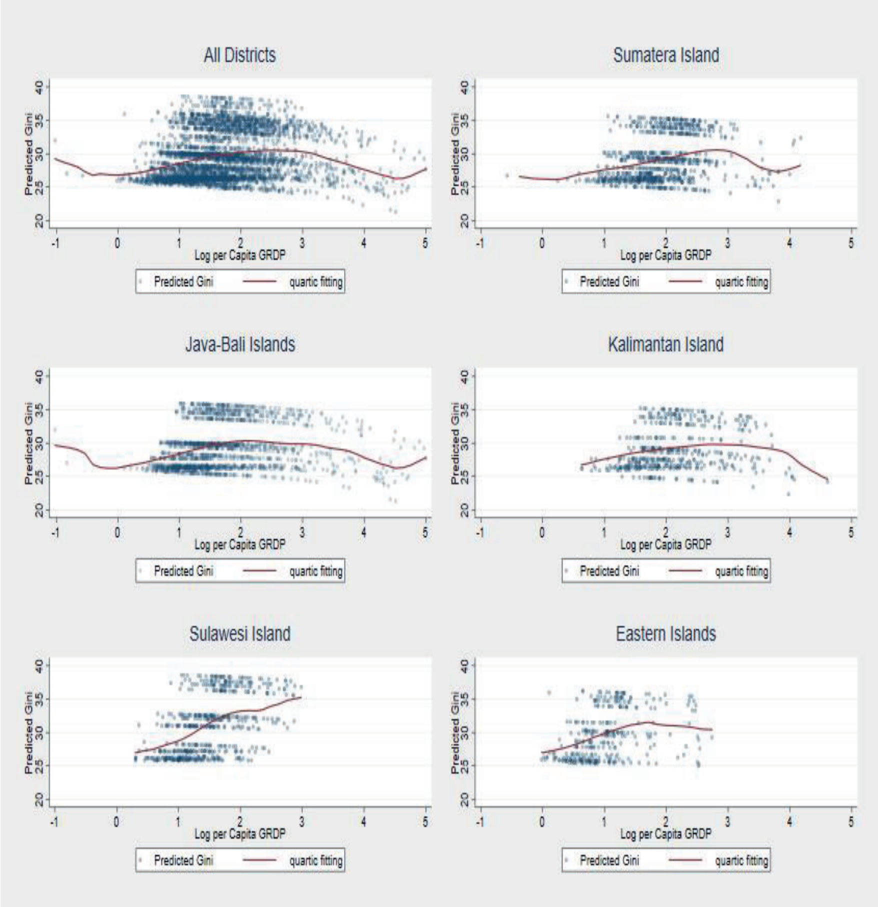
Appendix A2.7 Kuznets Estimation at the Province Level



Source: Authors' estimate.

Note: Java-Bali islands (Jakarta, Banten, West Java, Central Java, East Java, Yogyakarta, Bali); Sumatera (Aceh, North Sumatera, South Sumatera, West Sumatera, Riau, Riau islands, Bangka Belitung, Jambi, Bengkulu, Lampung); Kalimantan (West Kalimantan, East Kalimantan, Central Kalimantan, South Kalimantan); Sulawesi (North Sulawesi, South Sulawesi, South East Sulawesi, West Sulawesi, Central Sulawesi, Gorontalo), Eastern Islands (East Nusa Tenggara, West Nusa Tenggara, Maluku, North Maluku, Papua, West Papua).

Appendix A2.8 Kuznets Estimation at the District Level



Source: Authors' estimate.
Note: list of districts per island can be found in Appendix A2.1.

Appendix A2.9 Beta Convergence of Household Expenditure at Province Level 1999–2015

Source: Author's calculation from Susenas, BPS.
Note: ACH: Aceh; NSA: North Sumatera; WSA: West Sumatera; RIU: Riau; JBI: Jambi; SSA: South Sumatera; BKL: Bengkulu; LMP: Lampung; BBL: Bangka Belitung; RIS: Riau Islands; JKT: Jakarta; WJA: West Java; CJA: Central Java; YOG: Yogyakarta; EJA: East Java; BTN: Banten; BLI: Bali; WNT: West Nusa Tenggara; ENT: East Nusa Tenggara; WKL: West Kalimantan; CKL: Central Kalimantan; SKL: South Kalimantan; EKL: East Kalimantan; NSI: North Sulawesi; CSI: Central Sulawesi; SSI: South Sulawesi; TSI: Southeast Sulawesi; GOR: Gorontalo; WSI: West Sulawesi; MAL: Maluku; NML: North Maluku; WPA: West Papua; PAP: Papua.

Appendix A2.10 Province Mean Values for Different Income Ratio, 1987–2015

PROV	D_{10}/D_1		D_{10}/D_2		D_9/D_2		$D_{10}/(D_4–D_1)$	
	1987	2015	1987	2015	1987	2015	1987	2015
ACH	6.01	8.66	4.74	6.52	2.85	3.11	1.09	1.53
NSA	6.75	8.72	4.88	6.57	2.93	3.13	1.15	1.51
WSA	5.89	8.79	4.51	6.79	2.84	3.18	1.06	1.59
RIU	5.67	9.37	4.41	7.21	2.67	3.35	1.04	1.65
JBI	4.75	8.95	3.87	6.82	2.48	3.03	0.91	1.57
SSA	6.23	8.57	4.42	6.47	2.67	3.36	1.05	1.49
BKL	5.30	9.67	4.26	7.37	2.61	3.54	1.01	1.71
LMP	6.34	9.42	4.96	7.33	2.82	3.22	1.16	1.71
JKT	6.55	14.11	5.13	10.63	2.88	4.48	1.22	2.33
WJA	6.36	13.15	5.08	9.83	2.96	4.25	1.19	2.22
CJA	6.27	11.06	4.82	8.30	2.75	3.67	1.13	1.92
YOG	6.67	14.23	5.25	10.41	3.05	4.93	1.23	2.35
EJA	7.81	12.27	5.92	9.27	2.99	4.10	1.39	2.14
BLI	7.59	10.23	5.65	7.62	3.12	3.92	1.33	1.72
WNT	7.03	9.69	5.43	7.29	2.88	3.50	1.26	1.69
ENT	8.38	8.70	6.55	6.72	3.26	3.21	1.53	1.56
WKL	5.70	7.92	4.44	5.95	2.71	3.20	1.05	1.38
CKL	5.88	8.60	4.63	6.27	2.86	3.21	1.08	1.42
SKL	6.07	8.74	4.80	6.64	3.00	3.35	1.12	1.52
EKL	7.75	7.97	5.69	5.81	3.19	3.14	1.33	1.34
NSI	7.31	9.46	5.32	7.00	3.10	3.63	1.25	1.59
CSI	6.46	9.71	4.88	7.36	3.14	3.42	1.12	1.71
SSI	6.44	13.09	4.90	9.53	2.87	4.18	1.15	2.18
TSI	8.18	11.93	6.26	8.61	3.58	4.17	1.42	1.95
MAL	7.68	8.52	6.02	6.74	3.41	3.50	1.39	1.56
PAP	12.74	14.59	8.76	9.77	5.03	5.19	1.96	2.10

Source: Authors' calculation from Susenas, BPS.
Note: ACH: Aceh; NSA: North Sumatera; WSA: West Sumatera; RIU: Riau; JBI: Jambi; SSA: South Sumatera; BKL: Bengkulu; LMP: Lampung; BBL: Bangka Belitung; RIS: Riau Islands; JKT: Jakarta; WJA: West Java; CJA: Central Java; YOG: Yogyakarta; EJA: East Java; BTN: Banten; BLI: Bali; WNT: West Nusa Tenggara; ENT: East Nusa Tenggara; WKL: West Kalimantan; CKL: Central Kalimantan; SKL: South Kalimantan; EKL: East Kalimantan; NSI: North Sulawesi; CSI: Central Sulawesi; SSI: South Sulawesi; TSI: Southeast Sulawesi; GOR: Gorontalo; WSI: West Sulawesi; MAL: Maluku; NML: North Maluku; WPA: West Papua; PAP: Papua.

Appendix A2.11 Different Income Ratios by Province, 2015

PROV	D_{10}/D_1	D_{10}/D_2	D_9/D_2	$D_{10}/(D_4–D_1)$	$D_{10}/(D_9–D_5)$	$D_{10}/(D_9–D_7)$	D_{10}/D_9
ACH	8.66	6.52	3.11	1.53	0.61	0.87	2.10
NSA	8.72	6.57	3.13	1.51	0.59	0.85	2.10
WSA	8.79	6.79	3.18	1.59	0.62	0.89	2.14
RIU	9.37	7.21	3.35	1.65	0.63	0.90	2.15
JBI	8.95	6.82	3.03	1.57	0.63	0.91	2.25
SSA	8.57	6.47	3.36	1.49	0.56	0.80	1.93
BKL	9.67	7.37	3.54	1.71	0.63	0.89	2.08
LMP	9.42	7.33	3.22	1.71	0.68	0.96	2.28
JKT	14.11	10.63	4.49	2.33	0.74	1.02	2.37
WJA	13.15	9.83	4.25	2.22	0.73	1.01	2.31
CJA	11.06	8.30	3.67	1.92	0.69	0.96	2.26
YOG	14.23	10.41	4.93	2.35	0.69	0.93	2.11
EJA	12.27	9.27	4.10	2.14	0.69	0.96	2.26
BLI	10.23	7.62	3.92	1.72	0.58	0.81	1.95
WNT	9.69	7.29	3.50	1.69	0.62	0.87	2.08
ENT	8.70	6.72	3.21	1.56	0.61	0.87	2.10
WKL	7.92	5.96	3.20	1.38	0.54	0.77	1.86
CKL	8.60	6.27	3.21	1.42	0.54	0.79	1.96
SKL	8.74	6.64	3.35	1.52	0.57	0.82	1.98
EKL	7.97	5.81	3.14	1.34	0.52	0.75	1.85
NSI	9.46	7.00	3.63	1.59	0.56	0.79	1.93
CSI	9.71	7.37	3.42	1.71	0.64	0.91	2.16
SSI	13.09	9.53	4.18	2.18	0.71	0.98	2.28
TSI	11.93	8.61	4.17	1.95	0.65	0.91	2.07
MAL	8.52	6.74	3.50	1.56	0.57	0.80	1.93
PAP	14.59	9.77	5.19	2.10	0.60	0.82	1.88

Source: Authors' calculation from Susenas, BPS.

Note: ACH: Aceh; NSA: North Sumatera; WSA: West Sumatera; RIU: Riau; JBI: Jambi; SSA: South Sumatera; BKL: Bengkulu; LMP: Lampung; BBL: Bangka Belitung; RIS: Riau Islands; JKT: Jakarta; WJA: West Java; CJA: Central Java; YOG: Yogyakarta; EJA: East Java; BTN: Banten; BLI: Bali; WNT: West Nusa Tenggara; ENT: East Nusa Tenggara; WKL: West Kalimantan; CKL: Central Kalimantan; SKL: South Kalimantan; EKL: East Kalimantan; NSI: North Sulawesi; CSI: Central Sulawesi; SSI: South Sulawesi; TSI: Southeast Sulawesi; GOR: Gorontalo; WSI: West Sulawesi; MAL: Maluku; NML: North Maluku; WPA: West Papua; PAP: Papua.

Appendix A2.12 Variance per Capita Household Expenditure and per Capita Non-food Expenditure across Provinces and Islands by Income Class, 1987–2015

Source: Author's calculation from Susenas, BPS.

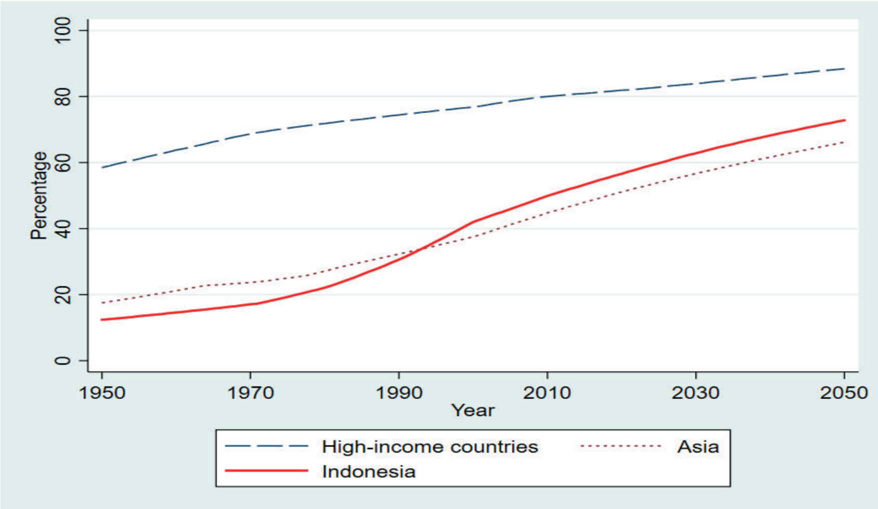
Chapter 3 City Size Distribution and Regional Income Inequality Dynamics in Indonesia

3.1 Introduction

In this chapter, we explore the relationship between income inequality and urbanization in Indonesia. In the previous chapter, we found that, over the last decades, the relatively strong economic growth in Indonesia has been associated with rapidly increasing income inequality. We also found evidence for regional convergence of inequality across islands and provinces, driven by the fact that the incomes of the rich in poor regions grew faster than those of the poor in rich regions. The middle class and especially the top incomes seem to benefit most from the economic growth dynamics. All of these findings may well relate to the fact that the relatively high GDP growth in Indonesia – around 5.5 per cent since the end of the recovery – is associated with rapid urbanization and the geographical concentration of economic activities in places like Jakarta. The urbanization rate in Indonesia rose continually from 42.0% in 2000 to 49.9% in 2010 and 55.3% in 2018 (United Nations 2018) due to either an increase in the number of cities or an increase in the size of the existing cities. Does urbanization in Indonesia associate with an increased inequality?

The urbanization rate in Indonesia has been faster than the average urbanization rate in Asia since 1993, when Indonesian's rate was 33.8 per cent and reached 50.0 per cent in 2010 (see Figure 3.1). Nearly all countries became at least 50.0 per cent urbanized before reaching middle-income status, and all high-income countries are 70–80 per cent urbanized. The high urbanization rate in Indonesia is associated with the implementation of an industrialization policy characterized by a transition from import substitution to export promotion, combined with liberalization and a banking reform in the mid-1990s. UN (2018) predicted more than 70 per cent of the world's population will live in cities in 2044. It seems that Indonesia needed only three decades to accomplish what today's industrialized countries took a century or more to achieve. Moreover, the number of urban agglomerations with at least 300 thousand inhabitants is expected to be 10 times higher than the number of urban agglomerations in 1950, and the number of cities with between 1 and 5 million inhabitants is expected to increase from 1 in 1950 to 9 in 2000 and 18 in 2035 (UN 2018).

Figure 3.1 Urbanization Rate in Indonesia and the Region, 1950–2050



Source: Authors' calculations from United Nations (2018).

Against this background, this chapter aims to explore whether and to what extent the number, size and distribution of cities might explain the evolution of inequality in Indonesia. Understanding the relationshiop between the size and distribution of cities and income inequality is important for policy makers in dealing with urban life and sustainable inclusive development. The impact of agglomeration on inequality relates to two main literatures, namely development economics and urban economics. A key question in this literature is what the relationship between inequality and development (including urbanization) looks like, inspired by the Kuznets hypothesis, which suggests a hump-shaped relationship. The Kuznets curve hypothesis has a spatial equivalent: the income gap between urban and rural areas first widens and then narrows. Important empirical contributions to the literature on inequality and economic development are the studies by Knight and Sabot (1983), Anand and Kanbur (1993) and Deininger and Squire (1997, 1998). Specific attention has also been paid to cities, but the size and income distribution of cities has not been studied much. Exceptions are the studies by Nord (1980), who focused on income inequality and city size; Sagala et al. (2014), who analysed the relationship between inequality and the process of urbanization; Baum-Snow and Pavan (2013), who studied wage inequality in larger cities; and Akita and Miyata (2008), who conducted an inequality decomposition analysis for urban and rural areas.

From the urban economics perspectives, the earlier literature suggested that the relationship between city size and inequality is negative, but recent studies have proposed that this relationship is positive. Due to a lack of data, most empirical studies are carried out at the country level, suggesting a U-shaped relationship between inequality and urbanization (Angeles 2010) or between inequality and city size (Castells-Quintana 2018). In urban development, agglomeration economies become more important over time, resulting in peripheral regions becoming urbanized (Henderson 2003) and cities that already exist becoming even bigger because of the important human capital accumulation and externalities (Henderson 2007). Furthermore, Baum-Snow and Pavan (2013) claimed that skill groups and industries that are disproportionately located in larger cities experience larger increases in their wage dispersion than those in smaller cities. Thus, these factors generate the city size-specific component of inequality growth. Ferré et al. (2012) found that inequality within large cities is not driven by a severe dichotomy between slum dwellers and others. However, very little evidence has confirmed that inequality in metropolitan areas is greater than inequality in smaller cities. In the case of Indonesia, Sagala et al. (2014) argued that the relationship between inequality and urbanization is an inverted U-shape and that it attains a peak if the urbanization rate is around 46–50 per cent. They claimed that inequality can be expected to decrease in Indonesia because it reached its peak at with an urbanization rate of around 50 per cent in 2010.

In this chapter, we focus specifically on exploring the relationship between average city size and inequality, which is a somewhat underexplored topic in the literature, as described above. More specifically, in this chapter, we investigate whether the average city size – which is in itself a product of agglomeration economies associated with economic growth – is correlated with the inequality levels across provinces in Indonesia. Recently, Castells-Quintana (2018) tested the relationship between inequality and city size across countries. He found a U-shaped relationship between average city size and income inequality: inequality first falls and then increases with average city size. However, we would like to argue that this kind of analysis is more interesting at the level of provinces rather than countries since each country has different regulations regarding urbanization. In addition, the mobility of people and goods across provinces within one country is in general easier and thus more intense than their mobility across countries. For example, people from rural regions tend to migrate to the cities within the same country rather than migrating to a different country. Therefore, in this chapter, we address this gap by empirically examining the question of whether the average size of a province's cities can explain inequality. Following Castells-Quintana (2018), we define the average agglomeration size for cities above 300 thousand inhabitants and employ this as an explanatory variable in the panel data regression analysis in 32 provinces over the period between 1990 and 2014.[6]

The remainder of this chapter is structured as follows. Sections 3.2 and 3.3 describe, respectively, the dynamic of urbanization in Indonesia and methodology that we use in this chapter. Section 3.4 then presents and discusses the key results of our analyses as well as several robustness checks. Finally, section 3.5 concludes.

3.2 City Size Dynamics

We are interested in the relationship between the size and distribution of cities and income inequality at the provincial level. This is a somewhat underexplored topic, also because defining a city is not straightforward. For that reason, we intend to work with a range of definitions, one that we will use to examine metropolitan regions and one that we will use to make a distinction between several cities that constitute metropolitan areas. We will use different samples later.

In this chapter, due to data availability, we calculate the average agglomeration size in each province from 1990 to 2014. The average city size across provinces is 615,000 inhabitants. Jakarta as one city accounts for the highest number with

[6] We use the term average agglomeration size interchangeably with average medium-sized city because those terms refer to the same definition.

8,884,000, followed by West Java with 912,000 and Banten with 793,000. However, if we unbundle Jakarta into five cities, the national average city size drops to 393,000 inhabitants and Jakarta's average size becomes 1,777,000. We also rank the average city size at the province level, and these figures seem to be consistent with Zipf's law, which states that the second-largest average city size of 912,000 (West Java) is about half of the largest average city size of Jakarta (1,777,000), and the third-largest average city size is 793,000 (Banten). Jakarta, as the capital city of Indonesia, outstrips the others' average agglomeration size, and its dominance has been increasing since 1980. A detailed size of the various cities that constitute integrated Jakarta is presented in Table 3.1.

Table 3.1 Metropolitan Region of Jakarta, 1980–2015

Area	Population (in millions)					Changes (%)		
	1980	1990	2000	2010	2015	2000–2010	2010–2015	2000–2015
Core	6.5	8.26	8.39	9.6	10.17	14.4	5.9	21.2
Jakarta*	6.5	8.26	8.39	9.6	10.17	14.4	5.9	21.2
Inner peripheries	n.a	n.a	4.93	7.22	8.36	46.5	15.8	69.6
City of Tangerang	n.a	n.a	1.33	1.8	2.04	35.3	13.3	53.4
City of S. Tangerang	n.a	n.a	0.8	1.29	1.53	61.3	18.6	91.3
City of Depok	n.a	n.a	1.14	1.75	2.09	53.5	19.4	83.3
City of Bekasi	n.a	n.a	1.66	2.38	2.7	43.4	13.4	62.7
Outer peripheries	5.41	8.88	7.31	11.2	13.09	53.2	16.9	79.1
City of Bogor	0.25	0.27	0.75	0.95	1.04	26.7	9.5	38.7
Tangerang Regency	1.53	2.77	2.02	2.84	3.36	40.6	18.3	66.3
Bekasi Regency	1.14	2.1	1.62	2.63	3.23	62.3	22.8	99.4
Bogor Regency	2.49	3.74	2.92	4.78	5.46	63.7	14.2	87.0
Metropolitan Region of Jakarta	11.91	17.14	20.63	28.02	31.62	35.8	12.8	53.3

Source: Authors' calculations based on Rukmana (2018).
* excluding district of Kepulauan Seribu from Jakarta's administration.

As a metropolitan region or *Jabotadebek*,[7] the outer peripheries grow faster than the inner peripheries as well as the core city of Jakarta, with corresponding figures of 21.2

[7] The peripheries of Jakarta consist of two jurisdictions of provinces, namely the City of Depok, City of Bogor, Regency of Bogor, City of Bekasi and Regency of Bekasi, which are within West Java province, and the City of Tangerang, Regency of Tangerang and City of South Tangerang, which constitute Banten province.

per cent and 53.3 per cent for the metropolitan region of Jakarta from 2000 to 2015. This high growth is partly the result of national policies to build large-scale industries in the peripheries of Greater Jakarta, which shifted manufacturing from the central city to the periphery, while Jakarta, as a core city, has disproportionately attracted investments in the service industry, such as finance and business, trade, and hotels and restaurants (Rukmana et al. 2018). A large migration flow from the urban core to the outskirts is an important factor explaining population growth in big cities in Indonesia. For instance, the share of the *Jabotadebek* population in Indonesia's population increased from 8.1 per cent in 1980 to 9.6 per cent in 1990 and 10.0 per cent in 2000 (Rukmana 2015) and its population density increased substantially from 25.5 persons per square hectare in 1990 to 37.6 in 2000 and 44.6 in 2010 (Firman 2014). Similar tendencies can be found in the second- and third-largest metropolitan areas, namely Surabaya as the capital city of East Java province, known as *Gerbangkertosusila*, and Bandung as the capital city of West Java province, called *Bandung Raya*. Industrial estates and most foreign investors are also located in the peripheries of those metropolitan regions. The peripheries' population is growing faster than the core's population, and the proportion of the core population in the total metropolitan areas is tending to decrease (Appendix A3.1). As a consequence, the concentration of skilled labour and entrepreneurs in the metropolitan regions will shape the socio-economic disparities.

If we look at the trend, the average agglomeration size across provinces is rising over time and the percentage of the population living in cities with at least 1 million inhabitants tends to increase during the period from 1990 to 2014 (Appendix A3.2). These trends translate into more and more people living in large cities.

Table 3.2 Urbanization Dynamics across Provinces, 1990–2014

Province	Average City Size 300k (in millions)			% Change in Ave. City Size 300k		
	1990	2000	2014	1990–2000	2000–2014	1990–2014
Aceh	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
North Sumatera	1.74	1.91	2.19	9.7	14.7	25.8
West Sumatera	0.63	0.71	0.89	12.5	24.1	39.7
Riau	0.40	0.59	1.00	45.1	71.3	148.6
Jambi	0.34	0.42	0.57	22.1	35.5	65.5
South Sumatera	1.15	1.22	1.55	5.9	27.3	34.8
Bengkulu	n.a.	n.a.	0.34	n.a.	n.a.	n.a.
Lampung	0.64	0.74	0.95	15.4	28.5	48.3
Bangka Belitung	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Riau Islands	0.34	0.46	1.13	36.1	146.7	235.9
Jakarta	8.54	8.61	10.02	0.8	16.4	17.3
Jakarta Metropolitan	*17.4*	*20.63*	*28.46*	*20.4*	*37.9*	*66.0*
West Java	1.21	1.18	1.20	–2.8	2.3	–0.6
Central Java	0.88	0.89	1.09	1.8	21.4	23.6
Yogyakarta	0.41	0.40	0.41	–3.6	2.2	–1.5
East Java	1.59	1.68	1.84	6.0	9.2	15.7
Banten	1.00	0.81	1.12	–18.6	38.2	12.5
Bali	0.40	0.52	0.86	31.2	64.8	116.3
West Nusa Tenggara	n.a.	0.32	0.44	n.a.	37.6	n.a.
East Nusa Tenggara	n.a.	n.a.	0.37	n.a.	n.a.	n.a.
West Kalimantan	0.40	0.47	0.59	18.3	25.8	48.8
Central Kalimantan	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
South Kalimantan	0.48	0.53	0.66	9.3	25.7	37.4
East Kalimantan	0.38	0.46	0.71	22.9	53.4	88.5
North Sulawesi	0.32	0.41	0.42	26.3	3.6	30.9
Central Sulawesi	n.a.	n.a.	0.36	n.a.	n.a.	n.a.
South Sulawesi	0.95	1.10	1.42	16.0	29.3	50.0
South East Sulawesi	n.a.	n.a.	0.33	n.a.	n.a.	n.a.
Gorontalo	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Maluku	n.a.	n.a.	0.39	n.a.	n.a.	n.a.
North Maluku	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
West Papua	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Papua	n.a.	n.a.	n.a.	n.a.	n.a.	n.a.
Average	1.15	1.17	1.23	2.1	5.4	7.5

Source: Authors' calculations from various editions of BPS and MOHA.

At the province level, Table 3.2 presents the average agglomeration size for cities with over 300 thousand residents across provinces and its change from 1990 to 2014. On average, the percentage of change in the agglomerated city size is 7.5 per cent for the last 24 years, but some provinces show an increasing city size which more than doubled, such as Bali, the Riau islands and Riau.[8] On the other hand, there are 6 provinces that do not have medium-sized cities of at least 300 thousand inhabitants (Bangka Belitung, Central Kalimantan, Gorontalo, North Maluku, Papua and West Papua). This table also indicates that urban inequality is greater if the urban population is concentrated in a few of the province's largest cities and much lower if the population is evenly distributed across large and small cities.

3.3 Methodology and Data

To seek the relationship between city size and income inequality, we take Kuznets's standard and perform fixed-effect panel data regressions to consider the variation within provinces over time based on a provincial panel dataset over the period from 1990 to 2014 with a total of 768 observations from 32 provinces. Due to missing data, 510 observations are used. In this chapter, we classify a city based on the administrative definition issued by the Ministry of Home Affairs[9] except for Jakarta. We treat Jakarta as one agglomerated city, not as five administrative cities, as in the Ministry's classification. Jakarta will be treated as five cities in our robustness check.

In our basic model, we restrict our sample to cities with at least 300 thousand inhabitants as this conforms to Duranton and Puga (2013). Following Castells-Quintana (2018), we use this average city size (above 300k) as the key explanatory variable because (i) this large city size sufficiently exposes agglomeration economies and congestion costs, (ii) according to Zipf's law, information on cities above 300k inhabitants should be enough to delineate the size of all cities. For province *i* in year *t*, we calculate the total number of people living in cities with at least 300 thousand

[8] After decentralization, some cities and regencies in Riau province formed a new province, Riau Island, in 2002, and all the cities in both provinces grew substantially afterwards.

[9] According to Law No. 32/2004 concerning Regional Administration, there are four levels of administrative divisions and each of them is led by a head of administration: (i) a province (*propinsi*) is headed by a governor, (ii) a regency or municipality (*kabupaten*) is headed by a *bupati* and a city (*kota*) is headed by a mayor, (iii) a sub-district (*kecamatan*) is headed by a camat or head of sub-district and (iv) a village (*desa*) is led by a head of village (*kepala desa*) or a *lurah* if the village is located in an urban area or a *kelurahan*. We use this city administration headed by a mayor in our sample.

inhabitants divided by the total number of cities to obtain the average medium city size[10] or *city300* as our operational variable. We expect inequality to have a U-shaped relationship with the average city size and an inverted U-shaped relationship with income. In addition, we hypothesize that a larger city size with inequality is less harmful if income increases; therefore, we interact city size and income. Our basic model is written as follows:

$$Ineq_{it} = \beta_0 + \beta_1 \cdot Citysize_{it} + \beta_2 \cdot [Citysize_{it}]^2 + \beta_3 \cdot Income_{it} + \beta_4 \cdot [Income_{it}]^2 + \beta_5 \cdot [Citysize_{it} \cdot Income_{it}] + \beta_6 \cdot X_{it} + \varepsilon_{it} , \quad 3.1$$

where $Ineq_{(it)}$ as the dependent variable denotes inequality expenditure as measured by the Gini coefficient in province i in year t. This coefficient is calculated based on household expenditure from various Susenas consumption modules (BPS). Our variables of interest are city size ($Citysize_{it}$), that is, the average agglomeration size in province i in year t and its square $[Citysize_{it}]^2$, income ($Income_{it}$), as measured by the logarithm of the per capita GRDP in province i in year t and its square $[Income_{it}]^2$, and the interaction of city size and income ($Citysize_{it} \cdot Income_{it}$). Hence, based on the proposition above, we expect that β_1 is negative, β_2 is positive, β_3 is positive, β_4 is negative and β_5 is negative.

As a robustness check, we exclude Jakarta because it is a capital and the most agglomerated city in Indonesia, with more than 8 million people, which is a much higher number than the median of the medium city size of 678 thousand; we also unbundle Jakarta into 5 cities to normalize the average city size and include all cities. For each province i in year t, we calculate the total number of people living in all cities divided by the number of all cities. We use *allcity* as our operational variable.

Further, we add $primacy_{it}$ and its square $[primacy_{it}]^2$, where $primacy_{it}$ is defined as the percentage of the urban population living in the largest city in each province i in year t. We hypothesize that primacy will correlate positively with inequality as higher returns to high skills in the largest city then correlate negatively

[10] Jakarta province consists of five cities; however, since Jakarta is the most agglomerated city, we merge all five cities into one city. Thus, in our sample, Jakarta province now has only one city and its size is huge. As a robustness check, we separate Jakarta's city size into five cities.

due to oversupplied highly skilled workers tending to reduce the wage premium and thus lower inequality (Chiswick 1968; Knight and Sabot 1983). Last, we replace the per capita GRDP with the initial GRDP in the year 1990 and its square to see how the stage of development relates with inequality. In our data, the GRDP for 1990 is the earliest one, so we use this variable to reduce the problems of reverse causality.

We apply the same control variables, that is, economic growth ($ecogorwth_{it}$), the share of investment (ki_{it}), the share of government consumption in the total GDP based on expenditure (kg_{it}), the net enrolment ratio in elementary school (ner_es_{it}), the firm density ($firmdens_{it}$), which we define as the number of manufacturing firms divided by the total population, the share of agricultural employment in the total employment ($shragrempl_{it}$) and the share of the provincial minimum wage in the national minimum wage ($shrwage_{it}$), in our basic model and robustness check regressions. Following Castells-Quintana (2018), all the right-hand-side variables are also presented in lag form to reduce endogeneity problems. We use Statistics Manufacture Industry to calculate the firm density and Provinces in Figures to obtain the key variables. Both datasets are provided by BPS.

3.3.1 Descriptive Statistics

The distribution of the city data is right skewed, as revealed by the big difference between the median and the mean. For instance, the average population of all cities is 1397.4 thousand while the median is around one-third or 538.7 thousand. Other city variables show similar distribution as presented in Table 3.3 and Kernel distribution in Appendix A3.3.

The average city size is around 318 thousand and 678 thousand for small and medium cities, respectively, while the average size of the largest city in each province is about 432 thousand. If we look further at the distribution of the population in big cities in each province, 83 per cent of the urban population is living in the largest city and 24 per cent of the population is living in the cities with more than 1 million inhabitants. The number of cities in each province ranges from 1 to 9, and most of the provinces have only 2 small cities and 1 medium city.

Table 3.3 Descriptive Statistics of the Key Variables

Variable	Obs.	Median	Mean	Std Dev.	Skew-ness	Min.	Max.
Inequality Expenditure	997	30.00	30.79	4.73	0.14	8.03	46.00
Average City Population (in thousands)	768	538.74	1397.41	2110.52	2.58	113.60	10020.02
Average Medium City Size in Each Province (in thousands)	510	677.91	1188.79	1774.99	3.85	300.35	9947.24
Average City Size in Each Province (in thousands)	768	318.04	614.83	1500.39	5.28	110.23	9947.24
Size of the Largest City in Each Province (in thousands)	768	432.24	951.63	1569.90	4.11	113.60	9947.24
% Pop. Living in Cities of >1 Million in Each Province	245	24.00	30.69	24.82	1.91	8.00	100.00
% Urban Pop. Living in the Largest City in Each Province	768	83.07	80.83	20.49	–0.79	25.22	100.00
Number of All Cities in Each Province	768	2.00	2.70	2.35	1.51	1.00	9.00
Number of Medium Cities in Each Province	510	1.00	1.47	1.22	3.45	1.00	8.00
Per Capita GRDP (in million Rupiah)	1199	35.27	56.67	67.18	2.80	0.86	455.91
Economic Growth (%)	1093	4.37	5.21	9.25	4.50	–26.36	123.14
Av. Share of Government Cons. in Total GRDP over 5 Years	579	12.48	13.46	6.92	0.73	2.07	36.63
Average Share of Investment in Total GRDP over 5 Years	579	21.65	22.75	9.31	1.13	3.08	78.29
Net Enrolment Ratio for Elementary School	635	92.57	91.90	3.52	–2.45	70.13	98.72
Share of Agriculture in Total Employment	514	47.64	45.66	16.80	–0.62	0.24	78.04
Share of Manufacturing in Total Employment	514	6.60	8.58	6.13	1.63	0.71	46.31
Share of Provincial Minimum Wage in National Minimum Wage	632	0.97	1.00	0.19	0.70	0.51	1.74
Share of Underemployment in Total Labour Force	528	31.43	30.76	9.13	–0.20	1.86	62.63
Share of Unemployment in Total Labour Force	528	6.56	7.33	3.28	0.91	1.83	19.16
Firm Density	1142	0.04	0.06	0.05	1.92	0.00	0.32

Source: BPS various editions, calculated by the author.

On average, the inequality across provinces is 30 while the highest and the lowest inequality are experienced by Gorontalo province in 2011 and in 2000[11]. Per capita GDP is around 35 million rupiah and its average annual growth is 4.5 per cent. Education seems to be left behind with a minimum net enrolment ratio for elementary school of only 70 per cent. This in line with the employment structure, in which almost half of the labour force works in the agricultural sector and less than 10 per cent works in the manufacturing sector. The firm density as a proxy for wage differentiation is very low – less than one – implying that the wage difference between manufacturing workers and other workers is very small (Table 3.3).

3.4 Empirical Results and Discussion

This section presents and discusses the results of the relationship between inequality and average agglomeration size and the relationship with income. We also discuss the results of the above models used as robustness checks, that is, excluding Jakarta, adding primacy, replacing income with the initial income in 1990 and unbundling Jakarta into five cities.

Table 3.4 presents a summary of the regression results of the basic model and the robustness checks. The first three columns of the basic model describe the relationship between the average medium city size (at least 300 thousand inhabitants) and inequality: column 1 includes agglomerated Jakarta, column 2 excludes Jakarta and column 3 includes unbundled Jakarta. The last two columns exhibit the relationship between all city sizes, including small cities, and inequality by including unbundled Jakarta and adding primacy, as shown in column 4 and column 5, respectively. Since our results confirm a U-shaped or inverted U-shaped, we are also interested in determining the bottom or the top of each model and calculating the percentage of observations left of the minimum or maximum value. The full regression results are reported in the Appendices.

[11] Province Gorontalo was one of the kabupaten in North Sulawesi province prior to 2000, when Susenas 2000 was carried out in 1999. This changing administrative status of Gorontalo may affect the Gini calculation in 2000.

The relationship between average agglomeration size and inequality across provinces in Indonesia has a U-shape, as discussed by Castells-Quintana (2018), only when we include Jakarta as one agglomerated city. Our results reveal that, when the average agglomeration size grows larger, inequality tends to decline until it reaches the bottom and then it increases as the average medium city size grows if we include Jakarta as an agglomerated city or as one bundled city (column 1). Bundled Jakarta has around 8 million inhabitants, making it an outlier compared with the median of the average agglomeration size across provinces of 678 thousand inhabitants. So, we downplay specification 2 without Jakarta and specification 3 with unbundled Jakarta. When we exclude Jakarta from our sample (column 2) or when we unbundle Jakarta into five cities (column 3), inequality first increases and then decreases with the average agglomeration size. In other words, inequality follows the Kuznets inverted U-curve with respect to the average agglomeration size, contradicting with Castell-Quintana (Figure 3.2 Panel A and Panel B). Without agglomerated Jakarta, an increase in inequality occurs with a larger urban agglomeration when the average agglomeration is already high. The city size with at least 300 thousand inhabitants may reflect agglomeration economies in which cities are more productive, benefit from more highly skilled workers and usually involve externalities or spillovers.

Sagala et al. (2014) suggested that the increasing inequality across provinces in Indonesia could be explained in the context of the dual economy consisting of the rural and the urban sector shifting from low income in rural areas to high income in urban areas, resulted in an inverted U-curve between inequality and urbanization following the classical theories by Lewis (1954) and Kuznets (1955).[12] It appears that the role of Jakarta as the most agglomerated city is essential in shaping the inequality pattern; inequality first increases with the average medium city size and income until the city size reaches 1383 thousand inhabitants, then it decreases while the cities keep growing along with the per capita GRDP (column 2). Our estimation also suggests that more cities are located to the left of the top point, implying that, without Jakarta, most of the provinces with an average medium city size are at the same stage of

[12] This inverted U-shaped relationship between inequality and urbanization does not exist in cross-country analysis (Angeles 2010).

development and follow the Kuznets path, whereby inequality increases with per capita GRDP increases.

Table 3.4 Summary of Regression Results:
Relationship between Inequality and Average City Size

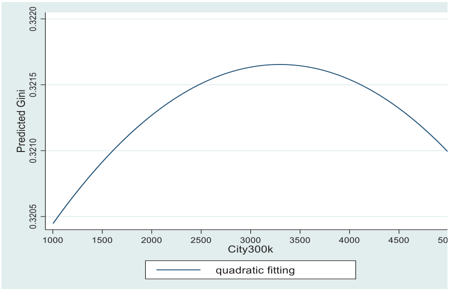
VARIABLES	Dependent Variable: $ineq_{it}$				
	(1) Bundled Jakarta	(2) Without Jakarta	(3) Unbundled Jakarta	(4) All cities	(5) Primacy
lag_city300	2.392	11.03***	3.862***		
	(1.89)	(3.26)	(1.15)		
lag_city300^2	0.523***	–3.987***	–0.250***		
	(0.16)	(1.39)	(0.08)		
lag_allcity				–12.20***	–4.262
				(4.46)	(5.171)
lag_allcity^2				10.53***	9.446***
				(3.44)	(3.556)
lag_lgrdpc	24.92***	20.48***		20.51***	18.48***
	(3.24)	(4.45)		(2.97)	(3.012)
lag_lgrdpc^2	–3.683***	–2.964***		–2.604***	–2.448***
	(0.41)	(0.62)		(0.37)	(0.371)
lag_primacy					0.412***
					(0.126)
lag_primacy^2					–0.003***
					(0.0008)
Constant	–15.65**	–10.22	26.23***	–7.153	–18.26**
	(7.26)	(8.41)	(1.25)	(5.99)	(7.567)
Shape Size~					
Turning Point	2288.12	1383.23	7724.06	579.46	72.89
% of Obs. Left of the Min. or Max. Value	95.29	86.83	98.24	88.41	91.28
Year FE	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES
Number of Observations	373	352	397	550	550
R-squared	0.807	0.815	0.725	0.734	0.742
F-test	YES	YES	YES	YES	YES
Number of Provinces	24	23	24	32	32

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1. ~ refers to the extreme values defined as $\beta_1/2\beta_2$.

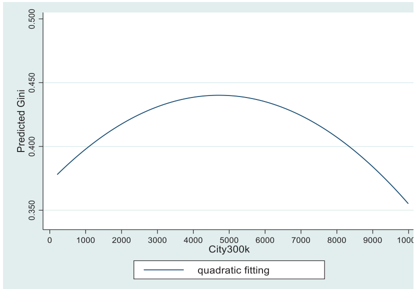
In addition, when we replace per capita GRDP with the initial GRDP in 1990 in our regression without Jakarta, we find a hump-shaped relationship, confirming the previous result. The full regression results can be found in Appendix A3.4.B3. This inverted U-shape is also apparent when we unbundle Jakarta into five cities (Figure

3.2 Panel B). Inequality keeps rising as the average agglomeration size enlarges until it passes 7,724 thousand inhabitants, after which it declines (column 3). In our data, the typical agglomeration size in Indonesia is a medium city with on average 318 thousand inhabitants, which tends to grow into a larger city and thus inequality is expected to increase.

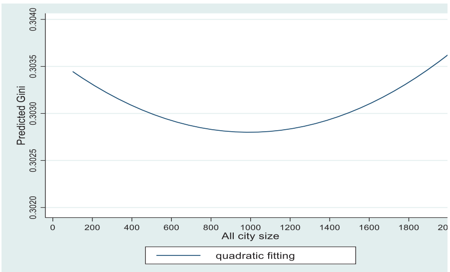
Figure 3.2 Kuznets Agglomeration Size and Robustness Check



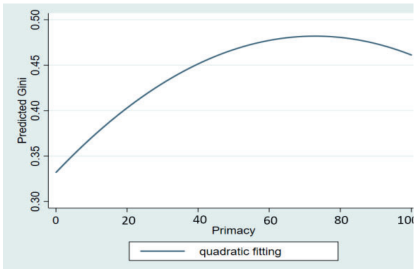
Panel A: The Relationship between Inequality and Average Medium City Size (excluding Agglomerated Jakarta)



Panel B: The Relationship between Inequality and Average Medium City Size (including Unbundled Jakarta)



Panel C: The Relationship between Inequality and Average All City Size (including Unbundled Jakarta)



Panel D: The Relationship between Inequality and Primacy (including Unbundled Jakarta)

With regard to urban inequality, it seems that the most agglomerated city, that is, Jakarta, has two extreme folds, specifically more opportunities that benefit low-income workers more strongly (Todaro 1976; Ferré et al. 2012) and higher returns to high skills (Baum-Snow and Pavan 2013; Behrens and Robert-Nicoud 2014), and the middle class in neighbouring cities levels up with increasing inequality (U-shaped). As discussed in the previous section, industrial estates located in the suburbs of

Jakarta are increasingly becoming specialized and intensified industries. The firm density, share of manufacturing employment and share of the non-government sector (manufacturing, finance and trade and hotels) in the neighbouring provinces of West Java and Banten are ranked in the top three provinces across Indonesia. This suggests that the average agglomeration size has association with firms' location and workers' location by industry. According to Baum-Snow and Pavan (2013), skill groups and industry, disproportionately located in larger cities, increase the wage dispersion more in larger cities than in small cities.

As a robustness check, we use all the city sizes, including small cities and unbundled Jakarta, and primacy to see how the urban population living in the largest city correlates with inequality (columns 4 and 5). Our results yield a negative for liniear term and positive for quadratic term suggesting that inequality first decreases when small cities grow and then increases when large cities grow. A decrease in inequality as the agglomeration size increases may be associated with the fact that larger cities provide more opportunities to earn more money for people with different abilities and skills because more high-paid as well as low-paid jobs, including jobs in the informal sector, are available. Therefore, a larger city is always attractive not only to more educated and highly skilled people who earn a high income but also to informal workers who earn from low-paid jobs as a result of high urbanization.

We find this U-shaped pattern is similar to the first specification, in which we estimate the average medium city size and include agglomerated Jakarta as one city. Inequality decreases with the average city size until the size touches the minimum value of 579 thousand inhabitants, then inequality increases as the average city size grows (column 4, Figure 3.2 Panel C). This turning point is much smaller than the turning point obtained using the average agglomeration size for unbundled Jakarta. It suggests that inequality tends to lower when the average city size keeps growing into the larger size; in other words, the gap will be smaller when the average city size is large enough to provide opportunities for unskilled labour due to urbanization.

Since the level of per capita GRDP is positively correlated with inequality and statistically significant, an increase in the average city size tends to lower inequality if the city residents can benefit from a higher GRDP per capita. However, when we add primacy to our regression, its coefficient positively correlates with inequality while the average city size shows the opposite (column 5). Our results in specification 5 indicate that the relationship between inequality and average city size is N-shaped; it is U-shaped with a small city size then inverted U-shaped with a primacy (Figure 3.2 Panel D). This pattern is in line with our findings for income.[13] The literatures suggest that productivity increases with the city size (Baum-Snow and Pavan 2013) and that urban primacy is advantageous to growth in low-income countries (Henderson 2003) while the largest city is also considered as the most unequal (Behrens and Robert-Nicoud 2014). Hence, high and increasing inequality of the largest city may help explain why the inverted U-shaped now has an N-shaped. Further, the highest level of inequality could be attained when the primacy is 72.89 per cent, and more observations are distributed to the left of the top value, implying that more people living in the largest city as well as higher per capita GRDP associates with higher inequality.

Moreover, part of the association between average city size and inequality may be explained by the association between city size and economic performance. Our results show that the interaction between city size and income is negatively correlated with inequality in all specifications. It suggests that the high inequality that is associated with the average city size could be lowered by an increasing income.

Our control variables have different correlations with inequality depending on the sample and definition of city size that we use. For instance, the share of the provincial minimum wage in the national minimum wage and the firm density are negatively associated with inequality and statistically significant when we unbundle Jakarta into five cities and define the city size with at least 300 thousand inhabitants (Appendix A3.4.C1). Intuitively, for the average medium city size, where the size is already agglomerated and large enough for the city to offer more opportunities, inequality decreases when the number of firms and the wages increase. On the other hand, if we use the full sample including the average small city size below 300 thousand residents, economic growth and education have a positive correlation with inequality (Appendix A3.4.D1). Our results suggest that, when the average city size is small, the enrolment in elementary school is also low, and then the average city size grows with economic performance, allowing people to access higher education. These

[13] See Chapter 2.

skilled people then benefit from the larger city and hence the gap with education and economic growth increases in a manner similar to the process postulated by Kuznets (1955). Recent literature on the impact of education on inequality has found that education affects the two tails of the distribution of income: education reduces the income share of the top earners and increases the share of the bottom earners (Abdullah 2015).

3.5 Conclusion

In this chapter, we investigated whether and to what extent the number, size and distribution of cities might explain the evolution of inequality across provinces in Indonesia. In a cross-country analysis Castells-Quintana (2018) found that beyond Kuznets' hypothesis there is a U-shaped relationship between average city size and inequality; inequality first falls and then increases with average city size. A key result of our analysis across provinces within one country is that this U-shaped relationship was only found if we included Jakarta as one bundled city in our sample. In contrast to Castells-Quintana (2018) this result is not robust. The U-shaped relationship between average city size and inequality did not hold when we excluded Jakarta or when we considered unbundled Jakarta. However, if we included small cities with less than 300 thousand inhabitants in our sample, and we add primacy, we found a U-shaped with a small city but inverted U-shaped with a primacy. This N-shaped where inequality first decreases with small city size, then increases with larger city size, and finally declines again with very large city size might be associated with the fact that larger cities provide more job opportunities for people with different abilities and skills. The interaction between city size and income is negative for all the specifications, suggesting that part of the association between average city size and inequality could be explained by the association between average city size and economic performance: larger cities are more productive, benefit from more highly skilled workers and usually involve spillover, and hence income increases. Our findings suggest that, if the current trend of an increasing average agglomeration size continues, we can expect inequality to climb further. This message implies that larger average agglomeration size may be desirable when cities are small, however, a very high average agglomeration size is undesirable so the medium city size may be more desirable. In the next chapter, we will examine the different dimensions of whether the change in inequality is correlated with the economic growth rate.

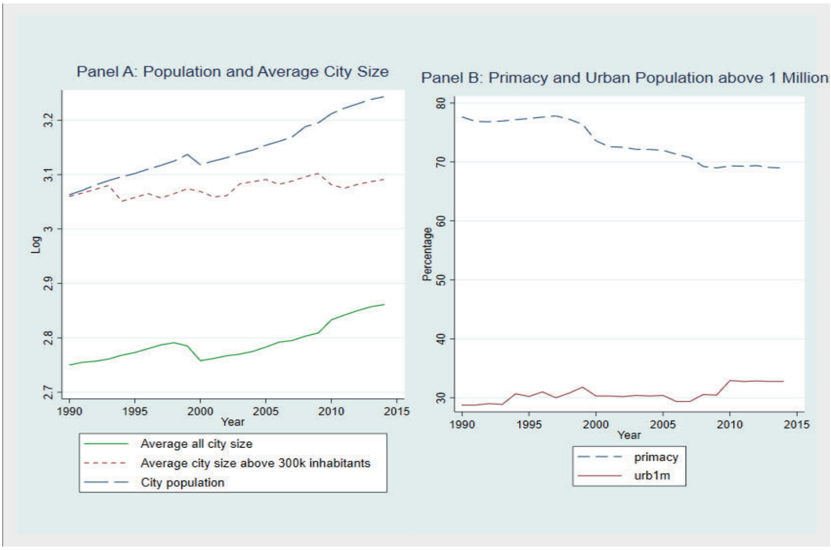
Appendices

Appendix A3.1 Core and Peripheries Population in Three Largest Metropolitan Areas, 1980–2015

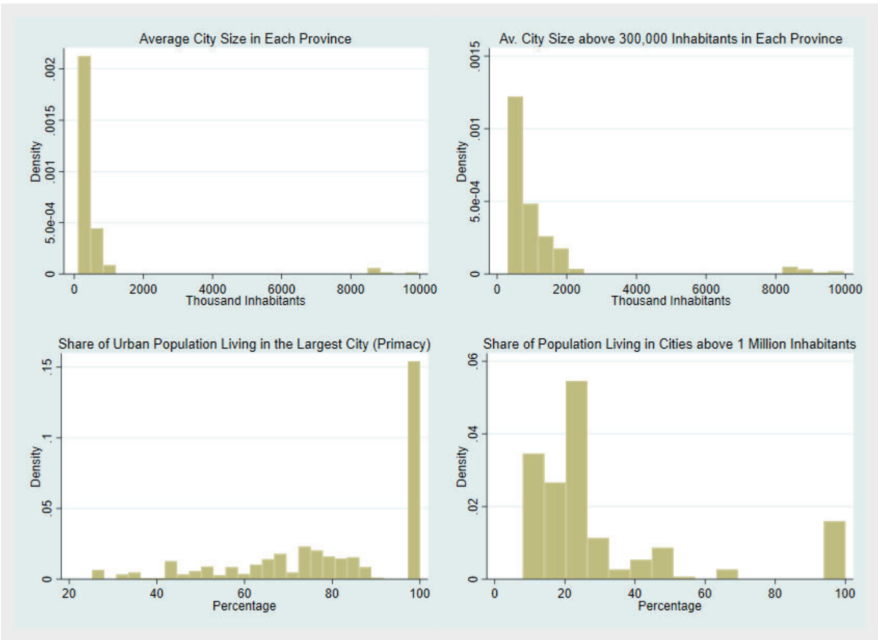
Area	Population (in millions)					Changes (%)		
	1980	1990	2000	2010	2015	2000–2010	2010–2015	2000–2015
Core								
Jakarta	6.50	8.26	8.39	9.60	10.17	14.4	5.9	21.2
Surabaya	2.02	2.47	2.60	2.77	2.84	6.6	2.6	9.4
Bandung	1.46	2.06	2.14	2.39	2.48	12.1	3.6	16.1
Peripheries								
Jakarta	5.41	8.88	12.24	18.42	21.45	50.5	16.4	75.2
Surabaya	4.09	4.76	5.57	6.37	6.71	14.3	5.4	20.5
Bandung	2.67	3.20	4.16	5.23	5.74	25.8	9.8	38.1
Metropolitan Region								
Jakarta	11.91	17.14	20.63	28.02	31.62	35.8	12.8	53.3
Surabaya	6.11	7.23	8.17	9.14	9.55	11.9	4.5	16.9
Bandung	4.13	5.26	6.29	7.62	8.22	21.1	7.8	30.6
	Share Core to Metropolitan Region (%)							
Jakarta	54.6	48.2	40.7	34.3	32.2	−15.8	−6.1	−20.9
Surabaya	33.0	34.2	31.8	30.3	29.8	−4.7	−1.8	−6.5
Bandung	35.4	39.1	33.9	31.4	30.2	−7.5	−3.9	−11.1

Source: Authors' calculations based on Rukmana (2018).

Appendix A3.2 Population and City across Provinces, 1990–2014



Appendix A3.3 Distribution of City Variables



Appendix A3.4.A1 Full Sample Including Bundled Jakarta (Basic Model)

Variables	(1) Model_a11 gini	(2) Model_a12 gini	(3) Model_a13 gini	(4) Model_a14 gini	(5) Model_a15 gini
lag_city300	–3.108**	–3.653**	2.392	7.770***	–1.525
	(1.31)	(1.81)	(1.89)	(2.01)	(8.08)
lag_city300^2		0.0644	0.523***	0.744***	0.196
		(0.15)	(0.16)	(0.18)	(0.58)
lag_lgrdpc			24.92***	0.0626	4.231
			(3.24)	(1.16)	(3.98)
lag_lgrdpc^2			–3.683***		
			(0.41)		
city300. grdpc*				–1.614***	0.456
				(0.21)	(1.75)
lag_ecogrowth					0.111
					(0.09)
lag_ki					0.0487
					(0.11)
lag_kg					–0.0182
					(0.20)
lag_ner_es					0.319
					(0.21)
lag_firmdens					–18.74
					(13.98)
lag_shrwage					–5.234*
					(2.99)
lag_shragrempl					0.0600
					(0.06)
Constant	33.39***	33.71***	–15.65**	24.34***	–14.57
	(1.58)	(1.74)	(7.26)	(4.77)	(25.71)
Year FE	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES
Number of Observations	397	397	373	373	171
R-squared	0.721	0.721	0.807	0.796	0.740
F-test	YES	YES	YES	YES	YES
Number of provinces	24	24	24	24	24

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Appendix A3.4.A2 Full Sample Including Bundled Jakarta (add Primacy as Robustness Check)

Variables	(1) model_a21 gini	(2) model_a22 gini	(3) model_a23 gini
lag_city300	2.183	7.076***	0.00590
	(1.922)	(2.056)	(8.162)
lag_city300^2	0.547***	0.764***	0.171
	(0.156)	(0.183)	(0.582)
lag_lgrdpc	24.51***	–0.0779	5.083
	(3.269)	(1.181)	(4.162)
lag_lgrdpc^2	–3.663***		
	(0.408)		
lag_primacy	0.0522	0.183	0.478
	(0.114)	(0.116)	(0.615)
lag_primacy^2	–0.000521	–0.00114	–0.00350
	(0.000685)	(0.000698)	(0.00533)
city300_grdpc		–1.597***	0.311
		(0.208)	(1.764)
lag_ecogrowth			0.0838
			(0.0911)
lag_ki			0.0890
			(0.116)
lag_kg			–0.0762
			(0.209)
lag_ner_es			0.290
			(0.218)
lag_firmdens			–21.21
			(14.17)
lag_shrwage			–5.227*
			(3.007)
lag_shragrempl			0.0701
			(0.0625)
Constant	–15.01*	18.59***	–31.78
	(8.061)	(6.517)	(30.88)
Year FE	YES	YES	YES
Province FE	YES	YES	YES
Observations	373	373	171
R-squared	0.809	0.798	0.744
F-test	YES	YES	YES
Number of provinces	24	24	24

Notes: Standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendix A3.4.A3 Full Sample Including Bundled Jakarta (Use Initial GRDP 1990 as Robustness Check)

Variables	(1) Model_a31 gini	(2) Model_a32 gini	(3) Model_a33 gini
lag_city300	1.040**	4.012***	–2.949
	(0.48)	(0.82)	(8.59)
lag_city300^2	–0.0561	0.294***	0.109
	(0.05)	(0.07)	(0.61)
log_grdp90	–9.464***	0.218	–2.121
	(3.02)	(0.25)	(17.20)
sqr_log_grdp90	1.113***		
	(0.37)		
city300 grdpc*		–1.136***	0.922
		(0.21)	(1.76)
lag_ecogrowth			0.0600
			(0.13)
lag_ki			0.0535
			(0.14)
lag_kg			0.0597
			(0.24)
lag_ner_es			0.341
			(0.25)
lag_firmdens			–13.98
			(14.84)
lag_shrwage			–3.507
			(3.43)
lag_shragrempl			0.0500
			(0.07)
Constant	48.32***	28.50***	4.396
	(5.82)	(1.20)	(59.29)
Year FE	YES	YES	YES
Province FE	NO	NO	NO
Number of Observations	353	334	151
R-squared	0.772	0.813	0.725
F-test	YES	YES	YES
Number of provinces	20	20	20

Notes: Standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendix A3.4.B1 Sample Excluding Bundled Jakarta (Basic Model)

Variables	(1) Model_b11 gini	(2) Model_b12 gini	(3) Model_b13 gini	(4) Model_b14 gini	(5) Model_b15 gini
lag_city300	–4.232***	4.224	11.03***	12.48***	22.55*
	(1.51)	(3.64)	(3.26)	(4.15)	(12.03)
lag_city300^2		–4.103**	–3.987***	–4.033**	–4.256
		(1.61)	(1.39)	(1.81)	(4.56)
lag_lgrdpc			20.48***	0.0379	7.313*
			(4.45)	(1.15)	(4.00)
lag_lgrdpc2			–2.964***		
			(0.62)		
city300 grdpc*				–0.504	–2.659
				(1.06)	(3.09)
lag_ecogrowth					0.113
					(0.09)
lag_ki					–0.00636
					(0.11)
lag_kg					0.0788
					(0.19)
lag_ner_es					0.324
					(0.21)
lag_firmdens					–29.85
					(19.91)
lag_shrwage					–4.849
					(2.97)
lag_shragrempl					0.0397
					(0.06)
Constant	32.06***	28.83***	–10.22	23.82***	–29.92
	(1.19)	(1.74)	(8.41)	(4.54)	(25.09)
Year FE	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES
Number of Observations	376	376	352	352	161
R-squared	0.757	0.762	0.815	0.802	0.719
F-test	YES	YES	YES	YES	YES
Number of provinces	23	23	23	23	23

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Appendix A3.4.B2 Sample Excluding Bundled Jakarta (Add Primacy as Robustness Check)

Variables	(1) model_b21 gini	(2) model_b22 gini	(3) model_b23 gini
lag_city300	10.49***	11.00**	23.38*
	(3.316)	(4.294)	(12.04)
lag_city300^2	–3.802***	–4.170**	–5.315
	(1.407)	(1.809)	(4.625)
lag_lgrdpc	20.30***	–0.0793	8.366**
	(4.624)	(1.177)	(4.186)
lag_lgrdpc^2	–2.963***		
	(0.651)		
lag_primacy	0.0569	0.181	0.461
	(0.114)	(0.115)	(0.606)
lag_primacy^2	–0.0004	–0.0011	–0.0033
	(0.0007)	(0.0007)	(0.0052)
int_city300_grdpc		–0.236	–2.206
		(1.080)	(3.131)
lag_ecogrowth			0.0823
			(0.0887)
lag_ki			0.0303
			(0.117)
lag_kg			0.0276
			(0.204)
lag_ner_es			0.287
			(0.216)
lag_firmdens			–35.37*
			(20.27)
lag_shrwage			–4.942*
			(2.973)
lag_shragrempl			0.0488
			(0.0616)
Constant	–10.92	17.56***	–47.32
	(8.757)	(6.325)	(30.32)
Year FE	YES	YES	YES
Province FE	YES	YES	YES
Observations	352	352	161
R-squared	0.816	0.803	0.725
F-test	YES	YES	YES
Number of provinces	23	23	23

Notes: Standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendix A3.4.B3 Sample Excluding Bundled Jakarta
(Use Initial GRDP 1990 as Robustness Check)

Variables	(1) Model_b31 gini	(2) Model_b32 gini	(3) Model_b33 gini
lag_city300	11.61***	9.436***	17.09
	(2.13)	(3.65)	(12.70)
lag_city300^2	–4.038***	–0.530	–3.644
	(1.20)	(1.72)	(5.39)
log_grdp90	18.29***	4.669	3.637
	(3.05)	(5.28)	(17.70)
sqr_log_grdp90	–3.474***		
	(0.95)		
city300 grdpc*		–2.009**	–1.549
		(0.96)	(3.36)
lag_ecogrowth			0.0808
			(0.12)
lag_ki			–0.00124
			(0.14)
lag_kg			0.170
			(0.24)
lag_ner_es			0.336
			(0.25)
lag_firmdens			–20.83
			(22.05)
lag_shrwage			–2.498
			(3.45)
lag_shragrempl			0.0213
			(0.07)
Constant		10.06	–18.41
		(16.06)	(60.37)
Year FE	YES	YES	YES
Province FE	NO	NO	NO
Number of Observations	332	313	141
R-squared	0.826	0.837	0.692
F-test	YES	YES	YES
Number of provinces	19	19	19

Notes: Standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendix A3.4.C1 Full Sample Including Unbundled Jakarta and above 300k Inhabitants (Basic Model)

Variables	(1) Model_c11 gini	(2) Model_c12 gini	(3) Model_c13 gini	(4) Model_c14 gini	(5) Model_c15 gini
lag_city300	0.746	3.862^{***}	–0.156	6.055^{***}	3.800
	(0.46)	(1.15)	(1.16)	(2.01)	(5.45)
lag_city300^2		-0.250^{***}	–0.00425	0.155^{*}	0.259^{*}
		(0.08)	(0.08)	(0.09)	(0.15)
lag_lgrdpc			14.64^{***}	-2.156^{*}	8.750^{**}
			(2.84)	(1.18)	(3.84)
lag_lgrdpc^2			-2.225^{***}		
			(0.32)		
*city300*grdpc*				-1.669^{***}	–1.468
				(0.41)	(1.17)
lag_ecogrowth					0.0814
					(0.09)
lag_ki					0.0992
					(0.12)
lag_kg					–0.195
					(0.20)
lag_ner_es					0.309
					(0.21)
lag_firmdens					-33.50^{***}
					(12.46)
lag_shrwage					-8.457^{***}
					(2.73)
lag_shragrempl					0.0919
					(0.062)
Constant	29.09^{***}	26.23^{***}	7.094	36.86^{***}	–24.90
	(0.80)	(1.25)	(6.62)	(4.61)	(25.01)
Year FE	YES	YES	YES	YES	YES
Province FE	YES	YES	YES	YES	YES
Number of Observations	397	397	373	373	171
R-squared	0.718	0.725	0.787	0.768	0.738
F-test	YES	YES	YES	YES	YES
Number of provinces	24	24	24	24	24

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Appendix A3.4.C2 Full Sample Including Unbundled Jakarta and above 300k Inhabitants (Add Primacy)

Variables	(1) model_c21 gini	(2) model_c22 gini	(3) model_c23 gini
lag_city300	0.685	6.783^{***}	6.412
	(1.416)	(2.216)	(6.069)
lag_city300^2	–0.0322	0.106	0.109
	(0.0905)	(0.0999)	(0.206)
lag_lgrdpc	14.18^{***}	$–2.107^{*}$	9.246^{**}
	(2.860)	(1.206)	(3.969)
lag_lgrdpc^2	$–2.161^{***}$		
	(0.326)		
lag_primacy	0.233	0.278^{*}	0.498
	(0.157)	(0.164)	(0.792)
lag_primacy^2	–0.0015	$–0.002^{*}$	–0.003
	(0.001)	(0.001)	(0.006)
city300_grdpc		$–1.572^{***}$	–1.412
		(0.422)	(1.194)
lag_ecogrowth			0.0654
			(0.0915)
lag_ki			0.0766
			(0.120)
lag_kg			–0.142
			(0.214)
lag_ner_es			0.249
			(0.223)
lag_firmdens			$–35.12^{***}$
			(12.61)
lag_shrwage			$–8.943^{***}$
			(2.777)
lag_shragrempl			0.0956
			(0.0622)
Constant	–1.541	25.08^{***}	–41.97
	(10.09)	(9.388)	(33.54)
Year FE	YES	YES	YES
Province FE	YES	YES	YES
Number of Observations	373	373	171
R-squared	0.789	0.770	0.740
F-test	YES	YES	YES
Number of provinces	24	24	24

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Appendix A3.4.C3 Full Sample Including Unbundled Jakarta and above 300k Inhabitants (Use Initial GRDP 1990)

Variables	(1) Model_c31 gini	(2) Model_c32 gini	(3) Model_c33 gini
lag_city300	0.673**	–2.501**	4.559
	(0.32)	(1.09)	(7.81)
lag_city300^2	–0.0402	–0.0302	0.0464
	(0.04)	(0.04)	(0.33)
log_grdp90	–7.062**	–0.609*	2.074
	(2.84)	(0.36)	(16.78)
sqr_log_grdp90	0.875**		
	(0.34)		
city300 grdpc*		0.749***	–0.867
		(0.28)	(1.23)
lag_ecogrowth			0.0410
			(0.13)
lag_ki			0.0517
			(0.14)
lag_kg			–0.0306
			(0.25)
lag_ner_es			0.345
			(0.25)
lag_firmdens			–29.33**
			(13.96)
lag_shrwage			–8.225***
			(3.19)
lag_shragrempl			0.0629
			(0.07)
Constant	43.18***	32.28***	–3.796
	(5.52)	(1.56)	(57.38)
Year FE	YES	YES	YES
Province FE	NO	NO	NO
Number of Observations	353	334	151
R-squared	0.772	0.753	0.711
F-test	YES	YES	YES
Number of provinces	20	20	20

Notes: Standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendix A3.4.D1 Full Sample Including Unbundled Jakarta and below 300k Inhabitants (Basic Model)

Variables	(1) Model_d11 gini	(2) Model_d12 gini	(3) Model_d13 gini	(4) Model_d14 gini	(5) Model_d15 gini
lag_allcity	–1.954	2.340	–12.20***	10.98**	–47.89***
	(2.71)	(4.80)	(4.46)	(4.55)	(16.21)
lag_allcity^2		–3.783	10.53***	17.63***	–5.906
		(3.49)	(3.44)	(4.28)	(9.06)
lag_lgrdpc			20.51***	3.344***	–0.370
			(2.97)	(0.96)	(2.44)
lag_lgrdpc^2			–2.604***		
			(0.37)		
allcity * *grdpc*				–7.276***	12.39**
				(1.15)	(5.09)
lag_ecogrowth					0.0747**
					(0.03)
lag_ki					0.0102
					(0.08)
lag_kg					0.145
					(0.11)
lag_ner_es					0.260**
					(0.10)
lag_firmdens					–10.71
					(10.61)
lag_shrwage					–2.374
					(2.03)
lag_shragrempl					0.00533
					(0.04)
Constant	30.05***	29.34***	–7.153	19.38***	9.318
	(1.14)	(1.32)	(5.99)	(3.53)	(13.25)
Number of Observations	588	588	550	550	254
R-squared	0.660	0.660	0.734	0.730	0.725
F-test	YES	YES	YES	YES	YES
Number of provinces	32	32	32	32	32

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Appendix A3.4.D2 Full Sample Including Unbundled Jakarta and below 300k Inhabitants (Add Primacy)

Variables	(1) model_d21 gini	(2) model_d22 gini	(3) model_d23 gini
lag_citysize	–4.262	18.66***	–42.14**
	(5.171)	(4.939)	(16.76)
lag_citysize^2	9.446***	17.91***	–4.509
	(3.556)	(4.483)	(9.556)
lag_lgrdpc	18.48***	2.404**	–0.783
	(3.012)	(0.976)	(2.504)
lag_lgrdpc^2	–2.448***		
	(0.371)		
lag_primacy	0.412***	0.520***	0.363*
	(0.126)	(0.127)	(0.215)
lag_primacy^2	–0.003***	–0.004***	–0.003
	(0.0008)	(0.0008)	(0.0016)
citysize_grdpc		–7.368***	11.56**
		(1.170)	(5.280)
lag_ecogrowth			0.0709**
			(0.0339)
lag_ki			0.0333
			(0.0803)
lag_kg			0.137
			(0.111)
lag_ner_es			0.236**
			(0.104)
lag_firmdens			–9.593
			(10.61)
lag_shrwage			–1.801
			(2.055)
lag_shragrempl			0.00375
			(0.0435)
Constant	–18.26**	2.031	0.204
	(7.567)	(5.940)	(14.90)
Year FE	YES	YES	YES
Province FE	YES	YES	YES
Number of Observations	550	550	254
R-squared	0.742	0.740	0.729
F-test	YES	YES	YES
Number of provinces	32	32	32

Notes: Standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendix A3.4.D3 Full Sample Including Unbundled Jakarta and below 300k Inhabitants (Use Initial GRDP 1990)

Variables	(1) Model_d31 gini	(2) Model_d32 gini	(3) Model_d33 gini
lag_allcity	–2.215	6.262*	–25.07
	(3.21)	(3.33)	(17.30)
lag_allcity^2	1.384	14.92***	1.916
	(2.04)	(2.55)	(9.59)
log_grdp90	–0.0287	0.651	–18.46***
	(13.26)	(1.13)	(2.93)
sqr_log_grdp90	–0.00835		
	(1.64)		
*citysize * grdpc*		–5.685***	6.795
		(0.74)	(5.35)
lag_ecogrowth			0.0741**
			(0.03)
lag_ki			–0.0219
			(0.09)
lag_kg			0.203*
			(0.11)
lag_ner_es			0.248**
			(0.11)
lag_firmdens			–11.41
			(11.71)
lag_shrwage			–2.422
			(2.29)
lag_shragrempl			0.0109
			(0.05)
Constant	30.45	29.41***	86.70***
	(25.79)	(4.12)	(15.04)
Year FE	YES	YES	YES
Province FE	NO	NO	NO
Number of Observations	481	457	206
R-squared	0.782	0.811	0.745
F-test	YES	YES	YES
Number of provinces	24	24	24

Notes: Standard errors in parentheses; *** p<0.01, ** p<0.05, * p<0.1.

Chapter 4 Regional Inequality, Growth and Institutional Quality in Indonesia

4.1 Introduction

In this chapter, we explore the relationship between income inequality and economic growth in Indonesia. The descriptive analysis in Chapter 2 showed that, over the last decades, relatively high economic growth in Indonesia has been associated with rapidly increasing income inequality. We also found that the incomes of the rich in poor regions grow faster than those of the poor in rich regions. In other words, the middle class and especially the top incomes seem to benefit most from the economic growth dynamics in Indonesia. In Chapter 3, we asked the question of whether and to what extent the number, size and distribution of cities might explain the evolution of inequality across provinces in Indonesia. Our examination of the relationship between income inequality and average agglomeration size at the province level revealed an inverted U-shaped relationship, suggesting that inequality can be expected to climb further if the average agglomeration size continues to grow.

The relationship between income inequality and economic growth is one of the most fundamental questions in economics. Hence, a large empirical literature has been devoted to the impact of inequality on growth. As noted by Banerjee and Duflo (2003, p. 267): "Many have felt compelled to try to say something about this very important question, braving the lack of reliable data and the obvious problems with identification". Indeed, it has proved to be very difficult to determine whether (rising) income inequality in a country is good or bad for its economic growth performance.

There are at least three main theoretical views on this relationship, namely (i) a negative (linear) inequality-growth relationship, (ii) a positive (linear) inequality-growth relationship and (iii) a non-linear relationship between inequality and growth. For a long time, most of the empirical research routinely imposed a linear structure on the data, whereby the different variants of the basic linear model (OLS, fixed effects and random effects) generated very different conclusions on the inequality-growth relationship (see, for example, Deininger and Squire 1996; Bruno et al. 1998; Li and Zou 1998; Barro 2000; Benabou 2000; Benhabib and Spiegel 2000; Forbes 2000; Nel 2008; Benjamin et al. 2011). In an important paper, Banerjee and Duflo (2003) developed a simple political economy model to argue that there are no *a priori* theoretical reasons to assume that the relationship between inequality and growth is monotonic, let alone linear, and that the omitted variable problem inherent in identifying a causal effect of inequality on growth can be solved by including a country fixed effect in a linear specification (as in Li and Zou 1998; Forbes 2000). In contrast, using a cross-country panel data analysis, they showed that imposing a linear structure on data for which there is no theoretical support can lead to serious misinterpretations. In particular, they found that changes in inequality (in any direction) are associated with lower future growth rates and that there is a non-linear relationship between inequality and the magnitude of changes in inequality. They also identified a negative relationship between growth rates and inequality lagged by one period. But, they cannot conclude that there is a causal relationship between inequality and growth in cross-country data because of identification problems. One of their key findings is that the non-linearity approach is sufficient to explain why previous estimates of the relationship between the level of inequality and growth are inconsistent from one to another.

In this chapter, we apply the cross-country panel data approach developed by Banerjee and Duflo (2003) to cross-district data within Indonesia for the period 1999–2014 to find the correlations between inequality and the growth rates. In Chapter 2, we found evidence for the existence of a typical Kuznets-like inverted U-shaped relationship between inequality and GDP at the district level in Indonesia: as the per capita GDP level increases, the level of inequality first increases and then decreases when the per capita GDP level increases beyond a certain threshold. Often, the implicit policy implication taken from this inverted U-shaped pattern is that inequality is a necessary and temporary part of the path to a higher level of GDP: it is the price that a society has to pay for the desired economic growth, while the problem (automatically) vanishes at higher welfare levels. However, in his path-breaking book, Piketty (2014) cast serious doubt on this optimistic view inherent to the Kuznets hypothesis. Based on unique long-term historical records, Piketty (2014) argued that inequality will not automatically disappear or reduce when a society accumulates wealth. The inverted U-shaped relationship between inequality and growth that has been observed so frequently may well be a historical exception, whereas the long-term regularity is that wealth accumulation is accompanied by rising inequality.

Against this background, in this chapter, we use the approach developed by Banerjee and Duflo (2003) to perform cross-district empirical tests in Indonesia to determine whether changes in inequality in the past associate with economic growth in the future. Dominicis et al. (2008) pointed out that the analysis of the impact of inequality on economic growth is more useful using single-country data at the regional level controlling economic, social and institutional characteristics. We also test for the existence of a non-linear relationship between inequality and the magnitude of changes in inequality and a relationship between growth rates and inequality lagged by one period. In doing so, we control for the role of institutional quality at the district level in our analyses. More precisely, we investigate whether institutional quality together with past inequality correlates with economic growth in the next period at the district level in Indonesia. Recent studies at the country level have shown that high institution quality leads to higher economic growth (Acemoglu and Robinson 2010) and tends to lower inequality (Chong and Gradstein 2007).

Evidently, increasing welfare in a society cannot be measured adequately using the per capita GDP alone. Many things other than income matter in improving people's life, such as the quality of and access to health care and the road, water, sanitation and electricity infrastructure. We therefore conclude our analysis by using the approach developed by Banerjee and Duflo (2003) to test empirically whether changes in inequality in the past correlate with the evolution of various development indicators in the future across districts in Indonesia.

The remainder of this chapter is organized as follows. Section 4.2 discusses the existing literature on the relationship between inequality, growth and institutional quality. In section 4.3, we present the econometric approach underlying our analyses. Section 4.4 describes the data set that we developed to test the hypotheses mentioned above. Section 4.5 presents and discusses the estimation results. Section 4.6 includes several robustness checks and sensitivity tests. Section 4.7 provides concluding remarks.

4.2 Literature Review

The relationship between inequality and growth has been examined for decades. The first seminal contribution was by Kuznets (1955), who argued that inequality first increases and later decreases during the process of economic development (the inverted U-shaped relationship). Inequality may increase again beyond a certain threshold, producing an N-shaped relationship. Piketty and Saez (2003) and Atkinson et al. (2011), in their studies of top incomes, implied that, if inequality is already high, the accumulation of wealth among the rich increases over time and strengthens and enhances inequality further. This high inequality could harm the future growth potential. Perotti (1996) also suggested that the relationship between income distribution and growth is negative because more equal societies are characterized by lower fertility rates and higher rates of investment in education, which are both reflected in higher growth rates.

Furthermore, Galor and Moav (2004) found that inequality is positively correlated with growth in poorer countries as it contributes to channelling resources to individuals with a higher marginal propensity to save, hence stimulating investments. While physical capital is a prime engine for growth in early stages of industrialization, in later stages, human capital becomes critical for economic growth and the relative return to physical capital decreases. Thus, the impact of inequality on growth switches from positive to negative (Galor and Moav 2004). In line with these findings, Barro (2000) confirmed that the impact of inequality on growth might be dependent on the stage of development, finding that inequality is negatively correlated with growth in a sample of poor countries but positively correlated with growth in a sample of rich countries. Li and Zou (1998) and Forbes (2000) also suggested that the relationship between the level of income inequality and the future economic growth is positive in the short run and negative in the long run.

Finding inconsistency in the results of the relationship between inequality and growth, Banerjee and Duflo (2003) searched for a non-linear relationship. They discovered that growth is an inverted U-shaped function of net changes in inequality. Changes in inequality (in any direction) are associated with reduced growth in the subsequent period. This implies that an economy's growth rate is maximized when there are no changes in inequality and that any deviation of inequality, in any direction, lowers growth.

Most of the literature on the relationship between institutional quality and economic growth has pointed at a positive correlation (Knack and Keefer 1995; Rodrik et al. 2004; Acemoglu and Robinson 2010). However, this positive association is not applicable to all regions or groups. It depends on the perception of the individual about the institutions, social norms and community rules (Nawaz et al. 2014). The role of the stages of economic development also determines the quality of institutions, then creating a virtuous circle between growth and institutional quality (Alonso and Garcimartín 2013). Hence, Valeriani and Peluso (2011) and Nawaz et al. (2014) concluded that institutions perform better in developed countries than in developing ones.

To understand the role of institutions in economic growth, rent-seeking (institutional mechanism) theory explores the situation in which the gap between the rich and the poor widens. This may result in a stronger temptation for the poor to engage in rent seeking or predatory activities at the expense of the rich (Benabou 1996). Other researchers have also proposed an institutional mechanism in which a rich elite will suppress democracy and equal rights before the law to preserve their privileged position (e.g. Bourguignon and Verdier 2000). Poor-quality institutions may slow down economic activities by providing room for economic agents to remain busy in redistributive politics with lower economic returns rather than growth-promoting economic activities (Murphy et al. 1993). Weak institutions divert resources from productive sectors to unproductive sectors, hence promoting rent-seeking activities (Iqbal and Daly 2014).[14] In contrast, strong institutions reduce the chances of rent-seeking activities and accelerate the economic growth process and productivity of reproducible factors.

On the basis of a cross-country analysis, Nel (2008) suggested that a high level of economic inequality undermines a country's growth potential, retards the development of social capital, and encourages corruption. Chong and Gradstein

[14] They defined a rent-seeking activity as an activity through which public power is exercised for private gain; this may involve misuse of public resources or, more generally, any attempted capture and commodification of state, social or commercial authority by politicians, public officials, elites and private interests.

(2007) studied that the double causality between institutional strength and a more equal distribution of income is empirically established using dynamic panel and linear feedback analysis. Countries with bad institutions also seem to be more likely to have high inequality. Hence, the economy may converge either to a steady state with high-quality institutions and minimal inequality or to low institutional quality and high inequality.

The current concern in Indonesia is much more that the increase in inequality is reflecting an increase in poverty. It is driven by the dynamics at the bottom of the distribution. From the limited studies that are available, we discuss some notable findings on inequality and growth in Indonesia. Timmer (2004) indicated that rapid pro-poor growth requires simultaneous and balanced interaction between growth and distribution processes. Using panel provincial data from 1993 to 2002, Resosudarmo and Vidyattama (2006) found that the disparity in provincial income per capita in Indonesia is relatively severe and that there is a conditional growth convergence in Indonesia. Moreover, Vidyattama (2013) showed that the inequality at the district level is considerably higher than that at the provincial level and that increasing inequality in the per capita GDP does not necessarily translate into increasing inequality in other development indicators. Van Leeuwen and Földvári (2016) found that rising inequality had a positive impact on the per capita GDP prior to the 1940s but that, during the period 1950–1980, when declining inequality was accompanied by an increased per capita GDP, poverty rates rapidly increased.

4.3 Empirical Strategy

We adopt the approach developed by Banerjee and Duflo (2003) to explain the basic correlation between income inequality and growth of the per capita GDP across districts using a non-linear function, and then we modify the model by introducing the quality of governance or institutions and exploiting different indicators as control variables (subject to data availability). To seek those relationships, we perform fixed-effect panel data regressions to consider the variation within districts over time based on a district panel dataset over the period 1999–2014.

4.3.1 Basic Model

We briefly discuss the theoretical model based on "hold-ups", developed by Banerjee and Duflo (2003), which we employ in our estimations. Suppose that there are two groups in society, A and B, that are competing as political groups, and assume that the economy is characterized by the sharing rule. In each period, this economy is presented with an opportunity that can lead to growth. However, the growth opportunity requires some policy adjustment to be implemented, which has to be agreed by both groups in society; that is, one group can block the realization of the growth opportunity and demand a "bigger share" of the growth. If one group blocks, then growth is reduced. Assume that group B has the chance to hold up the economy, whether or not it depends on how much additional "growth" it can extract for itself from group A. Group B will demand a transfer if the post-transfer portion of the growth is bigger than it would be without holding up the economy. Group A will agree to the transfer as long as the post-transfer share of new growth is at least as big as the old share. If group A disagrees with the transfer, the status quo is maintained and there is no growth.

Intuitively, each group in the economy will hold up the rest of the economy when its share of the output is low, which is when it has the smallest stake in the growth of the overall economy. Hence, hold-ups only happen when there are redistributive transfers that result from them. Assuming that the growth will be higher when the planned transfer is zero, we can expect a smooth inverted U-shaped relationship between expected growth and actual changes in inequality. The growth here does not have any direct distributional effect. Thus, the basic non-linear relationship in a reduced form appears as follows:[15]

$$(y_{it+a} - y_{it})/a = \alpha y_{it} + X_{it}\beta + k(g_{it} - g_{it-a}) + v_i + \varepsilon_{it} \qquad (4.1)$$

[15] A detailed discussion on the theoretical model can be found in Banerjee and Duflo (2003, pp. 269–278).

where y_{it} denotes the logarithm of the per capita GRDP in district i in year t, a is the length of the period that we choose,[16] $(y_{it+a} - y_{it}) / a$ is therefore the subsequent growth rate, X_{it} is a set of control variables, g_{it} is the Gini coefficient in district i in year t and $k(.)$ is a generic function. The error term is modelled as a district-specific time-invariant effect (v_i) and a time-varying error term (ε_{it}). y_{it} is included among the controls to capture the convergence effects, and X_{it} controls for possible sources of spurious correlation. Equation (4.1) is estimated with flexible specifications: (i) the first relationship relates to the square of the change in inequality to the level of inequality:

$$(g_{it+a} - g_{it})^2 = \alpha y_{it} + X_{it}\beta + h_2(g_{it-a}) + v_i + \varepsilon_{it}. \quad (4.2)$$

The second relationship is a "reduced-form relationship", which relates the level of inequality (lagged by one period) to the growth rate:

$$(y_{it+a} - y_{it})/a = \alpha y_{it} + X_{it}\beta + h(g_{it-a}) + v_i + \varepsilon_{it}. \quad (4.3)$$

Moreover, we consider occasional shocks that increase inequality: exogenous shocks that increase inequality and therefore reduce growth and endogenous reductions in inequality that are also associated with a fall in the growth rate. This means that measured changes in inequality in either direction will be associated with a fall in growth, suggesting that the equation to estimate should include both a direct effect of the level of inequality and an effect of changes in inequality. Hence, the final basic model is expressed in equation (4.4), where $(g_{it} - g_{it-a})$ is the change in inequality and $h(g_{it})$ is the initial level of inequality that matters for the growth rate in the short run:

$$(y_{it+a} - y_{it})/a = \alpha y_{it} + X_{it}\beta + k_1(g_{it} - g_{it-a}) + k_2(g_{it} - g_{it-a})^2 + + h(g_{it}) + v_i + \varepsilon_{it}. \quad (4.4)$$

[16] We choose three years (a = 3) as the length of the period due to data availability and to avoid seasonal effects. Our data period is from 1999 to 2014 for 242 districts throughout Indonesia. A detailed explanation of the data sources is presented in section 4.4.

We also follow the choice of control variables based on the Perotti model as employed by Banerjee and Duflo. However, due to data availability at the district level, we simplify the control variables of Perotti by employing only two out of four variables, that is, the initial level of per capita GDP y_t to express the convergence in the economy and the total education $educ_t$ as in Perotti specification without gender classification to capture human capital.

By replicating the above models, we estimate the growth of the per capita GDP (including and excluding oil) as a function of the initial level of inequality, the change in inequality, the square of this change to allow for non-monotonicity and a set control variables suggested by Perotti as shown in equation (4.4). Our basic model is now written in the following fashion:

$$g_{it} = \beta_0 + \beta_1 \cdot Gini_{it} + \beta_2 \cdot \Delta Gini_{it} + \beta_3 \cdot [\Delta Gini_{it}]^2 + \beta_i \cdot X_{it} + \varepsilon_{it}. \quad (4.5)$$

Our dependent variable is the growth rate (g_{it}), an annual average of the difference between the initial income and the three-year lagged income. Our interest variables are initial inequality ($Gini_{it}$), the change in inequality obtained by calculating the difference between the initial Gini and the three-year-lagged Gini in each period ($\Delta Gini_{it}$) and the square of this change ($\Delta Gini_{it})^2$. All of these equations are estimated using a fixed-effect specification to control for non-observed heterogeneity across districts that does not vary over time.

Furthermore, we estimate the growth of the per capita GDP (including and excluding oil) with the one-period-lagged inequality, the square of this lag to capture non-monotonicity and a lagged set of control variables suggested by Perotti, as shown in equation (4.3). We expect that the level of inequality (in the lagged form) is also negatively correlated with the growth, which implies that increasing inequality is associated with lower subsequent growth. Our lagged model is:

$$g_{it} = \beta_0 + \beta_1 \cdot Gini_{it-a} + \beta_2 \cdot [Gini_{it-a}]^2 + \beta_3 \cdot X_{it-a} + \varepsilon_{it}. \quad (4.6)$$

4.3.2 Modified Basic Model: Role of Institutional Quality

As a consequence of decentralization in Indonesia since 2001, public service delivery is now under local government jurisdiction, that is, provinces and districts/cities. Local elections have been implemented since 2004. These two events make the role of more than 500 districts in the governance process crucial and more strategic. Thus, the quality of institutions, as represented by audit performance, budget efficiency and local governance variables at the district level, also becomes an important factor in analysing the relationship between inequality and growth. Moreover, we hypothesize that increasing inequality is less harmful if proper institutions are in place.

We therefore introduce the quality of institutions into our basic model (4.5) as a variable to test whether the quality of institutions and its interactions with various forms of inequality correlate with the growth rate. Our modified model is shown below:

$$g_{it} = \beta_0 + \beta_1 \cdot Gini_{it} + \beta_2 \cdot \Delta Gini_{it} + \beta_i \cdot [\Delta Gini_{it}]^2 + \beta_4 \cdot Inst_{it} + \beta_5 \cdot (Gini_{it} \cdot Inst_{it}) + \beta_6 \cdot (\Delta Gini_{it} \cdot Inst_{it}) + \beta_7 \cdot ([\Delta Gini_{it}]^2 \cdot Inst_{it}) + \beta_i \cdot X_{it} + \varepsilon_{it}. \quad (4.7)$$

We exploit three institutional quality variables ($Inst_{it}$) in this model, namely audit performance (*audit*), budget efficiency (*budget*) and local regulation index (*locreg*). More details regarding the institutional quality indicators will be discussed in the next section. We also interact the initial level of inequality, the change in inequality and the square of the change in inequality with each of the institutional quality variable as follows: ($Gini_{it} \cdot Inst_{it}$), ($\Delta Gini_{it} \cdot Inst_{it}$) and ($\Delta Gini_{it}^2 \cdot Inst_{it}$). We apply the same control variables as in the basic model. We expect that institutional quality will have a positive correlation with the growth rate.

4.3.3 Development Outcome

To understand further the relationship between inequality and development outcomes, we test the basic and modified models with six development indicators, specifically access to health, asphalt roads, sanitation, clean water and electricity and poverty. We primarily estimate the basic model (4.5) and the model with institutional quality (4.7) with some modifications. In estimating the growth of development outcomes, we add initial development indicators as new control variables. For instance, if our dependent variable is the growth of access to health, then the control variables are initial health, initial education and initial per capita GDP (excluding oil), as displayed in Appendix A4.1.

4.4 Data Description

This research covers 242 districts throughout Indonesia. We exclude Aceh, Maluku and Papua due to a lack of data and then recode all the amalgamated districts back to the parent districts to obtain a more complete and consistent dataset (see Appendix A2.1 on constructing the true panel dataset). We utilize secondary data from various sources and levels, such as the Central Bureau of Statistics Indonesia (BPS) for the household and village levels and the Audit Board of Indonesia (BPK), the Ministry of Finance and Regional Autonomy Watch (KPPOD) for the district level. In this model, we define income inequality in various forms and institutional quality as interest variables for the right-hand side and growth as the dependent variable.

Per Capita GDP (Excluding Oil)

We use the deflated per capita GDP in log form to calculate growth. The earliest year of per capita GDP at the district level in our dataset is 1999, so we define 1999 as the initial year. We also use the per capita GDP excluding oil in our model to see how important oil is in the Indonesian economy, particularly in the relationship with inequality and development indicators. The per capita GDP and per capita GDP excluding oil are available at the district level in figures issued by the Indonesian Central Statistics Bureau (BPS) from 1999 to 2014.

Income Inequality

The Gini coefficient is widely used as a proxy for income inequality – a lower Gini means lower inequality. We calculate this coefficient from the total monthly household expenditures and aggregate it to the district level. The source of the expenditure data is the Susenas consumption module issued by BPS from 1999 until 2014. Like the per capita GDP, the earliest year for which the Gini coefficient is available at the district level in our dataset is 1999, so we define 1999 as the initial year.

Institutional Quality

We measure the quality of institutions through three indicators: audit performance, budget efficiency and local regulation index. These indicators come from different sources so the area and time coverage are different from one to another. For instance, audit results and budget efficiency are available since 2005 and 2002, repectively, while local regulation index are available from 2001 until 2011. Thus, missing values in particular years are unavoidable.

Audit Results (Audit): We use audit results as a proxy for institutional quality. The Audit Board of the Republic of Indonesia (BPK–RI) performs annual assessments on the compliance of all the government institutions throughout Indonesia in administering their expenditure and programme implementation. However, the results for local governments at the district level have only been released since 2005. The audit results are categorized into four opinions as follows: unqualified opinions as the best result, qualified opinions, adverse opinions and disclaimers as the worst result. We code these results from –1 for the best to –4 for the worst audit result. A higher audit result means better budget administration and programme implementation and thus better institutions.

Budget Efficiency (Budget): We measure budget efficiency from both the expenditure and the revenue side of the local government budget (APBD). The disbursement rate of the local government budget is the percentage of the total expenditure realizations of its plan, while the realization rate of own source revenue[17] is calculated as the

[17] Own source revenue is the revenue generated from within the local area itself. It consists of local taxes, including land and property tax, permits and licenses, fees for public services, fines from breaking rules and revenues from local state-owned enterprises (Law No. 28, 2009 on Local Taxes and Levies).

percentage of actual own source revenue of its target. These indicators show how efficient the local government is in spending its budget based on its plan and in collecting revenue from its own sources, such as local taxes and levies. A high disbursement rate (and high realization rate) could be perceived as indicating a better capacity of local governments to spend money (and earn money) as they planned (and targeted), and hence these indicators represent better institutions.

The source of these data is the Ministry of Finance from 2002 to 2014. However, there are outliers in our budget data as the maximum value of the disbursement or absorption rate should not exceed 100 per cent while our data record the disbursement and/or the absorption rate as much higher than 100 per cent – the maximum budget efficiency is 971.7 per cent (Appendix A4.2). These outliers are due to the budgeting system in Indonesia, which allows the government to reset its target for disbursement and absorption in the second half of the fiscal year. Thus, when the Ministry of Finance publishes these data, the disbursement or absorption rate refers to the status on 31 December or at the end of the fiscal year while the plan or the target refers to the status on 1 January or at the beginning of the fiscal year. Without outliers, the average budget disbursement rate is 90.7 per cent, with a minimum of 6.4 per cent.

Local Regulation Index (locreg). This index measures businesses' perception of the local regulation related to business climates, such as the types of licenses and permits required to start a business, cost, time spent, number of required documents to obtain a business license and so on. A higher index means that the local government has a greater ability to respond to what matters to the local businesses, so it captures a better government institution. This index is extracted from the survey on governance and the investment climate conducted by Regional Autonomy Watch (KPPOD) from 2001 until 2011. However, the coverage area of the survey varies in each round, so there are missing values in several districts.

Development Outcome

This variable consists of six indicators, namely access to health, sanitation, water, roads and electricity and the poverty rate. We calculate the percentage of births attended by certified health workers as the health indicator, the percentage of households with access to sanitation, the percentage of households with access to clean water, the percentage of households with access to electricity and the percentage of villages with asphalt roads. The source of data for all the indicators is extracted from the Susenas core module at the household level, except the percentage of villages with asphalt roads, which is calculated from Potential Village (Podes)[18] at the village level, and then all is aggregated at the district level. The poverty rate as the development outcome is obtained from District in Figures and is also provided by the Indonesian Central Statistics Bureau (BPS) from 1999 to 2014.

Furthermore, we use initial education as a control variable. We calculate the percentage of the population above 30 years old with a senior high school diploma as a proxy for education. These data are also available in our dataset from 1999, so we define 1999 as the initial year for education.

4.5 Results and Discussion

This section presents and discusses the results of the relationship between inequality and growth and the relationship with institutional quality. We also discuss the results of the above models regarding development outcomes. In general, the relationship between change in inequality and subsequent growth across districts in Indonesia supports the non-linearity hypothesis in the basic and the lagged model developed by Banerjee and Duflo (2003).

4.5.1 Basic Model

Table 4.1 presents the estimation results of the basic model. The initial level of inequality has a positive coefficient, while the change in inequality has the opposite correlation with the growth in the subsequent period. Both are statistically significant. This implies that decreasing current inequality associates with higher growth in the next period. If we add the square of the change in inequality to allow for non-monotonicity, the coefficient of the squared term is negative but statistically

[18] Since Podes is collected every three years, we interpolate the missing values of each variable within these three years.

insignificant; hence, the relationship between growth and lagged inequality growth shows an inverted U-shaped pattern, 95% CI [–1.81, 1.27], as shown in Figure 4.1.

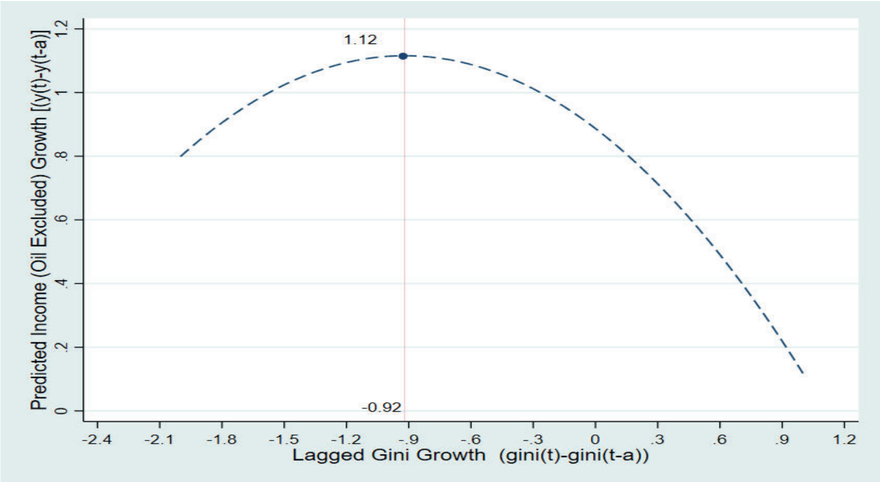
Table 4.1 Relationship between Inequality and Change in Inequality and Growth (Basic Model)

Variables	Per Capita GDP (Excluding Oil) Dependent Variable: g_{it}				
	1a	2a	3a	4a	5a
$Gini_t$	0.427^{***}	0.877^{***}	0.887^{***}	0.889^{***}	0.896^{***}
	(0.06)	(0.09)	(0.10)	(0.10)	(0.10)
$\Delta Gini_{it}$		-0.503^{***}	-0.498^{***}	-0.441^{***}	-0.732^{***}
		(0.07)	(0.07)	(0.09)	(0.24)
$(\Delta Gini_{it})^2$			–0.271	–0.469	–2.680
			(0.78)	(0.82)	(1.85)
$\Delta Gini_{(it)} > 0$				–0.007	–0.008
				(0.01)	(0.01)
$\Delta Gini_{it} * \Delta Gini_{it} > 0$					0.561
					(0.42)
y_t	-0.100^{***}	-0.155^{***}	-0.156^{***}	-0.156^{***}	-0.158^{***}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
$educ_t$	–0.002	–0.002	–0.002	–0.002	–0.002
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.091^{***}	0.064^{***}	0.063^{**}	0.067^{***}	0.062^{**}
	(0.02)	(0.02)	(0.02)	(0.03)	(0.03)
Number of Observations	1,210	968	968	968	968
R-squared	0.126	0.269	0.269	0.270	0.272
Number of Districts	242	242	242	242	242

Notes: Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

We are also interested in observing the difference between districts with positive and negative changes in inequality in relating with growth. However, when we introduce a dummy variable for positive change ($\Delta Gini_{it}>0$) and its interaction with the change in inequality, the results are not statistically significant. These results suggest that inequality at the district level in Indonesia is still an increasing function although at a decreasing rate (Figure 4.1). A negative change in inequality or decreasing current inequality compared with the previous period associates with an increase subsequent growth of 1.12 per cent until it reaches the value of lagged change in inequality of –0.92. After this top point, this negative change turns into a positive change in inequality and harms the growth. In other words, decreasing current inequality is positive correlated with growth up to a certain period and then this higher growth tends to increase inequality, which will be translated into a positive change in inequality as compared with the previous period; thus, the growth becomes slower. Hence, a larger change in inequality is associated with a larger decline in subsequent growth.

Figure 4.1 Relationship between Income Growth and Lagged Inequality Growth



Source: Authors' estimation.

As control variables, the initial per capita GDP (oil excluded) is a negative coefficient, suggesting a convergence story at the district level. A district with a low initial per capita GDP grows faster to catch up with its counterparts. We also use initial education as a control variable. It seems that education does not correlate with the subsequent growth as the coefficient is very small and not statistically significant. This is due to our definition of education, which only captures the share of the population aged over 30 that holds a high school diploma. Since our dataset covers only 15 years, while completing high school takes at least 12 years of schooling, in this case, education is not yet correlated with growth. Similar results are shown for the per capita GDP estimation, as presented in Appendix A4.3.

4.5.2 Lagged Model

Our estimation results for growth with one-period-lagged inequality are similar to those of the previous basic model of income (oil excluded) growth estimation, presented in Table 4.2. It implies that increasing inequality in the linear form is associated with high income (oil excluded) growth in the next period (the first two columns). These results seem to be consistent with the conclusions of Li and Zu (1998), Forbes (2000) and Galor and Moav (2004), suggesting that increasing inequality will boost economic growth without oil. In other words, their conclusions are still valid for the Indonesian case. If we introduce the square of lagged inequality, the negative coefficient exhibits an inverted U-shape, but it has a weak relationship with oil income growth. The relationship between lagged inequality and subsequent growth becomes less important if we use control variables in lagged form, as shown in the last two columns of the per capita GDP (excluded oil) and per capita GDP estimation.

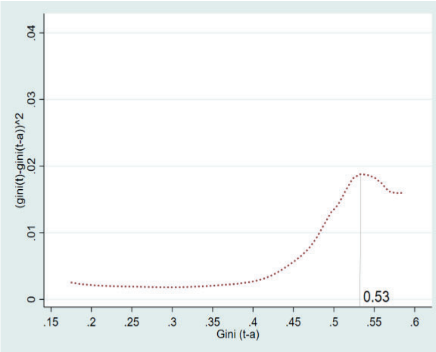
Table 4.2 Relationship between Lagged Inequality and Growth (Lagged Model)

Basic Model	Dependent Variable: g_{it}							
	Per Capita GDP (Excluding Oil)				Per Capita GDP			
	1a	2a	3a	4a	1b	2b	3b	4b
Gini(*t–a*)	0.495***	1.338***	0.208**	0.488	0.324***	0.841**	0.058	0.119
	(0.07)	(0.41)	(0.08)	(0.46)	(0.07)	(0.38)	(0.08)	(0.43)
Gini(*t–a*)^2		–1.406**		–0.466		–0.863		–0.103
		(0.67)		(0.76)		(0.63)		(0.69)
Control Var.	*X*(*t*)	*X*(*t*)	*X*(*t–a*)	*X*(*t–a*)	*X*(*t*)	*X*(*t*)	*X*(*t–a*)	*X*(*t–a*)
No. of Observations	968	968	964	964	968	968	968	968
R-squared	0.234	0.238	0.024	0.025	0.201	0.203	0.021	0.022
Number of Districts	242	242	242	242	242	242	242	242

Notes: Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

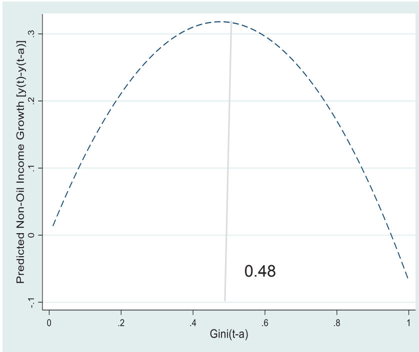
Figure 4.2 presents the non-linear relationship between the lagged inequality and the square of change in inequality, where the lagged inequality reaches the maximum around 0.53. If we look at Figure 4.3, the growth could decrease if the lagged inequality is larger than 0.48, which is close to the value at which the lagged inequality reaches its maximum, as shown in Figure 4.2. These lagged results as well as the non-linear models convince us that the rise in inequality is associated with a slowing of the economic growth.

Figure 4.2 Relationship between Lagged Gini and Square of Gini Changes (Equation 2)



Source: Authors' estimation

Figure 4.3 Reduced Form of the Lagged Model (Equation 3)



Source: Authors' estimation

4.5.3 Modified Model with Institutional Quality

Our estimation results from the modified model with institutional quality show the relationship between change in inequality and its square, and the income (oil excluded) growth in the next period remains the same as in the basic model's results. The positive coefficient for the initial level of inequality in all the specifications suggests that a high initial level of inequality benefits subsequent growth. The change in inequality has a negative correlation with growth, signalling that a decrease in the existing inequality would enhance growth.

With regard to institutional quality, only the local regulation index (*locreg*) has a positive correlation with growth and is statistically significant while audit performance (*audit*) and budget efficiency (*budget*) are positive and negative coefficients, respectively; neither is statistically significant. These suggest that districts with good audit results and budget efficiency do not guarantee high economic growth while the local regulation index could enhance economic growth. However, the interaction between the initial level of inequality and the institutional quality shows different relationship with growth; that is, interaction with audits results in decreasing growth while budget efficiency and the local regulation index support increasing growth (Table 4.3 column 1). Our findings indicate that a high inequality level alone correlates with an increasing growth but that, if the audit performance is poor, increasing inequality would harm growth.

On the other hand, budget efficiency and the local regulation index are important to increase growth when a high initial level of inequality exists. These results suggest that an increase in a level of inequality has positive relationship with subsequent economic growth if better institutional quality is in place, and each of the institutional quality indicators has a different association with the behaviour of inequality in the relationship with economic growth at the district level. We also interact institutional quality with change in inequality as well as its square of change in inequality, but none of the coefficients are statistically significant, as shown in the table below. Similar results are found for the growth of the per capita GDP with stronger coefficient for audit (Table 4.3 column 2).

All the specifications apply the same control variables, that is, initial per capita GDP and education. Similar to the previous model without institutions, a district with a low initial level of per capita GDP correlates with higher growth and education does not yet have a correlation with growth. In short, our results have interesting implications: (i) good institutions alone cannot boost growth; (ii) an initial level of inequality is still required to increase growth together with good institutions for a certain period; (iii) inequality and institutional quality correlate with economic growth depending on the type of institutional quality indicators; (iv) different districts require different sets of institutional quality; and (v) the role of economic development is important in determining growth.

Table 4.3 Summary Relationship between Subsequent Growth and Institutional Quality[19]

	Dependent Variable: $[y(t+a) - y(t)] / a$	
	(1) Per Capita GDP (Excl. Oil)	(2) Per Capita GDP
$audit_{(it)}$	0.002	0.008**
	(0.003)	(0.003)
$budget_{(it)}$	–0.0001	–3.11e–05
	(0.0002)	(0.0001)
$locreg_{(it)}$	0.016***	0.014***
	(0.01)	(0.01)
Interaction Variable		
$Gini_{(it)} * audit_{(it)}$	–0.106***	–0.049*
	(0.03)	(0.03)
$Gini_{(it)} * budget_{(it)}$	0.002***	0.001**
	(0.001)	(0.001)
$Gini_{(it)} * locreg_{(it)}$	0.129***	0.124***
	(0.04)	(0.04)
$\Delta Gini_{(t)} * audit_{(it)}$	–0.0004	–0.055
	(0.07)	(0.06)
$\Delta Gini_{(t)} * budget_{(it)}$	–0.0009	0.0006
	(0.003)	(0.003)
$\Delta Gini_{(t)} * locreg_{(it)}$	–0.064	–0.082
	(0.12)	(0.12)
$(\Delta Gini_{(t)})^2 * audit_{(it)}$	–0.188	–0.584
	(0.85)	(0.79)
$(\Delta Gini_{(t)})^2 * budget_{(it)}$	0.018	–0.006
	(0.04)	(0.04)
$(\Delta Gini_{(t)})^2 * locreg_{(it)}$	2.067	2.206
	(2.18)	(2.15)

Notes: Standard errors in parentheses; *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Nawaz et al. (2014), in their recent study, suggested that institutions are important in determining economic growth but that the impact of institutions on economic growth depends on the level of economic development. Moreover, they claimed that institutions are more effective in developed countries than in developing ones; hence,

[19] Full regression results are available on request.

different countries require different sets of institutions to promote economic growth. Their findings are relevant to our results at the district level in Indonesia.

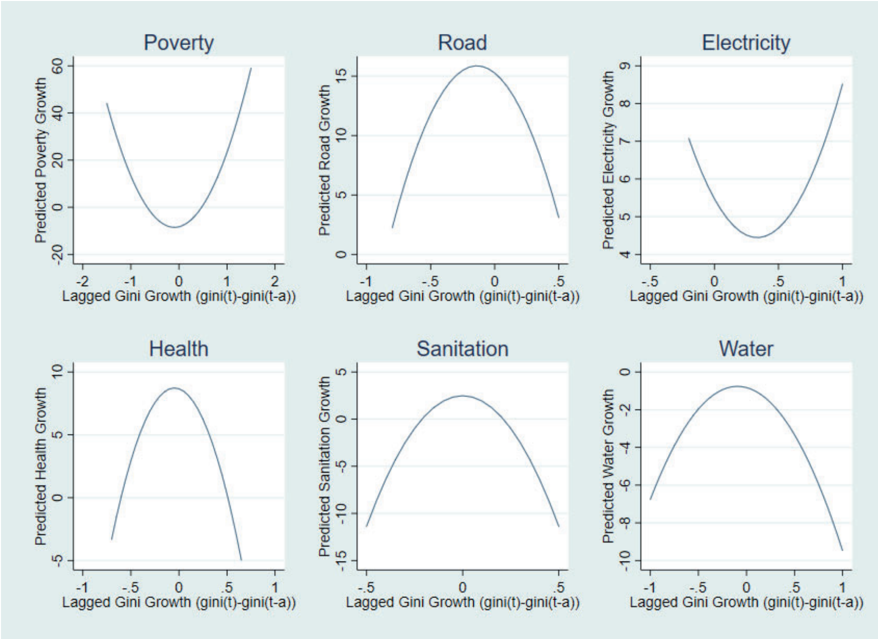
4.5.4 Development Outcomes

Following our basic and modified models in explaining growth from the change in inequality and institutional quality, we also apply these models to other development indicators. Our results suggest that a high initial level of inequality has a positive correlation with the growth of roads and reducing the growth of poverty. If we look at the change in inequality, existing inequality associates with high growth of roads and low growth of poverty, suggesting that high inequality is still expected to alleviate poverty for certain periods. In general, our regression results show an inverted U-shaped in the relationship with change in inequality except for poverty and electricity (Figure 4.4). These development indicators grow as the change in inequality increases but are negatively correlated after inequality reaches a certain point (the top of the curve). Thus, it could be expected that the relationship between existing inequality and the growth of these development indicators is positive at the district level.

High initial level of inequality associates with expanding roads and supporting poverty alleviation. Bourguignon (2004) explained this relationship through the PGI triangle (poverty, growth and inequality).[20] Our results also confirm the convergence story in which a district with a low initial level of development grows faster and the control variable of per capita GDP (excluding oil) plays an important role in speeding up the growth of roads and poverty alleviation (Appendix A4.4).

[20] This triangle highlights the dual routes through which the development strategy can reduce poverty and generate "pro-poor growth": economic growth and improvements in income distribution. These two processes indicate possible causal forces at work in each direction as follows: economic growth might affect income distribution, perhaps widening inequality in the way in which Kuznets (1955) hypothesized, while income distribution might affect economic growth in a negative direction, as Alesina and Rodrik (1994) and Easterly (2002) demonstrated, or in a positive direction, as Forbes (2000) showed.

Figure 4.4 Relationship between Development Growth and Lagged Inequality Growth



Source: Authors' estimation.

Moreover, we apply the modified model to see how institutional quality affects the growth of development outcomes. Our results work only for the growth of roads and poverty alleviation (Table 4.4). Like our basic model results, institutional quality solely has no statistically significant correlation with the growth of roads, but, if it is combined with an initial level of inequality, the relationship between institutional quality and growth of road varies depending on which indicator we use. For instance, if the initial level of inequality is high, poor audit performance (*audit*) negatively correlates with the growth of roads. However, this is not the case when the budget efficiency (*budget*) is high – a rise in the initial level of inequality positively relates with an increase the growth of roads.

Table 4.4 Summary Relationship between Subsequent Growth of Development Outcome and Institutional Quality[21]

	Dependent Variable: [y(t+a) – y(t)]/a					
	Health	Roads	Sanitation	Water	Poverty	Electricity
$audit_{(it)}$	0.277	–0.068	0.248	–0.157	–0.031	0.196
	(0.18)	(0.11)	(0.22)	(0.16)	(0.04)	(0.19)
$budget_{(it)}$	0.011	0.006	–0.001	–0.01	–0.004**	–0.012
	(0.01)	(0.01)	(0.01)	(0.01)	(0.002)	(0.01)
$locreg_{(it)}$	0.035	0.204	–0.322	–0.079	–0.464***	–0.783**
	(0.29)	(0.18)	(0.36)	(0.29)	(0.07)	(0.31)
Interaction Variable						
$Gini_{(it)} * audit_{(it)}$	–0.157	–2.812***	0.272	0.196	1.694***	3.254*
	(1.59)	(0.93)	(1.96)	(1.40)	(0.33)	(1.71)
$Gini_{(it)} * budget_{(it)}$	0.014	0.048**	–0.009	–0.015	–0.034***	–0.047
	(0.03)	(0.02)	(0.04)	(0.03)	(0.01)	(0.04)
$Gini_{(it)} * locreg_{(it)}$	0.911	1.261	1.711	–2.692	–0.930	–1.399
	(2.67)	(1.61)	(3.26)	(2.65)	(0.64)	(2.85)
$\Delta Gini_{(t)} * audit_{(it)}$	2.433	–0.870	4.415	1.883	1.051	–2.617
	(3.69)	(2.14)	(4.48)	(3.26)	(0.75)	(3.98)
$\Delta Gini_{(t)} * budget_{(it)}$	–0.121	–0.0368	0.145	–0.109	–0.0264	0.232
	(0.15)	(0.09)	(0.19)	(0.13)	(0.03)	(0.17)
$\Delta Gini_{(t)} * locreg_{(it)}$	–3.292	–3.023	–1.045	–1.317	–0.259	8.350
	(7.04)	(4.25)	(8.78)	(7.04)	(1.69)	(7.49)
$\Delta Gini_{(t)}^2 * audit_{(it)}$	90.36**	0.097	96.56*	37.99	8.738	–68.40
	(45.13)	(26.26)	(54.99)	(39.72)	(9.24)	(48.63)
$\Delta Gini_{(t)}^2 * budget_{(it)}$	0.793	–0.767	1.065	–0.319	–0.253	–1.764
	(2.06)	(1.19)	(2.49)	(1.77)	(0.42)	(2.21)
$\Delta Gini_{(t)}^2 * locreg_{(it)}$	–90.22	34.87	–83.46	–118.4	3.821	–141.2
	(127.9)	(76.96)	(155.6)	(127.1)	(30.56)	(136.0)

Notes: Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Unlike the relationship between quality of institutions and the growth of roads, budget efficiency and local governance (*locreg*) alone are highly correlated in reducing the growth of poverty. However, if we look at the interaction between audit results and

[21] Full regression results are available upon request from the author.

the initial level of inequality, its coefficient is positively associated with growth of poverty when a high initial level of inequality exists. It appears that audit results focus more on the mechanism through which a government programme's implementation complies with administrative procedures and follows the rules rather than measuring the effectiveness of the government programme in combatting poverty. On the other hand, the interaction between budget efficiency and initial level of inquality shows negative relationship with poverty growth. It seems that districts with the capacity and ability to collect tax revenue and to disburse government expenditure are expected to benefit from increasing inequality through support from anti-poverty programmes and the redistributive policy.

4.6 Robustness Checks and Sensitivity Test

In this section, we perform some checks to validate our empirical results. Initial inequality matters in our model; therefore, we include the square of initial inequality $(Gini_{(it)})^2$ in our basic model only for the per capita GDP (excluding oil) and use the same set of control variables.

$$g_{it} = \beta_0 + \beta_1 Gini_{it} + \beta_2 (Gini_{it})^2 + \beta_3 \Delta Gini_{it} + \beta_4 [\Delta Gini_{it}]^2 + \beta_5 X_{it} + \varepsilon_{it} \ . \quad (4.8)$$

The results of the robustness check from the basic model of growth of the per capita GDP (excluding oil) are consistent with our basic model, shown in Table 4.1. The initial level of inequality and the change in inequality are positively and negatively correlated with the subsequent growth, respectively, and the square of this change is a negative coefficient and is not statistically significant (Table 4.5). It is also consistent for the growth of the per capita GDP (Appendix A4.3). Hence, we conclude that our model is robust for the initial level of inequality, the change in inequality and its square.

In performing these non-linear regressions, we realize that our data on inequality at the district level are limited, covering only 15 years. Therefore, to test the sensitivity of our result to the change in inequality ($\Delta Gini_{(it)}$), we use the change in inequality within one period instead of between periods, as we did in our original basic model. Our basic model is now written as follows:

$$g_{it} = \gamma_0 + \gamma_1 Gini_{it} + \gamma_2 \Delta Gini_{it} + \gamma_3 [\Delta Gini_{it}]^2 + \gamma_4 X_{it} + \varepsilon_{it}. \quad (4.9)$$

As shown in Appendix A4.5, the change in inequality within one period has a positive relationship while in our previous basic model using a three-year lag for one period it has a negative coefficient. This suggests that our model is sensitive to the period of change in inequality. We also exclude education from our control variables in the basic model to see how the results are affected. Without education, our regression results are still consistent with the previous basic model (Appendix A4.6). We conclude that our model is sensitive to the change in inequality but not for education as a control variable.

Table 4.5 Robustness Check for the Basic Model's Relationship between Inequality, Change in Inequality and Growth

	Dependent Variable: g_{it}							
	Per Capita GDP (Excluding Oil)				Per Capita GDP			
	1a	2a	3a	4a	1b	2b	3b	4b
$Gini_{(t)}$	0.427***	1.013***	1.176**	1.156**	0.485***	0.613*	0.862*	0.788*
	(0.06)	(0.38)	(0.49)	(0.51)	(0.05)	(0.35)	(0.46)	(0.47)
$Gini_{(t)}^2$		–0.936	–0.469	–0.429		–0.203	–0.242	–0.093
		(0.59)	(0.76)	(0.79)		(0.55)	(0.71)	(0.74)
$\Delta Gini_{(t)}$			–0.504***	–0.501***			–0.334***	–0.325***
			(0.07)	(0.07)			(0.07)	(0.07)
$(\Delta Gini_{(t)})^2$				–0.151				–0.559
				(0.82)				(0.76)
y(t)	–0.100***	–0.102***	–0.156***	–0.156***	–0.149***	–0.149***	–0.155***	–0.155***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
educ(t)	–0.002	–0.002	–0.002	–0.002	–0.0001	–0.0002	–0.001	–0.001
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.091***	0.004	0.019	0.022	0.150***	0.131**	0.083	0.093
	(0.02)	(0.06)	(0.08)	(0.08)	(0.02)	(0.05)	(0.07)	(0.07)
Number of Observations	1,210	1,210	968	968	1,206	1,206	968	968
R-squared	0.126	0.128	0.270	0.270	0.324	0.324	0.243	0.244
Number of Districts	242	242	242	242	242	242	242	242

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

4.7 Conclusion and Policy Implications

This chapter basically studied the relationship between growth and inequality across districts in Indonesia by employing Banerjee-Duflo approach of non-linearity. More precise, we empirically tested whether actual changes in inequality associate with subsequent economic growth and whether this relationship is non-linear. We also examine the role of institutional quality in the relationship between past inequality and economic growth in the next period. Similar approach is applied to test whether changes in inequality in the past correlates with growth of development indicators in the future.

We consistently found an increase in a level of inequality has a significant and positive relationship with subsequent economic growth in various specifications. However, large changes in inequality in any direction were found to be associated with large decline in economic growth, as revealed by a hump-shaped relationship between economic growth and the actual change in inequality. These relationships held in all the specifications of the basic and the modified models, which included institutional quality.

Our findings are in line with our results in Chapter 2, in which a typical Kuznets-like inverted U-curve relationship was found at the district level in Indonesia, implying that inequality can be seen as a price to be paid for wealth. In this chapter, we indicated that high inequality in the past is associated with relatively high subsequent economic growth, this high growth is then associated with high inequality in the current period. If this current inequality is too high, then the change in inequality will be positive and large enough to reduce growth in the future. In other words, the high inequality in the past is associated with the slowing down of economic growth in the future and worsening the existing inequality. It seems that our results support Piketty's (2014) argument that inequality keeps increasing as societies accumulate wealth. With regard to the institutional quality, the type of institutional quality plays a significant role in shaping the relationship between past inequality and economic growth in the next period. This role appears more important when we interact institutional quality with the initial inequality, suggesting that a combination of a certain degree of inequality and institutional quality is positively correlated with economic growth.

In addition, our results show that a high initial level of inequality is highly associated with the growth of roads and the poverty reduction.The relationship with the change in inequality shows an inverted U-shaped for the growth of roads and a U-shaped for the poverty rate, suggesting that existing inequality is positive correlated with growth of roads while current inequality is negative correlated with growth of poverty until the this current inequality passes a threshold. If current inequality keeps increasing, it tends to reduce the growth of roads and increase the growth of poverty in the future. These findings indicate that the rising inequality may go at the expense of the accessibility of certain basic public services that we study in Chapter 5 and 6 of this thesis.

Appendices

Appendix A4.1 List of Variables in Basic Model for Development Indicators

	Development Outcome					
	Health	Road	Poverty	Sanitation	Clean Water	Electricity
Dependent Variable						
g_{it}	*g_health*	*g_road*	*g_poverty*	*g_sanitation*	*g_water*	*g_electricity*
Independent Variable						
$Gini_{(t)}$	$Gini_{(t)}$	$Gini_{(t)}$	$Gini_{(t)}$	$Gini_{(t)}$	$Gini_{(t)}$	$Gini_{(t)}$
$\Delta Gini_{(t)}$	$\Delta Gini_{(t)}$	$\Delta Gini_{(t)}$	$\Delta Gini_{(t)}$	$\Delta Gini_{(t)}$	$\Delta Gini_{(t)}$	$\Delta Gini_{(t)}$
$\Delta Gini_{(t)}^2$	$\Delta Gini_{(t)}^2$	$\Delta Gini_{(t)}^2$	$\Delta Gini_{(t)}^2$	$\Delta Gini_{(t)}^2$	$\Delta Gini_{(t)}^2$	$\Delta Gini_{(t)}^2$
Control Variable						
y(t)	*initial_health*	*initial_road*	*initial_poverty*	*initial_sanitation*	*initial_water*	*initial_electricity*
educ(t)	*initial_educ*	*initial_educ*	*initial_educ*	*initial_educ*	*initial_educ*	*initial_educ*
Y(t)	*initial_income*	*initial_income*	*initial_income*	*initial_income*	*initial_income*	*initial_income*

Appendix A4.2 Summary of Observed Variables

Variable	Number of Observations	Mean	Std. Dev.	Min	Max
Per Capita Income	3,868	1.69	0.72	–1.02	5.00
Per Capita Income (excl. oil)	3,872	1.68	0.70	–1.02	5.00
Gini	3,872	0.29	0.05	0.17	0.61
Audit	2,420	–2.11	0.74	–4.00	–1.00
Budget	3,076	108.09	34.22	6.35	971.73
Local Regulation	1,857	1.49	0.77	0.11	3.50
Education	3,872	2.87	1.55	0.11	8.67
Health	3,872	72.14	21.05	10.70	100.00
Road	3,872	70.88	24.99	2.44	100.00
Sanitation	3,763	49.59	24.26	0.35	99.41
Water	3,872	63.13	21.56	0.66	100.00
Electricity	3,872	79.21	19.97	11.49	100.00
Poverty	3,855	15.58	8.95	0.99	91.12

Source: BPS, authors' calculation.

Appendix A4.3 Relationship between Inequality, Change in Inequality, and Growth Per Capita GDP (Basic Model)

VARIABLES	Per Capita GDP				
	Dependent Variable: g_{it}				
	1b	2b	3b	4b	5b
Gini(t)	0.485***	0.709***	0.729***	0.729***	0.733***
	(0.05)	(0.09)	(0.09)	(0.09)	(0.09)
$\Delta Gini_{(it)}$		–0.334***	–0.325***	–0.336***	–0.533**
		(0.07)	(0.07)	(0.09)	(0.22)
$\Delta Gini_{(it)}^2$			–0.586	–0.545	–2.039
			(0.73)	(0.76)	(1.72)
$\Delta Gini_{(it)} > 0$				0.001	0.001
				(0.01)	(0.01)
$\Delta Gini_{(it)} * \Delta Gini_{(it)} > 0$					0.379
					(0.39)
y(t)	–0.149***	–0.154***	–0.155***	–0.155***	–0.156***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
educ(*t*)	–0.0001	–0.0008	–0.0008	–0.0007	–0.0006
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.150***	0.106***	0.102***	0.101***	0.0975***
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	1,210	968	968	968	968
R-squared	0.324	0.243	0.244	0.244	0.245
Number of Districts	242	242	242	242	242

Notes: Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendix A4.4 Relationship between Inequality, Change in Inequality, and Growth of Development Indicators (Basic Model)

Basic Model	Dependent Variable: g_{it}					
	Health			Road		
	1c	2c	3c	1d	2d	3d
Gini(*t*)	2.567	7.668	8.666	6.573***	14.12***	15.26***
	(2.83)	(5.16)	(5.38)	(1.85)	(3.08)	(3.20)
$\Delta Gini_{(it)}$		–3.087	–2.638		–9.191***	–8.710***
		(3.84)	(3.90)		(2.26)	(2.29)
$\Delta Gini_{(it)}^2$			–28.21			–31.23
			(42.02)			(24.33)
y(*t*)	–0.361***	–0.383***	–0.383***	–0.117***	–0.162***	–0.162***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
educ(*t*)	0.001	0.027	0.028	0.001	0.077	0.078
	(0.11)	(0.13)	(0.13)	(0.07)	(0.07)	(0.07)
Y(*t*)	–0.238	–0.647	–0.676	1.312***	1.332***	1.302***
	(0.39)	(0.52)	(0.52)	(0.27)	(0.31)	(0.31)
Constant	25.65***	26.33***	26.15***	5.009***	5.936***	5.724***
	(1.10)	(1.59)	(1.62)	(0.59)	(0.81)	(0.83)
Observations	1,210	968	968	1,210	968	968
R-squared	0.535	0.530	0.530	0.172	0.273	0.275
Number of Districts	242	242	242	242	242	242

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Continued

Basic Model	Dependent Variable: g_{it}					
	Poverty			Sanitation		
	1e	2e	3e	1f	2f	3f
Gini(*t*)	–6.053***	–7.285***	–8.284***	–0.970	0.449	2.478
	(0.87)	(1.13)	(1.17)	(3.56)	(6.47)	(6.74)
$\Delta Gini_{(it)}$		5.366***	4.969***		–0.880	–0.013
		(0.81)	(0.81)		(4.76)	(4.83)
$\Delta Gini_{(it)}^2$			26.57***			–55.37
			(8.58)			(51.73)
y(*t*)	–0.262***	–0.139***	–0.140***	–0.340***	–0.376***	–0.375***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
educ(*t*)	–0.028	–0.001	–0.001	–0.016	–0.088	–0.085
	(0.03)	(0.03)	(0.03)	(0.13)	(0.16)	(0.16)
Y(*t*)	–1.192***	–0.407***	–0.384***	0.121	0.362	0.296
	(0.13)	(0.11)	(0.11)	(0.50)	(0.66)	(0.67)
Constant	7.355***	4.325***	4.561***	16.98***	18.09***	17.67***
	(0.30)	(0.40)	(0.41)	(1.15)	(1.78)	(1.82)
Observations	1,203	967	967	1,165	936	936
R-squared	0.721	0.264	0.274	0.485	0.508	0.509
Number of Districts	242	242	242	242	242	242

Notes: Standard errors in parentheses. *** p<0.01, ** p<0.05, * p<0.1.

Continued

Basic Model	Dependent Variable: g_{it}					
	Clean Water			Electricity		
	1g	2g	3g	1h	2h	3h
Gini(*t*)	–0.601	–1.084	–0.825	1.433	5.802	5.476
	(2.49)	(4.53)	(4.71)	(3.04)	(5.56)	(5.79)
$\Delta Gini_{(it)}$		–1.474	–1.358		–6.009	–6.155
		(3.37)	(3.42)		(4.14)	(4.20)
$\Delta Gini_{(it)}^2$			–7.289			9.195
			(36.89)			(45.24)
y(*t*)	–0.374***	–0.408***	–0.408***	–0.371***	–0.404***	–0.404***
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
educ(*t*)	–0.098	–0.147	–0.147	–0.076	–0.091	–0.092
	(0.09)	(0.11)	(0.11)	(0.11)	(0.14)	(0.14)
Y(*t*)	–0.281	0.059	0.051	–1.062**	–1.419**	–1.410**
	(0.35)	(0.45)	(0.46)	(0.42)	(0.56)	(0.56)
Constant	24.51***	26.33***	26.29***	30.91***	32.89***	32.95***
	(1.01)	(1.44)	(1.45)	(1.26)	(1.83)	(1.85)
Observations	1,210	968	968	1,210	968	968
R-squared	0.522	0.550	0.550	0.535	0.538	0.538
Number of Districts	242	242	242	242	242	242

Notes: Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendix A4.5 Sensitivity Test for Basic Model Relationship between Inequality, Change in Inequality and Growth: Lagged One Year of Gini

`	Dependent Variable: g_{it}					
	Per Capita GDP (Excluding Oil)			Per Capita GDP		
	1a	2a	3a	1b	2b	3b
Gini(*t*)	0.427^{***}	0.806^{***}	0.818^{***}	0.485^{***}	0.795^{***}	0.810^{***}
	(0.06)	(0.08)	(0.08)	(0.05)	(0.07)	(0.07)
gini(*t*) – *gini*(*t*–1)		0.453^{***}	0.484^{***}		0.385^{***}	0.432^{***}
		(0.06)	(0.07)		(0.05)	(0.06)
$(gini(t) - gini(t-1))^2$			–0.733			-1.184^{*}
			(0.74)			(0.68)
y(*t*)	-0.100^{***}	-0.132^{***}	-0.133^{***}	-0.149^{***}	-0.172^{***}	-0.173^{***}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
educ(*t*)	–0.002	–0.002	–0.002	–0.0001	–0.001	–0.0004
	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)	(0.002)
Constant	0.0906^{***}	0.0214	0.0201	0.150^{***}	0.0901^{***}	0.0881^{***}
	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)	(0.02)
Observations	1,210	1,210	1,210	1,206	1,206	1,206
R-squared	0.126	0.178	0.178	0.324	0.359	0.361
Number of Districts	242	242	242	242	242	242

Notes: Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Appendix A4.6 Sensitivity Test for Basic Model Relationship between Inequality, Change in Inequality and Growth: Without Education as Control Variable

`	Dependent Variable: g_{it}					
	Per Capita GDP (Excluding Oil)			Per Capita GDP		
	1a	2a	3a	1b	2b	3b
$Gini_{(t)}$	0.427^{***}	0.878^{***}	0.887^{***}	0.485^{***}	0.709^{***}	0.729^{***}
	(0.06)	(0.09)	(0.10)	(0.05)	(0.09)	(0.09)
$\Delta Gini_{(t)}$		-0.503^{***}	-0.499^{***}		-0.334^{***}	-0.325^{***}
		(0.07)	(0.07)		(0.07)	(0.07)
$(\Delta Gini_{(t)})^2$			–0.276			–0.588
			(0.78)			(0.73)
y(*t*)	-0.100^{***}	-0.155^{***}	-0.156^{***}	-0.149^{***}	-0.155^{***}	-0.155^{***}
	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)	(0.01)
Constant	0.085^{***}	0.059^{**}	0.057^{**}	0.149^{***}	0.103^{***}	0.099^{***}
	(0.02)	(0.02)	(0.02)	(0.01)	(0.02)	(0.02)
Observations	1,210	968	968	1,206	968	968
R-squared	0.126	0.269	0.269	0.324	0.243	0.244
Number of Districts	242	242	242	242	242	242

Notes: Standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Chapter 5 Does Increased Access to Indonesian Health Care Invoke Ex Ante Moral Hazard?

5.1 Introduction

The decentralization and democratization processes in Indonesia have strongly impacted the entire society. In the previous chapters we analyzed how the Indonesian society changes in terms of the evolution of inequality in income and economic development across space and individuals and under influence of urbanization. In the remainder of this thesis, we analyze changes in the organization and accessibility of public services like health (this chapter) and electricity (Chapter 6). Did inequality in access to these services increase or decrease, and which factors drive the observed patterns of inequality? More specifically, in chapter 6 we analyse how spatial variation in population density, per capita GDP, geographical structure, industrial intensity and power supply have shaped access to electricity over time. In this chapter we analyze how changes in health insurance coverage – aimed to foster equality in acess to health services – have shaped individual healthcare consumption behaviour.

In 2014, the Indonesian government established the *Jaminan Kesehatan Nasional* (JKN), a single-payer comprehensive universal health care programme managed by the social security agency *Badan Pengelola Jaminan Sosial untuk Kesehatan* (BPJS),. This programme was designed as pooled insurance coverage with the aim of providing health coverage for all Indonesians, with the Government only paying a modest premium. As social insurance, the main source of finance for the JKN is premia. The Government, through the Ministry of Health and the local budget, pays the premia of poor households so that they can access free health care. This group is also categorized as subsidized people. In non-poor households with formal employment status, civil servants pay 5 per cent of their salaries, but only 2 per cent is deducted directly from their salary payment and the rest is borne by their employers. Non-civil servants pay 5 per cent of their income, with 0.5 per cent paid directly by the employees and the remaining 4.5 per cent paid by their employer. Those without employment status, that is, informal workers, pay premia based on their chosen inpatient ward level (Fuady 2019). As of 1 March 2019, the JKN serves 218,132,478 people, accounting for more than 80 per cent of the total population. It covers most treatments, ranging from simple diseases and symptoms to heart surgery, renal dialysis and cancer therapies, as well as maternity and child health services (BPJS 2020). However, currently the JKN is facing financial difficulties in paying health bills, affecting fiscal sustainability of the programme.

The health status of the Indonesian population is characterized by rapidly increasing life expectancy at birth, mainly through reductions in communicable, maternal, neonatal and nutritional (CMNN) diseases, but it is also facing a double burden of diseases whereby non-communicable diseases, such as ischemic heart diseases, cerebrovascular diseases and diabetes, are growing[22] due to unhealthy lifestyles, poor diets and a lack of physical activities, while at the same time communicable diseases, such as tuberculosis, diarrhoea and HIV/AIDS, remain the biggest challenges (Mboi et al. 2018).

After the introduction of the JKN, access to health care increased significantly, in particular for low-income groups and in rural areas in which the population previously could not afford the high costs of advanced treatments to cure non-communicable diseases. The frequency of visits to health facilities and inpatient care increased by 53 per cent and 38 per cent, respectively, from 2015 to 2017 (BPJS 2018). As a consequence, financing problems became unavoidable. The very low premia that are charged to the participants are far from sufficient to cover the existing curative costs. Furthermore, many participants just join the programme once they need further and/or high-cost treatment. Even worse, some participants stop paying their contribution once they have recovered from their illness. In other words, part of the population contributes very little but receives advanced high-cost treatment during a

[22] According to Global Burden of Disease Study 2016, these diseases are the first three leading causes of disability-adjusted life years in 2016, while diabetes ranked seventeenth, cerebrovascular diseases ranked eighth and ischemic heart diseases ranked fifth in 1990 (Global Burden of Disease Collaborative Network 2017).

short period of insurance. For instance, by making a compulsory[23] contribution of only $6 per month, a non-poor and informal participant with kidney failure could receive dialysis in a first-class room with average costs of around $600 per month (i.e. 100 times more than is paid; see Agustina et al. 2019). As a consequence, the BPJS, a social security agency managing and administering the JKN, encountered an escalating deficit from approximately $422.2 million in 2015 to $720.6 million in 2017 (an increase of 72 per cent in 2 years; BPJS 2018). To overcome this problem, the Ministry of Finance has bailed out the BPJS since 2015, but of course at the price of fiscal sustainability of the Indonesian economy, which can be harmed in the long run.

Despite the fact that the use and coverage of the JKN have increased considerably, the Government's allocation to health spending was consistently low, at around 1 per cent of the GDP, from 2007 to 2014 (World Bank 2020). To increase the health budget, the Ministry of Health has assigned at least 5 per cent of the national budget and 10 per cent of the local budgets to health spending. However, the percentage of government expenditure to GDP remained low, at 1.45 per cent, in 2017. Moreover, the total health expenditure (THE) in Indonesia was only 3.0 per cent of the GDP in 2017, which is lower than the expenditure in Cambodia, Vietnam and Thailand, with corresponding figures of 5.9 per cent, 5.5 per cent and 3.7 per cent, respectively (World Bank 2020). This financial problem also affects payment systems in which capitation systems (payments of premia on a per capita basis) and claim payments are only based on the number of members covered and the resource availability and do not consider the total facility burden and performance, so this system is perceived to provide too low and slow a claimant process given the required competency and service standard (Agustina et al. 2019). As a consequence, many providers are no longer willing to renew their contract with the BPJS, causing a shortage of health facilities and poor services.

In view of this state of affairs, we hypothesize that there might be a moral hazard problem for insured people in the sense that they tend not to pay their

---

[23] Compulsory means that non-poor and informal participants have to pay a monthly contribution to the BPJS based on their choice of inpatient ward. For non-poor people with formal employment status, a small part of their contribution is deducted directly from their salary and the rest is paid by the employee to the BPJS.

contribution after being treated and not to care sufficiently about their health situation that causes illness (ex ante moral hazard). This behaviour may be associated with increasing claims to the BPJS, so it is important to understand this behaviour to solve the BPJS's problem.

In this chapter, due to data availability, we only observe individual behaviour prior to the JKN era, when health insurance was provided by the central government from 1998.[24] We assume that the behaviour of insured people remains the same over time, so we use the latest household data prior to the JKN era to inform us about the effects of insurance schemes. From the individual perspective, there are two perceptions of health: investment and consumption. Individuals perceive health as an investment if they increase their health consumption to improve or maintain their health condition (high awareness). Hence, health insurance can be perceived as a subsidy to cover their preventive costs. We expect that an increasing frequency of visits to health facilities is for counselling. On the other hand, individuals perceive health as consumption if they increase their health consumption to cure their health problems. They tend not to pay attention to their health and are in a situation of low awareness. Thus, there might be a moral hazard in health consumption whereby they perceive that they will be covered by insurance regardless of their illness. In other words, there is no incentive to prevent a deteriorating health situation and this could harm the sustainability of the health insurance system in the long run. In short, health insurance could determine the probability of inpatient and outpatient treatment, but its impact is undetermined.

From the literature, it is apparent that health insurance coverage may affect consumers' health behaviour through multiple channels, so the net impact is theoretically ambiguous. Our research aims to contribute to our understanding of how people's behaviour towards health influences their demand for unhealthy consumption, which has an impact on outpatient and inpatient care and thus affects the sustainability and affordability of the BPJS programme. In particular, we try to

[24] As part of the coping strategy during the economic crisis, the central government developed a pro-poor policy known as the social safety net for health (JPS-BK) in 1998. This programme evolved and its name changed several times: the health financing scheme (JPK) in 2003, health security for the poor (JPK-MM) in 2005 and health insurance for the poor (*Askeskin*) and Community Health Security (*Jamkesmas*) since 2008 (Fuady 2019).

identify ex ante moral hazard for insured people to help the government to formulate policy recommendations.

Our main findings are as follows. Health insurance with or without subsidies improves accessibility to health care in terms of increasing the frequency of visits for outpatient care and the number of days spent receiving inpatient care. Subsidized people visit health facilities more frequently, but they spend fewer days in inpatient care. The latter might be due to a complex and bureaucratic procedure for the subsidy programme, lack of knowledge of the subsidized members about completing their paperwork and high transportation costs to reach public hospitals. To improve this situation, subsidized people need intermediaries or brokers in their local communities to help them to access free health care (Sambodho 2019). We also find that the probability of insured people, regardless of their type of insurance, consuming more excessive carbohydrates, protein, fat and calories is significantly higher than the probability of their uninsured counterparts. Furthermore, subsidized people consume more cigarettes than non-subsidized people. The probability of the insured, including subsidized people, consuming excessive fat also increases when they visit more frequently. These findings indicate that ex ante moral hazard exists among insured people and even the poor. As a consequence, insured people who consume excessive carbohydrates spend more days in inpatient care and subsidized people who consume excessive fat visit health facilities more frequently. However, there is no difference in the demand for inpatient care for the poor if they consume unhealthy food. The role of education is important not only to improve the health status of the insured households but also to access both outpatient and inpatient care, especially for subsidized households. Becoming older and richer is associated with better access to outpatient and inpatient treatment for insured households, but this is not the case for subsidized people, who spend fewer days as their age increases but visit more frequently as their income rises.

The remainder of this chapter is organized as follows. Section 5.2 presents a literature review on ex ante moral hazard in health care. In section 5.3, we discuss the empirical strategy to estimate demand functions for health care. Section 5.4 describes the dataset and the variables, and section 5.5 presents and interprets the estimation results. Section 5.6 provides concluding remarks and discusses avenues for future research.

5.2 Literature Review

In the health economics literature, moral hazard has been a central concept for decades. There are two types of moral hazard related to health insurance: ex ante moral hazard, which causes insured people to invest insufficiently in self-protection (Ehrlich and Becker 1972), and ex post moral hazard, which results in inefficiently high consumption of health care (Pauly 1968; Manning et al. 1987). Bhattacharya and Packalen (2011) identified a second form of ex ante moral hazard, known as "other" ex ante moral hazard, which arises through the impact that self-protection has on the reward for innovation. People are not aware that their decision to adopt a low level of self-protection, such as a lack of physical activities, poor diet and so on, which increases the incidence of diseases and their medical treatment, will benefit the innovators as positive externalities. This mechanism of "other" ex ante moral hazard leads people to select inefficiently into high levels of self-protection.

Most research on health insurance has focused on ex post moral hazard, whereby insured people demand unnecessary treatment that is covered by the insurance. It has provided fairly consistent evidence (cf. Zweifel and Manning 2000). This type of moral hazard could also induce moral hazard among providers, such as hospitals and pharmacies, to offer unnecessary treatment, examinations or medicines to obtain or secure their revenues in the long run. Evidence of ex post moral hazard that involves adverse health events is less consistent (Cutler and Zeckhauser 2000).

According to Zweifel and Breyer (1997), ex ante moral hazard depends on the opportunity cost of preventive effort, which is likely to be proportional to the wage rate. However, this moral hazard effect may be neutralized by risk aversion, which leads to ambiguity about the relationship between health insurance coverage and sick leave pay. Furthermore, Ehrlich and Becker (1972) suggested that insurance coverage may cause ex ante moral hazard whereby patients have less incentive to reduce their risky health behaviour because they will pay a relatively small amount of the financial cost of their future illness. For example, Dave and Kaestner (2009) found that

Medicare coverage increases the probability of daily alcohol consumption among men. However, health insurance does not reduce the non-financial consequences of illness, such as physical pain and suffering, which could limit the extent of ex ante moral hazard (Ehrlich and Becker 1972). Moreover, health insurance has significant incentive effects on lifestyle choices, increasing the propensity for heavy smoking, lack of exercise and obesity and decreasing the propensity for heavy drinking (Stanciole 2008). Rashad and Markowitz (2009) also presented evidence that being insured is associated with a higher body mass, particularly for those above a certain poverty threshold, and an increased probability of being overweight. There is no evidence that having insurance affects the probability of being obese.

Kenkel (1994) and Stanciole (2008) found that the use of preventive services decreases with age, which suggests adaptation to a shorter payoff period for investments in prevention. If insurance increases the coverage of curative treatment, the demand for preventive care also increases because the out-of-pocket costs are reduced (Courbage and de Coulon 2004; Pagán et al. 2007). This suggests that the two types of care are complements and may serve as substitutes for a patient's own preventive effort.

5.3 Empirical Strategy

In this section, we will discuss our empirical strategy to estimate the demand for unhealthy consumption, that is, cigarettes, alcohol and excessive or under-consumption of carbohydrates, protein, fat and calories, and the demand for health care in terms of outpatient and inpatient care. We observe the health spending behaviour of insured people by calculating their out-of-pocket health expenses from the Susenas consumption module 2014 (the most recent year prior to the JKN). Using this information, we estimate the demand for unhealthy consumption, number of days of inpatient care and frequency of visits to health facilities for outpatient care.

For insured households, if their out-of-pocket (OOP) payment for health increases, we expect that they will spend more on health for the following reasons: (i) their health expenditures are not or are only partially covered by their insurance; and (ii) their insurance only covers one person in their family, that is, the head of the family. From our dataset, we can see whether the frequency of visits to health facilities increases as the OOP increases. If this is the case, health is a complementary good and there is an income effect after implementing insurance. On the other hand, if people's OOP decreases, we expect that they allocate more to non-health expenditures. We can also see whether the number of days for inpatient treatment at health facilities increases as the OOP decreases.

We use a simple theoretical framework of individual preferences to understand individuals' behaviour regarding their health situation. If individuals value health as important, they will allocate more money to preventing illness. However, their allocation depends on their income. Rich people have enough money to pay their health expenses or can access the capital market to finance their health expenses; the preferences of the middle class create a market for their health-related needs, such as medicine, vitamins and health equipment. Within the class of low-income people, there is a group that is sufficiently aware of the possibilities, such as teachers, who do have knowledge about preventing illness, so they spend very little money on health.

Our challenging question concerns how to differentiate people who value health from those who do not care. Our hypothetical scenarios are as follows. First, people visit health facilities to cure their illnesses due to their lack of awareness about preventing them, so they spend more on outpatient or inpatient care, such as medicine, clinics and so on. Second, people do not visit health facilities because they are healthy due to their awareness of health. They spend more on health prevention, such as purchasing vitamins, attending medical check-ups and so on.

Our next challenging question regards how to classify people who spend more on self-treated expenditure. There are two possible methods: (i) patients' knowledge, which enables them to heal themselves, and (ii) a lack of access in terms of either the high OOP due to high user fees paid to clinics (affordability) or the physical access (i.e. the distance to the nearest hospital).

To control for the preferences, we use the incidence of illness, frequency of visits, consumption of cigarettes, alcohol, carbohydrates, protein, fat and calories, education, employment status, having children and living in an urban environment. The health preventive costs, which are calculated from the Susenas consumption module, consist of expenditures on pregnancy examination, immunization, birth control, medical check-ups and vitamins.

There will be heterogeneity in health preferences among the poor and the near-poor group. We classify people into four groups as follows: (i) people with high awareness of health and willingness to spend more on health prevention and thus having low expenditure for curative purposes; (ii) people who have high awareness of health but who are not willing to spend more on health prevention (if they also spend more on curative treatment, they could face catastrophic payments with their expenditure on health exceeding 40 per cent of their income remaining after spending on subsistence needs (Xu et al. 2003)); (iii) people with low awareness of health (and with a poor lifestyle), who spend a large amount on curative treatment (a group that is likely to experience catastrophic payments); and (iv) people who have low awareness of health but do not spend on curative treatments (the members of this group probably cannot access health care because they are too poor).

5.3.1 Basic Model

In our data, there are clear indications about the relevance of moral hazards. Insured people spend more on curative care than on preventive care in terms of the share of the total health expenditure and per capita. These patterns may relate to co-payment or additional costs for uncovered illness. For instance, diseases derived from smoking are associated with unhealthy behaviour that correlates with increasing curative costs. This unhealthy behaviour of insured people is classified as ex ante moral hazard. They also experience more frequent illnesses and health disruptions and thus more days of health disruptions. As a consequence, they visit health facilities more frequently for outpatient care and spend more days receiving inpatient care. Further characteristics of insured people will be discussed in section 5.5.

It seems that many insured people in Indonesia insure themselves not for risk protection but to acquire cheap access to health services. Once they are insured, they tend to consume as much health care as possible without paying much attention to their own health status. Insurance also seems to trigger an unhealthy lifestyle (Stanciole 2008; Dave and Kaestner 2009; Rashad and Markowitz 2009). In other words, it resembles the behaviour of "all you can eat" restaurants where people pay a fixed amount irrespective of how much they eat. The study conducted by Rashad and Markowitz (2009) described how insurance is associated with a higher body mass, particularly for those above the poverty threshold, and an increased probability of being overweight.

In this chapter, we try to identify this ex ante moral hazard phenomenon through the expenditure function. However, we do not have information on the price of health services because Susenas only recorded the monthly average curative and preventive expenditure for the last three months (Appendix A5.A1). Therefore, we use control variables that affect the price, such as the location and quality of housing and its characteristics.

5.3.2 Hypotheses

Our objectives are (i) to identify ex ante moral hazard by estimating the demand for unhealthy consumption, that is, cigarettes, alcohol and excessive consumption of carbohydrates, protein, fat and calories, and (ii) to estimate the demand for health care under insurance programmes, that is, the frequency of visits to health facilities for outpatient care and the number of days spent in inpatient care.

Thus, our hypotheses are as follows. Compared with uninsured people, insured people display more unhealthy consumption patterns and consume more outpatient as well as inpatient care. Moreover, if insured people are subsidized, they tend to pay less attention to their lifestyles by increasing their unhealthy consumption compared with their non-subsidized counterparts. In our estimations, we control for income, education, age, housing quality, employment status, family composition and characteristics of the place where people live.

People with a poor health status, reflected by a high frequency of illness, are expected to spend more days on inpatient care. Curative costs, as a proxy for the price of health services, are expected to have a negative correlation with outpatient and inpatient care. Meanwhile, preventive costs as well as medical check-ups could improve people's health status, resulting in them visiting health facilities less frequently and spending fewer days receiving inpatient care.

People's marital status is an important determinant of their health status (Rashad and Markowits 2009). Married people have a greater incentive to stay healthy because of the responsibilities associated with family live, which tends to result in selecting healthy consumption and hence reduced spending on outpatient and inpatient care. We expect that income, education and housing quality will correlate negatively with outpatient care but positively with inpatient care. The latter reflects that inpatient care is expensive and tends to be accessible only if people have sufficient income and are sufficiently knowledgeable about how to obtain it. Having children and living in urban areas are also important determinants of visits to health facilities and spending on inpatient care. Children are likely to be ill more frequently in urban areas, while at the same time urban areas offer more health facilities. After privatization in 1988, a significant number of hospitals and private clinics emerged throughout Indonesia, especially in urban areas and big cities (Heywood and Harahap 2009).

Thus, we have three main basic models explaining the variation in unhealthy consumption, outpatient care and inpatient care. The first basic model estimates the demand for unhealthy consumption of household i of consumption type j (UC_{ij}), where j = 1,…,6, namely cigarettes, alcohol and excessive consumption of carbohydrates, protein, fat and calories. Ex ante moral hazard can be identified by looking at whether insured people with or without subsidies increase their demand for unhealthy consumption compared with the uninsured group. An increasing insurance premium that correlates with increasing unhealthy consumption could also indicate that ex ante moral hazard prevails. In other words, they consume more unhealthy food because they think they do not need to protect their health and that, if they fall ill due to their unhealthy lifestyle and poor diet, the associated expenses will be covered by the insurance. Thus, we expect that increasing health insurance premia are positively correlated with unhealthy consumption:

$$
\begin{aligned}
UC_{ij} = \gamma_0 + \gamma_1 \cdot ln\ Hins_i + \gamma_2 \cdot OP_i + \gamma_3 \cdot pct\ Married_i \\
+ \gamma_4 \cdot (Married_i \cdot Income_i) + \gamma_5 \cdot d\_ins_i + \gamma_6 \cdot d\_sub_i \\
+ \gamma_7 \cdot (d\_ins_i \cdot OP_i) + \gamma_8 \cdot (d\_sub_i \cdot OP_i) \\
+ \gamma_9 \cdot (Hins_i \cdot OP_i) + \gamma_{10} \cdot Xs_i + \varepsilon_{ij}.
\end{aligned}
\qquad (5.1)
$$

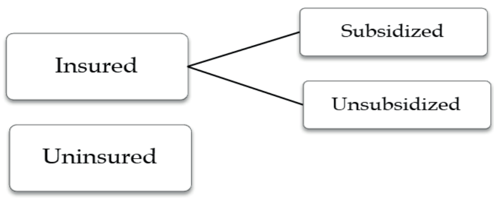
Visiting health facilities gives people a chance to contact health professionals and obtain indirect benefits, such as counselling on healthy lifestyles (Dave and Kaestner 2009). Therefore, we use the frequency of visits in this model: increasing visits to health facilities are expected to have an association with decreasing unhealthy consumption.

Moreover, we add an interaction between insurance and outpatient care to see whether insured people change their demand for unhealthy consumption if they visit health facilities. We expect that this interaction has a negative impact on unhealthy consumption if the insured people contact health professionals and receive counselling on healthy food consumption during their visit.[25]

As we discussed earlier, marital status and its interaction with income could associate with more healthy lifestyles. Hence, we expect that marital status will negatively relate to unhealthy consumption, while its interaction with income could be positive or negative, as discussed above. All the control variables are expected to have a negative impact on unhealthy consumption.

We introduce an insurance dummy to determine how people's preferences regarding health are affected by being insured and receiving a subsidy. Therefore, we generate two insurance dummies, namely insurance (*d_ins*) and subsidy (*d_sub*), in our model (Figure 5.1). The latter is a dummy variable among insured people indicating whether they pay the non-subsidy premium or receive a subsidy card that allows the card holders to access with free health care in the insurance system. The card holders' subsidy premia are paid by the central and the local government.

Figure 5.1 Possible States of Insurance



[25] Many insurance programmes require their members to visit health facilities at least once to examine their health status.

We then interact each insurance dummy with age ($d\_ins_i \cdot Age_i$ and $d\_sub_i \cdot Age_i$), education ($d\_ins_i \cdot Educ_i$ and $d\_sub_i \cdot Educ_i$) and income ($d\_ins_i \cdot Income_i$ and $d\_sub_i \cdot Income_i$) to examine whether the impact of insurance is conditional on age, education and income. In all three models, we routinely control for seven variables (X_i), viz. income, education, age, housing quality, a dummy for employment status, a dummy for having children and a dummy for living in urban areas.

As discussed previously, the demand for outpatient care (OP_i) and inpatient care (IP_i) can be explained by the per capita curative cost ($ln\ Cure_i$), per capita preventive cost ($ln\ Prev_i$), per capita medical check-up expense (ln $Mdcheck_i$) and other variables as written in the first model, that is, per capita premium health insurance ($ln\ Hins_i$), married household ($pct\ Married_i$) and its interaction with income, insurance dummies (d_ins_i and d_sub_i) and their interaction with age, education and income and a set of control variables (Xs_i). Our second model of outpatient care (OP_i) is written as follows:

$$\begin{aligned} OP_i = {} & \alpha_0 + \alpha_1 \cdot lnCure_i + \alpha_2 \cdot lnPrev_i + \alpha_3 \cdot lnMdcheck_i + \alpha_4 \cdot lnHins_i \\ & + \alpha_5 \cdot pctMarried_i + \alpha_6 \cdot (Married_i \cdot Income_i) + \alpha_7 \cdot d\_ins_i \\ & + \alpha_8 \cdot d\_sub_i + \alpha_9 \cdot (d\_ins_i \cdot Age_i) + \alpha_{10} \cdot (d\_ins_i \cdot Educ_i) \\ & + \alpha_{11} \cdot (d\_ins_i \cdot Income_i) + \alpha_{12} \cdot (d\_sub_i \cdot Age_i) + \alpha_{13} \cdot (d\_sub_i \cdot Educ_i) \\ & + \alpha_{14} \cdot (d\_sub_i \cdot Income_i) + \alpha_{15} \cdot Xs_i + \varepsilon_{ij}. \end{aligned} \quad (5.2)$$

In addition, we use the frequency of illness per capita ($Illness_{(i)}$) as a proxy for health status to explain the demand for inpatient care (IP_i). Increasing frequency of illness shows poorer health status, which is associated with increasing inpatient care. Our third model of inpatient care (IP_i) is expressed in the equation below:

$$\begin{aligned} IP_i = {} & \beta_0 + \beta_1 \cdot Illness_{(i)} + \beta_2 \cdot lnCure_{(i)} + \beta_3 \cdot lnPrev_{(i)} + \beta_4 \cdot lnMdcheck_{(i)} \\ & + \beta_5 \cdot lnHins_{(i)} + \beta_6 \cdot pctMarried_{(i)} + \beta_7 \cdot (Married_{(i)} \cdot Income_{(i)}) \\ & + \beta_8 \cdot d\_ins_i + \beta_9 \cdot d\_sub_i + \beta_{10} \cdot (d\_ins_i \cdot Age_i) + \beta_{11} \cdot (d\_ins_i \cdot Educ_i) \\ & + \beta_{12} \cdot (d\_ins_i \cdot Income_i) + \beta_{13} \cdot (d\_sub_i \cdot Age_i) + \beta_{14} \cdot (d\_sub_i \cdot Educ_i) \\ & + \beta_{15} \cdot (d\_sub_i \cdot Income_i) + \beta_{16} \cdot Xs_i + \varepsilon_t. \end{aligned} \quad (5.3)$$

The interaction between marital status and income in the second and third models could be a positive or negative correlation with the demand for outpatient and inpatient care. If this interaction has a negative coefficient, decreasing outpatient or inpatient care is due more to people's awareness of the need to keep healthy after marriage rather than to their income. However, if this is not the case, an increasing income after marriage could link to increasing unhealthy consumption due to being a poor household. A higher income after marriage indicates that the household can afford more consumption, including unhealthy forms, which relate to diseases that requires more outpatient or inpatient care.

We also expect the coefficient of the interaction between insurance dummies (d_ins_i and d_sub_i) and age as well as income to be positive. The elderly are likely to experience numerous health complaints that need more outpatient or inpatient care, and richer people can easily access health facilities either for outpatient or for inpatient care. The interaction between insurance and education is expected to be a negative correlation, meaning that more educated people understand how to protect their health status and reduce their outpatient and inpatient care.

Detailed explanations for each variable and the data sources used are provided in section 5.4. In addition, we modify the basic model of outpatient and inpatient care by introducing each type of unhealthy consumption as a variable of interest and then interact unhealthy consumption with the insurance dummies (d_ins_i and d_sub_i) to see how outpatient or inpatient care changes when unhealthy consumption changes under a particular insurance scheme.

5.4 Data Description

We conduct a cross-sectional analysis to explain the variation in the demand for health as a function of the insurance status of individuals. The main source of our data is Susenas 2014, provided by the BPS, which comprises consumption at the household level and a core module at the individual level. This survey covers a 0.34 per cent sample of the total population (around 1.1 million individuals). We exclude individuals from Aceh, Maluku, West Papua and Papua from our dataset because there are too many missing responses. Our unit of analysis is the household level. The total sample consists of 202,924 households. We extract expenditure information, such as curative costs, preventive costs, health insurance premium and household expenditure, from the Susenas consumption module. Exploiting Susenas, and especially the expenditure data, has limitations. For instance, health consumption is financed by out of pocket and/or by subsidies especially for the poor. The subsidy recipient might receive health services free of charge so health expenditure as out-of-pocket is likely underestimated (Johar, et al. 2018). A more detailed explanation of data limitations in Susenas is presented in Appendix A5.A1 and Section 2.4. For variables with high variability and heterogeneity, we control the heteroscedasticity by transforming our variables into log form, including per capita curative and preventive costs, medical check-ups, health insurance premium, consumption on cigarettes and alcohol. Accordingly, we can interpret the results more easily as (semi-) elasticities. A detailed description of each variable that we use is presented below.

Unhealthy Consumption

Given the limitations of the consumption module in Susenas, we classify unhealthy consumption as follows: cigarettes, alcohol and unstandardized carbohydrate, protein, fat and calorie intake. The last 4 consumption types are based on the nutrient standards issued by the Ministry of Health of the Government of Indonesia (*Peraturan Menteri Kesehatan No 41 Tahun 2014;* Ministry of Health 2014), which state that Indonesians' daily consumption, to be healthy, should be 300–500 grams/capita of carbohydrates, 55–65 grams/capita of protein, 20–35 grams/capita of fat and 1,200–2,000 calories/capita. Therefore, we generate four dummy variables (*d_carbo*, *d_protein*, *d_fat* and *d_calorie*), which are equal to 1 if the consumption is on the unhealthy side of the threshold and 0 otherwise. We calculate the consumption of cigarettes (*log_cigarcap*) and alcohol (*log_alcoholcap*) in quantities, 1 pack of cigarettes consisting of 10 pieces and 1 bottle of alcohol being 620 ml. Both amounts are transformed into logs.

Inpatient and Outpatient Care

Inpatient care per capita (*IP*) is calculated by adding up the number of days spent in all health facilities, such as hospitals, integrated health centres (Puskesmas), health workers' practices and traditional healers' practices, in the last year and then dividing the total by the number of household members. Outpatient care per capita (*OP*) is calculated as the frequency of visits to health facilities during the last month, again on a per capita basis. Hence, we sum up the total number of visits to hospitals, clinics, integrated health centres and their support (Puskesmas/Pustu), health workers' practices and traditional healers' and birth attendants' practices and divide the total by the number of household members.

Frequency of Illness

The frequency of illness per capita (*illness_cap*) shows the individual health status. The more frequent people's illnesses are, the poorer their health status is. We add up the incidence of fever, cough, influenza, asthma, diarrhoea, headache and toothache reported by each household member in the last month and divide this figure by the number of household members.

Health Expenditure

Health expenditure is a monthly average calculated based on three months of expenditure taken from the Susenas consumption module. We distinguish two main expenditure types, namely curative and preventive costs. Curative costs (*log_curecap*) are the per capita monthly average expenditures to cure illness (i.e. fees for hospitals, clinics, health workers' practices, traditional healers' and birth attendants' practices, medicines, plaster cast, wheelchairs and glasses). Preventive costs (*log_prevcap*) are the per capita monthly average expenditures for prevention (i.e. pregnancy examination, immunization, birth control, medical check-up, vitamin, fitness, etc.). We transform these expenditures into log form.

Health Insurance

We use two variables for health insurance: the insurance premium and a dummy capturing whether a household is insured. The insurance premium (*log_premiumcap*) is the monthly average of three months of health premium expenditure divided by the number of household members calculated from the consumption module of Susenas. To understand people's behaviour towards health once they are insured, we classify our sample into three groups, viz. uninsured, insured and subsidized groups (see Figure 5.1). We therefore use two insurance dummies, namely *d_ins*, which is one if at least one family member is insured and zero if no member is insured, and *d_sub*, which equals one if at least one family member is insured and holds a subsidy card and zero otherwise.

Using the core module of Susenas 2014,[26] we classify insurance based on its premium, namely insurance with a subsidized premium and a non-subsidized premium. About 67 per cent of insured households have insurance with a subsidized premium paid by the central or local government, that is, health insurance for the poor (*Jamkesmas*), local health insurance (*Jamkesda*) and maternity insurance (*Jampersal*).[27] The remaining 33 per cent of insured households are registered as non-subsidized insurance members whose premium payment varies depending on their coverage, that is, insurance for civil servants or the army or retirees (*Askes*), insurance for workers (*Jamsostek*)[28] and private insurance. We also transform this premium expenditure into log form.

Married Household

This variable is calculated based on the percentage of married people within a household (*pct_married*).

Other Control Variables

In this chapter, we use seven control variables for our regressions: the per capita income calculated from household expenditure (*log_incomecap*), education of the

---

[26] The data collection was carried out early in 2014, when the JKN had just been introduced, so the BPS did not include the JKN in the Susenas core questionnaire in 2014.

[27] Prior to 2014, members of the national insurance for the poor (*Jamkesmas* and *Askeskin*) held an insurance card provided by the central government through the Ministry of Social Affairs, giving them free access to health facilities. Since the JKN started in 2014, all the members of national and local insurance for the poor were transferred to the BPJS and were known as contribution beneficiaries – *Penerima Bantuan Iuran* (PBI); they received a new health card – the *Kartu Indonesia Sehat* (KIS). The BPS recorded these subsidized people after being validated and verified by the Ministry of Social Affairs; however, some problems remained to be solved, for example misclassification between poor and non-poor households and an outdated database (Fuady 2019).

[28] Jamsostek is the insurance programme for workers (white and blue collar), the benefit package of which is similar to Askes (the insurance programme for civil servants); the premium is deducted directly from workers' salary.

household head (*educ_hh*), age of the household head (*log_age_hh*), housing quality (*housequality*), a dummy for the employment status of the household head (*d_empl_hh*), a dummy for having children (*d_kids*) and a dummy for urban (*d_urban*).

Education is defined as the total years spent to obtain the highest diploma held by the household head. For instance, obtaining an elementary diploma takes at least six years, junior high school nine years, senior high school twelve years, college fourteen to fifteen years, university seventeen years and a master's or doctoral degree at least twenty-two years. If the household head does not hold any diploma, we categorize the household as no school or zero years.

We construct an index with equal weighting for housing quality based on housing characteristics, that is, wall, roof, floor, toilet facility and source of water for drinking, cooking and washing. The type of wall consists of concrete, wood or bamboo, of which concrete is a higher quality than wood and bamboo, respectively; the type of roof ranges from concrete to tile, shingled, tin, asbestos and leaves, a concrete roof being the highest quality; the type of floor is marble, ceramic, granite, wood, terrazzo, cement or soil, of which marble is the highest quality; the type of toilet facility is classified based on the ownership type, that is, private, shared, public and none, private ownership being perceived as the highest quality; and the water source for drinking, cooking and washing can be tap water, which is categorized as the highest quality, followed by artesian or protected wells, unprotected wells, rivers and rain, respectively. We rank each feature and set 100 as the highest quality; we then add up each of them and divide the total by five to obtain the average housing index. Thus, the higher the value of the index is, the higher the quality of the house is.

We use a dummy for the employment status of the household head, in which one is the formal status or white collar and zero is the non-formal status or blue collar. The formal status includes workers, employees or staffs with employment contract, self-owned businesses with permanent or paid workers, and freelancers. The dummy for having children equals one when a household has children and zero otherwise; the dummy for urban is one for households living in urban areas and zero otherwise.

5.5 Descriptive Statistics

As discussed previously, we classify our Susenas 2014 into an insured and an uninsured group. Most of the difference in the means across the insured and uninsured groups is significant, with a p-value<0.01, except that the share of preventive expenditure in the total expenditure, the share of medical check-ups in the total health expenditure and the medical check-up expenditure are not significant. In general, the means of these household and health characteristics are higher for insured people than for uninsured people (Table 5.1).

On average, the household head is almost 50 years old and spent around 7 years on education, holding an elementary diploma. Insured people are slightly more educated and about 46 per cent work with a formal status as compared with uninsured people. More than 80 per cent of our sample is married, and the average size of the household is 3.95 and 3.6 for insured and uninsured households, respectively. In terms of housing quality, the uninsured group has better quality but its per capita household expenditure as a proxy for income is lower than that of the insured group.

In terms of health status, insured people reported illness and health disruption of 0.34 and 0.36, respectively, while their counterparts reported 0.31 and 0.33. They also could not work for almost 2 days due to health disruption in the last month. They prefer visiting public health facilities, that is, hospitals and clinics, to private ones. On average, insured people visit health facilities more frequently than uninsured people. In our sample, about 67 per cent of insured people are subsidized by the Government because they are poor. According to the BPJS guidelines, patients under the subsidy programme require referral from public primary health care every time they need further examination or treatment in higher public health facilities, that is, local, provincial or national public hospitals. Moreover, the consultation time is limited to 10 minutes per patient in public facilities. If patients need further examination, such as a blood test or chest X-ray they should obtain an additional referral. If this is the case, the patients need to travel back and forth at least three times not only due to the referral system but also due to the time limit for consultation during outpatient care. In addition, there is widespread anecdotal experience that patients with diabetes mellitus or hypertension, which require regular control at primary public health facilities, are advised by the local health office (*Dinas Kesehatan*) to meet the doctor every one or two weeks, while in fact this is not necessary and it could be every four to eight weeks.[29] As a consequence, this procedure, along with anecdotal experiences, contributes to the higher number of visits for insured people with subsidies.

For inpatient care, patients with insurance prefer to choose public hospitals, private hospitals and primary health care facilities, respectively, and they spend a longer time there than uninsured patients. As our sample consists of more than half of insured people who are subsidized, they opt for public health facilities. In addition, higher income groups spend more days receiving inpatient care than lower income groups.

If we break down the total household health expenditure, it shows that the share of curative expenditure is around double the share of preventive expenditure for both insured and uninsured groups. This implies that insured people still need to pay to cure their illnesses because the insurance package does not cover all types of illness or a co-payment method applies in the insurance scheme. In terms of per capita expenditure, the curative cost is five times higher than the preventive cost. This suggests that the cost of curing illnesses, such as hospital fees, health treatment or health examination, medicine and so on, is still expensive for both insured and uninsured people.

[29] This advice is mostly derived to ensure that the patients obey the routine control and take their medicine properly as ordered, but it places a higher burden on the doctor.

Table 5.1 Descriptive Statistics

Variable	Definition	Insured[30]	Uninsured	Difference
Age	Age of household head in years	49.12	48.01	−1.11***
Married	Percentage married	81.75	80.31	−1.44***
Education	Years of schooling of household head	7.71	7.22	−0.49***
Formal Employment	Formal employment status of household head	46.16	39.02	−7.14***
Working Activity	Percentage worked in the last week	49.49	51.49	2.00***
Household Size	Number of household members	3.95	3.60	−0.35***
Housing Quality	Index of housing material (the higher the index, the better the quality)	67.40	68.34	0.94***
Household Expenditure	Per capita monthly household expenditure (thousand rupiah)	1,629.24	1,599.17	−30.07***
People Reported Illness	Number of household members reporting illness in the last month	0.34	0.31	−2.61***
People Reported Health Disruption	Number of household members above 5 y.o. who reported health disruption in the last month	0.36	0.33	−2.93***
Days of Health Disruption	Per capita (above 5 y.o.) days of health disruption in the last month	1.94	2.08	0.15***
People Taking Outpatient Care	Number of household members visiting health facilities in the last month	0.33	0.30	−0.03***
People Taking Inpatient Care	Number of household members taking inpatient care over the last 12 months	0.02	0.03	−0.01***
Frequency of Outpatient Visits	Per capita visits to health facilities in the last month	0.26	0.21	−0.04***
Days of Inpatient Care	Per capita days spent at health facilities over the last 12 months	0.19	0.11	−0.08***
Curative Expenditure	Per capita monthly expenditure to cure illness (rupiah)	34,579	28,521	−6,057.86***
Preventive Expenditure	Per capita monthly expenditure to prevent illness (rupiah)	6,214	5,617	−597.29***
Medical Check-Up Expenditure	Per capita monthly medical check-up (rupiah)	18,878	16,369	−2,509.46
Share of Curative Expenditure in Total Health Expenditure	Percentage of curative to total health expenditure (%)	67.61	64.00	−3.61***
Share of Preventive Expenditure in Total Health Expenditure	Percentage of preventive to total health expenditure (%)	32.39	36.00	3.61***

[30] Among insured households, 71,547 and 34,981 households are registered as subsidy insurance and non-subsidy insurance holders, respectively.

Continued

Share of Medical Check-Up Expenditure in Total Health Expenditure	Percentage of medical check-up to total health expenditure (%)	0.44	0.40	−0.04
Share of Curative Expenditure in Total Expenditure	Percentage of curative to total household expenditure (%)	1.38	1.21	-0.17^{***}
Share of Preventive Expenditure in Total Expenditure	Percentage of preventive to total household expenditure (%)	0.35	0.35	0.00
Immunization	Number of immunizations received by children under 5 y.o.	1.19	1.00	-0.19^{***}
Giving birth	Giving birth with a certified birth attendant	1.13	1.11	-0.02^{***}
Days of Breastmilk	Number of days consuming breastmilk	152.15	137.50	-14.65^{***}
Cigarette Consumption	Per capita adult weekly cigarette consumption (pack @ 10 pieces)	1.98	2.25	0.27^{***}
Alcohol Consumption	Per capita adult weekly alcohol consumption (bottle @ 620 ml)	0.86	1.21	0.35^{***}
Carbohydrate Consumption	Per capita daily carbohydrate consumption (grams)	515.67	472.55	-43.12^{***}
Protein Consumption	Per capita daily protein consumption (grams)	88.24	83.24	-5.00^{***}
Fat Consumption	Per capita daily fat consumption (grams)	62.64	61.22	-1.41^{***}
Calorie Consumption	Per capita daily calorie consumption (cal)	2826.61	2652.49	-174.12^{***}
Number of Observations		106,530	96,394	

Notes: *** p<0.01.

In our data, medical check-ups are part of preventive care, and the composition of preventive care expenditure is as follows: medical check-ups with about 38 per cent followed by pregnancy examination, self-preventive, immunization and birth control with 26 per cent, 16 per cent, 14 per cent and 6 per cent, respectively. This expenditure rises as age, education and income increase for all households.

Moreover, in terms of mother–child health, insured households show their awareness of health. For instance, children under 5 years old received more immunization and mothers opted to give birth with a certified birth attendant and breastfeed their babies for much longer than their counterparts from uninsured households. However, if we look at mothers' educational background, highly educated mothers tend to breastfeed their babies for shorter periods than low educated mothers. This might be due to their limited maternity leave, requiring educated mothers to return to work immediately.[31] On the other hand, the children of low educated mothers and poor families have more complete immunization than the children of their counterparts. This might be due to the results of the national immunization programme, which focuses on poor families (as well as low educated mothers), while fewer immunized children of highly educated mothers and richer families might be due to their perception of the religious belief against the immunization programme (Wombwell et al. 2015).

If we look at the pattern of preventive and curative expenditure based on having children, age and education, the ratio of curative to preventive expenditure is higher for families with children, suggesting that having children is associated with higher curative costs than for families with no children. Nevertheless, preventive care decreases while curative care increases as people age, so the ratio of curative to preventive care for old people is more than double that for young people. We also consider how education plays an important role in shaping health spending. Preventive care expenditure increases while curative care decreases as the education level rises. The ratio of curative to preventive care is higher with a higher education level, suggesting that curing illnesses, such as selecting medical treatment or the type of medicine, is associated with knowledge. These patterns are also shown for all households.

In addition, with regard to unhealthy consumption, the insured group consumes less alcohol and fewer cigarettes than the uninsured group. Indonesia is now facing high-risk smoking behaviour whereby almost 30 per cent of the population above 10 years old smokes. The prevalence of smoking in the school age group (10 to 18 years old) increased from 7.2 per cent in 2013 to 9.1 per cent in 2018 (Dartanto et al. 2019). This young generation is likely to be part of the uninsured group, which consumes 2.25 packs per week.

[31] According to the Ministry of Manpower Government of Indonesia, the total maternity leave for a mother is three months, which could be used one month prior to the due date and two months afterwards.

Moreover, insured people consume more excessive carbohydrates, protein, fat and calories than their uninsured counterparts. It appears that the insured group pays less attention to the diet nutrients suggested by the Ministry of Health; thus, we can indicate that the insured group displays the ex ante moral hazard phenomenon. This unhealthy behaviour also leads to severe illnesses, such as diabetes, high blood pressure, heart attacks and so on, which might partly be covered by the insurance package and hence have serious implications for the financing burden as well as the health status

.

5.6 Results and Discussion

This section presents our results from (i) the basic model of the demand for unhealthy consumption, outpatient care and inpatient care and (ii) the modified model for outpatient and inpatient care.

5.6.1 Basic Model

5.6.1.1 Unhealthy Consumption

Following our hypothesis that "all-you-can-eat" behaviour exists for insured people in Indonesia, as discussed in the previous section, our results show different patterns for the consumption of cigarettes, alcohol and excessive carbohydrates, protein, fat and calories between (i) the uninsured and the insured group either with or without a premium subsidy and (ii) insured people with and without a premium subsidy.

As we discussed in the previous section, the demand for unhealthy consumption is estimated to determine whether ex ante moral hazard exists for insured households either with or without a subsidy. We also introduce the interaction between insurance and frequency of visits to gauge whether visiting health facilities could improve the awareness of a healthy lifestyle when insurance exists. In estimating the demand for unhealthy consumption, we use OLS for cigarette and alcohol consumption and logit for dummy excessive carbohydrates, protein, fat and calories.[32]

The results from our basic model of unhealthy consumption are summarized in Table 5.2, in which the last four columns denote the marginal effects. This table exhibits how the type of insurance is associated with the demand for unhealthy consumption. Compared with uninsured households, insured households have an increased probability of excessively consuming carbohydrates, protein and calories but decreased cigarette and alcohol consumption. However, when insured people receive a subsidy from the Government, they increase their cigarette consumption but decrease their probability of consuming excessive fat compared with those without a subsidy.

Table 5.2 Relationship between Insurance and Unhealthy Consumption

Variable	*Dependent Variable: Unhealthy Consumption (UCi)*					
	Cigarettes	Alcohol	Dummy Carbohydrates	Dummy Protein	Dummy Fat	Dummy Calories
d_insurance	–0.0617***	–0.0047***	0.0178***	0.0105***	0.0001	0.0158***
	(–0.0492)	(–0.0137)	(0.0025)	(0.0018)	(0.0015)	(0.0021)
d_subsidy	0.0973***	0.0008	–0.0026	0.0005	–0.0209***	0.0061
	(0.0750)	(0.0021)	(0.0047)	(0.0033)	(0.0029)	(0.0038)
Interaction						
*d_insurance * visit*	–0.0004	–0.0005	–0.0002	0.0037	0.0067***	0.0034
	(–0.0003)	(–0.0012)	(0.0046)	(0.0032)	(0.0021)	(0.0036)
*d_subsidy * visit*	–0.0104	–0.0020	–0.0093	0.0003	0.0089***	0.0032
	(–0.0075)	(–0.0049)	(0.0069)	(0.0049)	(0.0034)	(0.0052)

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standardized beta coefficients in parentheses for cigarettes and alcohol. Standard error in parentheses for the marginal effects of carbohydrates, protein, fat and calories. This table is a summary of Appendices 5.B1–5.B2.

These results indicate that there is ex ante moral hazard in consuming excessive carbohydrates, protein and calories for insured people and consuming more cigarettes for subsidized people. For instance, the probability for insured people to consume excessive carbohydrates is 0.02 per cent higher than for those without insurance, and subsidized people, who are likely to be poor people, consume 0.1 per cent more

[32] We also use OLS to estimate demand for unhealthy consumption in continuous form, and the results have similar patterns.

cigarettes but have a lower probability of consuming excessive fat than those without a subsidy (Table 5.2).

Unhealthy consumption could also be avoided by visiting health facilities or having contact with professional health workers, ensuring that patients are well informed about the consequences of unhealthy consumption (Dave and Kaestner 2009). Our results confirm this hypothesis as more frequent visits to health facilities are associated with lower levels of unhealthy consumption (Appendix A5.B1–A5B.2). However, when we add the interaction between insurance and frequency of visits, our results are no longer statistically significant in supporting this hypothesis, suggesting that the frequency of visits does not affect both uninsured and non-subsidized households in cutting unhealthy consumption. Moreover, we find a positive and statistically significant coefficient for excessive fat, suggesting that, when insured and subsidized households visit health facilities more frequently, their probability of consuming excessive fat becomes higher compared with the non-insured and non-subsidized groups (Table 5.2). The typical favourite Indonesian meals or snacks are deep-fried or fat-based meals (*gorengan*). Intuitively, patients with insurance or subsidies prefer these meals to others during their visits to health care facilities, maybe because these meals are easy to find at an affordable price.

We are also interested in investigating how the insurance premium and its interaction with the frequency of visits and marital status and its interaction with income correlate with unhealthy consumption (Appendix A5.B1–A5.B2). The insurance premium has a negative correlation with the demand for cigarettes, but it turns positive when we interact it with frequency of visits. Further, a higher premium tends to increase the probability of consuming excessive carbohydrates but decrease the probability of consuming excessive calories. Despite small changes, the marginal effect is slightly larger for subsidized households than for insured households. These findings suggest that there is ex ante moral hazard when insurance and subsidies exist.

Marital status is also important in determining the demand for unhealthy consumption. Our results indicate that married households consume more cigarettes and alcohol but that their probability of consuming excessive carbohydrates, protein, fat and calories is lower than that of unmarried households. However, if we add the interaction between marital status and income, it presents the opposite correlation with the demand for the above types of consumption. These findings imply that an increasing probability of consuming unhealthy food is likely to be due to the affordability. It appears that income after marriage or joint income could foster a higher probability of consuming those foods as there is more money available to spend in the household. On the other hand, the lower demand for cigarettes and alcohol after interacting income and marital status could be interpreted as showing increasing awareness after marriage.

We control our estimation with income, education, age, housing quality and dummy variables for employment status, children and urban area. Our results show that increasing income per capita and formal employment status are positively associated with the demand for cigarettes, alcohol and excessive fat consumption and negatively associated with the probability of consuming excessive carbohydrates, protein and calories (Appendix A5.B1–A5B.2).

Households in which the head of the family spent more years on schooling have a higher probability of consuming more excessive carbohydrates, protein, fat and calories but consume fewer cigarettes and less alcohol. These findings imply that the level of education, where the average number of years of schooling is below 8 years, is not sufficient to enhance the awareness of healthy consumption. Age is also associated with an increasing probability of consuming excessive carbohydrates, protein and calories.

Households that have children tend to increase their probability of consuming excessive protein, fat and calories but decrease their probability of consuming carbohydrates, cigarettes and alcohol. Our findings imply that households with children are lack knowledge of the recommended consumption amounts of diet nutrients issued by the Ministry of Health. Although protein, fat and calories are important nutrients for children, excessive consumption of those meals could relate with their health in the future. On the other hand, the decreasing cigarette and alcohol consumption of households with children indicates an awareness of health.

Housing quality as a proxy for wealth and the urban dummy are strongly positively correlated with alcohol consumption, while the rest of the unhealthy forms of consumption have a negative correlation. This clearly explains that alcohol is consumed by relatively rich households and that those who are living in urban areas consume more alcohol than their counterparts.

5.6.1.2 Outpatient and Inpatient Care

If we look at the pattern of outpatient and inpatient care after introducing insurance, households visit health facilities 0.03 times more and spend 0.07 more days at health facilities than those without insurance. Among the insured households, people with a subsidy visit 0.04 times more but spend 0.02 fewer days there than households with no subsidy (Table 5.3). These findings confirm that insurance could expand accessibility to health care (Agustina et al. 2019). Since subsidy insurance is part of the poverty alleviation programme, the more frequent visits for outpatient care but fewer days for inpatient care for insured people with a subsidy have two implications: poor people have greater awareness after visiting health facilities more frequently, so they experience fewer severe illnesses that need inpatient treatment, or they are too poor or too old to access inpatient treatment, as shown in the negative and statistically significant coefficient of the interaction with age.

Our results also show that subsidized households increase their frequency of visiting health care as their income increases. It seems that visiting health care more frequently could be perceived as substituting inpatient treatment for poor people. The latter is in line with the BPJS (2018), in which the number of JKN members visiting hospital for outpatient care increased from 39.8 million cases in 2015 to 64.4 million cases in 2017 or approximately 62 per cent more in 2 years. However, the more frequent visits for outpatient care made by subsidized people could also reflect the referral system in accessing public health facilities and the consultation time limit, as discussed in section 5.4.

If we look further at the interaction of this insurance with income, the coefficient exhibits the strongest correlation with both outpatient and inpatient care, as shown in its standardized beta coefficients. As income increases, insured households visit health facilities more frequently and spend more days receiving inpatient care. This pattern holds with age but not with education. They visit less frequently and spend fewer days as inpatients as their years of schooling increase. This suggests that education could improve the health status of insured people.

On the other hand, when insured households are subsidized by the Government, they visit more frequently as well as spending more days receiving inpatient care as their level of education rises. If we add 1 more year of schooling, we expect subsidized households to increase their frequency of visits by 0.04 times and spend 0.05 days in hospital compared with their non-subsidized counterparts (Table 5.3, standardized beta coefficient). This suggests that poor people with subsidized insurance could access health facilities if they are more knowledgeable, that is, able to understand the procedures to obtain free health care at clinics, including preparing the paperwork that is required to access free inpatient care at a hospital or free advanced examinations.

Table 5.3 Relationship between Insurance and Health Care (Basic Model)

Variable	Frequency of Visits as Outpatient	Days Spent as Inpatient
d_insurance	0.0321***	0.0680***
	(0.0318)	(0.0364)
d_subsidy	0.0386***	–0.0155*
	(0.0360)	(–0.0071)
Interaction Insurance Dummy		
d_insurance ∗ age	0.0149**	0.0736***
	(0.0568)	(0.1523)
d_insurance ∗ educ	–0.0010**	–0.0015*
	(–0.0109)	(–0.0088)
d_insurance ∗ income	0.0109***	0.0350***
	(0.1415)	(0.2451)
Interaction Subsidy Dummy		
d_subsidy ∗ age	–0.0028	–0.0864***
	(–0.0099)	(–0.1533)
d_subsidy ∗ educ	0.0042***	0.0114***
	(0.0358)	(0.0478)
d_subsidy ∗ income	0.0472***	–0.0172
	(0.5630)	(–0.1008)

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standardized beta coefficients in parentheses. This table is a summary of Appendices 5.C1–5.C2.

The complex and bureaucratic procedure for subsidized patients to access free health care might explain these findings. Sambodho (2019) described the complicated procedure that a subsidized (poor) patient should follow to gain a bed in a hospital.

The patient should bring at least three documents,[33] including an identification card (ID and family card), subsidized insurance card and referral letter from a primary health care provider, every time he or she needs inpatient treatment. However, those documents do not guarantee that the subsidized patient can automatically obtain a bed in a public hospital, mostly because of the lack of bed availability. To overcome this situation, the subsidized patient usually seeks help from the intermediary or broker services in his or her local community to argue successfully with the administrative staff and finally gain a bed in a hospital in time (Sambodho 2019). Thus, our findings suggest that education plays an important role in accessing both outpatient and inpatient care for the poor while being older and being richer are associated with more access to health care for households with insurance.

In addition, we examine the health care cost, that is, the cost of curative and preventive treatment and medical check-ups, insurance premium, marital status and its interaction with income, and unhealthy consumption in estimating the frequency of visits for outpatient care and days spent in inpatient care. Our results indicate that the curative cost and medical check-ups are positively correlated with outpatient and inpatient care while the preventive cost shows the opposite sign. A high insurance premium is associated with less frequent visits to outpatient care but more days spent receiving inpatient care. Marital status and its interaction with income shows a positive and negative correlation, respectively, with outpatient care and neither is statistically significant in the inpatient model. These findings suggest that the curative cost as the price of health care and medical check-ups arises during visits to health facilities. If we look at married households, it appears that they visit more frequently for outpatient services, but this depends on their income. A lower income for married households tends to reduce the frequency of visits.

[33] Many poor (subsidized) patients do not even have basic documents, such as ID and/or a family card and subsidized insurance card. When they are ill, they do not know how to obtain a referral from the primary health care facility. If this is the case, the patient should obtain additional documents from various offices, specifically a poverty reference letter from the village office, official proof and stamp on that letter from the sub-district office and a fee waiver letter from the welfare office at the district level. Obtaining letters from those different offices is not an easy task and is time consuming, so most subsidized people seek help from the local broker of their community and pay to cover at least the broker's transportation cost (Sambodho 2019).

As we expect, our control variables per capita income and age confirm a positive correlation while education, formal employment status and having children show negative correlations with both the frequency of visits for outpatient care and the number of days spent receiving inpatient care. Housing quality as a proxy for wealth and living in an urban area reveal a positive correlation with days spent on inpatient treatment. Our results suggest that inpatient care is more accessible for wealthier and elderly people and those who live in urban areas, where the number of health facilities is much higher than that in rural areas. Being more educated is important to improve people's health status, as presented by the decreasing frequency of visits and number of days spent in hospital. Households that have children tend to follow a healthy lifestyle, reducing both the frequency of visits and the number of days in inpatient care. In addition, households in which the head of the family has formal employment status are associated with a lower demand for both outpatient and inpatient care. The full regression output of the basic outpatient and inpatient care model can be found in Appendices A5.C1–A5.C2.

5.6.2 Modified Model

We modify our basic model by introducing unhealthy consumption, specifically cigarettes, alcohol and excessive carbohydrates, protein, fat and calories, to understand how these types of consumption affect the frequency of visits and number of days spent receiving inpatient care if insurance exists.

5.6.2.1 Frequency of Visits for Outpatient Care

Table 5.4 presents a summary of the modified outpatient model in which we introduce unhealthy consumption and its interaction with insurance. The consumption of cigarettes and excessive fat is associated with less frequent visits to health care facilities. When we interact unhealthy consumption with insurance, our results show there is no statistical difference in the frequency of visits between insured and uninsured households if they consume cigarettes, alcohol or excessive carbohydrates, protein and fat. This pattern is not shown for subsidized households. Compared with the non-subsidized group, subsidized households visit health facilities less frequently when they consume cigarettes but multiply the frequency of visits if they consume excessive fat. The latter denotes a moral hazard phenomenon among subsidized households. It appears that consuming excessive fat tends to worsen the health status of subsidized people and is more severe than cigarette consumption, which exerts an effect on smokers in the long run. The role of insurance and subsidies and their interactions with age, education and income in outpatient care are consistent with the results of the basic model in the previous section. The full regression output of modified outpatient care can be found in Appendices A5.D1–A5.D6.[34]

[34] In addition, we use variable of unhealthy consumption in continuous form to estimate the frequency of visits for outpatient care and these results show a similar pattern.

Table 5.4 Relationship between Insurance and Frequency of Visits for Outpatient Care with Unhealthy Consumption (Modified Model)

Variable	*Dependent Variable: Frequency of Visits (OPi)*					
	Unhealthy Consumption					
	Cigarettes	Alcohol	Dummy Carbohydrates	Dummy Protein	Dummy Fat	Dummy Calories
Unhealthy Consumption	–0.0191***	–0.0019	0.0001	–0.0009	–0.0113***	–0.0007
	(–0.0238)	(–0.0006)	(0.0001)	(–0.0006)	(–0.0068)	(–0.0006)
d_insurance	0.0310***	0.0321***	0.0321***	0.0321***	0.0321***	0.0321***
	(0.0307)	(0.0318)	(0.0318)	(0.0318)	(0.0318)	(0.0318)
d_subsidy	0.0404***	0.0386***	0.0386***	0.0386***	0.0385***	0.0386***
	(0.0377)	(0.0360)	(0.0360)	(0.0360)	(0.0359)	(0.0360)
Interaction Insurance Dummy						
*d_insurance * age*	0.0153**	0.0149**	0.0149**	0.0148**	0.0150**	0.0148**
	(0.0586)	(0.0569)	(0.0568)	(0.0568)	(0.0572)	(0.0568)
*d_insurance * educ*	–0.0010**	–0.0010**	–0.0010**	–0.0010**	–0.0010**	–0.0010**
	(–0.0113)	(–0.0109)	(–0.0109)	(–0.0109)	(–0.0110)	(–0.0109)
*d_insurance * income*	0.0105***	0.0109***	0.0109***	0.0109***	0.0108***	0.0109***
	(0.1356)	(0.1415)	(0.1414)	(0.0086)	(0.1396)	(0.1412)
*d_insurance * unhealthycons*	0.0016	0.0030	0.0051	0.0088	–0.0025	–0.0084*
	(0.0015)	(0.0008)	(0.0046)	(0.0086)	(–0.0025)	(–0.0081)
Interaction Subsidy Dummy						
*d_subsidy * age*	–0.0026	–0.0028	–0.0028	–0.0028	–0.0027	–0.0028
	(–0.0095)	(–0.0099)	(–0.0049)	(–0.0099)	(–0.0098)	(–0.0100)
*d_subsidy * educ*	0.0045***	0.0042***	0.0042***	0.0042***	0.0042***	0.0042***
	(0.0387)	(0.0358)	(0.0355)	(0.0359)	(0.0359)	(0.0358)
*d_subsidy * income*	0.0514***	0.0472***	0.0474***	0.0473***	0.0472***	0.0472***
	(0.6132)	(0.5630)	(0.5590)	(0.5643)	(0.5633)	(0.5630)
*d_subsidy * unhealthycons*	–0.0142***	0.0014	–0.0016	–0.0038	0.0393***	–0.0012
	(–0.0145)	(0.0004)	(–0.0017)	(–0.0037)	(0.0380)	(–0.0011)

Notes: *** $p<0.01$, ** $p<0.05$, * $p<0.1$. Standardized beta coefficients in parentheses. This table is a summary of Appendices 5.D1–5.D6.

5.6.2.2 Days Spent Receiving Inpatient Care

We also look at the correlation of each form of unhealthy consumption with the length of time spent receiving inpatient care when insurance exists by interacting these two variables. The consumption of excessive carbohydrates, protein and calories is positively correlated with inpatient care and statistically significant, while cigarette and alcohol consumption show the opposite. When we interact unhealthy consumption with insurance, our results indicate that there is ex ante moral hazard where insured people, compared with the uninsured group, spend more days receiving inpatient care when they consume excessive carbohydrates. It seems that illnesses derived from excessive carbohydrate consumption require hospitalization. From the health literature, we find that excessive carbohydrate consumption mostly causes diabetes, which leads to potential complications such as heart disease, stroke and kidney and nerve damage. Once these conditions combine, they can lead to a life-threatening stage that requires inpatient care. Nevertheless, people spend fewer days receiving inpatient treatment when they consume excessive fat and calories. It appears that excessive consumption of fat and calories, which likely causes hypertension and obesity, leads to severe complications in the long run. In addition, there is no statistical difference in inpatient care between subsidized and non-subsidized households if they consume cigarettes, alcohol or excessive carbohydrates, protein, fat and calories (Table 5.5).

The role of insurance and subsidies and their interactions with age, education and income in inpatient care are consistent with the results of the basic model in the previous section. The full regression output of the modified inpatient care model can be found in Appendices A5.E1–A5.E6.[35]

[35] In addition, we use variable of unhealthy consumption in continuous form to estimate the number of days spent receiving inpatient care and these results show a similar pattern.

Table 5.5 Relationship between Insurance and Days Spent in Inpatient Care and Unhealthy Consumption (Modified Model)

Variable	*Dependent Variable: Days Spent (IPi)*					
	Unhealthy Consumption					
	Cigarettes	Alcohol	Dummy Carbohyd rates	Dummy Protein	Dummy Fat	Dummy Calories
Unhealthy Consumption	–0.0447***	–0.0204*	0.0118***	0.0174***	–0.0084	0.0094*
	(–0.0301)	(–0.0037)	0.0063	(0.0067)	(–0.0027)	(0.0044)
d_insurance	0.0655***	0.0679***	0.0678***	0.0678***	0.0680***	0.0679***
	(0.0351)	(0.0364)	(0.0363)	(0.0363)	(0.0364)	(0.0364)
d_subsidy	–0.0111	–0.0155*	–0.0154*	–0.0155*	–0.0158*	–0.0155*
	(–0.0051)	(–0.0071)	(–0.0071)	(–0.0071)	(–0.0073)	(–0.0071)
<u>Interaction Insurance Dummy</u>						
d_insurance ∗ age	0.0747***	0.0737***	0.0737***	0.0739***	0.0737***	0.0737***
	(0.1545)	(0.1524)	(0.1523)	(0.1527)	(0.1524)	(0.1525)
d_insurance ∗ educ	–0.0016*	–0.0015*	–0.0015*	–0.0015*	–0.0015*	–0.0015*
	(–0.0092)	(–0.0087)	(–0.0088)	(–0.0089)	(–0.0088)	(–0.0088)
d_insurance ∗ income	0.0339***	0.0350***	0.0352***	0.0351***	0.0349***	0.0351***
	(0.2376)	(0.2453)	(0.2467)	(0.2458)	(0.2443)	(0.2462)
d_insurance ∗ unhealthycons	–0.0054	0.0063	0.0184**	0.0053	–0.0375***	–0.0239**
	(–0.0027)	(0.0009)	(0.0090)	(0.0028)	(–0.0201)	(–0.0126)
<u>Interaction Subsidy Dummy</u>						
d_subsidy ∗ age	–0.0860***	–0.0864***	–0.0867***	–0.0865***	–0.0859***	–0.0864***
	(–0.1526)	(–0.1533)	(–0.1538)	(–0.1534)	(–0.1523)	(–0.1532)
d_subsidy ∗ educ	0.0122***	0.0114***	0.0115***	0.0114***	0.0114***	0.0114***
	(0.0511)	(0.0478)	(0.0482)	(0.0479)	(0.0478)	(0.0478)
d_subsidy ∗ income	–0.0084	–0.0172	–0.0156	–0.0166	–0.0169	–0.0171
	(–0.0492)	(–0.1008)	(–0.0917)	(–0.0975)	(–0.0993)	(–0.1005)
d_subsidy ∗ unhealthycons	0.0052	0.0167	–0.0134	–0.0143	0.0038	–0.0047
	(0.0026)	(0.0026)	(–0.0063)	(–0.0068)	(0.0018)	(–0.0023)

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standardized beta coefficients in parentheses. Note: This table is a summary of Appendices 5.E1–5.E6.

5.7 Conclusion and Policy Implications

Given the high disparity in access to health care, one of the goals of the Indonesian universal health care programme (JKN) is to ensure that all Indonesians have access to health care. However, our findings show that this policy regime has invoked ex ante moral hazard whereby insured and subsidized households tend to adopt relatively unhealthy lifestyles by increasing their unhealthy consumption when insured. The results in this chapter show that insured people behave differently from uninsured people, and within the insured groups, the subsidized people behave differently from the non-subsidized people. These different behaviours may also appear from the implementation of the JKN in 2014 onwards.

Compared with uninsured households, the probability of consuming excessive carbohydrates, protein and calories is significant higher for their insured counterparts. Subsidized people, who are likely to be poor, even increase their cigarette consumption compared with unsubsidized people. Further, we found unhealthy consumption is positively correlated with demand for health care after introducing insurance. Insured people who consume excessive carbohydrates spend more days for inpatient care compared with their uninsured counterparts while subsidized people who consume excessive fat visit health facilities more frequent compared with the non-subsidized group. These findings suggest that ex ante moral hazard among insured and subsidized people exists and tends to increase demand for health care, and hence raises insurance claims.

We confirm that the insurance programme, with or without subsidies, improves the accessibility of health care. Compared with the uninsured group, insured households visit outpatient care more frequently and spend more days in inpatient care. This holds with age and income but not with education. More knowledgeable people visit less frequently and spend fewer days receiving inpatient care. Among insured people, subsidized people visit more frequently and spend more days there if they are more knowledgeable than non-subsidized members. This also holds for income for outpatient but not for inpatient care. It appears that the poor prefer to access outpatient care than inpatient care. This result has two implications: (i) poor people have greater awareness after visiting health facilities more frequently, so they experience less severe illnesses that do not require inpatient treatment, or (ii) they are too poor and too old to access inpatient treatment. The latter could be perceived as indicating that poor people tend to substitute inpatient care with visiting health care facilities more frequently. However, this high frequency of visits for outpatient care by subsidized people could also reflect the referral system in accessing public health facilities and the consultation time limit.

As we expected, education turns out to play an important role in improving health status. Becoming older and becoming richer are associated with greater use of inpatient treatment for insured households but being knowledgeable is more important for subsidized people to understand the procedure for accessing free health care. Our findings suggest that giving more access to health care solely by providing universal health care is inadequate. People should be made aware that it is and will remain crucial to maintaining their health status. Therefore, the insurance programme should cover preventive care as well as curative care. Moreover, an incentive system is necessary to change the behaviour of insurance members to follow a healthy lifestyle.

In addition, complex and bureaucratic procedures in accessing free health care for the poor should be eliminated. For the middle and upper income classes, the Government and the Parliament could amend National Social Security (SJSN) Law No. 40, which was passed in 2004, and BPJS Law No. 24, passed in 2011, which limit co-payment to obtain a higher class of inpatient care. The BPJS could determine the base provision and optional services such as MRIs, complete medical check-ups and so on depending on their willingness to pay the premium.

Since moral hazards could exist on both sides – insured people and providers – the next avenue for future research is to identify incentives for both parties to control the moral hazard effects in the health care system. Further, the JKN has been implemented for at least five years, so it remains interesting to test our hypothesis among JKN members to determine whether they show ex ante moral hazard.

Appendices

Appendix A5.A1 Data Limitations

Susenas, a socio-economic national survey, is now conducted every year throughout Indonesia by the Central Statistics Bureau BPS. Susenas consists of two modules, namely the consumption and core modules. The consumption survey records all expenditures, including health expenses, at the household level on a weekly and monthly basis for food and non-food expenditures, respectively. In this chapter, we therefore define health expenditure based on the types of health expenditures listed in the Susenas consumption questionnaire. We classify health expenditure into curative and preventive expenditure. Curative expenditure consists of payments for public hospitals, private hospitals, primary health care (*puskesmas*), private clinics, health worker practices, traditional healer practices, traditional birth attendants, prescribed medicine, traditional medicine and self-treatment (including purchasing non-prescribed medicine). Preventive expenditure covers pregnancy examination, immunization, family planning or birth control, health prevention, medical check-ups and purchasing health devices such as glasses, plaster cast, wheelchairs and so on. We calculate these expenditures as an average for the last three months and deflate them into real values. To analyse health expenditure, we calculate these expenditures in terms of the per capita share in the total health expenditure and the total expenditure. The frequency of outpatient visits and the number of days spent receiving inpatient care are extracted from the Susenas core module. The BPS records the frequency of visits for outpatient care for the last month and the number of days spent in inpatient care for the last year in each health facility. We calculate the per capita visits for outpatient care and the per capita days for inpatient care.

Despite the wide survey coverage and large sample size, Susenas, which we use to perform several regression analyses, has some limitations. Susenas is not a special survey designated for measuring detailed health expenditure and health services. We do not know exactly how much the households pay for each health-related service that they receive because the relevant questions are asked in separate modules, as discussed above. The respondents cannot give explicit answers about the payment for administration or laboratory or other related health expenses paid by the households. These typical questions can only be answered satisfactorily by exit inpatients or outpatients survey who are interviewed directly on site (see Hidayat et al 2015). Instead, Susenas's respondents are general households interviewed at their homes and answering questions on expenditures made during the last week, the last month and the last 3 months, the frequency of visits in the last month and the number of days spent as inpatients over the last 12 months. As a consequence, Susenas's respondents have difficulty in answering the questions related to health expenditure precisely – their memory recall might be biased. This implies that the three variables used in this paper – (a) curative expenditure, (b) preventive expenditure and (c) other health expenditure – are much lower than the corresponding figures shown in other reports that use exit patient surveys to calculate health expenditure.[36] Similar problems occur when calculating the frequency of visits, which refers to outpatients a month ago, and days spent in health facilities, which refer to inpatients a year ago.

Further, health expenditure in Susenas could be misinterpreted. Health consumption is financed by out-of-pocket and/or health subsidy. However, calculating the size of health subsidy is difficult because health goods and services vary greatly in both type and intensity, and medical fees can be very expensive. Susenas relied on appraisal value from the respondent, but people in general do not know the exact price until they receive a bill. The problem occurred when the subsidy recipient do not receive the bill. Thus, health cost is likely to be underestimated especially among subsidized households with high medical needs (Johar et al. 2018). Having those limitations, the health figures calculated from Susenas should be interpreted carefully.

[36] Further information on the patient exit survey is available through the following link: https://health.bmz.de/what_we_do/Universal-Health-Coverage/Indonesia_on_the_way_to_universal_health_coverage/Policy_Brief_GIZ_SPP_OOP_Spending_in_Indonesian_Health_Insurance.pdf.

Appendix A5.B1 Relationship between Insured Households and Unhealthy Consumption

Variable	*Dependent Variable: unhealthy consumption (UCi)*					
	Log Cigarette	Log Alcohol	Dummy Carbohydrate	Dummy Protein	Dummy Fat	Dummy Calorie
d_insurance	–0.0617***	–0.0047***	0.0178***	0.0105***	0.0001	0.0158***
	(–0.0492)	(–0.0137)	(0.0025)	(0.0018)	(0.0015)	(0.0021)
freq_visit	–0.0547***	–0.0010	–0.002	–0.0038*	–0.0072***	–0.0081***
	(–0.0439)	(–0.0029)	(0.0033)	(0.0023)	(0.0015)	(0.0026)
log_premiun	–0.0152***	0.0001	0.0008***	0.0000	-0.0001	–0.0007***
	(–0.0459)	(0.0015)	(0.0003)	(0.0002)	(0.0001)	(0.0002)
pct_married	0.0038***	0.0009***	–0.0137***	–0.0038***	–0.0061***	–0.0109***
	(0.1693)	(0.1533)	(0.0008)	(0.0006)	(0.0004)	(0.0006)
Interaction						
*d_insurance * visit*	–0.0004	–0.0005	–0.0002	0.0037	0.0067***	0.0034
	(–0.0003)	(–0.0012)	(0.0046)	(0.0032)	(0.0021)	(0.0036)
*married * income*	–0.0001	–0.0001***	0.0009***	0.0002***	0.0005***	0.0007***
	(–0.0459)	–(0.1567)	(–0.0001)	(0.000)	(0.0000)	(0.0000)
*premium * visit*	0.0029*	0.0001	0.002	–0.0001	–0.001	–0.0005
	(0.0047)	(0.0003)	(0.0013)	(0.0009)	(0.0007)	(0.0010)
Control						
log_income	0.2877***	0.0048***	–0.9571***	–0.0600*	0.2212***	–0.6200***
	(0.3165)	(0.0193)	(0.0445)	(0.0336)	(0.0227)	(0.0366)
educ_hhh	–0.0168***	–0.0002**	0.0013***	0.0007***	0.0022***	0.0005**
	(–0.1339)	(–0.0065)	(0.0003)	(0.0002)	(0.0002)	(0.0002)
log_age_hhh	–0.2601***	–0.0034**	0.2054***	0.1722***	–0.0095	0.2347***
	(–0.1192)	(–0.0058)	(0.0165)	(0.0117)	(0.0092)	(0.0135)
housequality	–0.0027***	0.0006***	–0.0033***	–0.0014***	–0.0001*	–0.0030***
	(-0.0489)	(0.0388)	(0.0001)	(0.0001)	(0.0001)	(0.0001)
d_empl_hhh	0.0301***	0.0031***	-0.0434***	–0.0113***	0.0015	–0.0224***
	(0.0239)	(0.0092)	(0.0024)	(0.0017)	(0.0014)	(0.0021)
d_kids	–0.0038	–0.0133***	–0.1220***	0.0039*	0.1587***	0.2052***
	(–0.0025)	(–0.0319)	(0.0030)	(0.0022)	(0.0015)	(0.0021)
d_urban	–0.0740***	0.0058***	–0.0257***	–0.0165***	–0.0019	–0.0797***
	(-0.0590)	(0.0171)	(0.0025)	(0.0018)	(0.0015)	(0.0021)

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standardized beta coefficients in parentheses for Cigarette and Alcohol. Standard error in parentheses for marginal effect of Carbohydrate, Protein, Fat and Calorie.

Appendix A5.B2 Relationship between Subsidized Households and Unhealthy Consumption

Variable	*Dependent Variable: unhealthy consumption (UCi)*					
	Log Cigarette	Log Alcohol	Dummy Carbohydrate	Dummy Protein	Dummy Fat	Dummy Calorie
d_subsidy	0.0973***	0.0008	–0.0026	0.0005	–0.0209***	0.0061
	(0.0750)	(0.0021)	(0.0047)	(0.0033)	(0.0029)	(0.0038)
freq_visit	–0.0496***	–0.0000	0.0067	0.001	–0.0062**	–0.0048
	(–0.0410)	(–0.0001)	–0.0059	–0.0041	–0.003	–0.0044
log_premiun	–0.0105***	0.0002	0.0014***	–0.0001	–0.0004	–0.0010**
	(–0.0411)	(0.0020)	–0.0005	–0.0003	–0.0002	–0.0004
pct_married	0.0085***	0.0009**	–0.0175***	–0.0036***	–0.0066***	–0.0112***
	(0.3654)	(0.1291)	–0.001	–0.0007	–0.0005	–0.0008
Interaction						
*d_subsidy * visit*	–0.0104	–0.0020	–0.0093	0.0003	0.0089***	0.0032
	(–0.0075)	(–0.0049)	(0.0069)	(0.0049)	(0.0034)	(0.0052)
*married * income*	–0.0004***	–0.0001**	0.0012***	0.0002***	0.0005***	0.0008***
	(–0.2495)	(–0.1242)	–0.0001	–0.0001	0.0000	–0.0001
*premium * visit*	0.0021	–0.0001	0.0009	0.0002	–0.0008	0.0000
	(0.0044)	(–0.0009)	–0.0015	–0.001	–0.0007	–0.0011
Control						
log_income	0.3058***	0.0039**	–1.2224***	–0.0942**	0.0645**	–0.7711***
	(0.3593)	(0.0157)	(0.0612)	(0.0450)	(0.0307)	(0.0495)
educ_hhh	–0.0144***	–0.0001	0.0023***	0.0015***	0.0018***	0.0016***
	(–0.1239)	(–0.0038)	(0.0004)	(0.0003)	(0.0002)	(0.0003)
log_age_hhh	–0.2433***	0.0002	0.2046***	0.1640***	–0.0013	0.2397***
	(–0.1099)	(0.0003)	(0.0238)	(0.0167)	(0.0130)	(0.0193)
housequality	–0.0011***	0.0006***	–0.0033***	–0.0014***	–0.0001	–0.0028***
	(–0.0205)	(0.0399)	(0.0002)	(0.0001)	(0.0001)	(0.0001)
d_empl_hhh	0.0290***	0.0059***	–0.0585***	–0.0134***	0.0034*	–0.0199***
	(0.0234)	(0.0164)	(0.0034)	(0.0024)	(0.0020)	(0.0029)
d_kids	–0.0020	–0.0131***	–0.1093***	0.0133***	0.1510***	0.2063***
	(–0.0013)	(–0.0284)	(0.0043)	(0.0030)	(0.0020)	(0.0030)
d_urban	–0.0636***	0.0073***	–0.0377***	–0.0169***	–0.0017	–0.0800***
	(–0.0515)	(0.0202)	(0.0035)	(0.0025)	(0.0020)	(0.0029)

Notes: *** p<0.01, ** p<0.05, * p<0.1. Standardized beta coefficients in parentheses for Cigarette and Alcohol. Standard error in parentheses for marginal effect of Carbohydrate, Protein, Fat and Calorie.

Appendix A5.C1 Basic Model Frequency of Visit for Outpatient Care

`	Frequency Visit Outpatient Care							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
log_curecap	0.0531***	0.0549***	0.0531***	0.0531***	0.0531***	0.0549***	0.0548***	0.0548***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_prevcap	–0.0032***	–0.0041***	–0.0032***	–0.0032***	–0.0032***	–0.0041***	–0.0041***	–0.0040***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_medcheckcap	0.0041***	0.0033*	0.0041***	0.0042***	0.0041***	0.0033*	0.0034*	0.0037*
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
log_premium	–0.0022***	–0.0014**	–0.0022***	–0.0020***	–0.0024***	–0.0014**	–0.0012*	–0.0008
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
pct_married	0.0023***	0.0044***	0.0023***	0.0023***	0.0023***	0.0044***	0.0043***	0.0038***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
*married * income*	–0.0002***	–0.0003***	–0.0002***	–0.0002***	–0.0002***	–0.0003***	–0.0003***	–0.0003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_insurance	0.0321***		–0.0248	0.0395***	–0.1102***			
	(0.002)		(0.028)	(0.004)	(0.041)			
d_subsidy		0.0386***				0.0492	0.0008	–0.5842***
		(0.004)				(0.044)	(0.008)	(0.065)
*d_insurance * age*			0.0149**					
			(0.007)					
*d_insurance * educ*				–0.0010**				
				(0.000)				
*d_insurance * income*					0.0109***			
					(0.003)			
*d_subsidy * age*						–0.0028		
						(0.011)		
*d_subsidy * educ*							0.0042***	
							(0.001)	
*d_subsidy * income*								0.0472***
								(0.005)
log income	0.0132***	0.0319***	0.0130***	0.0134***	0.0072**	0.0320***	0.0330***	0.0023
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.005)
educ_hhh	–0.0032***	–0.0024***	–0.0033***	–0.0027***	–0.0033***	–0.0024***	–0.0049***	–0.0021***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)
log_age_hhh	0.0475***	0.0500***	0.0402***	0.0482***	0.0466***	0.0518***	0.0538***	0.0501***
	(0.004)	(0.006)	(0.005)	(0.004)	(0.004)	(0.009)	(0.006)	(0.006)
housequality	–0.0002**	–0.0004***	–0.0002**	–0.0002**	–0.0003**	–0.0004***	–0.0004***	–0.0005***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_empl_hhh	–0.0065***	–0.0067**	–0.0062***	–0.0062***	–0.0068***	–0.0066**	–0.0048	–0.0069**
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
d_kids	–0.0975***	–0.1165***	–0.0974***	–0.0975***	–0.0979***	–0.1166***	–0.1173***	–0.1173***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
d_urban	–0.0011	0.0053	–0.0012	–0.0012	–0.0009	0.0053	0.0047	0.0060*
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)
Constant	–0.2808***	–0.5206***	–0.2511***	–0.2911***	–0.1969***	–0.5282***	–0.5155***	–0.1140
	(0.045)	(0.063)	(0.047)	(0.045)	(0.051)	(0.071)	(0.063)	(0.076)
Observations	190,571	99,748	190,571	190,571	190,571	99,748	99,748	99,748
R-squared	0.167	0.174	0.167	0.167	0.167	0.174	0.175	0.175

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix A5.C2 Basic Model Days Spent Receiving Inpatient Care

Variable	Days Spent (Inpatient Care)							
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)
illness_cap	–0.0058*	–0.0037	–0.0058*	–0.0059*	–0.0058*	–0.0034	–0.0036	–0.0036
	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)
log_curecap	0.0358***	0.0427***	0.0358***	0.0358***	0.0358***	0.0426***	0.0425***	0.0427***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_prevcap	–0.0052***	–0.0058***	–0.0052***	–0.0052***	–0.0052***	–0.0059***	–0.0060***	–0.0058***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_medcheckcap	0.004	0.004	0.004	0.004	0.0039	0.0038	0.0043	0.0039
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)
log_premium	0.0038***	0.0023	0.0038***	0.0040***	0.0031***	0.0024*	0.0029**	0.002
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
pct_married	–0.0008	0.0017	–0.0009	–0.0008	–0.0008	0.0022	0.0014	0.0019
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)
*married * income*	0.0001	–0.0001	0.0001	0.0001	0.0001	–0.0001	–0.0001	–0.0001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_insurance	0.0680***		–0.2137***	0.0790***	–0.3879***			
	(0.004)		(0.056)	(0.008)	(0.082)			
d_subsidy		–0.0155*				0.3162***	–0.1177***	0.2108
		(0.009)				(0.096)	(0.017)	(0.143)
*d_insurance * age*			0.0736***					
			(0.015)					
*d_insurance * educ*				–0.0015*				
				(0.001)				
*d_insurance * income*					0.0350***			
					(0.006)			
*d_subsidy * age*						–0.0864***		
						(0.025)		
*d_subsidy * educ*							0.0114***	
							(0.002)	
*d_subsidy * income*								–0.0172
								(0.011)
log income	0.0783***	0.0942***	0.0775***	0.0786***	0.0591***	0.0966***	0.0970***	0.1050***
	(0.006)	(0.010)	(0.006)	(0.006)	(0.007)	(0.010)	(0.010)	(0.012)
educ_hhh	–0.0009*	–0.0023***	–0.0010*	–0.0000	–0.0011**	–0.0025***	–0.0092***	–0.0024***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_age_hhh	0.0704***	0.0795***	0.0345***	0.0715***	0.0676***	0.1354***	0.0897***	0.0794***
	(0.008)	(0.013)	(0.011)	(0.008)	(0.008)	(0.021)	(0.013)	(0.013)
housequality	0.0004**	0.0003	0.0004**	0.0004**	0.0004*	0.0003	0.0002	0.0004
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_empl_hhh	–0.0334***	–0.0476***	–0.0320***	–0.0329***	–0.0344***	–0.0442***	–0.0425***	–0.0476***
	(0.005)	(0.007)	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)
d_kids	–0.0489***	–0.0599***	–0.0482***	–0.0488***	–0.0501***	–0.0635***	–0.0622***	–0.0596***
	(0.006)	(0.009)	(0.006)	(0.006)	(0.006)	(0.009)	(0.009)	(0.009)
d_urban	0.0207***	0.0346***	0.0199***	0.0204***	0.0210***	0.0345***	0.0331***	0.0343***
	(0.005)	(0.007)	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)
Constant	–1.3392***	–1.5239***	–1.1921***	–1.3545***	–1.0705***	–1.7642***	–1.5102***	–1.6717***
	(0.089)	(0.139)	(0.094)	(0.090)	(0.101)	(0.155)	(0.139)	(0.167)
Observations	190,571	99,748	190,571	190,571	190,571	99,748	99,748	99,748
R-squared	0.031	0.035	0.031	0.031	0.031	0.035	0.035	0.035

Standard errors in parentheses
*** p<0.01, ** p<0.05, * p<0.1

Appendix A5.D1 Modified Model Frequency of Visit for Outpatient Care: Cigarette

Variable	Frequency Visit Outpatient Care									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
log_curecap	0.0529***	0.0547***	0.0529***	0.0529***	0.0529***	0.0529***	0.0547***	0.0546***	0.0546***	0.0547***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_prevcap	–0.0032***	–0.0040***	–0.0032***	–0.0032***	–0.0032***	–0.0032***	–0.0040***	–0.0041***	–0.0039***	–0.0040***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_medcheckcap	0.0039***	0.003	0.0039***	0.0039***	0.0039***	0.0039***	0.003	0.0031	0.0034*	0.003
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
log_premium	–0.0024***	–0.0016**	–0.0024***	–0.0023***	–0.0026***	–0.0024***	–0.0016**	–0.0014**	–0.0009	–0.0016**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
pct_married	0.0024***	0.0046***	0.0024***	0.0024***	0.0024***	0.0024***	0.0046***	0.0045***	0.0040***	0.0048***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
*married * income*	–0.0002***	–0.0003***	–0.0002***	–0.0002***	–0.0002***	–0.0002***	–0.0003***	–0.0003***	–0.0003***	–0.0004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_cigarcap	–0.0178***	–0.0195***	–0.0179***	–0.0179***	–0.0178***	–0.0187***	–0.0195***	–0.0205***	–0.0221***	–0.0100**
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.004)
d_insurance	0.0310***		–0.0277	0.0387***	–0.1054**	0.0305***				
	(0.002)		(0.028)	(0.004)	(0.041)	(0.002)				
d_subsidy		0.0404***					0.0505	-0.0003	-0.6376***	0.0448***
		(0.004)					(0.044)	(0.008)	(0.066)	(0.004)
*d_insurance * age*			0.0153**							
			(0.007)							
*d_insurance * educ*				–0.0010**						
				(0.000)						
*d_insurance * income*					0.0105***					
					(0.003)					
*d_insurance * cigarette*						0.0016				
						(0.003)				
*d_subsidy * age*							–0.0026			
							(0.011)			
*d_subsidy * educ*								0.0045***		
								-0.001		
*d_subsidy * income*									0.0514***	
									(0.005)	
*d_subsidy * cigarette*										–0.0142***
										(0.005)
log income	0.0184***	0.0380***	0.0182***	0.0186***	0.0126***	0.0184***	0.0381***	0.0394***	0.0065	0.0389***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)
educ_hhh	–0.0036***	–0.0026***	–0.0036***	–0.0030***	–0.0036***	–0.0036***	–0.0026***	–0.0054***	–0.0024***	–0.0026***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
log_age_hhh	0.0430***	0.0454***	0.0355***	0.0437***	0.0421***	0.0429***	0.0471***	0.0493***	0.0449***	0.0459***
	(0.004)	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)	(0.009)	(0.006)	(0.006)	(0.006)
housequality	–0.0003***	–0.0004***	–0.0003***	–0.0003***	–0.0003***	–0.0003***	–0.0004***	–0.0005***	–0.0005***	–0.0004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_empl_hhh	–0.0059***	–0.0062*	–0.0057**	–0.0056**	–0.0063***	–0.0060***	–0.0061*	–0.0041	–0.0063*	–0.0063*
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
d_kids	–0.0975***	–0.1164***	–0.0974***	–0.0974***	–0.0979***	–0.0975***	–0.1165***	–0.1173***	–0.1173***	–0.1160***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
d_urban	–0.0024	0.0041	–0.0025	–0.0026	–0.0023	–0.0024	0.0041	0.0034	0.0047	0.0041
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Constant	–0.3212***	–0.5755***	–0.2906***	–0.3319***	–0.2406***	–0.3206***	–0.5828***	–0.5728***	–0.1399*	–0.5940***
	(0.045)	(0.064)	(0.047)	(0.045)	(0.051)	(0.045)	(0.071)	(0.064)	(0.076)	(0.064)
Observations	190,571	99,748	190,571	190,571	190,571	190,571	99,748	99,748	99,748	99,748
R-squared	0.167	0.175	0.167	0.167	0.167	0.167	0.175	0.175	0.176	0.175

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix A5.D2 Modified Model Frequency of Visit for Outpatient Care: Alcohol

Variable	Frequency Visit Outpatient Care									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
log_curecap	0.0531***	0.0549***	0.0531***	0.0531***	0.0531***	0.0531***	0.0549***	0.0548***	0.0548***	0.0549***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_prevcap	–0.0032***	–0.0041***	–0.0032***	–0.0032***	–0.0032***	–0.0032***	–0.0041***	–0.0041***	–0.0040***	–0.0041***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_medcheckcap	0.0041***	0.0033*	0.0041***	0.0042***	0.0041***	0.0041***	0.0033*	0.0034*	0.0037*	0.0033*
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
log_premium	–0.0022***	–0.0014**	–0.0022***	–0.0020***	–0.0024***	–0.0022***	–0.0014**	–0.0012*	–0.0008	–0.0014**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
pct_married	0.0023***	0.0044***	0.0023***	0.0023***	0.0023***	0.0023***	0.0044***	0.0043***	0.0038***	0.0044***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
*married * income*	–0.0002***	–0.0003***	–0.0002***	–0.0002***	–0.0002***	–0.0002***	–0.0003***	–0.0003***	–0.0003***	–0.0003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_alcoholcap	–0.0006	0.0001	–0.0007	–0.0006	–0.0007	–0.0024	0.0001	0.0002	0.0002	–0.0010
	(0.006)	(0.008)	(0.006)	(0.006)	(0.006)	(0.010)	(0.008)	(0.008)	(0.008)	(0.017)
d_insurance	0.0321***		–0.0248	0.0395***	–0.1102***	0.0321***				
	(0.002)		(0.028)	(0.004)	(0.041)	(0.002)				
d_subsidy		0.0386***					0.0492	0.0008	–0.5842***	0.0386***
		(0.004)					(0.044)	(0.008)	(0.065)	(0.004)
*d_insurance * age*			0.0149**							
			(0.007)							
*d_insurance * educ*				–0.0010**						
				(0.000)						
*d_insurance * income*					0.0109***					
					(0.003)					
*d_insurance * alcohol*						0.003				
						(0.013)				
*d_subsidy * age*							–0.0028			
							(0.011)			
*d_subsidy * educ*								0.0042***		
								(0.001)		
*d_subsidy * income*									0.0472***	
									(0.005)	
*d_subsidy * alcohol*										0.0014
										(0.020)
log income	0.0132***	0.0319***	0.0130***	0.0134***	0.0072**	0.0132***	0.0320***	0.0330***	0.0023	0.0319***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)
educ_hhh	–0.0032***	–0.0024***	–0.0033***	–0.0027***	–0.0033***	–0.0032***	–0.0024***	–0.0049***	–0.0021***	–0.0024***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
log_age_hhh	0.0474***	0.0500***	0.0402***	0.0482***	0.0466***	0.0474***	0.0518***	0.0538***	0.0501***	0.0500***
	(0.004)	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)	(0.009)	(0.006)	(0.006)	(0.006)
housequality	–0.0002**	–0.0004***	–0.0002**	–0.0002**	–0.0003**	–0.0002**	–0.0004***	–0.0004***	–0.0005***	–0.0004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_empl_hhh	–0.0065***	–0.0067**	–0.0062***	–0.0062***	–0.0068***	–0.0065***	–0.0066**	–0.0048	–0.0069**	–0.0067**
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
d_kids	–0.0975***	–0.1165***	–0.0974***	–0.0975***	–0.0979***	–0.0975***	–0.1166***	–0.1173***	–0.1173***	–0.1165***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
d_urban	–0.0011	0.0053	–0.0012	–0.0012	–0.0009	–0.0011	0.0053	0.0047	0.0060*	0.0053
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Constant	–0.2808***	–0.5206***	–0.2511***	–0.2912***	–0.1969***	–0.2808***	–0.5282***	–0.5155***	–0.1140	–0.5205***
	(0.045)	(0.063)	(0.047)	(0.045)	(0.051)	(0.045)	(0.071)	(0.063)	(0.076)	(0.063)
Observations	190,571	99,748	190,571	190,571	190,571	190,571	99,748	99,748	99,748	99,748
R-squared	0.167	0.174	0.167	0.167	0.167	0.167	0.174	0.175	0.175	0.174

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix A5.D3 Modified Model Frequency of Visit for Outpatient Care: Carbohydrate (Discrete)

Variable	Frequency Visit Outpatient Care									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
log_curecap	0.0531***	0.0549***	0.0531***	0.0531***	0.0531***	0.0531***	0.0549***	0.0548***	0.0548***	0.0549***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_prevcap	–0.0032***	–0.0040***	–0.0032***	–0.0032***	–0.0032***	–0.0032***	–0.0041***	–0.0041***	–0.0040***	–0.0041***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_medcheckcap	0.0042***	0.0033*	0.0041***	0.0042***	0.0041***	0.0041***	0.0033*	0.0034*	0.0037*	0.0033*
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
log_premium	–0.0022***	–0.0014**	–0.0022***	–0.0020***	–0.0024***	–0.0022***	–0.0014**	–0.0012*	–0.0008	–0.0014**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
pct_married	0.0023***	0.0044***	0.0023***	0.0023***	0.0023***	0.0024***	0.0044***	0.0043***	0.0039***	0.0044***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
*married * income*	–0.0002***	–0.0003***	–0.0002***	–0.0002***	–0.0002***	–0.0002***	–0.0003***	–0.0003***	–0.0003***	–0.0003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_carbo	–0.0005	0.0006	–0.0004	–0.0005	–0.0003	–0.0031	0.0006	0.0009	0.0019	0.0016
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)
d_insurance	0.0321***		–0.0248	0.0395***	–0.1101***	0.0292***				
	(0.002)		(0.028)	(0.004)	(0.041)	(0.003)				
d_subsidy		0.0386***					0.0492	0.0008	–0.5861***	0.0394***
		(0.004)					(0.044)	(0.008)	(0.065)	(0.005)
*d_insurance * age*			0.0149**							
			(0.007)							
*d_insurance * educ*				–0.0010**						
				(0.000)						
*d_insurance * income*					0.0109***					
					(0.003)					
*d_insurance * d_carbo*						0.0051				
						(0.004)				
*d_subsidy * age*							–0.0028			
							(0.011)			
*d_subsidy * educ*								0.0042***		
								(0.001)		
*d_subsidy * income*									0.0474***	
									(0.005)	
*d_subsidy * d_carbo*										–0.0016
										(0.006)
log income	0.0132***	0.0320***	0.0130***	0.0134***	0.0072**	0.0133***	0.0321***	0.0330***	0.0024	0.0320***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)
educ_hhh	–0.0032***	–0.0024***	–0.0033***	–0.0027***	–0.0033***	–0.0032***	–0.0024***	–0.0049***	–0.0021***	–0.0024***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
log_age_hhh	0.0475***	0.0500***	0.0402***	0.0482***	0.0466***	0.0475***	0.0518***	0.0538***	0.0500***	0.0500***
	(0.004)	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)	(0.009)	(0.006)	(0.006)	(0.006)
housequality	–0.0002**	–0.0004***	–0.0002**	–0.0002**	–0.0003**	–0.0002**	–0.0004***	–0.0004***	–0.0005***	–0.0004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_empl_hhh	–0.0065***	–0.0067**	–0.0062***	–0.0062***	–0.0068***	–0.0064***	–0.0066**	–0.0048	–0.0068**	–0.0067**
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
d_kids	–0.0976***	–0.1164***	–0.0975***	–0.0975***	–0.0980***	–0.0976***	–0.1165***	–0.1172***	–0.1172***	–0.1164***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
d_urban	–0.0011	0.0053	–0.0012	–0.0013	–0.0010	–0.0010	0.0053	0.0048	0.0061*	0.0053
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Constant	–0.2800***	–0.5217***	–0.2503***	–0.2904***	–0.1964***	–0.2799***	–0.5294***	–0.5173***	–0.1165	–0.5217***
	(0.045)	(0.064)	(0.047)	(0.045)	(0.051)	(0.045)	(0.071)	(0.064)	(0.076)	(0.064)
Observations	190,571	99,748	190,571	190,571	190,571	190,571	99,748	99,748	99,748	99,748
R-squared	0.167	0.174	0.167	0.167	0.167	0.167	0.174	0.175	0.175	0.174

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix A5.D4 Modified Model Frequency of Visit for Outpatient Care: Protein (Discrete)

Variable	Frequency Visit Outpatient Care									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
log_curecap	0.0531***	0.0549***	0.0531***	0.0531***	0.0531***	0.0531***	0.0549***	0.0548***	0.0549***	0.0549***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_prevcap	–0.0032***	–0.0040***	–0.0032***	–0.0032***	–0.0032***	–0.0032***	–0.0040***	–0.0041***	–0.0040***	–0.0040***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_medcheckcap	0.0041***	0.0033*	0.0041***	0.0042***	0.0041***	0.0041***	0.0033*	0.0034*	0.0037*	0.0033*
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
log_premium	–0.0022***	–0.0014**	–0.0022***	–0.0020***	–0.0024***	–0.0022***	–0.0014**	–0.0012*	–0.0008	–0.0014**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
pct_married	0.0023***	0.0044***	0.0023***	0.0023***	0.0023***	0.0023***	0.0044***	0.0043***	0.0038***	0.0044***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
*married * income*	–0.0002***	–0.0003***	–0.0002***	–0.0002***	–0.0002***	–0.0002***	–0.0003***	–0.0003***	–0.0003***	–0.0003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_protein	–0.0016	0.0029	–0.0015	–0.0015	–0.0015	–0.0060	0.0029	0.0032	0.0039	0.0054
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.007)
d_insurance	0.0321***		–0.0247	0.0395***	–0.1101***	0.0246***				
	(0.002)		(0.028)	(0.004)	(0.041)	(0.005)				
d_subsidy		0.0386***					0.0492	0.0007	–0.5856***	0.0418***
		(0.004)					(0.044)	(0.008)	(0.065)	(0.009)
*d_insurance * age*			0.0148**							
			(0.007)							
*d_insurance * educ*				–0.0010**						
				(0.000)						
*d_insurance * income*					0.0109***					
					(0.003)					
*d_insurance * d_protein*						0.0088				
						(0.006)				
*d_subsidy * age*							–0.0028			
							(0.011)			
*d_subsidy * educ*								0.0042***		
								-0.001		
*d_subsidy * income*									0.0473***	
									(0.005)	
*d_subsidy * d_protein*										–0.0038
										(0.009)
log income	0.0132***	0.0319***	0.0130***	0.0134***	0.0072**	0.0132***	0.0320***	0.0330***	0.0023	0.0319***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)
educ_hhh	–0.0032***	–0.0024***	–0.0033***	–0.0027***	–0.0033***	–0.0033***	–0.0024***	–0.0049***	–0.0021***	–0.0024***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
log_age_hhh	0.0475***	0.0499***	0.0403***	0.0482***	0.0466***	0.0475***	0.0517***	0.0537***	0.0499***	0.0499***
	(0.004)	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)	(0.009)	(0.006)	(0.006)	(0.006)
housequality	–0.0002**	–0.0004***	–0.0002**	–0.0002**	–0.0003**	–0.0002**	–0.0004***	–0.0004***	–0.0005***	–0.0004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_empl_hhh	–0.0065***	–0.0067**	–0.0062***	–0.0062***	–0.0068***	–0.0065***	–0.0066**	–0.0048	–0.0068**	–0.0067**
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
d_kids	–0.0975***	–0.1165***	–0.0974***	–0.0975***	–0.0979***	–0.0976***	–0.1167***	–0.1174***	–0.1174***	–0.1165***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
d_urban	–0.0011	0.0053	–0.0012	–0.0013	–0.0010	–0.0011	0.0053	0.0048	0.0061*	0.0053
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Constant	–0.2795***	–0.5231***	–0.2498***	–0.2899***	–0.1957***	–0.2762***	–0.5308***	–0.5183***	–0.1165	–0.5239***
	(0.045)	(0.063)	(0.047)	(0.045)	(0.051)	(0.045)	(0.071)	(0.063)	(0.076)	(0.063)
Observations	190,571	99,748	190,571	190,571	190,571	190,571	99,748	99,748	99,748	99,748
R-squared	0.167	0.174	0.167	0.167	0.167	0.167	0.174	0.175	0.175	0.174

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix A5.D5 Modified Model Frequency of Visit for Outpatient Care: Fat (Discrete)

Variable	Frequency Visit Outpatient Care									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
log_curecap	0.0530***	0.0549***	0.0531***	0.0530***	0.0530***	0.0531***	0.0549***	0.0548***	0.0549***	0.0549***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_prevcap	−0.0032***	−0.0041***	−0.0032***	−0.0032***	−0.0032***	−0.0032***	−0.0041***	−0.0041***	−0.0040***	−0.0041***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_medcheckcap	0.0041***	0.0033*	0.0041***	0.0042***	0.0041***	0.0041***	0.0033*	0.0034*	0.0037*	0.0033*
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
log_premium	−0.0022***	−0.0014**	−0.0022***	−0.0021***	−0.0024***	−0.0022***	−0.0014**	−0.0012*	−0.0008	−0.0014**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
pct_married	0.0022***	0.0044***	0.0022***	0.0022***	0.0022***	0.0022***	0.0044***	0.0043***	0.0038***	0.0045***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
*married * income*	−0.0002***	−0.0003***	−0.0002***	−0.0002***	−0.0002***	−0.0002***	−0.0003***	−0.0003***	−0.0003***	−0.0003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_fat	−0.0114***	−0.0016	−0.0114***	−0.0114***	−0.0112***	−0.0101**	−0.0016	−0.0018	−0.0022	−0.0313***
	(0.004)	(0.005)	(0.004)	(0.004)	(0.004)	(0.005)	(0.005)	(0.005)	(0.005)	(0.010)
d_insurance	0.0321***		−0.0252	0.0396***	−0.1083***	0.0344***				
	(0.002)		(0.028)	(0.004)	(0.041)	(0.007)				
d_subsidy		0.0385***					0.049	0.0008	−0.5846***	0.0023
		(0.004)					(0.044)	(0.008)	(0.065)	(0.011)
*d_insurance * age*			0.0150**							
			(0.007)							
*d_insurance * educ*				−0.0010**						
				(0.000)						
*d_insurance * income*					0.0108***					
					(0.003)					
*d_insurance * d_fat*						−0.0025				
						-0.007				
*d_subsidy * age*							−0.0027			
							(0.011)			
*d_subsidy * educ*								0.0042***		
								(0.001)		
*d_subsidy * income*									0.0472***	
									(0.005)	
*d_subsidy * d_fat*										0.0393***
										(0.012)
log income	0.0132***	0.0319***	0.0130***	0.0134***	0.0073**	0.0132***	0.0320***	0.0329***	0.0023	0.0322***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)
educ_hhh	−0.0032***	−0.0023***	−0.0032***	−0.0026***	−0.0033***	−0.0032***	−0.0024***	−0.0049***	−0.0021***	−0.0023***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
log_age_hhh	0.0472***	0.0500***	0.0399***	0.0480***	0.0463***	0.0472***	0.0518***	0.0538***	0.0500***	0.0505***
	(0.004)	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)	(0.009)	(0.006)	(0.006)	(0.006)
housequality	−0.0002**	−0.0004***	−0.0002**	−0.0002**	−0.0003**	−0.0002**	−0.0004***	−0.0004***	−0.0005***	−0.0004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_empl_hhh	−0.0065***	−0.0067**	−0.0062***	−0.0062***	−0.0068***	−0.0065***	−0.0066**	−0.0048	−0.0069**	−0.0063*
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
d_kids	−0.0950***	−0.1161***	−0.0949***	−0.0950***	−0.0955***	−0.0950***	−0.1163***	−0.1169***	−0.1169***	−0.1159***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
d_urban	−0.0011	0.0053	−0.0012	−0.0013	−0.0010	−0.0011	0.0053	0.0047	0.0060*	0.0052
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Constant	−0.2720***	−0.5191***	−0.2420***	−0.2824***	−0.1894***	−0.2730***	−0.5266***	−0.5138***	−0.1117	−0.4967***
	(0.045)	(0.063)	(0.047)	(0.045)	(0.051)	(0.045)	(0.071)	(0.063)	(0.076)	(0.064)
Observations	190,571	99,748	190,571	190,571	190,571	190,571	99,748	99,748	99,748	99,748
R-squared	0.167	0.174	0.167	0.167	0.167	0.167	0.174	0.175	0.175	0.174

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix A5.D6 Modified Model Frequency of Visit for Outpatient Care: Calorie (Discrete)

Variable	Frequency Visit Outpatient Care									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
log_curecap	0.0531***	0.0549***	0.0531***	0.0531***	0.0531***	0.0531***	0.0549***	0.0548***	0.0548***	0.0549***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_prevcap	–0.0032***	–0.0041***	–0.0032***	–0.0032***	–0.0032***	–0.0032***	–0.0041***	–0.0041***	–0.0040***	–0.0041***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_medcheckcap	0.0042***	0.0033*	0.0041***	0.0042***	0.0041***	0.0042***	0.0033*	0.0034*	0.0037*	0.0033*
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
log_premium	–0.0022***	–0.0014**	–0.0022***	–0.0020***	–0.0024***	–0.0022***	–0.0014**	–0.0012*	–0.0008	–0.0014**
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
pct_married	0.0023***	0.0044***	0.0023***	0.0023***	0.0023***	0.0023***	0.0044***	0.0043***	0.0038***	0.0044***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
*married * income*	–0.0002***	–0.0003***	–0.0002***	–0.0002***	–0.0002***	–0.0002***	–0.0003***	–0.0003***	–0.0003***	–0.0003***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_calorie	–0.0014	–0.0005	–0.0013	–0.0013	–0.0012	0.0028	–0.0005	–0.0005	–0.0003	0.0002
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)	(0.006)
d_insurance	0.0321***		–0.0247	0.0395***	–0.1099***	0.0383***				
	(0.002)		(0.028)	(0.004)	(0.041)	(0.004)				
d_subsidy		0.0386***					0.0492	0.0008	–0.5841***	0.0394***
		(0.004)					(0.044)	(0.008)	(0.065)	(0.007)
*d_insurance * age*			0.0148**							
			(0.007)							
*d_insurance * educ*				–0.0010**						
				(0.000)						
*d_insurance * income*					0.0109***					
					(0.003)					
*d_insurance * d_calorie*						–0.0084*				
						(0.005)				
*d_subsidy * age*							–0.0028			
							(0.011)			
*d_subsidy * educ*								0.0042***		
								(0.001)		
*d_subsidy * income*									0.0472***	
									(0.005)	
*d_subsidy * d_calorie*										–0.0012
										(0.007)
log income	0.0131***	0.0319***	0.0130***	0.0133***	0.0072**	0.0131***	0.0320***	0.0329***	0.0023	0.0319***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.004)	(0.003)	(0.004)	(0.004)	(0.005)	(0.004)
educ_hhh	–0.0032***	–0.0024***	–0.0033***	–0.0027***	–0.0033***	–0.0033***	–0.0024***	–0.0049***	–0.0021***	–0.0024***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.001)	(0.000)	(0.000)
log_age_hhh	0.0475***	0.0501***	0.0403***	0.0483***	0.0466***	0.0474***	0.0519***	0.0539***	0.0501***	0.0500***
	(0.004)	(0.006)	(0.005)	(0.004)	(0.004)	(0.004)	(0.009)	(0.006)	(0.006)	(0.006)
housequality	–0.0003**	–0.0004***	–0.0003**	–0.0002**	–0.0003**	–0.0003**	–0.0004***	–0.0004***	–0.0005***	–0.0004***
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_empl_hhh	–0.0065***	–0.0067**	–0.0062***	–0.0062***	–0.0068***	–0.0065***	–0.0066**	–0.0048	–0.0069**	–0.0067**
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
d_kids	–0.0972***	–0.1163***	–0.0971***	–0.0971***	–0.0976***	–0.0972***	–0.1165***	–0.1172***	–0.1173***	–0.1163***
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
d_urban	–0.0012	0.0052	–0.0013	–0.0013	–0.0010	–0.0012	0.0052	0.0047	0.0060*	0.0052
	(0.002)	(0.003)	(0.002)	(0.002)	(0.002)	(0.002)	(0.003)	(0.003)	(0.003)	(0.003)
Constant	–0.2792***	–0.5199***	–0.2496***	–0.2896***	–0.1957***	–0.2808***	–0.5276***	–0.5149***	–0.1137	–0.5197***
	(0.045)	(0.063)	(0.047)	(0.045)	(0.051)	(0.045)	(0.071)	(0.063)	(0.076)	(0.063)
Observations	190,571	99,748	190,571	190,571	190,571	190,571	99,748	99,748	99,748	99,748
R-squared	0.167	0.174	0.167	0.167	0.167	0.167	0.174	0.175	0.175	0.174

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix A5.E1 Modified Model Days Spent Receiving Inpatient Care: Cigarette

Variable	Days Spent (Inpatient Care)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
illness_cap	−0.0067**	−0.0047	−0.0066**	−0.0067**	−0.0066**	−0.0067**	−0.0044	−0.0046	−0.0047	−0.0047
	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)
log_curecap	0.0355***	0.0423***	0.0355***	0.0355***	0.0354***	0.0355***	0.0422***	0.0421***	0.0423***	0.0423***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_prevcap	−0.0051***	−0.0057***	−0.0051***	−0.0051***	−0.0051***	−0.0051***	−0.0058***	−0.0059***	−0.0057***	−0.0057***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_medcheckcap	0.0034	0.0034	0.0034	0.0035	0.0034	0.0034	0.0032	0.0037	0.0033	0.0034
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
log_premium	0.0032***	0.0018	0.0032***	0.0034***	0.0025**	0.0032***	0.0019	0.0025*	0.0017	0.0018
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
pct_married	−0.0006	0.0021	−0.0008	−0.0006	−0.0006	−0.0006	0.0026	0.0018	0.0022	0.002
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
*married * income*	0.0001	−0.0001	0.0001	0.0001	0.0001	0.0001	−0.0001	−0.0001	−0.0001	−0.0001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_cigarcap	−0.0420***	−0.0463***	−0.0422***	−0.0421***	−0.0418***	−0.0393***	−0.0462***	−0.0489***	−0.0459***	−0.0498***
	(0.004)	(0.006)	(0.004)	(0.004)	(0.004)	(0.005)	(0.006)	(0.006)	(0.006)	(0.009)
d_insurance	0.0655***		−0.2204***	0.0770***	−0.3764***	0.0672***				
	(0.004)		(0.056)	(0.008)	(0.082)	(0.005)				
d_subsidy		−0.0111					0.3191***	−0.1202***	0.0994	−0.0127
		(0.009)					(0.096)	(0.017)	(0.144)	(0.010)
*d_insurance * age*			0.0747***							
			(0.015)							
*d_insurance * educ*				−0.0016*						
				(0.001)						
*d_insurance * income*					0.0339***					
					(0.006)					
*d_insurance * cigarette*						−0.0054				
						(0.007)				
*d_subsidy * age*							−0.0860***			
							(0.025)			
*d_subsidy * educ*								0.0122***		
								(0.002)		
*d_subsidy * income*									−0.0084	
									(0.011)	
*d_subsidy * cigarette*										0.0052
										(0.011)
log income	0.0906***	0.1085***	0.0898***	0.0909***	0.0719***	0.0906***	0.1109***	0.1123***	0.1137***	0.1082***
	(0.006)	(0.010)	(0.006)	(0.006)	(0.007)	(0.006)	(0.010)	(0.010)	(0.012)	(0.010)
educ_hhh	−0.0016***	−0.0029***	−0.0017***	−0.0007	−0.0018***	−0.0016***	−0.0032***	−0.0104***	−0.0030***	−0.0030***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_age_hhh	0.0597***	0.0685***	0.0232**	0.0609***	0.0571***	0.0598***	0.1241***	0.0789***	0.0686***	0.0683***
	(0.008)	(0.013)	(0.011)	(0.008)	(0.008)	(0.008)	(0.021)	(0.013)	(0.013)	(0.013)
housequality	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0003	0.0001	0.0003	0.0003
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_empl_hhh	−0.0321***	−0.0464***	−0.0307***	−0.0316***	−0.0332***	−0.0321***	−0.0430***	−0.0408***	−0.0464***	−0.0463***
	(0.005)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)
d_kids	−0.0489***	−0.0599***	−0.0483***	−0.0488***	−0.0502***	−0.0489***	−0.0634***	−0.0623***	−0.0597***	−0.0600***
	(0.006)	(0.009)	(0.006)	(0.006)	(0.006)	(0.006)	(0.009)	(0.009)	(0.009)	(0.009)
d_urban	0.0176***	0.0317***	0.0168***	0.0173***	0.0179***	0.0176***	0.0316***	0.0300***	0.0316***	0.0317***
	(0.005)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)
Constant	−1.4336***	−1.6534***	−1.2845***	−1.4498***	−1.1726***	−1.4357***	−1.8924***	−1.6460***	−1.7244***	−1.6466***
	(0.089)	(0.140)	(0.094)	(0.090)	(0.102)	(0.089)	(0.156)	(0.140)	(0.168)	(0.140)
Observations	190,571	99,748	190,571	190,571	190,571	190,571	99,748	99,748	99,748	99,748
R–squared	0.032	0.035	0.032	0.032	0.032	0.032	0.035	0.036	0.035	0.035

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix A5.E2 Modified Model Days Spent Receiving Inpatient Care: Alcohol

Variable	Days Spent (Inpatient Care)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
illness_cap	–0.0059*	–0.0037	–0.0058*	–0.0059*	–0.0058*	–0.0059*	–0.0034	–0.0036	–0.0036	–0.0037
	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)
log_curecap	0.0358***	0.0427***	0.0358***	0.0358***	0.0358***	0.0358***	0.0426***	0.0425***	0.0427***	0.0427***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_prevcap	–0.0052***	–0.0058***	–0.0051***	–0.0052***	–0.0052***	–0.0052***	–0.0059***	–0.0060***	–0.0058***	–0.0058***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_medcheckcap	0.004	0.004	0.004	0.004	0.0039	0.004	0.0038	0.0043	0.0039	0.004
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
log_premium	0.0038***	0.0023	0.0038***	0.0040***	0.0031***	0.0038***	0.0024*	0.0029**	0.002	0.0023
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
pct_married	–0.0007	0.0017	–0.0009	–0.0008	–0.0008	–0.0007	0.0022	0.0014	0.0019	0.0017
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
married ∗ income	0.0001	–0.0001	0.0001	0.0001	0.0001	0.0001	–0.0001	–0.0001	–0.0001	–0.0001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
log_alcoholcap	–0.0178	–0.0167	–0.0180	–0.0177	–0.0180	–0.0215	–0.0166	–0.0165	–0.0167	–0.0297
	(0.012)	(0.018)	(0.012)	(0.012)	(0.012)	(0.019)	(0.018)	(0.018)	(0.018)	(0.038)
d_insurance	0.0679***		–0.2141***	0.0789***	–0.3883***	0.0680***				
	(0.004)		(0.056)	(0.008)	(0.082)	(0.004)				
d_subsidy		–0.0155*					0.3161***	–0.1177***	0.2108	–0.0153*
		(0.009)					(0.096)	(0.017)	(0.143)	(0.009)
d_insurance ∗ age			0.0737***							
			(0.015)							
d_insurance ∗ educ				–0.0015*						
				(0.001)						
d_insurance ∗ income					0.0350***					
					(0.006)					
d_insurance ∗ alcohol						0.0063				
						(0.025)				
d_subsidy ∗ age							–0.0864***			
							(0.025)			
d_subsidy ∗ educ								0.0114***		
								(0.002)		
d_subsidy ∗ income									–0.0172	
									(0.011)	
d_subsidy ∗ alcohol										0.0167
										(0.043)
log income	0.0784***	0.0943***	0.0776***	0.0787***	0.0591***	0.0784***	0.0966***	0.0970***	0.1050***	0.0943***
	(0.006)	(0.010)	(0.006)	(0.006)	(0.007)	(0.006)	(0.010)	(0.010)	(0.012)	(0.010)
educ_hhh	–0.0009*	–0.0023***	–0.0010*	–0.0001	–0.0011**	–0.0009*	–0.0025***	–0.0092***	–0.0024***	–0.0023***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_age_hhh	0.0704***	0.0795***	0.0344***	0.0714***	0.0675***	0.0704***	0.1354***	0.0898***	0.0795***	0.0795***
	(0.008)	(0.013)	(0.011)	(0.008)	(0.008)	(0.008)	(0.021)	(0.013)	(0.013)	(0.013)
housequality	0.0004**	0.0004	0.0004**	0.0005**	0.0004**	0.0004**	0.0003	0.0002	0.0004	0.0004
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_empl_hhh	–0.0333***	–0.0475***	–0.0319***	–0.0329***	–0.0344***	–0.0333***	–0.0441***	–0.0424***	–0.0475***	–0.0476***
	(0.005)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)
d_kids	–0.0491***	–0.0602***	–0.0485***	–0.0490***	–0.0504***	–0.0491***	–0.0637***	–0.0624***	–0.0599***	–0.0602***
	(0.006)	(0.009)	(0.006)	(0.006)	(0.006)	(0.006)	(0.009)	(0.009)	(0.009)	(0.009)
d_urban	0.0208***	0.0347***	0.0201***	0.0205***	0.0211***	0.0208***	0.0346***	0.0333***	0.0344***	0.0347***
	(0.005)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)
Constant	–1.3408***	–1.5257***	–1.1935***	–1.3560***	–1.0719***	–1.3408***	–1.7659***	–1.5120***	–1.6735***	–1.5254***
	(0.089)	(0.139)	(0.094)	(0.090)	(0.101)	(0.089)	(0.155)	(0.139)	(0.167)	(0.139)
Observations	190,571	99,748	190,571	190,571	190,571	190,571	99,748	99,748	99,748	99,748
R–squared	0.031	0.035	0.031	0.031	0.031	0.031	0.035	0.035	0.035	0.035

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix A5.E3 Modified Model Days Spent Receiving Inpatient Care: Carbohydrate (Discrete)

Variable	Days Spent (Inpatient Care)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
illness_cap	−0.0058*	−0.0036	−0.0058*	−0.0059*	−0.0058*	−0.0058*	−0.0033	−0.0035	−0.0036	−0.0037
	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)
log_curecap	0.0358***	0.0427***	0.0358***	0.0358***	0.0358***	0.0358***	0.0426***	0.0425***	0.0427***	0.0427***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_prevcap	−0.0051***	−0.0057***	−0.0051***	−0.0051***	−0.0052***	−0.0051***	−0.0058***	−0.0059***	−0.0058***	−0.0058***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_medcheckcap	0.0039	0.0039	0.0039	0.004	0.0038	0.0039	0.0037	0.0042	0.0038	0.0039
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
log_premium	0.0038***	0.0022	0.0038***	0.0039***	0.0030***	0.0038***	0.0023	0.0028**	0.002	0.0022
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
pct_married	−0.0006	0.002	−0.0008	−0.0007	−0.0007	−0.0006	0.0025	0.0018	0.0022	0.002
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
*married * income*	0.0001	−0.0001	0.0001	0.0001	0.0001	0.0001	−0.0001	−0.0001	−0.0001	−0.0001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_carbo	0.0106**	0.0206***	0.0107**	0.0106**	0.0110**	0.0011	0.0207***	0.0217***	0.0202***	0.0292***
	(0.004)	(0.007)	(0.004)	(0.004)	(0.004)	(0.006)	(0.007)	(0.007)	(0.007)	(0.011)
d_insurance	0.0678***		−0.2141***	0.0788***	−0.3910***	0.0576***				
	(0.004)		(0.056)	(0.008)	(0.082)	(0.006)				
d_subsidy		−0.0154*					0.3175***	−0.1185***	0.1906	−0.0084
		(0.009)					(0.096)	(0.017)	(0.144)	(0.012)
*d_insurance * age*			0.0737***							
			(0.015)							
*d_insurance * educ*				−0.0015*						
				(0.001)						
*d_insurance * income*					0.0352***					
					(0.006)					
*d_insurance * d_carbo*						0.0184**				
						(0.008)				
*d_subsidy * age*							−0.0867***			
							(0.025)			
*d_subsidy * educ*								0.0115***		
								(0.002)		
*d_subsidy * income*									−0.0156	
									(0.011)	
*d_subsidy * d_carbo*										−0.0134
										(0.014)
log income	0.0790***	0.0960***	0.0782***	0.0793***	0.0597***	0.0793***	0.0984***	0.0989***	0.1057***	0.0957***
	(0.006)	(0.010)	(0.006)	(0.006)	(0.007)	(0.006)	(0.010)	(0.010)	(0.012)	(0.010)
educ_hhh	−0.0009*	−0.0023***	−0.0010*	−0.0001	−0.0011**	−0.0009*	−0.0025***	−0.0093***	−0.0024***	−0.0023***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_age_hhh	0.0700***	0.0786***	0.0340***	0.0710***	0.0671***	0.0700***	0.1347***	0.0889***	0.0786***	0.0788***
	(0.008)	(0.013)	(0.011)	(0.008)	(0.008)	(0.008)	(0.021)	(0.013)	(0.013)	(0.013)
housequality	0.0005**	0.0004	0.0005**	0.0005**	0.0004**	0.0005**	0.0004	0.0003	0.0004	0.0004
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_empl_hhh	−0.0329***	−0.0464***	−0.0315***	−0.0324***	−0.0340***	−0.0327***	−0.0430***	−0.0412***	−0.0464***	−0.0465***
	(0.005)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)
d_kids	−0.0477***	−0.0579***	−0.0470***	−0.0476***	−0.0489***	−0.0478***	−0.0614***	−0.0600***	−0.0577***	−0.0575***
	(0.006)	(0.009)	(0.006)	(0.006)	(0.006)	(0.006)	(0.009)	(0.009)	(0.009)	(0.009)
d_urban	0.0209***	0.0354***	0.0202***	0.0207***	0.0213***	0.0210***	0.0353***	0.0339***	0.0351***	0.0352***
	(0.005)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)
Constant	−1.3571***	−1.5643***	−1.2098***	−1.3723***	−1.0872***	−1.3567***	−1.8056***	−1.5524***	−1.6979***	−1.5646***
	(0.089)	(0.140)	(0.094)	(0.090)	(0.102)	(0.089)	(0.156)	(0.139)	(0.168)	(0.140)
Observations	190,571	99,748	190,571	190,571	190,571	190,571	99,748	99,748	99,748	99,748
R-squared	0.031	0.035	0.031	0.031	0.031	0.031	0.035	0.035	0.035	0.035

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix A5.E4 Modified Model Days Spent Receiving Inpatient Care: Protein (Discrete)

Variable	Days Spent (Inpatient Care)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
illness_cap	–0.0058*	–0.0036	–0.0057*	–0.0058*	–0.0057*	–0.0058*	–0.0033	–0.0035	–0.0035	–0.0036
	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)
log_curecap	0.0358***	0.0427***	0.0358***	0.0358***	0.0358***	0.0358***	0.0426***	0.0425***	0.0427***	0.0427***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_prevcap	–0.0051***	–0.0058***	–0.0051***	–0.0051***	–0.0052***	–0.0051***	–0.0059***	–0.0060***	–0.0058***	–0.0058***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_medcheckcap	0.004	0.004	0.004	0.004	0.0039	0.004	0.0038	0.0042	0.0038	0.004
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
log_premium	0.0038***	0.0023	0.0038***	0.0040***	0.0031***	0.0038***	0.0024*	0.0029**	0.002	0.0023
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
pct_married	–0.0007	0.0017	–0.0009	–0.0007	–0.0007	–0.0007	0.0022	0.0015	0.0019	0.0017
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
*married * income*	0.0001	–0.0001	0.0001	0.0001	0.0001	0.0001	–0.0001	–0.0001	–0.0001	–0.0001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_protein	0.0159***	0.0203**	0.0161***	0.0160***	0.0162***	0.0133	0.0203**	0.0211**	0.0200**	0.0294*
	(0.006)	(0.009)	(0.006)	(0.006)	(0.006)	(0.008)	(0.009)	(0.009)	(0.009)	(0.015)
d_insurance	0.0678***		–0.2147***	0.0789***	–0.3894***	0.0633***				
	(0.004)		(0.056)	(0.008)	(0.082)	(0.011)				
d_subsidy		–0.0155*					0.3163***	–0.1181***	0.2036	–0.0034
		(0.009)					(0.096)	(0.017)	(0.143)	(0.019)
*d_insurance * age*			0.0739***							
			(0.015)							
*d_insurance * educ*				–0.0015*						
				(0.001)						
*d_insurance * income*					0.0351***					
					(0.006)					
*d_insurance * d_protein*						0.0053				
						(0.012)				
*d_subsidy * age*							–0.0865***			
							(0.025)			
*d_subsidy * educ*								0.0114***		
								(0.002)		
*d_subsidy * income*									–0.0166	
									(0.011)	
*d_subsidy * d_protein*										–0.0143
										(0.019)
log income	0.0783***	0.0943***	0.0775***	0.0786***	0.0590***	0.0783***	0.0966***	0.0970***	0.1047***	0.0939***
	(0.006)	(0.010)	(0.006)	(0.006)	(0.007)	(0.006)	(0.010)	(0.010)	(0.012)	(0.010)
educ_hhh	–0.0009*	–0.0023***	–0.0010*	–0.0000	–0.0011**	–0.0009*	–0.0025***	–0.0093***	–0.0024***	–0.0023***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_age_hhh	0.0698***	0.0787***	0.0337***	0.0709***	0.0669***	0.0698***	0.1347***	0.0890***	0.0787***	0.0787***
	(0.008)	(0.013)	(0.011)	(0.008)	(0.008)	(0.008)	(0.021)	(0.013)	(0.013)	(0.013)
housequality	0.0005**	0.0004	0.0005**	0.0005**	0.0004**	0.0005**	0.0003	0.0002	0.0004	0.0004
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_empl_hhh	–0.0332***	–0.0474***	–0.0318***	–0.0327***	–0.0343***	–0.0332***	–0.0439***	–0.0422***	–0.0473***	–0.0474***
	(0.005)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)
d_kids	–0.0490***	–0.0604***	–0.0484***	–0.0490***	–0.0503***	–0.0490***	–0.0639***	–0.0626***	–0.0600***	–0.0603***
	(0.006)	(0.009)	(0.006)	(0.006)	(0.006)	(0.006)	(0.009)	(0.009)	(0.009)	(0.009)
d_urban	0.0209***	0.0349***	0.0202***	0.0207***	0.0213***	0.0209***	0.0349***	0.0335***	0.0347***	0.0349***
	(0.005)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)
Constant	–1.3527***	–1.5418***	–1.2053***	–1.3682***	–1.0834***	–1.3507***	–1.7822***	–1.5287***	–1.6846***	–1.5447***
	(0.089)	(0.139)	(0.094)	(0.090)	(0.101)	(0.089)	(0.155)	(0.139)	(0.168)	(0.139)
Observations	190,571	99,748	190,571	190,571	190,571	190,571	99,748	99,748	99,748	99,748
R-squared	0.031	0.035	0.031	0.031	0.031	0.031	0.035	0.035	0.035	0.035

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix A5.E5 Modified Model Days Spent Receiving Inpatient Care: Fat (Discrete)

Variable	Days Spent (Inpatient Care)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
illness_cap	-0.0059^{*}	-0.0038	-0.0059^{*}	-0.0060^{*}	-0.0059^{*}	-0.0059^{*}	-0.0035	-0.0037	-0.0037	-0.0038
	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)
log_curecap	0.0358^{***}	0.0427^{***}	0.0358^{***}	0.0358^{***}	0.0358^{***}	0.0358^{***}	0.0426^{***}	0.0425^{***}	0.0427^{***}	0.0427^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_prevcap	-0.0052^{***}	-0.0058^{***}	-0.0052^{***}	-0.0052^{***}	-0.0052^{***}	-0.0052^{***}	-0.0059^{***}	-0.0060^{***}	-0.0058^{***}	-0.0058^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_medcheckcap	0.004	0.004	0.004	0.004	0.0039	0.004	0.0038	0.0043	0.0039	0.004
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
log_premium	0.0038^{***}	0.0023	0.0038^{***}	0.0040^{***}	0.0031^{***}	0.0038^{***}	0.0023^{*}	0.0029^{**}	0.002	0.0023
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
pct_married	-0.0008	0.0014	-0.0010	-0.0009	-0.0009	-0.0009	0.0019	0.0012	0.0016	0.0014
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
married ∗ income	0.0001	-0.0001	0.0001	0.0001	0.0001	0.0001	-0.0001	-0.0000	-0.0001	-0.0001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_fat	-0.0087	-0.0213^{*}	-0.0088	-0.0087	-0.0080	0.01	-0.0208^{*}	-0.0218^{*}	-0.0211^{*}	-0.0241
	(0.007)	(0.011)	(0.007)	(0.007)	(0.007)	(0.010)	(0.011)	(0.011)	(0.011)	(0.022)
d_insurance	0.0680^{***}		-0.2140^{***}	0.0791^{***}	-0.3865^{***}	0.1016^{***}				
	(0.004)		(0.056)	(0.008)	(0.082)	(0.013)				
d_subsidy		-0.0158^{*}					0.3137^{***}	-0.1182^{***}	0.2071	-0.0193
		(0.009)					-0.096	-0.017	-0.143	-0.025
d_insurance ∗ age			0.0737^{***}							
			(0.015)							
d_insurance ∗ educ				-0.0015^{*}						
				(0.001)						
d_insurance ∗ income					0.0349^{***}					
					(0.006)					
d_insurance ∗ d_fat						-0.0375^{***}				
						(0.014)				
d_subsidy ∗ age							-0.0859^{***}			
							(0.025)			
d_subsidy ∗ educ								0.0114^{***}		
								(0.002)		
d_subsidy ∗ income									-0.0169	
									(0.011)	
d_subsidy ∗ d_fat										0.0038
										(0.025)
log income	0.0783^{***}	0.0939^{***}	0.0775^{***}	0.0786^{***}	0.0592^{***}	0.0782^{***}	0.0963^{***}	0.0967^{***}	0.1045^{***}	0.0939^{***}
	(0.006)	(0.010)	(0.006)	(0.006)	(0.007)	(0.006)	(0.010)	(0.010)	(0.012)	(0.010)
educ_hhh	-0.0009^{*}	-0.0022^{***}	-0.0009^{*}	-0.0000	-0.0011^{**}	-0.0009^{*}	-0.0024^{***}	-0.0092^{***}	-0.0023^{***}	-0.0022^{***}
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_age_hhh	0.0702^{***}	0.0791^{***}	0.0342^{***}	0.0713^{***}	0.0674^{***}	0.0703^{***}	0.1347^{***}	0.0894^{***}	0.0791^{***}	0.0791^{***}
	(0.008)	(0.013)	(0.011)	(0.008)	(0.008)	(0.008)	(0.021)	(0.013)	(0.013)	(0.013)
housequality	0.0004^{**}	0.0003	0.0004^{**}	0.0004^{**}	0.0004^{*}	0.0004^{**}	0.0003	0.0002	0.0004	0.0003
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_empl_hhh	-0.0334^{***}	-0.0476^{***}	-0.0320^{***}	-0.0329^{***}	-0.0344^{***}	-0.0332^{***}	-0.0442^{***}	-0.0425^{***}	-0.0476^{***}	-0.0476^{***}
	(0.005)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)
d_kids	-0.0470^{***}	-0.0553^{***}	-0.0463^{***}	-0.0469^{***}	-0.0484^{***}	-0.0471^{***}	-0.0589^{***}	-0.0574^{***}	-0.0550^{***}	-0.0553^{***}
	(0.006)	(0.009)	(0.006)	(0.006)	(0.006)	(0.006)	(0.009)	(0.009)	(0.009)	(0.009)
d_urban	0.0207^{***}	0.0346^{***}	0.0199^{***}	0.0204^{***}	0.0210^{***}	0.0207^{***}	0.0345^{***}	0.0331^{***}	0.0343^{***}	0.0346^{***}
	(0.005)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)
Constant	-1.3325^{***}	-1.5031^{***}	-1.1850^{***}	-1.3478^{***}	-1.0651^{***}	-1.3479^{***}	-1.7423^{***}	-1.4889^{***}	-1.6490^{***}	-1.5010^{***}
	(0.089)	(0.139)	(0.094)	(0.090)	(0.101)	(0.089)	(0.156)	(0.139)	(0.168)	(0.140)
Observations	190,571	99,748	190,571	190,571	190,571	190,571	99,748	99,748	99,748	99,748
R-squared	0.031	0.035	0.031	0.031	0.031	0.031	0.035	0.035	0.035	0.035

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Appendix A5.E6 Modified Model Days Spent Receiving Inpatient Care: Calorie (Discrete)

Variable	Days Spent (Inpatient Care)									
	(1)	(2)	(3)	(4)	(5)	(6)	(7)	(8)	(9)	(10)
illness_cap	–0.0057*	–0.0037	–0.0057*	–0.0058*	–0.0057*	–0.0057*	–0.0033	–0.0035	–0.0036	–0.0037
	(0.003)	(0.005)	(0.003)	(0.003)	(0.003)	(0.003)	(0.005)	(0.005)	(0.005)	(0.005)
log_curecap	0.0358***	0.0427***	0.0358***	0.0358***	0.0358***	0.0358***	0.0426***	0.0425***	0.0427***	0.0427***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_prevcap	–0.0051***	–0.0058***	–0.0051***	–0.0051***	–0.0052***	–0.0052***	–0.0059***	–0.0060***	–0.0058***	–0.0058***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_medcheckcap	0.004	0.004	0.004	0.004	0.0039	0.004	0.0038	0.0043	0.0038	0.004
	(0.003)	(0.004)	(0.003)	(0.003)	(0.003)	(0.003)	(0.004)	(0.004)	(0.004)	(0.004)
log_premium	0.0038***	0.0023	0.0039***	0.0040***	0.0031***	0.0038***	0.0024*	0.0029**	0.002	0.0023
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
pct_married	–0.0007	0.0017	–0.0009	–0.0007	–0.0007	–0.0007	0.0022	0.0015	0.0019	0.0017
	(0.001)	(0.002)	(0.001)	(0.001)	(0.001)	(0.001)	(0.002)	(0.002)	(0.002)	(0.002)
*married * income*	0.0001	–0.0001	0.0001	0.0001	0.0001	0.0001	–0.0001	–0.0001	–0.0001	–0.0001
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_calorie	0.0079	0.0049	0.008	0.0079	0.0083*	0.0199***	0.0048	0.005	0.0048	0.0077
	(0.005)	(0.008)	(0.005)	(0.005)	(0.005)	(0.007)	(0.008)	(0.008)	(0.008)	(0.012)
d_insurance	0.0679***		–0.2142***	0.0789***	–0.3901***	0.0857***				
	(0.004)		(0.056)	(0.008)	(0.082)	(0.008)				
d_subsidy		–0.0155*					0.3160***	–0.1178***	0.2101	–0.0121
		(0.009)					-0.096	-0.017	-0.143	-0.015
*d_insurance * age*			0.0737***							
			(0.015)							
*d_insurance * educ*				–0.0015*						
				(0.001)						
*d_insurance * income*					0.0351***					
					(0.006)					
*d_insurance * d_calorie*						–0.0239**				
						(0.010)				
*d_subsidy * age*							–0.0864***			
							(0.025)			
*d_subsidy * educ*								0.0114***		
								(0.002)		
*d_subsidy * income*									–0.0171	
									(0.011)	
*d_subsidy * d_calorie*										–0.0047
										(0.016)
log income	0.0786***	0.0944***	0.0778***	0.0789***	0.0593***	0.0784***	0.0968***	0.0972***	0.1052***	0.0943***
	(0.006)	(0.010)	(0.006)	(0.006)	(0.007)	(0.006)	(0.010)	(0.010)	(0.012)	(0.010)
educ_hhh	–0.0009*	–0.0023***	–0.0010*	–0.0000	–0.0011**	–0.0009*	–0.0025***	–0.0092***	–0.0024***	–0.0023***
	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)	(0.001)
log_age_hhh	0.0699***	0.0792***	0.0339***	0.0710***	0.0670***	0.0696***	0.1351***	0.0894***	0.0792***	0.0790***
	(0.008)	(0.013)	(0.011)	(0.008)	(0.008)	(0.008)	(0.021)	(0.013)	(0.013)	(0.013)
housequality	0.0005**	0.0004	0.0005**	0.0005**	0.0004**	0.0005**	0.0003	0.0002	0.0004	0.0004
	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)	(0.000)
d_empl_hhh	–0.0332***	–0.0476***	–0.0318***	–0.0328***	–0.0343***	–0.0332***	–0.0441***	–0.0424***	–0.0475***	–0.0477***
	(0.005)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)
d_kids	–0.0508***	–0.0612***	–0.0502***	–0.0508***	–0.0523***	–0.0509***	–0.0647***	–0.0635***	–0.0609***	–0.0611***
	(0.006)	(0.009)	(0.006)	(0.006)	(0.006)	(0.006)	(0.009)	(0.009)	(0.009)	(0.009)
d_urban	0.0213***	0.0350***	0.0206***	0.0210***	0.0217***	0.0213***	0.0349***	0.0335***	0.0347***	0.0350***
	(0.005)	(0.007)	(0.005)	(0.005)	(0.005)	(0.005)	(0.007)	(0.007)	(0.007)	(0.007)
Constant	–1.3482***	–1.5299***	–1.2010***	–1.3636***	–1.0788***	–1.3527***	–1.7699***	–1.5163***	–1.6771***	–1.5289***
	(0.089)	(0.139)	(0.094)	(0.090)	(0.101)	(0.089)	(0.155)	(0.139)	(0.168)	(0.139)
Observations	190,571	99,748	190,571	190,571	190,571	190,571	99,748	99,748	99,748	99,748
R-squared	0.031	0.035	0.031	0.031	0.031	0.031	0.035	0.035	0.035	0.035

Standard errors in parentheses

*** p<0.01, ** p<0.05, * p<0.1

Chapter 6 What Drives Electrification Inequality AcrossTime and Space in Indonesia?

6.1 Introduction

In this chapter, we analyse the spatial patterns of inequality in access to electricity and its supply provision in Indonesia. It is our aim to identify the determinants of spatial variation in the timing of electricity diffusion across Indonesian provinces. Several studies on technology adoption or diffusion have been carried out at the country level. For example, Comin et al. (2012) suggested that understanding technology diffusion over space is crucial to understanding the speed of technology diffusion. They found that countries located far away from the adoption leaders benefit less rapidly from the technology diffusion. Other studies have suggested that income per capita, openness, human capital and the type of regime across countries are associated with the speed of technology adoption (Comin and Hobijn 2004). Inspired by these considerations, in this chapter, we study the pattern of technology (electricity) diffusion within a country (Indonesia) at the province level.

Evidently, Indonesia is a particularly interesting country to study the spatial diffusion pattern of electrification. Indonesia is an archipelagic country where people are unevenly distributed across five big islands and hundreds of small islands. Unlike the situation in non-archipelagic countries, where new connections can be provided by extending the existing lines, Indonesia needs huge investments to build new lines to bring electricity to people on the different islands, and the up-front costs are very high. Therefore, electrification in Indonesia is spatially heterogeneous due to the geographical barriers, meaning that many islands have to rely on autonomous self-contained electricity systems.

The electrification of Indonesia started in the early 1900s in the colonialization era, especially at the centre of economic activities in most provinces. Evidently, in its initial phase, the diffusion of electricity mainly served the economic interests of the colonial elite. Interestingly, the diffusion process was very slow until about the 1980s, after which the electrification ratio started to grow rapidly in some provinces while remaining low in other provinces – thus implying an increasing spatial variation in access to electricity over time. For example, in 1975, the electrification ratio was 20 per cent in Jakarta and only 3 per cent in Papua. The most recent figures show that the percentage of households with access to electricity in Papua has increased to 35 per cent while in Jakarta it equals 100 per cent. Similarly, in 1975, the total installed electricity generation capacity was 1,066 MW, of which 771 MW was located in Java–Bali and only 16 MW was allocated to the Eastern islands. Since then, the national electricity generation capacity has increased to 57,822 MW (the 2018 figure), of which 65 per cent is located on the Java–Bali islands and only 3 per cent is to be found in Eastern Indonesia (PLN 2019).

Of course, this high level of disparity is mainly caused by spatial variation in the location of households, the stage of economic development, proximity to the centres of economic activity, institutional quality and economic incentives (Barnes and Foley 2004; Mulder and Tembe 2008; Jimenez 2017). Foley (1992) showed that electricity is a derived demand occurring only when an area has reached a certain level of development. As a consequence, the timing and the level of resources that should be committed to it at any particular moment of time or level of economic development varies. Moreover, spatial heterogeneity in electrification ratios is partly driven by a response to political conditions, such as post-independence until the early 1950s to gain sovereign, the cleaning-up communism ideology in the 1960s and other internal conflicts in Aceh, Maluku, Papua and so on. Hence, it is obvious that connecting people, including those living on the small islands, is not easy and has potentially large social implications. Furthermore, the story of the electricity uptake provides lessons for regional disparities throughout Indonesia and how to accelerate the development aimed at closing the gap between the more and the less developed provinces.

To understand the determinants of temporal and spatial variation in the speed of electricity diffusion, of course, we first need to describe the pattern of electricity diffusion across Indonesia by estimating the access to electricity at the province level over time. This is not an easy step to take, however, given the limitations to publicly available and consistent electricity access data series for a longer time period across provinces. Therefore, we built a consistent database describing the spatial and temporal variation of electrification in Indonesia. To achieve this aim, we make use of unpublished statistics kept by PLN, which, as the only company providing electricity throughout Indonesia, plays an important role in shaping the electricity diffusion pattern across Indonesia. Hence, an important contribution of this chapter is also to document the complexities of electricity data in Indonesia and to describe the methodological choices that we must make in constructing a consistent dataset. Given the limitations in data availability, we are able to construct consistent time series for the period 1975 to 2018. We complemented our dataset with information on per capita GDP, population density, firm density, landscape complexity (measured as the share of households living in flat areas) and per capita installed electricity generation capacity across provinces and islands.

Using this rich dataset, we develop a non-linear least square estimation to estimate the speed of the diffusion process within an S-curve framework of access to electricity. More specifically, we first regress the electrification ratio in various provinces over time on a set of explanatory variables, including per capita GDP, population density, firm density, landscape complexity and per capita installed electricity generation capacity – thus exploring the spatial and time variation in our data. Next, we translate the non-linear least square regression outcomes into the β parameter that defines the speed of diffusion in the S-curve framework. By plugging the estimated β_i into our sample, we obtain the predicted electrification ratios for each province in terms of an S-curve over time. Finally, we calculate for each province the gap between its own electrification performance and that of the leading island (Java-Bali) and use our regression outcomes to decompose this gap into the respective contribution of the driving forces of electrification mentioned above, which differ across space.

Our main hypothesis in this chapter is that spatial variation in the speed of electricity diffusion can be explained by variation in the population density, per capita GDP, geographical structure, industrial intensity and power supply. To test this hypothesis, we adopt a non-linear S-shaped diffusion curve framework and then use a pooled regression analysis to estimate the parameters that determine the speed of diffusion within such a framework. Finally, we identify those of the above-mentioned factors that determine the pattern of diffusion and simulate the number of years that are required to electrify at least 50 per cent of households.

The remainder of this chapter is organized as follows. In section 6.2, we discuss the construction of long-term data series of electricity access ratios in Indonesia. Section 6.3 provides a short overview of the history of electricity diffusion in Indonesia based on our newly constructed dataset as well as existing data. Section 6.4 describes the empirical strategy that we develop to estimate the diffusion of electricity access based on historical data, including a description of the data employed. In section 6.5, we simulate electricity diffusion patterns for each province by integrating the results of our regression analysis into a simple theoretical S-shaped diffusion framework. In section 6.6, we identify for each province the key determinants that drive the speed of electricity diffusion, using a combination of model simulation and decomposition techniques. Section 6.7 provides concluding remarks and discusses avenues for future research.

6.2 Constructing a Long-term Electricity Dataset

In this section, we will present the unique electricity dataset that we have constructed for the research presented in this chapter. As noted, this dataset is constructed on the basis of unpublished statistics kept by PLN, Indonesia's main electricity provider. We use the PLN annual electricity statistics at the region and distribution levels throughout Indonesia to obtain the number of customers and installed capacity as our main variables. We collect these statistics from various sources, specifically PLN's website, PLN's headquarters and PLN's archives for data after 2000, during the 1990s and before 1990, respectively. Moreover, we directly consult a resource person from PLN[37] to understand the technical aspects and the context behind the statistics. Nevertheless, most of the statistics are available in the form of hard copies that we had to to digitize, enter and clean. We finally manage to gather 44 years of electricity data covering the period from 1975 to 2018 at the province level. Some challenges are encountered in constructing the electricity dataset, especially regarding the nomenclature of business units, area coverage and tariff classification. Below we

[37] We are grateful to Amir Rosidin for his expert judgement in preparing the electricity dataset.

summarize the main features of our dataset and the methodological choices involved in processing the raw data. For more details we refer to the Appendix A6. in this chapter.

6.2.1 Nomenclature of Business Units

To serve customers throughout Indonesia, PLN established two main business units, *Kantor Wilayah* (the regional office) and *Kantor Distribusi* (the distribution office). *Kantor Wilayah* is responsible for managing not only the customer but also the production side, while *Kantor Distribusi* is mainly in charge of the customer side. In particular regions, such as Maluku and Papua, *Kantor Wilayah* performs three main functions, that is, production, transmission and distribution. Meanwhile, in Java–Bali, where all the systems are already interconnected, each function is carried out by an independent entity; that is, the production is under subsidiary companies, the transmission is carried out by the Transmission Load and Dispatch Centre (*Pusat Pengatur dan Penyaluran Beban –P3B*) and the distribution is performed by *Kantor Distribusi*. These task divisions are carried out for efficiency and reliability purposes. Due to this nature, PLN frequently changes its business units to accommodate the electricity development. This company's transformation affects our data collection process. For instance, PLN business units in 1975 comprised 1 generating unit, 2 *Kantor Distribusi* and 13 *Kantor Wilayah*, and since 2018 number of business units have become 15 *Kantor Wilayah*, 7 *Kantor Distribusi* and 10 related generating units. As a consequence, customer and installed capacity data are recorded and stored in different PLN's business units following the new organization. Therefore, the consistency in the annual PLN statistics needs to be taken into consideration. Historical dynamics of PLN business units is presented in Appendix A6.1.

6.2.2 Area Coverage

PLN distinguishes the areas for customers and production where the boundaries differ from those of the BPS's provinces. One PLN area can consist of more than one BPS province. To compile all the electricity data from the annual PLN statistics, we first match the PLN area coverage within the PLN statistics over time to obtain a coherent dataset at the province level and then we reclassify the matched PLN provinces into the BPS's provinces. For the period from 1990 onwards, the PLN statistics are available at the regional level with a provincial breakdown, so we can easily match them with the BPS province data. However, the name of PLN regions might not reflect the true coverage area, so we need to reconfirm this with PLN during this reclassification process.

Some PLN regions cover several provinces. For the purpose of coding thirty-three provinces, we recode these regions and distributions as follows. Prior to 1990, for some regions with more than one province, their data are recoded to the province where the regional office was located. This pragmatic strategy is chosen because, according to PLN, the location of the regional office was selected based on the number of PLN customers. In other words, the capital province of a region represents more electrified households than other provinces within the region. For example, Region III consists of West Sumatera, Riau and Riau Island, so data were recorded as West Sumatera province because the office of Region III is located in Padang, the capital city of West Sumatera province. Thus, seven regional provinces were considered as one province according to this pragmatic strategy, namely West Sumatera (Region III), South Sumatera (Region IV), South Kalimantan (Region VI), North Sulawesi (Region VII), South Sulawesi (Region VIII), Bali (Region XI) and Central Java (Distribution Central Java and Yogyakarta). However, it should be noted that there is a potential jump in the number of customers due to discontinuity in the definition of the regions. Detailed coverage area PLN business units and BPS can be found in Table A6.2. For new business units that are established to response electricity development in some provinces, we reclassify these new business units back to the parent regions or provinces to make our dataset consistent with the BPS's classification at the province level.

6.2.3 Installed Capacity and Power Plants

In this chapter, we use installed capacity (megawatt) as a proxy of power supply, which is to be understood as the arrival of power supply in that area (island). According to PLN, the installed capacity is the capacity of one generating unit as written on the generator's name plate. In other words, the installed capacity is the maximum output of electricity that a generator can produce under ideal conditions. PLN records this information based on the location of the power plants, including the ones owned by private companies, and the interconnection system in which the capacity of the entire area should be managed (Appendix A6.1). As a consequence, the division of regions for installed capacity becomes more complicated because it should be managed as one system and in line with its transmission. For instance, the generator system of Northern Sumatera covers Aceh, North Sumatera, West Sumatera, Riau and Riau Island, while the generator system of Southern Sumatera comprises South Sumatera, Bangka Belitung, Jambi, Bengkulu and Lampung. As a result, electricity can be transmitted across provinces throughout Sumatera island. This interconnected system is also known as autonomous self-contained electricity system in which the capacity could be supplied, transmitted and distributed within one island via this system.

Having this complexity and as suggested by PLN, we prepare the capacity dataset at the island level, and then we use the weighted population to obtain the installed capacity per province. We categorize islands into five groups as follows: (i) Sumatera, (ii) Java–Bali, (iii) Kalimantan, (iv) Sulawesi and (v) Eastern. Unlike the first three main islands, the latter consists of one big island (Papua) and many small islands (Maluku and Nusatenggara). The weighted population ($wpop_{ij}$) for each province is calculated as the share of the population of province i on island j (pop_{ij}) in the total population of island j (pop_j), where i is an index for provinces (i = 1, …, 33) and j is an index for islands (j = 1, …, 5):

$$wpop_{ij} = \frac{pop_{ij}}{pop_j}$$

The calculation of the installed capacity also includes the amount of installed capacity from the private sector or the independent power producer (IPP) as part of the power supply since 2014. In sum, the quite frequent changes in PLN's organization structure require us not only to consult PLN directly but also to prepare our electricity dataset carefully, especially due to the fact that the power plants belong to different units in different regions.

6.2.4 Tariff Classification

As discussed, PLN issued annual statistics in different formats, including a tariff classification. We reclassify tariffs to calculate the number of customers as follows: (i) prior to 1980, tariffs A, B and C are classified as residential, industry and business, respectively, and tariffs D, E and F and *Khusus* are classified as “other”; (ii) in the period from 1980/1981 onwards, tariff R is classified as residential, I as industry, U and H as “business” and S, G, J and M as “other”. We then translate customers who held a residential tariff as household customers. Thus, in general, we have four types of customers: households, industry, business and other (Table A6.3). From 1989 onwards, PLN issued not only the number of customers based on tariffs but also the number of customers based on types, namely residential, industry, business, social, government office and public street lighting. To make our classification consistent with previous classifications, we categorize the last three types as “other”. In this chapter, we only use household customers to determine the number of connections.

6.2.5 Electrification Ratio

The main variable in this chapter is the electrification ratio, which we use as the dependent variable in our diffusion model. According to the Asia Development Bank (ADB 2016), the *electrification ratio* is *defined* as the *ratio* of the number of households with access to electricity relative to the total number of households. This ratio is important for measuring the extent to which people have benefited from the development, that is, electricity. As discussed previously, PLN only issues annual statistics without the total number of households. Therefore, the electrification ratio is calculated based on two different sources: (i) the number of households with access to electricity as the nominator is recorded by PLN as its customers and the Ministry of Energy and Mineral Resources (MEMR) for non-PLN customers[38]; and (ii) the total number of households as the denominator is recorded by the BPS based on the family card. In this chapter, we only focus on electrified households based on PLN data

[38] These electrified households are part of rural electrification program under Ministry of Energy and Mineral Resources.

because the share of non-PLN electrified households is less than 5 per cent of the total number of electrified households.

Electrified households are recorded by PLN based on the number of meters installed at the customer's house – one meter is counted as one household customer. There are two potential measurement issues. First, in rural areas where several families live in one house, the electrification ratio measured could be lower than the effective ratio. For instance, the electrification ratio calculated by PLN is 56.3 per cent while the author's calculation using the BPS data results in a ratio of 58.5 per cent.[39]

Second, in the big cities like Jakarta, where housing is also available in the form of apartments or flats, pavilions (part of a main house that can be rented) and rental rooms inside a house (*kos-kosan*), the name of the owner of those properties could be the same as the name of the PLN customer who lives in the rental property where the individual meter is installed. Thus, the effective number of PLN household customers will be larger than the number of actual households because an owner's name could be registered for several meters.

During our data collection, we could not find the PLN statistics from 1984 to 1988, even when we searched the archives at the PLN headquarters in Jakarta. To fill in the missing values, we interpolate two groups: (i) 1984 to 1985, using data from 1979 to 1983, and (ii) 1986 to 1988, using data from 1989–1999. After interpolating them, we consult PLN to check that the figures are reasonable.

6.3 A Short History of Electricity Diffusion in Indonesia

The supply and diffusion of electricity in Indonesia started in the period of Dutch colonization, and Jakarta was the earliest to benefit, in 1897 (Table 6.1). Other provinces with substantial economic activity, such as mining, sugar and tea factories, plantations, trading and so on, also started electrification in the early 1900s (McCawly 1971). After independence, three main private Dutch electricity companies, namely *Nederlandsch Indische Gasmaatschappij* or the Netherland Indies Gas Company

[39] In several PLN statistics, we also find that the denominator for calculating the electrification ratio was three years behind while the nominator was calculated from the actual or current data, so the ratio might be too high compared with the effective one.

(NIGM), *Gemeentelijk Electriciteitsbedrijf Bandoeng en Omstreken* or the Municipal Electricity Works for Bandung and Surroundings (GEBEO) and *Algemeene Nederlandsch-Indische Electriciteits-Maatschappij* or the Netherlands Indies General Electricity Company (ANIEM), were nationalized by the Government of Indonesia in the early 1950s, but the process was fraught with difficulties. This tough transition was followed by political instability and hyperinflation in the 1960s. Together with a lack of funds, engineers and qualified management, it resulted in PLN running the electricity company with poor operating conditions and a lack of expansion possibilities (McCawly 1971). After several organizational changes, PLN became the only state-owned electricity company that was responsible for connecting people throughout Indonesia from 1972 onwards.[40] Therefore, the available PLN data that we can collect reach back to 1975.

If we look at the area coverage of electricity (the third column of Table 6.1), about 75 per cent of districts in Indonesia were already connected prior to 1970. However, the electrification ratio after 4 decades was only 83 per cent in 2018. This suggests that the electrification ratio prior to the 1970s was very low and electricity connections were concentrated in a few areas only. In Figure 6.1, we show the variation in the electrification ratios across provinces. It can be seen that only a few people in particular provinces had access to electricity in 1975 (Panel A) and that this picture had changed considerably by 2018 (Panel B).

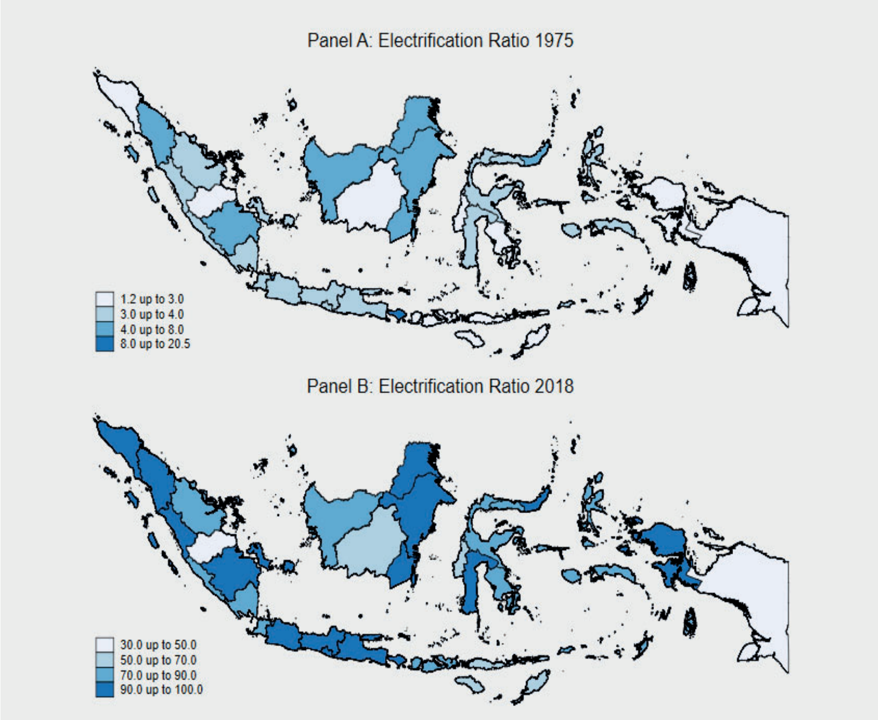
[40] On 27 October 1945, Sukarno, the first Indonesian president, established *Jawatan Listrik dan Gas* (the gas and electricity company) under the Ministry of Public Works and Power with an installed capacity of 157.5 MW. In the period from 1 January 1961 until 1 January 1965, this electricity company was changed into *Badan Pimpinan Umum Perusahaan Listrik Negara* or the General Leader Agency PLN (BPU-PLN) and its capacity almost doubled to almost 300 MW. In 1972, the Government of Indonesia issued a Government Regulation, which stated that PLN was the only company to serve electricity throughout Indonesia. Since 1994, PLN has been a state-owned company and it operates under the Ministry of State Owned Companies and is supported by the Ministry of Mining and Energy.

Table 6.1 Electrification in Indonesia

Province Name	First Time Electrified	Percentage of Districts Electrified Prior to 1970*
Nanggroe Aceh Darussalam	1900	96
North Sumatera	1899	79
West Sumatera	1900	100
Riau	1900	100
Jambi	1900	82
South Sumatera	1920	53
Bengkulu	1911	20
Lampung	1899	27
Bangka Belitung Islands	1917	43
Riau Islands	1900	100
Jakarta	1897	83
West Java	1900	100
Central Java	1902	63
Yogyakarta	1906	100
East Java	1899	100
Banten	1901	88
Bali	1950	33
West Nusatenggara	1916	60
East Nusatenggara	1900	14
West Kalimantan	1921	100
Central Kalimantan	1899	93
South Kalimantan	1899	100
East Kalimantan	1900	60
North Sulawesi	1901	100
Central Sulawesi	1900	100
South Sulawesi	1921	83
Southeast Sulawesi	1950	79
Gorontalo	1905	100
West Sulawesi	1950	83
Maluku	1903	45
North Maluku	1900	100
West Papua	1900	54
Papua	1900	45
Indonesia	1897	75

Source: McCawly (1971), PT PLN (1995), author's calculation from the PLN database.
* This column is calculated based on the first time electricity was available, recorded by PLN in the particular district, regardless of the number of connections. As a consequence, a district is counted as an electrified district even if there is only one customer or one electric pole.

Figure 6.1 Electrification Ratio in Indonesia in 1975 and 2018 based on PLN Regions



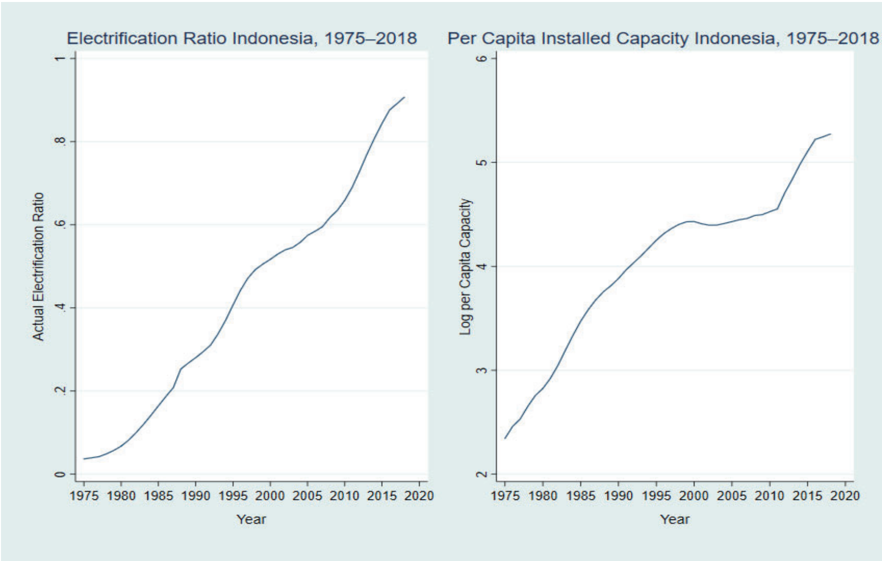
Source: PLN, produced by the authors.

PLN expanded steadily after the 1970s, then began to accelerate after 2010, as is apparent in the increasing electrification ratio as well as the installed capacity, as shown in Figure 6.2. However, the Asian Financial Crisis in 1998 contributed to some delays in the electrification programme, as we can see in the slowing down of the capacity expansion and the substantially smaller electrification ratio between 1998 and 2010. During the crisis, many infrastructure projects, including private power plants, were postponed while, at the same time, PLN encountered financial difficulties and a lack of investment funds (Sambodo 2016).[41] After some renegotiation and

[41] As discussed by Sambodo (2016), in the 1990s, domestic money to finance investments in the power sector was limited, while borrowing money from international institutions was flexible. Therefore, PLN preferred to obtain long-term loans in foreign currencies as a source of financing. Moreover, to accelerate the electrification programme, PLN signed 26 power purchase agreements (PPA) with private investors to develop power generation. However, the economic crisis that hit Indonesia in 1997/1998 caused PLN to suffer not only from huge

rescheduling with the investors, PLN has gradually increased its capacity since 2005 and intensified it from 2010 onwards through fast-track programmes followed by Jokowi's 35,000 MW programme. As a result, the electrification ratio has also increased slightly. This acceleration programme also includes rural electrification in 433 villages located in Eastern Indonesia, where villages have remained completely dark until now (Ministry of Energy and Mineral Resources 2020).

Figure 6.2 Electrification Ratio and Per Capita Installed Capacity in Indonesia, 1975–2018

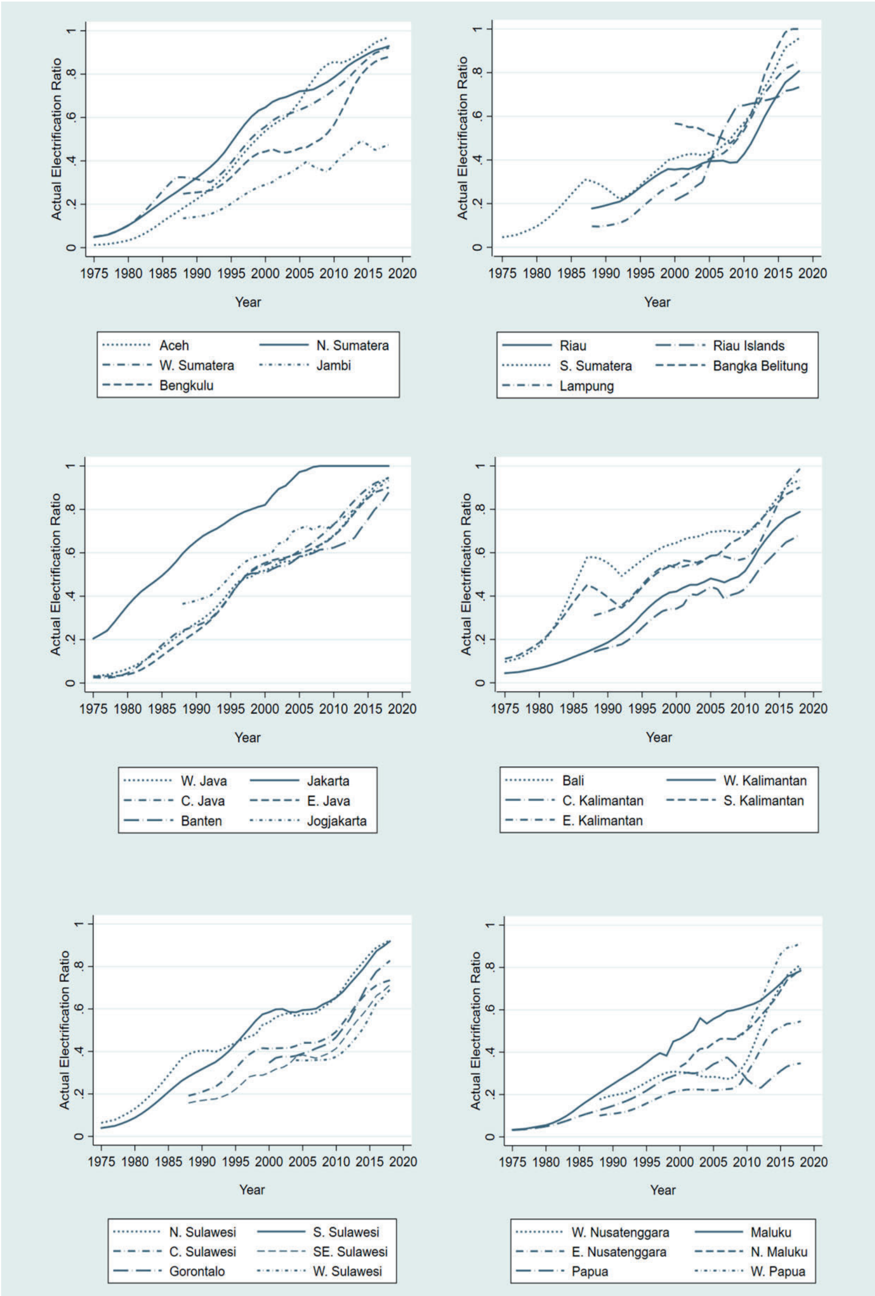


Source: PLN statistics, calculated by the authors.

Interestingly, the huge expansion of the national electricity generation capacity did not substantially change the spatial distribution of this capacity. The share of islands in the national capacity has remained roughly the same over time, that is, Java–Bali with the largest share (70 per cent) followed by Sumatera with 20 per cent, Kalimantan and Sulawesi with around 6 per cent and the Eastern islands with around 3 per cent.

operating losses due to currency depreciation and very high interest rates but also from an international arbitration lawsuit that forced it to pay a penalty for postponed or cancelled contracts of PPAs. These financial problems lasted for 6 years after the crisis.

Figure 6.3 Actual Electrification Ratio of All Provinces in Indonesia, 1975–2018



Source: PLN statistics, calculated by the authors.
Note: Jakarta excludes the Seribu Island district in the electrification ratio calculation.

Using our new dataset, we can calculate that, on average, the electrification ratio in Indonesia increased at 6 per cent per year from 6 per cent in 1975 to 83 per cent in 2018. However, this ratio varies across provinces and its average is 45 per cent (Table 6.3). Only Jakarta (excluding the Seribu Islands district) has reached a 100 per cent electrification ratio since 2008, followed by Bangka Belitung Island in 2017, while the three lowest electrification ratios in 2018 are found in Papua, Jambi and East Nusatenggara with corresponding figures of 35 per cent, 47 per cent and 55 per cent, respectively (Figure 6.3).

Table 6.2 shows how fast electricity has diffused across provinces since 1975. For instance, only two provinces achieved at least 10 per cent access to electricity in 1975 and at least 50 per cent in 1990, and only one province was fully electrified in 2010. Both Bali and Jakarta provinces achieved at least a 50 per cent electrification ratio in 1986, while Jakarta and Bangka Belitung were fully electrified in 2008 and 2017, respectively. It appears that Bali is not as fast as Bangka Belitung in reaching full electrification as Bali only achieved 93.3 per cent in 2018.

Table 6.2 Number of Provinces Achieving the Minimum Electrification Ratio, 1975–2018

Year	Minimum Electrification Ratio				
	≥10%	≥25%	≥50%	≥75%	100%
1975	2	0	0	0	0
1980	6	1	0	0	0
1985	14	5	0	0	0
1990	25	13	2	0	0
1995	26	21	2	1	0
2000	31	29	15	1	0
2005	32	31	16	1	0
2010	33	33	23	3	1
2015	33	33	31	19	1
2018	33	33	31	25	2

Source: PLN statistics, calculated by the authors.

The electrification trend began to accelerate in 2000; at least half of the households in 15 provinces were connected to the electricity grid in 2000 and at least 75 per cent electrification was achieved by 19 provinces in 2015. This relatively quick diffusion is due to the national programme to accelerate electrification, that is, the fast-track programmes followed by the 35,000 MW programme. If we set the minimum electrification ratio as 10 per cent, most of the provinces reached this threshold in the 1990s. It seems that the minimum of 10 per cent electrification acted as a threshold from which to take off to achieve sustained connection. Further discussion will be presented in the next section.

6.4 Empirical Strategy and Data Description

6.4.1 Empirical Strategy

Our data show the non-linear pattern of the electrification ratio throughout Indonesia (Figure 6.3). The diffusion electricity process follows an S-curve, as discussed by Comin et al. (2012). In this chapter, we define the diffusion of electricity S_t as the share of electrified households in the total households; hence, the value of S is between 0 and 1 or between 0 and 100 on the percentage scale. The diffusion of electricity (S_{ti}) at time *t* for province *i* depends on the slope β_i of several factors (X_i), as presented in equation (6.1). Parameter β_i determines the speed of the parameter switches due to variations in X_i. As we discussed earlier, each province started electrification differently, so the diffusion pattern is non-linear. We assume that the diffusion process can be described by a continuous logistic function between 0 and 1, given by[42]:

$$S_{ti} = \frac{1}{1 + e^{-\sum \beta_i X_i + \varepsilon_{t(ei)}}} \qquad (6.1)$$

where S_{ti} denotes the diffusion of electricity and the X_i's refer to the same set of explanatory variables, which are measured in logs whenever the variables are not the share or the percentage. Therefore, we apply a non-linear least square estimation for parameter β_i, which determines the speed of the diffusion process within the S-curve framework of access to electricity (Stoneman 2002).

[42] We use this simple equation following Lankhuizen et al. (2011).

Figure 6.4 Hypothetical Diffusion of Technology

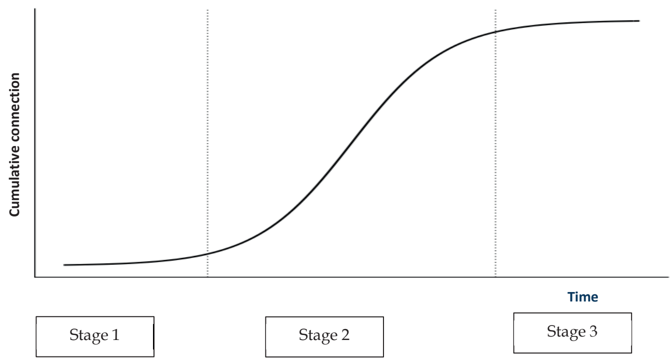


Figure 6.5 describes how electricity diffuses in a hypothetical province. We adopt the typical diffusion of technology curve to explain electrification in Indonesia. In brief, the electricity connection is very low and increases slowly in the early period (stage 1), then it becomes accessible more quickly as time passes, known as the take-off period (stage 2). The last stage occurs when the electricity connection reaches 100 per cent and the speed remains the same (stage 3).

Thus, our questions include the following: what is the global pattern of electricity adoption in Indonesia, how is it underpinned by regional- and national-level processes, which provinces adopted electricity first and do the latecomers catch up with them? To answer these questions with regard to the electrification ratio in Indonesia, we define the concept of "electricity take-off", which occurs when the share of electrified households reaches at least 10 per cent of the total households and marks the transition to a sustained connection. Our actual data show that only two provinces reached an electrification ratio of at least 10 per cent in 1975, fourteen provinces in 1985 and all provinces in 2010 (Table 6.2). Due to a common pattern of new technology deployment, the so-called "S-curve" (Stoneman 2002), the take-off time can largely explain the subsequent growth of electrification. In the Indonesian context, the story of electricity uptake provides lessons for regional disparity throughout Indonesia and how access to electricity could accelerate the closing of this gap.

Given the limited data across provinces in Indonesia, our hypotheses are as follows: the electricity diffusion process, presented as the electrification ratio in province *i* in year *t* ($electratio_{it}$), correlates with internal factors or province characteristics *i* in year *t*, such as economic activities measured in per capita GDP (log_gdpcap_{it}), population density ($log\_popdens_{it}$), geographic structure measured as the number of people per island within one province ($log\_isldist_{it}$) and share of households living in flat areas (pct_hhflat_{it}) as well as the firm density defined as the number of firms per 1000 population ($firmdens_{it}$). Provinces with more economic activities, more densely populated areas, a less scattered population living across islands, more people living in flatter areas and more firms operating are expected to have a positive correlation with the electricity diffusion.

We also expect the diffusion process to correlate positively with the power supply as an external factor. To electrify an island area, PLN has to build an island-specific infrastructure, including generation capacity transmission lines and distribution sub-stations; Indonesia comprises a series of autonomous self-contained electricity systems, as the islands are not connected in terms of electricity infrastructure (except for Java-Bali). Hence, cross-island differences in the timing of the arrival of power supply (i.e. investment in the first power plant by PLN) is in this context an obvious determinant of cross-islands speeds in the speed of electricity diffusion; after all PLN did not install firt-time electricity generation capacity in the same year across all islands. We use PLN's definition of the electricity supply, namely installed capacity as the capacity of one generating unit as written on the generator's plate or prime mover. In this chapter, we calculate the per capita installed capacity in watts (log_watcap_{it}) as our control variable. A detailed description of the calculation of the installed capacity can be found in section 6.2.

Following the discussion of Figure 6.4, each province has a different speed of diffusion; therefore, we include time ($time_{it}$) as our independent variable to capture time variation in the diffusion process. Thus, our specification to estimate the electrification ratio with pooled regression is presentedn in equation (6.2). In this model, there are two main dimensions of the diffusion process, that is, cross-sectional variation and time variation. We are interested in understanding which is the typical province that first adopted electricity and how long it took to reach a certain electrification ratio. Further, we translate the non-linear least square regression outcomes into the β parameter in the S-curve. By plugging the estimated β_i into our sample, we obtain the estimated electrification ratio to produce an S-curve over time for each province.

$$electratio_{it} = 1 / (1 + \exp(-1 \cdot (\beta_0 + \beta_1 \cdot log\_gdpcap_{it} + \beta_2 \cdot log\_popdens_{it} + \beta_3 \cdot log\_isldist_{it} + \beta_4 \cdot pct\_hhflat_{it} + \beta_5 \cdot time_{it} + \beta_6 \cdot firmdens_{it} + \beta_7 \cdot log\_watcap_{it}))) + \varepsilon_{it} \quad (6.2)$$

To smooth out the short-term fluctuations in our dataset, we use five-year moving averages; that is, we take two years in the past and two years ahead because these factors in the expectation and we assume that people at time *t* have an expectation that materializes two years ahead. This strategy helps us to eliminate the odd pattern of the predicted electrification ratio.

In our next discussion, we perform a simulation using two thresholds of electrification rate, that is, 50 per cent as the mid-point of the S-curve and 25 per cent as an arbitrary level of the electrification ratio above the take-off level of at least 10 per cent. As a benchmark, we use Jakarta's characteristics in our simulations because Jakarta (excluding Seribu Island district) is the most and the earliest developed province, where its electrification ratio reached 100 per cent in 2008. Hence, we calculate how long province *i* took to reach its electrification ratios of at least 50 per cent and 25 per cent, respectively, by comparing the predicted pattern of provinces with the hypothetical pattern if the provinces add the power supply of Jakarta. Then, we can decompose the gap to derive the predicted path by adding the per capita GDP, landscape and population density of Jakarta one by one to see which part of the gap is contributed by internal and external factors. We argue that provinces located in Eastern Indonesia need a longer time to reach those thresholds, mostly due to island barriers.

Our unit of analysis is the province level, so we aggregate all the variables to the province level. We use pooled data from 1975 to 2018 for 33 provinces, including

7 new provinces following the BPS classification[43] and 5 islands, namely Sumatera, Java–Bali, Kalimantan, Sulawesi and Eastern islands. The latter consist of one big island and numerous small islands, specifically East Nusatenggara, West Nusatenggara, Maluku, North Maluku, Papua and West Papua.

In our regression framework, we estimate the electrification ratio – i.e. the percentage of electrified households to the total number of households – as function of a set of explanatory variables: per capita GRDP, population density, geography and landscape, and per capita capacity (watt), to capture the power supply of electricity. We use the geography and landscape variables to see how these natural conditions correlate with the diffusion of electricity in an archipelagic country like Indonesia. The geography variable is calculated as the average distribution of people per island in each province, while the landscape variable is calculated as the average percentage of households living in the flat areas in each village and then aggregated from the village to the province level.

6.4.2 Data Description

Our electricity data mainly originate from various statistics sources by PLN, as described in section 6.2, supplemented with data from BPS. However, aggregating data at the province level for long time series is challenging. The choice of variables is subject to the data availability at the province level for the last 44 years. The variables per capita GRDP and population density from 1975 to 2018 are drawn from various editions of Statistics Indonesia and Provinces in Figures, and geography and landscape are taken from Podes (village potential) 1986–2018. We assume that the number of islands per province and the percentage of villages located in flat areas has

[43] Aceh, North Sumatera, West Sumatera, Riau, Jambi, South Sumatera, Bengkulu, Lampung, Bangka Belitung Islands (a new province splitting from South Sumatera in 1999), Riau Islands (a new province splitting from Riau in 2002), Jakarta, West Java, Central Java, Jogjakarta, East Java, Banten (a new province splitting from West Java in 1999), Bali, West Nusatenggara, East Nusatenggara, West Kalimantan, Central Kalimantan, South Kalimantan, East Kalimantan (including the new province North Kalimantan, established in 2014), North Sulawesi, Central Sulawesi, South Sulawesi, Southeast Sulawesi, Gorontalo (a new province splitting from North Sulawesi in 1999), West Sulawesi (a new province splitting from South Sulawesi in 2004), Maluku, North Maluku (a new province splitting from Maluku in 1999), West Papua (a new province splitting from Papua in 2003) and Papua.

not changed much during the last 10 years, so we interpolate the missing data for geography and landscape from 1975 to 1985. The per capita capacity (watt) at the province level is calculated from the installed capacity of the PLN statistics 1975–2018, which we weight by the population (see section 6.2 for a detailed explanation).

Among the explanatory variables, population density has the highest variation, followed by per capita installed capacity (Table 6.3). This suggests that people living in scattered areas, mostly inhabited by poor communities, in line with the distribution of people per island and the variation in capacity, could explain the difficulties in connecting electricity throughout Indonesia, especially in rural areas. Barnes (2007) associated rural electrification with a lower number of connections per kilometre of line, which leads to high capital and operating costs, a low level of consumption due to poor consumers, the lack of an industrial load, a heterogeneous landscape, a lack of motivation for private investors and interventions from politicians insisting on favoured constituents. Table 6.4 shows that all the explanatory variables are statistically correlated with the electrification ratio. Among those variables, per capita installed capacity has the highest correlation, followed by per capita GDP and population density. This table also indicates that there is no multicollinearity as none of explanatory variables are strongly correlated with each other.

Table 6.3 Descriptive Statistics: Key Variables

Variable	Obs.	Mean	SD	Min.	Max.
Electrification Ratio	1125	0.45	0.24	0.01	1
Per Capita GDP (in Million Rupiah)	1282	7.66	8.8	0.18	54.3
Population Density	1286	648.8	2,280	3.3	15,589
People Distribution per Island (in Thousands)	1286	43.44	62.82	0.3	366
Share of Households Living in Flat Areas	1257	0.66	0.19	0.09	1
Firm Density	1286	0.06	0.05	0.00	0.3
Per Capita Installed Capacity (in Watts)	1286	76.88	73.17	2.85	650.07
Log per Capita GDP	1282	15.42	0.9	12.1	17.81
Log Population Density	1286	4.67	1.66	1.19	9.65
Log People Distribution	1286	9.58	1.66	5.71	12.81
Log per Capita Installed Capacity	1286	3.96	0.94	1.05	6.48

Table 6.4 Correlation Matrix: Key Variables

	(1)	*(2)*	*(3)*	*(4)*	*(5)*	*(6)*	*(7)*
1. Electrification Ratio	1						
2. Per Capita GDP (in Million Rupiah)	0.39*	1					
3. Population Density	0.32*	0.41*	1				
4. People Distribution per Island (in Thousands)	0.09*	–0.12*	0.09*	1			
5. Share of Households Living in Flat Areas	0.29*	0.11*	0.29*	0.19*	1		
6. Firm Density	0.31*	0.29*	0.61*	0.38*	0.27*	1	
7. Per Capita Installed Capacity (in Watts)	0.68*	0.16*	0.15*	0.27*	0.25*	0.35*	1

Note: $^{*}p<0.01$.

6.5 Estimating the Diffusion Patterns

This section presents the econometric results to identify the relative contribution of the different factors in driving the diffusion speed of electricity in the context of the S-curve i.e. level of the economy, population, geographical barriers, landscape, firm density, time and installed capacity of power supply. As previously noted, the latter can be considered as an exogenous factor in this context because it essentially captures cross-island differences in the timing of the arrival of power supply (i.e. investment in the first power plant by PLN). The island structure defines the electricity system of Indonesia as a series of autonomous systems (the main islands are not connected in terms electricity infrastructure), the first year of investment in power supply by PLN differs across islands and the investment size was obviously not driven by actual electricity demand.

Table 6.5 Identification of the Relative Contribution of Several Potential Determinants of the Speed of Electricity Diffusion

Variable (1)	Expected Sign of β (2)	Operational Variable (3)	Predicted β (4)
Dependent Variable			
electrification ratio		*electratio*	
log per capita GDP	+	*lgdpcap*	0.112***
			(0.02)
log population density	+	*lpopdens*	0.180***
			(0.02)
log island distribution	–	*lisl_distr*	–0.159***
			(0.01)
share pop in flat area	+	*pct_hhflat*	0.429***
			(0.09)
time	+	*time*	0.0457***
			(0.002)
firm density	+	*firmdens*	0.176
			(0.60)
log per capita capacity	+	*lw_watcap*	0.717***
			(0.04)
Constant			–5.715***
			(0.39)
Observations			1,112
R-squared			0.967

Notes: Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

Our results show that an increasing per capita GDP, higher population density, less scattered population living across islands, larger share of households living in flat areas, higher firm density and higher per capita capacity will speed up the diffusion of electricity. These coefficients are statistically significant except for firm density (Table 6.5). The latter result might be due to the fact that firms' location is concentrated only in a few areas within provinces; many big factories have their own generators to ensure that they can operate uninterruptedly, so the concentration of firms is not statistically significant in the relationship with electrification ratio. This implies that connecting people involves more pressure than providing electricity to firms.

Table 6.6 Marginal Effects of he Determinants of the Speed of Electricity Diffusion

Variable	Marginal Effect
log per capita GDP	0.0217^{***}
	(0.005)
log population density	0.0349^{***}
	(0.003)
log island distribution	-0.0308^{***}
	(0.002)
share pop in flat area	0.0832^{***}
	(0.02)
time	0.0089^{***}
	(0.001)
firm density	0.0341
	(0.12)
log per capita capacity	0.1390^{***}
	(0.01)
Observations	1,112
R-squared	0.964

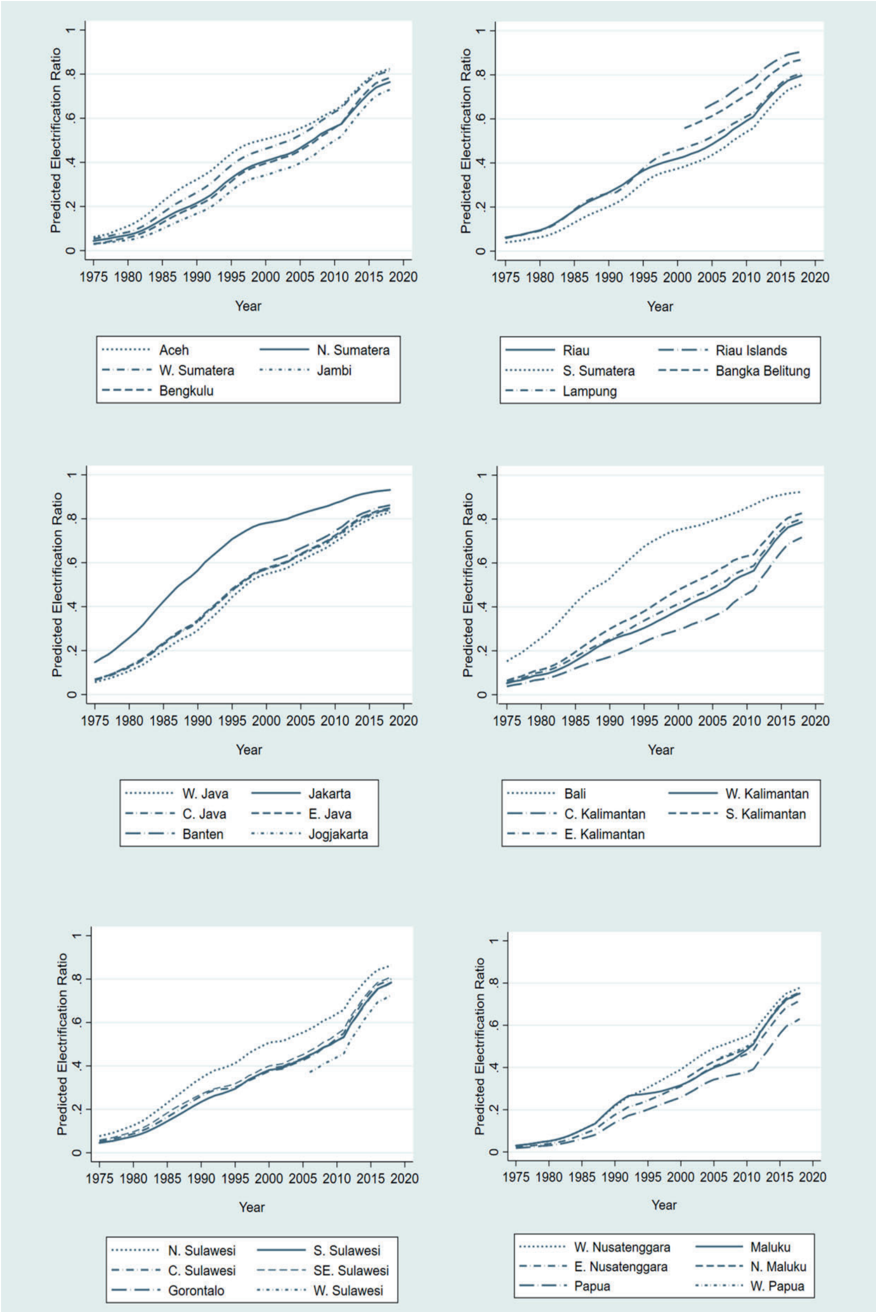
Notes: Robust standard errors in parentheses. *** $p<0.01$, ** $p<0.05$, * $p<0.1$.

The magnitude of the respective variables' correlation with the electricity ratio is reported in Table 6.6. As expected, the installed capacity of power supply at an island is an important driver of the speed up of electricity diffusion. If we increase the installed capacity per capita by 1 watt, then we expect the share of electrified households to increase by 0.139 percentage points. Since the electricity supply is an autonomous self-contained system on each island, and the timing of investments in installed power supply capacity by PLN differs across islands, this suggests that the geographic island barrier is a main determinant of the (lack of) speed of electricity diffusion. Our analysis also shows that the geographic structure, as presented by the share of the population living in flat areas, contributes to speeding up the electricity diffusion process. Comin et al. (2012) suggested that geography plays a significant role in determining technology diffusion across countries. For instance, to increase the electrification ratio, off-grid technologies have been successfully supplemented in the on-grid programme in remote areas in Bangladesh (Rahman et al. 2013), in Sub-Saharan Africa (Dagnachew et al. 2017) and in China (Bhattacharyya and Ohiare 2012).

Indonesia has a high variation in population density, ranging from 3.3 to 15,589 people per square kilometre (Table 6.3). Our results show that, if the population density increases by 1 point, we expect the electrification ratio to rise by 0.035 percentage points. Time also contributes to speed up the electrification. For every additional 10 years, the electrification ratio increases by 0.089 percentage points. This could explain why the provinces in Java and on other islands that have already been connected since the early 1900s also experience a higher electrification ratio.

If we look at the diffusion pattern at the province level for the last 44 years, the electrification ratio tends to increase at different rates. As the electrification ratio is calculated from the number of households subscribing to PLN, the predicted connections could only increase in this research (Figure 6.5).

Figure 6.5 Predicted Electrification Ratio All Provinces, 1975–2018



Source: Authors' estimation

Our results show that Jakarta, Central Java, Yogyakarta, East Java and Bali reached the minimum of 50 per cent electrification around the mid-1990s, while Papua, East Nusatenggara, Central Kalimantan, Maluku and Jambi reached that threshold in 2010 or later. This is a similar pattern to the first provinces, which achieved the minimum 25 per cent access around the mid-1980s, but some provinces seem to have been very slow and found it hard to catch up with others, namely Bengkulu and Jambi in the early 1990s, Central Kalimantan and East Nusatenggara around the mid-1990s and Papua in 2000. Details of the predicted time to reach the minimum 50 per cent and 25 per cent electrification ratios in each province are presented in Table 6.8 and Table 6.9.

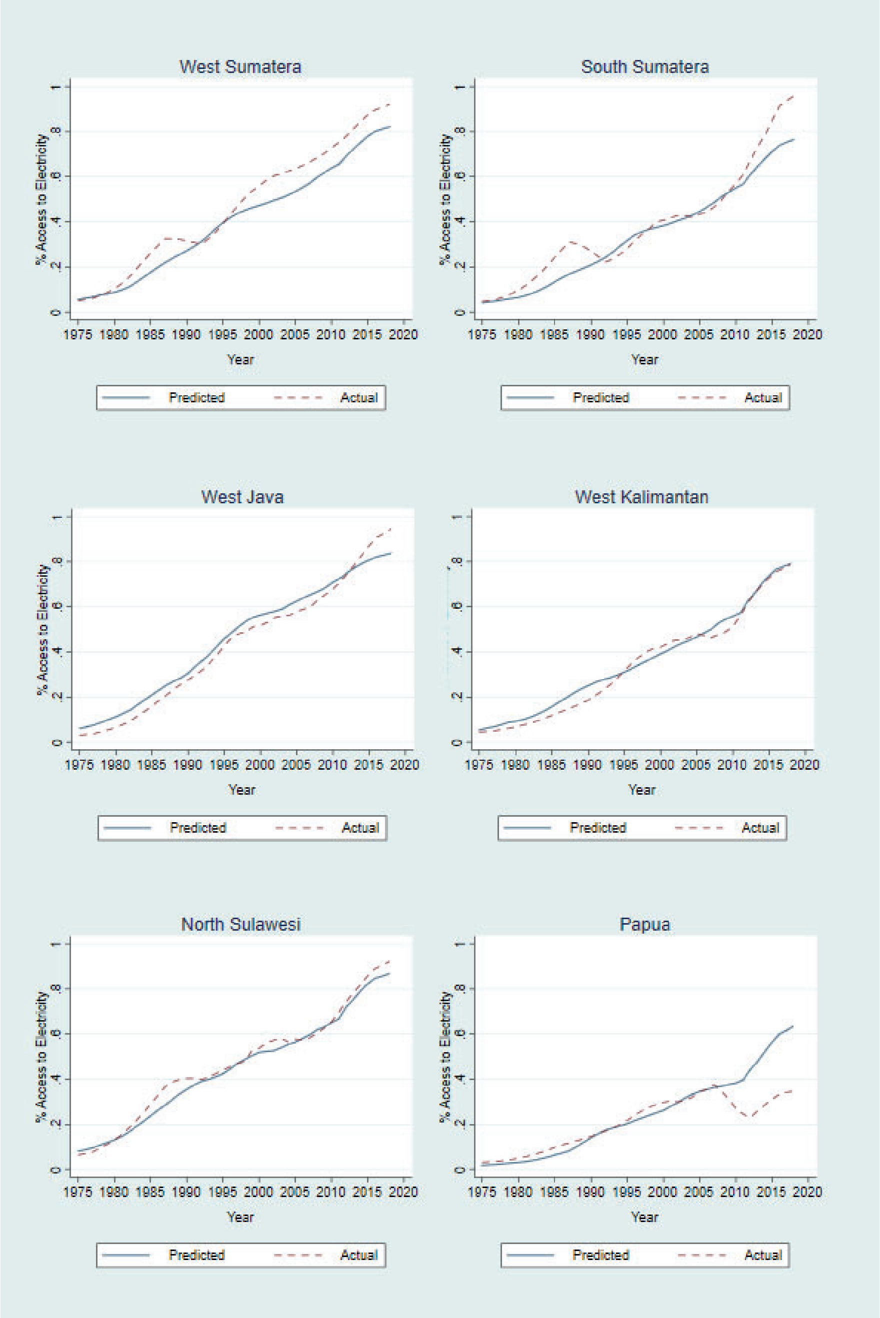
However, if we compare the predicted with the actual patterns, which show some spikes in particular provinces and years, Figure 6.6 reveals the problem in our data. First, after decentralization, several provinces opted to split from their parent provinces. Those new provinces are Bangka Belitung islands splitting from South Sumatera, Banten from West Java, Gorontalo from North Sulawesi and West Papua from Papua.[44] This change, however, cannot be automatically responded to by PLN because PLN has its own regional system to record its customers. As a consequence, some fluctuations occurred in the transition period.

Second, the PLN regional coverage crossed several provinces, suggesting that we should apply a pragmatic strategy to calculate individual provinces, as discussed previously in section 6.2. For instance, West Sumatera and South Sumatera show that their actual patterns have been slightly higher than the predicted ones since the early years. This gap might be due to the PLN regional classification, which recorded West Sumatera as Region III, which also included Riau and Riau island provinces, and South Sumatera as Region IV, which covered South Sumatera, Bangka Belitung, Jambi, Lampung and Bengkulu provinces. Consequently, the actual electrification

[44] The actual electrification ratio of Papua dropped in 2013 due to the termination of the rural electrification programme for Eastern Indonesia. This programme was initiated by the Ministry of Energy and Mineral Resources (MEMR) and collaborated with banks. The households who received a connection had to pay their monthly installment collectively or directly to the cooperatives, then the cooperatives channelled this instalment to the banks. However, this mechanism was not sustainable because many households could not pay regularly; hence, the banks were threatened by increasing non-performing loans. An increase in access to electricity from 2014 onwards is due to the national programme of acceleration electrification (fast-track programme).

ratio of West Sumatera and South Sumatera could be slightly overvalued. Another possible explanation is the model specification, which could not capture unobservable variables, such as the centre of activities during colonialization, development stages, urbanization and so on due to data limitations, leading our results to tend to be underestimated.

Figure 6.6 Predicted and Actual Diffusion of Electricity Pattern for Selected Provinces, 1975–2018



Source: Authors' calculation and estimation

6.6 Drivers of Diffusion Patterns

This section discusses how fast the provinces increase their electrification ratio if we conduct a simulation using Jakarta's characteristics, that is, the power supply, per capita GDP, landscape and population density. We apply two scenarios: (i) at least 50 per cent and (ii) at least 25 per cent access to electricity for all provinces. At the island level, we calculate the average time across provinces within islands to reach those thresholds. For instance, if provinces in the Eastern islands have all of Jakarta's characteristics, they would reach the threshold of 50 per cent access to electricity more quickly. On average, they reach it 26 years faster than the provinces on other islands. Among other factors, the supply availability as in Jakarta contributes 45 per cent followed by the population density, per capita GDP and landscape of Jakarta, with the corresponding figures of 34 per cent, 12 per cent and 9 per cent, respectively, to accelerate the access in the Eastern islands (Table 6.7).

The same pattern is also apparent for the second simulation with at least 25 per cent electrification. This suggests that the availability of the power supply varies across islands and represents an island barrier in slowing down the diffusion process, especially in the Eastern islands. The slow diffusion of electricity in the Eastern islands is in line with the findings of Comin et al. (2012). They concluded that technology diffuses more slowly to locations that are farther away from the adoption leaders. As we can see from the map (Figure 6.1), the Eastern islands are much farther away from the Java islands as the adoption leader.

Sinaga et al. (2019) identified five key barriers to electrifying Eastern Indonesia, namely: (i) difficulties in reaching the location and its geographical conditions, (ii) poor inter-sectoral coordination, (iii) a lack of government funding, (iv) difficulties in land acquisition and (v) inadequacy in equipment, material and human resources. Moreover, they found that the lack of road and bridge infrastructure, the long time required to obtain permits and social resistance from the community exacerbate the existing initial barriers in the system. To overcome these barriers, inter-sectoral coordination among ministries in association with the local government and the community leaders is required.

Other islands show a different pattern whereby the population density plays a more important role in reaching at least 25 per cent electricity access compared with the first threshold of at least a 50 per cent electrification ratio. It seems that the population density becomes the main consideration to electrify provinces with a very low electrification ratio.

Table 6.7 Average Time Required and Its Decomposition to Accelerate 50 Per Cent and 25 Per Cent Electricity Access by Islands

	Average Time Acceleration (Years)				Decomposition of Average Time Acceleration			
	Sim4	*Sim3*	*Sim2*	*Sim1*	*Pop. Density JKT*	*Landscape JKT*	*GDP/ Cap. JKT*	*Supply JKT*
At Least 50% Access								
Sumatera	18	10	9	6	47%	7%	16%	31%
Java Bali	7	3	2	0	64%	11%	25%	0%
Kalimantan	19	8	6	5	54%	12%	7%	26%
Sulawesi	21	13	11	9	38%	9%	9%	44%
Eastern	26	17	15	12	34%	9%	12%	45%
At Least 25% Access								
Sumatera	11	4	3	2	61%	9%	13%	17%
Java Bali	5	2	2	0	65%	4%	31%	0%
Kalimantan	13	4	3	2	71%	4%	8%	18%
Sulawesi	12	5	5	4	60%	0%	9%	32%
Eastern	18	11	10	8	42%	5%	8%	44%

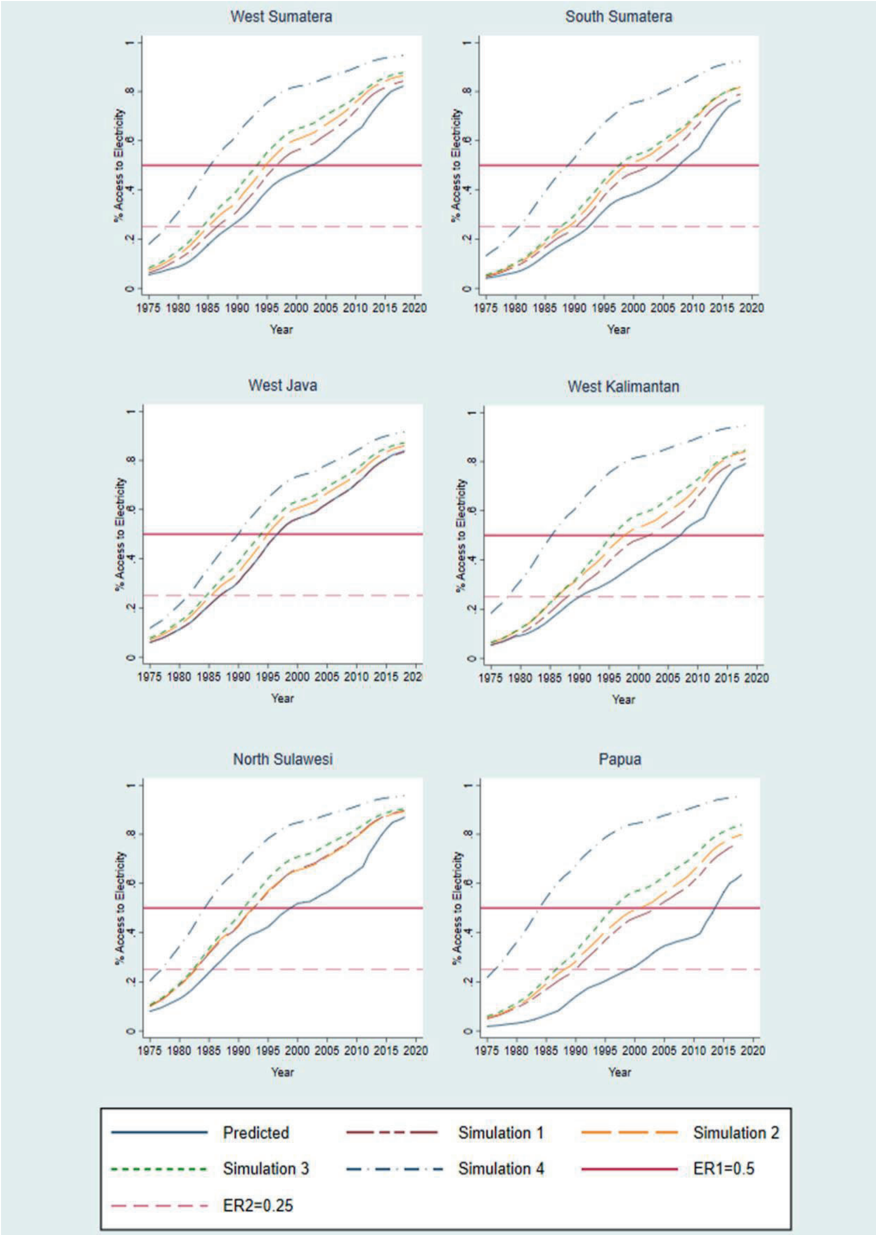
Note: Authors' calculation from Table 6.8 and Table 6.9.
Sim1 (with supply of Jakarta), Sim2 (with supply and per capita GDP of Jakarta), Sim3 (with supply, per capita GDP and landscape of Jakarta) and Sim4 (with supply, per capita GDP, landscape and population density of Jakarta).

At the province level, Table 6.8 and Table 6.9 present two simulation results as well as the factor decomposition in contributing to the gap. Figure 6.7 describes the gap from each simulation for selected provinces representing each island. For instance, Papua, located in the Eastern islands, shows the largest gap in Simulation 4 (using all Jakarta's characteristics) in reaching the thresholds of all the provinces. It would reach the minimum 50 per cent of electricity access 30 years faster if it had all Jakarta's characteristics. If we simulate Papua using only the power supply of Jakarta (Simulation 1), the power supply and per capita GDP of Jakarta (Simulation 2) or the power supply, per capita GDP and landscape of Jakarta (Simulation 3), then Papua will take 10, 13 and 17 years faster to reach the threshold of having at least half of its households electrified (Table 6.8). For the second scenario, Papua will take 9, 11, 13 and 23 years faster if we apply Simulations 1, 2, 3 and 4, respectively, to reach at least 25 per cent access (Table 6.9). In contrast, West Java, as part of Java Island, where the power supply system is the same as that of Jakarta and which has been electrified since 1900, shows the smallest gap in our simulation (Figure 6.7). To reach at least 50 per cent electrification, West Java will take 7, 3 and 2 years faster if we apply Simulation 4, Simulation 3 and Simulation 2, respectively. The population density contributes 57 per cent to speeding up the accessibility (Table 6.8).

From these simulations, it appears that the population density as an internal factor and the power supply availability as an external factor mainly contribute to speeding up the electrification ratio in most provinces. A different pattern is shown in the provinces located in the Eastern islands, where the availability of the power supply as in Jakarta becomes the first factor to speed up the access to electricity to at least 50 per cent. Provinces in the eastern part of Indonesia, that is, Sulawesi (except North Sulawesi), Nusa Tenggara, Maluku and Papua islands, will reach the minimum 50 per cent electrification 10 to 14 years faster if they are supplied as Jakarta's capacity. However, this is not the case when the minimum threshold is 25 per cent electrification, whereby those provinces will take a shorter time of less than 10 years. These results suggest that the availability of the power supply could be translated into an island barrier due to an autonomous self-contained electricity system, and the availability of the supply is more important when the population density has reached a sufficient number to access electricity.

Moreover, our findings imply that the heterogeneity at the province level contributes to shaping the diffusion pattern. Connecting the last unserved populations is more challenging than past electrification efforts, especially as many of them live in remote areas where there is no transmission and distribution line. The investment cost per customer in this typical area outside Java island is very expensive up to $10,000 while it costs about $700 and $100 in non-remote areas outside Java and in Java island, respectively (PLN 2017). As of 31 December 2019, about 79,041 villages or 94.2 per cent of the total villages, including those in remote areas, have been electrified through the Rural Electrification Programme. The remaining non-electrified villages consist of 325 in Papua, 102 in West Papua, 5 in East Nusatenggara and 1 in Maluku (Ministry of Energy and Mineral Resources 2020). For small islands in developing countries, Surroop et al. (2018) found that the inability to extend grids, high upfront costs for rural electrification, inadequate policy mechanism, limited knowledge base and continued high dependence on energy imports are the main problems in energy access.

Figure 6.7 Simulations to Reach Electrification Ratios of at Least 50 Per Cent and 25 Per Cent



Note: Simulation 1 (with the supply of Jakarta), Simulation 2 (with the supply and per capita GDP of Jakarta), Simulation 3 (with the supply, per capita GDP and landscape of Jakarta) and Simulation 4 (with the supply, per capita GDP, landscape and population density of Jakarta).

Table 6.8 Simulation 1: At least 50 Per Cent access to Electricity

Province	Year When at Least 50 Per Cent Access to Electricity					Changes between Predicted and Simulations (Years)				Decomposition of Percentage of Year Changes				
	Predicted	Sim1	Sim2	Sim3	Sim4	Sim4	Sim3	Sim2	Sim1	*All JKT*	*Pop. Density JKT*	*Landscape JKT*	*Per Capita GDP JKT*	*Supply JKT*
Sumatera Island														
Aceh	1999	1995	1994	1994	1985	14	5	5	4	100%	64%	0%	7%	29%
North Sumatera	2008	2001	1996	1994	1987	21	14	12	7	100%	33%	10%	24%	33%
West Sumatera	2003	1997	1995	1994	1986	17	9	8	6	100%	47%	6%	12%	35%
Riau	2006	1999	1998	1996	1987	19	10	8	7	100%	47%	11%	5%	37%
Jambi	2010	2006	2002	1999	1989	21	11	8	4	100%	48%	14%	19%	19%
South Sumatera	2008	2003	1999	1998	1989	19	10	9	5	100%	47%	5%	21%	26%
Bengkulu	2007	2001	1997	1996	1987	20	11	10	6	100%	45%	5%	20%	30%
Lampung	2003	1997	1995	1994	1987	16	9	8	6	100%	44%	6%	13%	38%
Java-Bali Island														
West Java	1997	1997	1995	1994	1990	7	3	2	0	100%	57%	14%	29%	0%
Central Java	1996	1996	1994	1993	1989	7	3	2	0	100%	57%	14%	29%	0%
Yogyakarta	1996	1996	1994	1993	1989	7	3	2	0	100%	57%	14%	29%	0%
East Java	1996	1996	1994	1994	1989	7	2	2	0	100%	71%	0%	29%	0%
Bali	1995	1995	1994	1993	1987	8	2	1	0	100%	75%	13%	13%	0%
Kalimantan Island														
West Kalimantan	2007	2002	1998	1996	1986	21	11	9	5	100%	48%	10%	19%	24%
Central Kalimantan	2012	2008	2007	2002	1988	24	10	5	4	100%	58%	21%	4%	17%
South Kalimantan	2001	1996	1995	1994	1986	15	7	6	5	100%	53%	7%	7%	33%
East Kalimantan	2006	1999	1999	1997	1985	21	9	7	7	100%	57%	10%	0%	33%
Sulawesi Island														
North Sulawesi	1999	1993	1993	1991	1985	14	8	6	6	100%	43%	14%	0%	43%
Central Sulawesi	2009	1999	1996	1994	1984	25	15	13	10	100%	40%	8%	12%	40%
South Sulawesi	2009	1998	1996	1994	1987	22	15	13	11	100%	32%	9%	9%	50%
Southeast Sulawesi	2008	1998	1995	1993	1984	24	15	13	10	100%	38%	8%	13%	42%
Eastern Islands														
West Nusatenggara	2005	1995	1993	1991	1985	20	14	12	10	100%	30%	10%	10%	50%
East Nusatenggara	2012	1998	1995	1993	1984	28	19	17	14	100%	32%	7%	11%	50%
Maluku	2011	1998	1993	1992	1984	27	19	18	13	100%	30%	4%	19%	48%
Papua	2014	2004	2001	1997	1984	30	17	13	10	100%	43%	13%	10%	33%

Note: Sim1 (with Supply Jakarta), Sim2 (with Supply and Per Capita GDP Jakarta), Sim3 (with Supply, Per Capita GDP, Landscape Jakarta), Sim3 (with Supply, Per Capita GDP, Landscape, Population Density Jakarta)

Table 6.9 Simulation 2: At least 25 Per Cent Access to Electricity

Province	Year When at Least 25 Per Cent Access to Electricity					Changes between Predicted and Simulations (Years)				Decomposition of Percentage of Year Changes				
	Predicted	Sim1	Sim2	Sim3	Sim4	Sim4	Sim3	Sim2	Sim1	*All JKT*	*Pop. Density JKT*	*Landscape JKT*	*Per Capita GDP JKT*	*Supply JKT*
Sumatera Island														
Aceh	1986	1984	1984	1984	1978	8	2	2	2	100%	75%	0%	0%	25%
North Sumatera	1992	1990	1986	1985	1979	13	7	6	2	100%	46%	8%	31%	15%
West Sumatera	1989	1987	1986	1985	1978	11	4	3	2	100%	64%	9%	9%	18%
Riau	1989	1987	1987	1987	1979	10	2	2	2	100%	80%	0%	0%	20%
Jambi	1994	1993	1991	1989	1981	13	5	3	1	100%	62%	15%	15%	8%
South Sumatera	1993	1991	1989	1988	1981	12	5	4	2	100%	58%	8%	17%	17%
Bengkulu	1993	1991	1988	1986	1979	14	7	5	2	100%	50%	14%	21%	14%
Lampung	1988	1986	1986	1985	1980	8	3	2	2	100%	63%	13%	0%	25%
Java-Bali Island														
West Java	1987	1987	1986	1985	1982	5	2	1	0	100%	60%	20%	20%	0%
Central Java	1986	1986	1984	1984	1981	5	2	2	0	100%	60%	0%	40%	0%
Yogyakarta	1986	1986	1984	1984	1981	5	2	2	0	100%	60%	0%	40%	0%
East Java	1986	1986	1985	1985	1981	5	1	1	0	100%	80%	0%	20%	0%
Bali	1986	1986	1984	1984	1980	6	2	2	0	100%	67%	0%	33%	0%
Kalimantan Island														
West Kalimantan	1990	1988	1987	1987	1978	12	3	3	2	100%	75%	0%	8%	17%
Central Kalimantan	1996	1993	1991	1990	1980	16	6	5	3	100%	63%	6%	13%	19%
South Kalimantan	1988	1986	1985	1985	1978	10	3	3	2	100%	70%	0%	10%	20%
East Kalimantan	1990	1988	1988	1987	1977	13	3	2	2	100%	77%	8%	0%	15%
Sulawesi Island														
North Sulawesi	1986	1983	1983	1983	1977	9	3	3	3	100%	67%	0%	0%	33%
Central Sulawesi	1990	1986	1985	1985	1977	13	5	5	4	100%	62%	0%	8%	31%
South Sulawesi	1991	1987	1985	1985	1979	12	6	6	4	100%	50%	0%	17%	33%
Southeast Sulawesi	1989	1985	1984	1984	1976	13	5	5	4	100%	62%	0%	8%	31%
Eastern Islands														
West Nusatenggara	1992	1985	1983	1983	1977	15	9	9	7	100%	40%	0%	13%	47%
East Nusatenggara	1996	1987	1985	1984	1977	19	12	11	9	100%	37%	5%	11%	47%
Maluku	1992	1985	1985	1984	1976	16	8	7	7	100%	50%	6%	0%	44%
Papua	2000	1991	1989	1987	1977	23	13	11	9	100%	43%	9%	9%	39%

Note: Sim1 (with Supply Jakarta), Sim2 (with Supply and Per Capita GDP Jakarta), Sim3 (with Supply, Per Capita GDP, Landscape Jakarta), Sim3 (with Supply, Per Capita GDP, Landscape, Population Density Jakarta)

6.7 Conclusion and Policy Implications

The disparity in access to electricity and its supply is high across Indonesia. Despite the fact that electrification started in the early 1900s and is today nearly 100% in paet of country, more than one hundred years later, the electrification ratio in some regions in Indonesia is still low and the speed of electricity technology diffusion slow. What drives this inequality in diffusion of access to electricity across time and space? Of course, Indonesia's particular geography shaped by its archipelagic nature implies that the country needs huge investments to build new lines to bring electricity to people on the different islands, making the up-front costs of electrification especially very high. Unlike the situation in non-archipelagic countries, new connections across space can often not be provided by extending existing lines – the different islands are in fact autonomous self-contained electricity systems. How important is this geographic feature in explaining spatial variation in the speed of electrification over time?

To answer this question, we first had to construct a dataset with information on electricity access a across time and space on the basis of unpublished statistics kept by PLN, Indonesia's national electricity utility company. We complemented our dataset with information on per capita GDP, population density, firm density, landscape complexity (measured as the share of households living in flat areas) and per capita installed electricity generation capacity across provinces and islands. Using this rich dataset, we then developed a non-linear least square estimation to estimate the speed of the diffusion process within an S-curve framework of access to electricity. To this aim, we regress the electrification ratio in various provinces over time on a set of explanatory variables. Next, we translated the non-linear least square regression outcomes into the β parameter that defines the speed of diffusion in the S-curve framework. By plugging the estimated β_i into our sample, we obtain the predicted electrification ratios for each province in terms of an S-curve over time. Finally, we calculate for each province the gap between its own electrification performance and that of the leading island (Java-Bali) and use our regression outcomes to decompose this gap into the respective contribution of the driving forces of electrification mentioned above, which differ across space.

Our main findings are as follows. The relatively fast electricity diffusion process in the provinces located at the Java–Bali islands is mainly driven by their relatively high population density as well as the power supply availability. In contrast, the provinces on the Eastern islands have difficulties in catching up to the electrification levels in other provinces. We find that lack of power supply facilities is the main factor that explains the delay in electrification at the Eastern while the population density plays an important role to explain slower electrification patterns at Sumatera, Kalimantan and Sulawesi islands. Since the spatial variation in power supply availability is of course shaped by the geographical constraints inherent to Indonesia's island structure, our results imply that spatial heterogeneity at the province level contributes substantially to shaping the diffusion pattern.

These findings suggest that, in order to speed up the electrification ratio the Indonesian government may need to consider different policy instruments to stimulate the electricity diffusion across the islands that are underserved sofar. For instance one can think of a fiscal incentive that stimulates the use of small stand-alone generators in remote areas or on small islands in the Eastern islands, strengthening the transmission line infrastructure between large islands, namely Sumatera, Kalimantan and Sulawesi, and regulations to improve the quality of the electricity on Java–Bali islands or in areas where access is much higher than on other islands.

Understanding the role of regional interconnections across provinces is a first prerequisite to improve the accessibility of electricity. We suggest that future research in this area for Indonesia focusses on the question whether people with an electricity connection can actually afford to use the energy that they really need, how they benefit from the electricity and so on. Further, it remains interesting to modify and implement our model at the district or the village level to understand the variation in the diffusion pattern at the lower level of aggregation as the second avenue for future research.

Appendix A6 – Constructing the Electricity Dataset

A.6.1. Dynamics of PLN's Business Units

In this section, we will discuss the practicalities and problems encountered in constructing electricity dataset from PLN statistics. As a state company and main electricity provider served customers throughout Indonesia, the structure of PLN's organization is determined by the Ministry of State Owned Enterprises (MSOE), and the business units are established accordingly. Prior to 2008, PLN did not have any regional directors, only a functional management. As a consequence, all the strategic decisions concerning the regions should be taken at the PLN headquarters in Jakarta. This of course affected the decision-making process. From 2008 onwards, the ministry added regional directors so that regional problems could be settled faster by the regional directors without waiting too long for decisions to be made at the headquarters in Jakarta.[45] Given changing in organization structure and the needs to accommodate the electricity development, PLN frequently changes its business units that affect our data collection processes.

Prior to 1981/1982, PLN had one *Pembangkit* (generation unit), two *Kantor Distribusi* (the distribution offices) in Jakarta and West Java and thirteen *Kantor Wilayah* (the regional offices), namely (1) Aceh, (2) North Sumatera, (3) West Sumatera, Riau and Riau Island, (4) South Sumatera, Jambi, Lampung, Bengkulu and Bangka Belitung, (5) West Kalimantan, (6) Central Kalimantan, South Kalimantan and East Kalimantan, (7) North Sulawesi, Central Sulawesi and Gorontalo, (8) South Sulawesi, Southeast Sulawesi and West Sulawesi, (9) Maluku and North Maluku, (10) Papua and West Papua, (11) Bali, West Nusatenggara and East Nusatenggara, (12) East Java and (13) Central Java. The last two regional offices have been altered to the distribution offices of East Java and Central Java since 1982/1983, so the number of PLN business units became 11 *Kantor Wilayah* and 4 *Kantor Distribusi* until 2001.

[45] The number of regional directors was amended based on PLN's needs. For example, two regions distinguished from 2008 until 2015 expanded into seven regions from 2016 to 2019 and were then merged back into the current four regions.

In addition, since 1982/1983, two generation units in Eastern and Western Java have been established, and since 1985 they have been integrated with their transmission units as *Pembangkitan dan Penyaluran (Kitlur)*: generation and transmission in Eastern and Western Java. Moreover, PLN established two subsidiary companies, *PT Pembangkit Jawa Bali* (*PT PJBI* and *PT PJB II*), focusing on power generation in Java and Bali, in 1995. Since then, those generation units have been handed over as part of subsidiary companies while the transmission lines have been managed separately by *Pusat Pengatur dan Penyaluran Beban* (P3B) or the Transmission Load and Dispatch Centre Java Bali. Following Java–Bali interconnected system, PLN has operated this system in Sumatera since 2014 and built this system in Sulawesi and Kalimantan since 2018.

PLN split its regional offices by establishing six new *Kantor Wilayah* and one *Kantor Distribusi* in 2002, namely the *Kantor Wilayah* Riau, Bangka Belitung, Lampung, East Kalimantan, East Nusatenggara and West Nusatenggara and the new *Kantor Distribusi* Bali. Furthermore, since 2012, *Kantor Wilayah* Lampung has been altered to *Kantor Distribusi*, and, since 2016, the Banten branch has split from the *Kantor Distribusi* West Java and established an independent *Kantor Distribusi* Banten; thus PLN currently has 15 *Kantor Wilayah* and 7 *Kantor Distribusi*.

Table A6.1 describes the dynamics of PLN's business units, which have been split into specialized and independent distribution, transmission and production units. The latter refers to the installed capacity that we use in this chapter. We apply the same names of *Kantor Wilayah* and *Kantor Distribusi* in both the connection and the installed capacity dataset. Table A6.1 also shows how the electricity data were recorded and stored in several PLN units. Therefore, the consistency in the annual PLN statistics needs to be taken into consideration.

Further, PLN has established several new development units (IUPs) to respond to President Jokowi's programme of 35,000 MW, which was launched in 2015. After the recovery from the Asian Financial Crisis, the former president of Indonesia Susilo Bambang Yudhoyono (SBY) initiated two fast-track programmes (FTPs) aiming to accelerate the development of electricity; specifically, FTP I 10,000 MW commenced in 2006 to substitute diesel or fuel power plants for coal-fired power plants and FTP II 17,400 MW was launched in 2010 to increase the use of renewable power plants.

Almost 100 per cent of FTP I was completed, but only 28 per cent of FTP II materialized, which was then carried over to the 35,000 MW programme. To operate and manage the new completed power plants resulting from the FTPs, PLN formed *Unit Pembangkitan Jawa Bali* (UPJB) or generation unit Java–Bali in 2012 (Table A6.1). This unit was also responsible for handing over these power plants to PLN subsidiary companies (PT PJB and PT IP) in 2016, and it was dissolved afterwards.

Table A6.1 Summary of Business Units Related to Customers and Production 1975–2018[*]

Period	Number of Business Units Related to Customers (Connection)	Number of Business Units Related to Production (Installed Capacity)
1975/1976–1981/1982	13 regions 2 distributions	13 regions, 2 distributions, 1 generation (K)
1982/1983	11 regions 4 distributions	11 regions, 4 distributions, 2 generations (K_I and K_{II})
1989/1990–1990/1991	11 regions 4 distributions	11 regions, 4 distributions New units: generation of Eastern and Western Java
1992–2001	11 regions 4 distributions Batam (as a special region in 1993 and as a subsidiary company, PT PLN Batam, in 2000)	11 regions, 4 distributions, generation of Eastern and Western Java New units: • Batam (as a special region in 1993 and as a subsidiary company, PT PLN Batam, in 2000) • Transmission Load and Dispatch Centre Java Bali or *Pusat Pengatur dan Penyaluran Beban* (P3B) to expand the former load dispatch centre (P2B) • Subsidiary generation company: o *PT Pembangkit Jawa Bali I* in 1995, which then changed to *PT Indonesia Power* (PT IP) in 2000 o *PT Pembangkit Jawa Bali II* in 1995, which then changed to *PT Pembangkit Jawa Bali* (PT PJB) in 2000 • Generation of Northern and Southern Sumatera (1997)
2002–2011	16 regions 5 distributions PT PLN Batam	16 regions, 5 distributions, PT PLN Batam, PT IP, PT PJB, generation of Northern and Southern Sumatera New units: • Subsidiary company PT PLN Tarakan (2004) • Power plants (Muara Tawar in 2005, Tanjung in Jati B 2006, Lontar Unit 1 in 2010, Indramayu Unit 1 in 2011)

Continued

2012–2013	15 regions 6 distributions PT PLN Batam PT PLN Tarakan	15 regions, 6 distributions, PT PLN Batam, PT IP, PT PJB, generation of Northern and Southern Sumatera, PT PLN Tarakan, Muara Tawar, Tajung Jati B, Indramanyu Unit 1, Lontar Unit 1 New units: • *Unit Pembangkitan Jawa Bali* (UPJB) or the generation unit of Jawa–Bali focused on managing and operating power plants under Fast Track Program (FTP) I, namely Indramayu Unit 2 and Unit 3 (2012) and Lontar Unit 2 (2013)
2014–2015	15 regions 6 distributions PT PLN Batam PT PLN Tarakan	15 regions, 6 distributions, PT PLN Batam, PT IP, PT PJB, generation Northern and Southern Sumatera, PT PLN Tarakan, Tanjung Jati B, UPJB (Indramayu and Lontar) New units: • *Pusat Pengatur dan Penyaluran Beban (P3B)* or Transmission Load Dispatch Centre Sumatera • Development Units (UIP) Eastern Kalimantan, Nusatenggara, Eastern Java and Bali (2014), Sumatera and Centre of Kalimantan (2015)
2016–2018	15 regions 7 distributions PT PLN Batam	15 regions, 7 distributions, PT PLN Batam, PT IP, PT PJB, generation in Northern and Southern Sumatera, Tanjung Jati B, P3B Sumatera, P2B Jawa–Bali, IUP (Eastern Java–Bali, Sumatera, Centre of Kalimantan, Eastern Kalimantan). PT PLN Tarakan was closed and merged back with Region East Kalimantan in 2017 New units: • Generation of Western Java, Centre of Java, Eastern Java and Bali (2016) and Nusatenggara (2018) • Development Unit (IUP) Southern Sulawesi (2017), Western Kalimantan, and Papua and West Papua (2018) • *Unit Induk Pembangkitan dan Penyaluran* (UIKL) or Unit of Generation and Transmission Kalimantan and Sulawesi (2018)

Notes: * We use the terms generation, region and distribution to refer to PLN's business unit names in Bahasa, that is, *Pembangkit*, *Kantor Wilayah* and *Kantor Distribusi*, respectively.

After reclassifying PLN business units to obtain customers and production data, we recode the PLN's area into BPS province as described in Table A6.2. In response to PLN's needs and electricity development, PLN established new business units in some provinces. In those cases, we reclassify new PLN business units back to the parent regions or provinces, as we did for PLN Batam and PT PLN Tarakan, to make our dataset consistent at the province level overtime. PLN Batam was established in 1993, when Batam was treated as a special economic zone and turned into a subsidiary company, PT PLN Batam, in 2000. On the other hand, Riau Island province, where PLN Batam is located, was established in 2002 after splitting from Riau province.

Thus, we add PLN Batam to Riau province from 1993 to 2001, then, after 2002, we add PLN Batam to Riau Island province.

Table A6.2 Coverage Area of PLN Business Units, BPS Provinces and Islands

PLN Business Unit	BPS Province	Island
Region I (Banda Aceh)	Nanggroe Aceh Darussalam	Sumatera
Region II (Medan)	North Sumatera	Sumatera
Region III (Padang)*	West Sumatera, Riau, Riau Island	Sumatera
Region IV (Palembang)*	South Sumatera, Bangka Belitung, Jambi, Lampung, Bengkulu	Sumatera
Region V (Pontianak)	West Kalimantan	Kalimantan
Region VI (Bnjarmasin)*	Central Kalimantan, South Kalimantan, East Kalimantan	Kalimantan
Region VII (Manado)	North Sulawesi, Central Sulawesi, Gorontalo	Sulawesi
Region VIII (Makassar)	South Sulawesi, Southeast Sulawesi, West Sulawesi	Sulawesi
Region IX (Ambon)	Maluku, North Maluku	Eastern
Region X (Jayapura)	Papua, West Papua	Eastern
Region XI (Denpasar)*	Bali and West Nusa Tenggara, East Nusatenggara	Java-Bali, Eastern
PT PLN Batam	Riau island	Sumatera
Dist. East Java (Surabaya)	East Java	Java-Bali
Dist. Central Java (Semarang)	Central Java, Jogjakarta	Java-Bali
Dist. W. Java and Banten (Bandung)*	West Java, Banten	Java-Bali
Dist. Jaya and Tangerang (Jakarta)	Jakarta, Banten (*Kota* Tangerang, *Kabupaten* Tangerang, *Kota* South Tangerang)	Java-Bali

Parentheses refer to the capital city where the region or distribution office is located.
* Riau, Bangka Belitung, Lampung, East Kalimantan, West Nusatenggara, and East Nusatenggara have been established as independent regional office (*Kantor Wilayah*) and Bali as a distribution office (*Kantor Distribusi*) since 2002. Lampung's regional office was changed to Lampung's distribution office in 2012, while Banten's branch was established as an independent distribution office in 2016.

Moreover, East Kalimantan's region was established in 2002 after splitting from Region VI while PT PLN Tarakan, as a subsidiary company located in East Kalimantan province, was established in 2004. Therefore, we combine PT PLN

Tarakan with East Kalimantan province from 2004 to 2016, then we add the new province North Kalimantan back to its parent province East Kalimantan after 2014. Other regional offices or distribution offices are coded as they are as long as they represent the province where these business units are located. For example, Lampung's regional unit, which changed into a distribution unit in 2012, is still coded as Lampung province in our dataset. Through this procedure, our dataset becomes more consistent with the BPS's classification at the province level overtime.

A.6.2 Area Coverage in Kantor Distribusi Jakarta, West Java and Banten

Prior to 2015, the Banten's branch was part of *Kantor Distribusi Jawa Barat dan Banten* or the distribution office of DJBB. Due to an increasing demand for electricity, especially from large-scale industrial customers, and the growing development of Banten province, PLN separated the Banten's branch from DJBB in 2015 and the Banten's branch became an independent distribution office (*Kantor Distribusi Banten*). As a consequence, the Tangerang's branch, which was originally part of *Kantor Distribusi Jakarta Raya dan Tangerang* or the distribution office of Disjaya, should be recorded not as Disjaya but as *Kantor Distribusi Banten* by reclassifying it at the branch level and recalculating the number of connections and the capacity of these three distribution offices of Jakarta, West Java and Banten.

Without this procedure, our data would be misleading. For instance, when we calculate access to electricity as a percentage of electrified households in the total households, the figure is much higher for Jakarta province and much lower for Banten province due to different definitions of Tangerang's coverage between BPS and PLN. The total number of households as the denominator is calculated by the BPS according to the administrative definition of the Ministry of Home Affairs; on the other hand, the number of electrified households as the nominator is recorded by PLN based on the technical area of PLN, and the discrepancy due to this area definition is significant. Below is a detailed description of the coverage area of the BPS and PLN.

Administrative area of the BPS

Our dataset at the province level following the BPS coding is in line with the definition of the Ministry of Home Affairs: Jakarta province, West Java province and Banten province. The latter was established in 1999 after splitting from West Java province. Hence, prior to 1999, Banten was part of West Java province. Banten province comprises four cities (*kota*) and four districts (*kabupaten*), and three out of those eight autonomous regions are neighbours of Jakarta province, namely *Kota* Tangerang, *Kabupaten* Tangerang and *Kota* South Tangerang.

Technical area of PLN

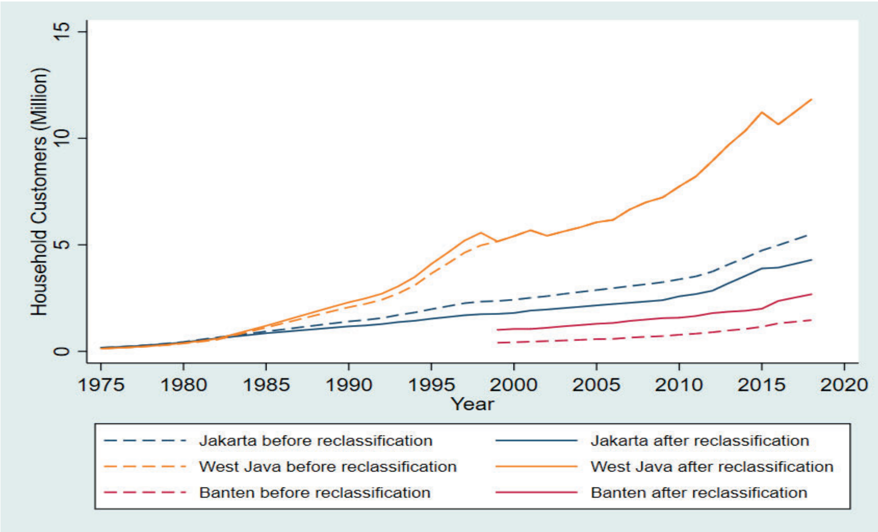
The lowest level in the distribution office is the branch, which is formed on the base of technical aspects and is likely to cross the administrative area of the BPS. Tangerang branch was part of *Kantor Distribusi Jakarta Raya dan Tangerang*, while Tangerang branch itself covered four administrative areas of the BPS, namely (i) *Kota* Tangerang (sub-branch Cikokol), (ii) *Kota* South Tangerang (sub-branch Serpong), (iii) *Kabupaten* Tangerang (sub-branch Sepatan, Teluk Naga, Curug, Cisoka and Cikupa) and (iv) *Kota* West Jakarta (sub-branch Cengkareng and Kalideres). It is obvious that Tangerang branch crossed two provinces i.e. Banten and Jakarta.

The main implication of those different area definitions is that the electrification ratio could be overestimated for Jakarta province but underestimated for West Java province for the period 1975–1998 and for Banten province for the period 1999–2018. Due to this problem, we reclassify our household customer dataset at the branch level. We unbundle the Tangerang branch from Jakarta province and then tally it up with West Java province for the period 1975–1998 and with Banten province for the period 1999–2018. For incomplete years, we interpolate using the moving average. Hence, our new dataset for Jakarta, West Java and Banten provinces is now more sensible.

Figure A6.1 presents the number of PLN household customers before and after reclassification, as we discussed. The new Jakarta blue line after excluding the Tangerang branch is lower than the old Jakarta blue dashed line, while the new Banten maroon line after including the Tangerang branch is higher than the old Banten maroon dashed line; the new West Java orange line is slightly higher than the old West

Java orange dashed line before 1999 or when Banten was still part of West Java, then, after 1999, the two lines of West Java appear to overlap. Spikes are shown in two transition periods: (i) around 1999, when Banten administration split from West Java province and was established as a new province, and (ii) around 2015, when Banten area split from the distribution office of West Java and Banten (DJBB) and became a new independent distribution office.

Figure A6.1 Number of Electrified Households in Jakarta, West Java and Banten before and after Reclassification, 1975–2018



Source: Calculated by the authors.

In addition, high spikes in the number of customers Jakarta in 2011 was partly due to the programme *Gerakan Sehari Sejuta Sambungan* (GRASS), or 1 million customers per day, launched in 2011, in which 1 house was allowed to have more than 1 meter at that time. As a result, there are 4,040,310 household customers in Jakarta (PLN Statistics 2018) while the number of households issued by the BPS is 3,393,014; hence, if we calculate the electrification ratio for Jakarta, the figure is 119 per cent. This measurement error leads to an incorrect picture of the electrification programme. To overcome this problem, we recalculate the denominator in 2011 by dividing the total population of Jakarta by 3, the average number of household members in Jakarta so the electrification ratio is corrected accordingly.

A.6.3 Tariff Transformation

PLN has recorded number of customers based on tariff classification and changed this classification regularly as electricity develops. Table A6.3 describes the transformation of electricity tariff. We encountered some problems in calculating number of household customers as follows. The very poor households without a meter were sometimes categorized as S_1 (social users) in PLN statistics, so we reclassify them as households after consulting PLN. This specific problem was relevant to the Tangerang branch. Therefore, we sum up the number of customers under tariffs S_1, R_1, R_2, R_3 and R_4 to arrive at the total number of household connections in Tangerang in 1992, 1993 and 1997.

Table A6.3 Tariff Classification 1975–2018

Period	Tariff Classification/Customer Type			
	Household	Industry	Business	Other
Before 1980/1981	A_1 and A_2	B_1 and B_2	C_1 and C_2	D: gov. office E: social F: street lighting *Khusus*: special tariff
1980/1981 –1998	R_1: simple house R_2: small house R_3: medium house R_4: large house	I_1: micro industry I_2: small industry I_3: medium industry I_4: large industry	U_1: small enterprise U_2: medium enterprise U_3: large enterprise H_1: simple and small accommodation H_2: hotel industry	S_1: very small user S_2: social foundation G_1: medium gov. office G_2: large gov. office J : street lighting M: multipurpose (special tariff)
1999–onwards	R_1: simple house R_2: small house R_3: medium/large house	I_1: micro industry I_2: small industry I_3: medium/ large industry	B_1: small business B_2: medium business B_3: large business	S_1: very small user S_2: small social foundation S_3: medium social foundation P_1: medium gov. office P_2: large gov. office P_3: street lighting T : *traksi* (train, MRT) C : *curah* (special tariff) M : multipurpose

Source: PLN statistics

Regarding the payment method, PLN initially applied a post-paid method, but, since 2007, it has introduced a prepaid method to reduce the cost of collection.[46] A mistake occurred in PT PLN Batam that required reclassification before calculating the electrification ratio. PT PLN Batam recorded prepaid household customers as business type customers due to the multipurpose (*Multiguna*) tariffs from 2011 to 2016, but in fact those customers were household customers who merely changed their payment method. Therefore, we subtract prepaid customers from "business" and add them to "household" for the period 2011–2016. In some cases, we also find two different numbers of customers, by tariff as well as by type, in the PLN statistics. If this is the case, we choose the number of customers by type rather than by tariff, which involves more technical aspects.

[46] The number of prepaid customers is continuing to increase and was almost half of the total number of PLN household customers in 2019. However, this prepaid method will no longer be implemented in the future. According to PLN, it is now preparing to replace the existing meter with a smart meter that could reduce the collection fee and implement the post-paid method much more easily.

Chapter 7 Conclusion

In this thesis we developed a series of empirical studies to provide insight into the nature, causes and consequences of inequality dynamics in Indonesia during the past decades. These decades have been characterized by rapid and far-reaching decentralization and democratization processes that strongly impacted the entire society. These institutional reforms have had broad consequences for, amongst others, the level and evolution of inequality across individuals and regions as well as the organization and accessibility of public services like health and electricity.

The different chapters of this thesis implicitly asked the question as to how inclusive Indonesia is today by looking at various dimensions. We analyzed the relationship between inequality and economic development across regions and the evolution of the accessibility of basic public services, i.e. health and electricity. Social inclusiveness is a central notion in the Sustainable Development Goals (SDG), and as such the thesis has been motivated by the pursuit of SDG in Indonesia as a rapidly changing and emerging economy. The archipelagic nature of Indonesia of course make it especially relevant to assess the spatial dimension of inequality and inclusiveness – with the vast territory of Indonesia people are unevenly distributed across five big islands and hundreds of small islands, which makes that national aggregate scores on the various SDGs often mask substantial regional differences. In this concluding chapter we reflect upon the main findings of our analyses and suggest some avenues for future research.

7.1 On Income Inequality and Development

We started this thesis with a descriptive analysis of inequality dynamics in Indonesia over time and over space, distinguishing three levels of spatial aggregation namely (i) the national level (in an international perspective); (ii) the regional level (islands and provinces); (iii) the individual level (income classes). Chapter 2 analyzed at the (cross–)country level and at the province level covers the period 1961–2015, while our analysis for income classes is based on micro data (from Susenas) for the period

1987–2015. This approach allows us to present new insights into the inequality dynamics in Indonesia. The key messages of this chapter are as follows: over the last decades relatively high economic growth in Indonesia is associated with rapidly increasing income inequality. Regional convergence of inequality across islands and provinces is driven by the fact that higher income class in the relatively poor regions grew faster than the lower income class in the relatively rich regions. In other words, the middle class and especially the top incomes seem to benefit most from the economic growth dynamics in Indonesia.

Chapter 3 investigated whether and to what extent the number, size and distribution of cities might explain the evolution of inequality across provinces in Indonesia. In a cross-country analysis Castells-Quintana (2018) found that beyond Kuznets' hypothesis there is a U-shaped relationship between average city size and inequality; inequality first falls and then increases with average city size. A key result of our analysis across provinces within one country is that this U-shaped relationship was only found if we include Jakarta as one bundled city in our sample. In contrast to Castells-Quintana (2018), this result did not hold when we excluded Jakarta or when we considered unbundled Jakarta. However, if we included small cities with less than 300 thousand inhabitants in our sample, and we add urban primacy, we found a U-shaped relationship for the case of small cities but an inverted U-shaped for the case of urban primacy. The N-shaped relationship where inequality first decreases with small city size, then increases with larger city size, and finally declines again with extra large city size, might be associated with the fact that larger cities provide more job opportunities for people with different abilities and skills. The interaction between city size and income is negative for all the specifications, suggesting that part of the association between average city size and inequality could be explained by the association between average city size and economic performance: larger cities are more productive, benefit from more highly skilled workers and usually involve spillover, and hence income increase. The main message drawn from this chapter is that, if the current trend of an increasing average agglomeration size continues, we can expect inequality to climb further. This message implies that larger average agglomeration size may be desirable when cities are small, however, a very high average agglomeration size is undesirable, so the medium city size may be more desirable.

In Chapter 4 we basically examined the relationship between economic growth and inequality across districts in Indonesia. A large empirical literature has been devoted to the impact of inequality on growth but it is very difficult to answer the question as to whether rising inequality in a country is good or bad for its economic growth performance. For a long time, most of the empirical research routinely imposed a linear structure on the data, whereby the different variants of the basic linear model generated very different conclusions on the inequality–growth relationship. Finding inconsistency in the results of the relationship between inequality and growth, Banerjee and Duflo (2003) searched for a non-linear relationship by developing a simple political-economy model to argue that there are no *a priori* theoretical reasons to assume that the relationship between inequality and growth is monotonic. Their results show that growth is inverted U-shaped function of net changes in inequality. Changes in inequality (in any direction) are associated with lower future growth rates. It implies that an economy's growth rate is maximized when there are no changes in inequality and that any deviation of inequality, in other way, lowers growth.

Adopting Banerjee and Duflo's approach, we empirically tested whether actual changes in inequality in the past associate with subsequent economic growth and whether this relationship is non-linear. We also examine the role of institutional quality in the relationship between past inequality and economic growth in the next period. Similar approach is applied to test whether changes in inequality in the past correlates with growth of development indicators in the future. Our results are in line with Banerjee and Duflo (2013). We consistently found an increase in a level of inequality has a significant and positive relationship with subsequent growth in various specifications. However, large changes in inequality in any direction were found to be associated with a large decline in economic growth, as revealed by a hump-shaped relationship between economic growth and the actual change in inequality. These relationships held in all the specifications of the basic and the modified models, which included institutional quality. In this thesis, we are not aiming to search for the causal relationship but the correlation.

Our main messages from this chapter are as follows: high inequality in the past is associated with relatively high subsequent growth, this high growth is then associated with high inequality in the current period. If this current inequality is too high, then the change in inequality will be positive and large enough to reduce growth in the future. In other words, the high inequality in the past is associated with the slowing down the economic growth and worsening the existing inequality. It seems that our results support the Piketty's (2014) argument that inequality keeps increasing as societies accumulate wealth. With regard to the institutional quality, the type of institutional quality plays a significant role in shaping the relationship between past inequality and economic growth in the next period. This role appears more important if we interact institutional quality with the initial inequality, suggesting that a combination of a certain degree of inequality and institutional quality is positively correlated with economic growth.

In addition, our results show that a high initial level of inequality is highly associated with the growth of roads and the poverty reduction.The relationship with the change in inequality shows an inverted U-shaped for the growth of roads and a U-shaped for the poverty rate, suggesting that existing inequality is positive correlated with growth of roads while current inequality is negative correlated with growth of poverty until the this current inequality passes a threshold. If current inequality keeps increasing, it tends to reduce the growth of roads and increase the growth of poverty in the future. These findings indicate that the rising inequality may go at the expense of the accessibility of certain basic public services that we study in Chapters 5 and 6 of this thesis.

7.2 On Accessibility

The second part of this thesis delved deeper into the questions regarding the accessability of basic public services, namely health and electricity. In Chapter 5 we hypothesize that there might be an ex ante moral hazard problem for insured people in the sense that they tend not to care sufficiently of their health situation that causes illness. This behavior may be associated with demand for health treatment, and hence increase the amount of claim to social security agency (BPJS). So, understanding the behavior of insured people is essential to seek the solution for BPJS problem. In this chapter, due to data availability, we only observe individual behavior prior to the universal healthcare program (JKN) era or when health insurance has been provided by the central government since 1998. We assume that the behavior of insured people remains the same over time, so we use the latest household data prior to the JKN era to inform us about the effects of insurance schemes. Our research aims to contribute to our understanding of how people's behavior towards health influences their demand for unhealthy consumption, which has an impact on outpatient and inpatient care and thus correlates with the sustainability and affordability of the BPJS programme. In particular, we try to identify ex ante moral hazard among insured people to help the government to formulate policy recommendations.

From the results in Chapter 5, given the health reform promoting inclusiveness in health access, we can conclude that insured people behave differently than uninsured people, and within the insured group, the subsidized people behave differently from the non-subsidized people. Our findings show that there is ex ante moral hazard whereby people do not prevent health status as they consume more unhealthy food once insurance exists. Insurance program with or without subsidy improves accessibility of health care. Getting older and richer are associated with more use of inpatient treatment for the insured households, but being knowledgeable is more important for the subsidized people to understand the procedure of free healthcare. Further, education turns out to play an important role in improving health status. Since moral hazards could exist on both sides – insured people and providers – the next avenue for future research is to identify the incentives for both parties to control moral hazard effects in health care system. Further, the JKN has been implemented for at least five years, so it remains interesting to test our hypothesis among JKN members to determine whether they show ex ante moral hazard. These future empirical research avenues could generate policy implications to improve universal health care in Indonesia.

Chapter 6 analyzed the spatial patterns of inequality in access to electricity and its supply provision. Evidently, Indonesia is a particular interesting country to study the spatial diffusion pattern of electrification. Indonesia is an archipelagic country where people are unevenly distributed across five big islands and hundreds of small islands. Unlike the situation in non-archipelagic countries, where new connections can be provided by extending the existing lines, Indonesia needs huge investments to build new lines to bring electricity to people on the different islands, and the up-front costs are very high. Therefore, electrification in Indonesia is an exceptionally spatially heterogeneous process due to an extreme form of discrete electricity planning caused by the geographical island barriers, with the main islands essentially being autonomous self-contained electricity systems. In this chapter, we aim to identify the relative contribution of several potential determinants of electrifitation to the the speed of electricity diffusion across Indonesian provinces.

From the results in Chapter 6, given the complexity of connecting all people including from all the relative small islands with huge disparities in terms of income and geography, we can conclude that the disparity in access to electricity and its supply is high across Indonesia. The electrification ratio remains low and the speed of technology diffusion is slow. A faster electricity diffusion process is experienced by the provinces in Java–Bali islands. The electrification ratio at the province level is positively correlated with population density as an internal factor and installed capacity of power supply availability as an external factor mainly contribute to increasing electrification ratio at the province level. We find that lack of power supply facilities is the main factor that explains the delay in electrification at the Eastern while the population density plays an important role to explain slower electrification patterns at Sumatera, Kalimantan and Sulawesi islands. Since the spatial variation in power supply availability is of course shaped by the geographical constraints inherent to Indonesia's island structure, our results imply that spatial heterogeneity at the province level contributes substantially to shaping the diffusion pattern.

To improve accessibility on electricity, the first prerequisite is to better understand the connection pattern across provinces. The next avenue for future research is to study whether people with an electricity connection can really afford to use the energy that they really need, how they benefited from the electricity and so on. A second promising avenue is to analyze diffusion pattern at the lower level of aggregation i.e. at the district or the village level. Analyzing diffusion patterns as well as energy consumption patterns at the household level would be very interesting as the third future research. Those future empirical studies could have essential policy implications especially regarding accessibility of key determinants of human well-being.

References

Abdullah, A. (2015). Does Education Reduce Income Inequality? A Meta-Regression Analysis. *Journal of Economic Surveys*, 29(2), 301–316.

Acemoglu, D. and J. Robinson (2010). The Role of Institutions in Growth and Development. *Review of Economics and Institutions*, 1(2), 1–33.

Aghion, P., E. Caroli and C. Garcia-Penalosa (1999). Inequality and Economic Growth: The Perspective of the New Growth Theories. *Journal of Economic Literature*, 37(4), 1615–1660.

Agustina, R., T. Dartanto, R. Sitompul, K.A. Susiloretni, Suparmi, E.L. Achadi, A. Taher, F. Wirawan, S. Sungkar, P. Sudarmono, A.H. Shankar and H. Thabrany (2019). Universal Health Coverage in Indonesia: Concept, Progress, and Challenges. *The Lancet*, 393(10166), 75–102.

Agustina, R.D., W. Fengler and G. Schulze (2012). The Regional Effects of Indonesia's Oil and Gas Policy: Options for Reform. *Bulletin of Indonesian Economic Studies*, 48(3), 369–397.

Akita, T. and A.S. Alisjahbana (2002). Regional Income Inequality in Indonesia and the Initial Impact of the Economic Crisis. *Bulletin of Indonesian Economic Studies*, 38(2), 201–222.

Akita, T. (2003). Decomposing Regional Income Inequality in China and Indonesia using Two-Stage Nested Theil Decomposition Method. *Annals of Regional Science*, 37(1), 55–77.

Akita, T., P.A. Kurniawan and S. Miyata (2011). Structural Changes and Regional Income Inequality in Indonesia: A Bi-dimensional Decomposition Analysis. *Asian Economic Journal*, 25(1), 55–77.

Akita, T. and R.A. Lukman (1995). Interregional Inequalities in Indonesia: A Sectoral Decomposition Analysis for 1975–92. *Bulletin of Indonesian Economic Studies*, 31(2), 61–80.

Akita, T. and S. Miyata (2008). Urbanization, Educational Expansion, and Expenditure Inequality in Indonesia in 1996, 1999, and 2002. *Journal of the Asia Pacific Economy*, 13(2), 147–167.

Alesina, A. and D. Rodrik (1994). Distributive Politics and Economic Growth. *Quarterly Journal of Economics*, 109, 465–490.

Alonso, J.A. and C. Garcimartín (2013). The Determinants of Institutional Quality. More on the Debate. *Journal of International Development*, 25(2), 206–226.

Anand, S. and S.M.R. Kanbur (1993). The Kuznets Process and the Inequality–Development Relationship. *Journal of Development Economics*, 40(1), 25–52.

Angeles, L. (2010). An Alternative Test of Kuznets' Hypothesis. *Journal of Economic Inequality*, 8, 463–473.

Atkinson, A.B. and A. Brandolini (2001). Promise and Pitfalls in the Use of "Secondary" Data-Sets: Income Inequality in OECD Countries as a Case Study. *Journal of Economic Literature*, 39(3), 771–799.

Atkinson, A.B., T. Piketty and E. Saez (2011). Top Incomes in the Long Run History. *Journal of Economic Literature*, 49(1), 3–71.

Asia Development Bank (2012). *Asian Development Outlook 2012: Confronting Rising Inequality in Asia*. Manila: Asia Development Bank.

Asia Development Bank (2016). *Achieving Universal Electricity Access in Indonesia*. Manila: Asia Development Bank.

Azis, I.J. (2015). Integration, Contagion and Income Distribution. In: P. Nijkamp, A. Rose and K. Kourtit (eds), *Regional Science Matters*. Cham: Springer International Publishing Switzerland, pp. 131–161.

Badan Penyelenggara Jaminan Sosial (BPJS) (2018). *Ringkasan Eksekutif Laporan Pengelolaan Program dan Laporan Keuangan Jaminan Sosial Kesehatan Tahun 2017*. Jakarta: BPJS Kesehatan.

Badan Penyelenggara Jaminan Sosial (BPJS) (2020). *Ringkasan Eksekutif Laporan Pengelolaan Program dan Laporan Keuangan Jaminan Sosial Kesehatan Tahun 2019*. Jakarta: BPJS Kesehatan.

Badan Pusat Statistik (BPS) (2020). *Statistik Indonesia 2020*. Jakarta: BPS

Baland, J.M. and P. Francois (2000). Rent-Seeking and Resource Booms. *Journal of Development Economics*, 61, 527–542.

Banerjee, A.V. and E. Duflo (2003). Inequality and Growth: What Can the Data Say? *Journal of Economic Growth*, 8(3), 267–299.

Barnes, D. (2007). *The Challenge of Rural Electrification Strategies for Developing Countries*. London: Routledge.

Barnes, D. and G. Foley (2004). *Rural Electrification in the Developing World: A Summary of Lessons from Successful Programs*. Washington, DC: UNDP/World Bank (ESMAP).

Barro, R.J. (2000). Inequality and Growth in a Panel of Countries. *Journal of Economic Growth*, 5(1), 5–32.Barro, R.J. and X. Sala-i-Martin (1992). Convergence. *Journal of Political Economy*, 100(2), 223–251.

Baum-Snow, N. and R. Pavan (2013). Inequality and City Size. *Review of Economics and Statistics*, 95(5), 1535–1548.

Behrens, K. and F. Robert-Nicoud (2014). Survival of the Fittest in Cities: Urbanization and Inequality. *Economic Journal*, 124(581), 1371–1400.

Benabou, R. (1996). Inequality and Growth. In: B. Bernanke and J. Rotemberg (eds), *NBER Macroeconomics Annual 1996*. Cambridge, MA: MIT Press, pp. 11–92.

Benabou, R. (2000). Unequal Societies: Income Distribution and the Social Contract. *American Economic Review*, 90(1), 96–129.

Benhabib, J. and M.M. Spiegel (2000). The Role of Financial Development in Growth and Investment. *Journal of Economic Growth*, 5, 341–360.

Benjamin, D., L. Brandt and J. Giles (2011). Did Higher Inequality Impede Growth in Rural China? *Economic Journal*, 121(557), 1281–1309.

Bhattacharya, J. and M. Packalen (2011). The Other Ex Ante Moral Hazard in Health. *Journal of Health Economics*, 31(1), 135–146.

Bhattacharyya, S.C. and S. Ohiare (2012). The Chinese Electricity Access Model for Rural Electrification: Approach, Experience and Lessons for Others. *Energy Policy*, 49, 676–687.

Boediono (1990). Growth and Equity in Indonesia. *Singapore Economic Review*, 35(1), 84–101.

Booth, A. (1992). Income Distribution and Poverty. In A. Booth (ed.), *The Oil Boom and After: Indonesian Economic Policy and Performance in the Soeharto Era*. Singapore: Oxford University Press, pp. 323–366.

Booth, A. (2000). Poverty and Inequality in the Soeharto Era: An Assessment. *Bulletin of Indonesian Economic Studies*, 36(1), 73–104.

Booth, A. and R.M. Sundrum (1981). Income Distribution. In: A. Booth and P. McCawly (eds), *The Indonesian Economy during the Soeharto Era*. Kuala Lumpur: Oxford University Press, pp. 181–217.

Bourguignon, F. (2004). *The Poverty–Growth–Inequality Triangle*. Indian Council for Research on International Economic Relations, New Delhi Working Papers, No. 125. New Delhi: Indian Council for Research on International Economic Relations (ICRIER).

Bourguignon, F. and T. Verdier (2000). Oligarchy, Democracy, Inequality and Growth. *Journal of Development Economics*, 62(2), 285–313.

Bruno, M., M. Ravallion and L. Squire (1998). Equity and Growth in Developing Countries: Old and New Perspectives on the Policy Issues. In: V. Tanzi and K. Chuk (eds), *Income Distribution and High-Quality Growth*. Cambridge, MA: MIT Press, pp. 117–146.

Burke, P.J. and B.P. Resosudarmo (2012). Survey of Recent Developments. *Bulletin of Indonesian Economic Studies*, 48(3), 299–324.

Castells-Quintana, D. (2018). Beyond Kuznets: Inequality and the Size and Distribution of Cities. *Journal of Regional Science*, 58, 564–580.

Chiswick, B.R. (1968). The Average Level of Schooling and the Intra-Regional Inequality of Income: A Clarification. *American Economic Review*, 58(3), 495–500.

Chong, A. and M. Gradstein (2007). Inequality and Institutions. *Review of Economics and Statistics*, 89(3), 454–465.

Chongvilaivan, A. (2013). *Individual Income Inequality and Its Drivers in Indonesia: A Theil Decomposition Reassessment*. ISEAS Working Paper No. 2. Singapore: Institute of Southeast Asian Studies.

Comin, D., M. Dmitriev and E. Rossi-Hansberg (2012). *The Spatial Diffusion of Technology*. NBER Working Papers, No. 18534. Cambridge, MA: National Bureau of Economic Research.

Comin, D. and B. Hobijn (2004). Cross-Country Technology Adoption: Making the Theories Face the Facts. *Journal of Monetary Economics*, 51(1), 39–83.

Courbage, C. and A. de Coulon (2004). Prevention and Private Health Insurance in the U.K. *Geneva Papers on Risk and Insurance – Issues and Practice*, 29(4), 719–727.

Cutler, D. and R. Zeckhauser (2000). The Anatomy of Health Insurance. In: A. Culyer and J. Newhouse (eds), *Handbook of Health Economics*, Vol. 1. North Holland: Elsevier, pp. 563–643.

Dagnachew, A.G., P.L. Lucas, A.F. Hof, D.E.H.J. Gernaat, H.-S. de Boer and D.P. van Vuuren (2017). The Role of Decentralized Systems in Providing Universal Electricity Access in Sub-Saharan Africa – A Model-Based Approach. *Energy*, 139, 184–195.

Dartanto, T., C.H. Siregar, A.U. Lumbanraja, Usman, H. Bintara, W. Pramono and N.K. Sholihah (2019). *Enrolment of Informal Sector Workers on the National Health Insurance System in Indonesia: A Qualitative Analysis*. Working Paper LPEM-FEBUI No. 033. Jakarta: Institute for Economic and Social Research Faculty of Economics and Business Universitas Indonesia.

Dave, D. and R. Kaestner (2009). Health Insurance and Ex Ante Moral Hazard: Evidence from Medicare. *International Journal of Health Care Finance and Economics*, 9, 367–390.

Davies, J.B., S. Sandström, A. Shorrocks and E. Wolff (2009). The Global Pattern of Household Wealth. *Journal of International Development*, 21(8), 1111–1124.

Deininger, K. and L. Squire (1996). A New Data Set Measuring Income Inequality. *World Bank Economic Review*, 10(3), 565–591.

Deininger, K. and L. Squire (1997). Economic Growth and Income Inequality: Re-Examining the Links. *Finance and Development*, March, 38–41.

Deininger, K. and L. Squire (1998). New Ways of Looking at Old Issues: Inequality and Growth. *Journal of Development Economics*, 57(2), 259–287.

Dick, H., V.J.H. Houben, J.T. Lindblad and T.K. Wie (2002). *The Emergence of a National Economy: An Economic History of Indonesia, 1800–2000*. Singapore: NUS Press.

Dollar, D. and A. Kraay (2002). Growth is Good for the Poor. *Journal of Economic Growth*, 7(3), 195–225.

Dominicis, L., R.J.G.M. Florax and H.L.F. de Groot (2008). A Meta-Analysis on the Relationship between Income Inequality and Economic Growth. *Scottish Journal of Political Economy*, 55(5), 654–682.

Duranton, G. and D. Puga (2013). The Growth of Cities. In: V. Henderson and W. Strange (eds), *Handbook of Economic Growth*, Volume 2A. Amsterdam: North-Holland, pp. 781–853.

Easterly, W. (2002). *Inequality Does Cause Underdevelopment: New Evidence*. Center for Global Development Working Paper No. 1. Washington, DC: Center for Global Development.

Easterly, W. (2007). Inequality Does Cause Underdevelopment: Insights from a New Instrument. *Journal of Development Economics*, 84(2), 755–776.

Edward, P. and A. Sumner (2015). *New Estimates of Global Poverty and Inequality: How Much Difference Do Price Data Really Make?* Center for Global Development Working Paper No. 403. Washington, DC: Center for Global Development.

Ehrlich, I. and G.S. Becker (1972). Market Insurance, Self-Insurance, and Self-Protection. *Journal of Political Economy*, 80, 623–648.

Ferré, C., F.H.G. Ferreira and P. Lanjouw (2012). Is There a Metropolitan Bias? The Relationship between Poverty and City Size in a Selection of Developing Countries. *World Bank Economic Review*, 26(3), 351–382.

Firman, T. (2014). The Dynamics of Jabotadebek Development: the Challenge of Urban Governance. In Hill, H. (ed.), *Regional Dynamics in a Decentralized Indonesia*. Singapore: Institute of Southeast Asian Studies, pp. 368-385.

Foley, G. (1992). Rural Electrification in the Developing World. *Energy Policy*, 20, 145–253.

Forbes, K.J. (2000). A Reassessment of the Relationship between Inequality and Growth. *American Economic Review*, 90(4), 869–887.

Frankema, E. and D. Marks (2009). Was It Really “Growth with Equality” under Soeharto? A Theil Analysis of Indonesian Income Inequality, 1961–2002. *Economics and Finance in Indonesia*, 57(1), 48–80.

Frankema, E. and D. Marks (2010). Growth, Stability, and Equality? Re-assessing Indonesian Inequality from a Comparative Perspective. *Economic History of Developing Regions*, 25(1), 75–104.

Fuady, A. (2019). *Arsitektur Jaminan Kesehatan Indonesia*. Jakarta: Sagung Seto Publisher.

Galor, O. (2011). *Unified Growth Theory*. Princeton: Princeton University Press.

Galor, O. and O. Moav (2004). From Physical to Human Capital Accumulation: Inequality and the Process of Development. *Review of Economic Studies*, 71(4), 1001–1026.

Galor, O. and Zeira, J. (1993) Income Distribution and Macroeconomics. *Review of Economic Studies 60(1), 35–52*.

Global Burden of Disease Collaborative Network (2017). *Global Burden of Disease Study 2016 (GBD 2016) Results*. Seattle, United States: Institute for Health Metrics and Evaluation (IHME).

Goldberg, P.K. and N. Pavcnik (2007). Distributional Effects of Globalization in Developing Countries. *Journal of Economic Literature*, 45(1), 39–82.

Haughton, J.H. and S.R. Khandker (2009). *Handbook on Poverty and Inequality*. Washington, DC: World Bank Publications.

Henderson, V. (2003). The Urbanization Process and Economic Growth: The So-What Question. *Journal of Economic Growth*, 8, 47–71.

Henderson, V. (2007). Understanding Knowledge Spillovers. *Regional Science and Urban Economics*, 37, 497–508.

Heywood, P. and N.P. Harahap (2009). Health Facilities at the District Level in Indonesia. *Australia and New Zealand Health Policy*, 6(13), 1–11.

Hidayat, B., Mundiharno, J. Němec, V. Rabovskaja, C.S. Rozanna and J. Spatz (2015). *Financial Sustainability of the National Health Insurance in Indonesia: A First Year Review*. Policy Brief. Jakarta: The Indonesian–German Social Protection Programme.

Hill, H., B.P. Resosudarmo and Y. Vidyattama (2008). Indonesia's Changing Economic Geography. *Bulletin of Indonesian Economic Studies*, 44(3), 407–435.

Howes, S. and R. Davies (2014). Survey of Recent Developments. *Bulletin of Indonesian Economic Studies*, 50(2), 157–183.

Iqbal, N. and V. Daly (2014). Rent Seeking Opportunities and Economic Growth in Transitional Economies. *Economic Modelling*, 37, 16–22.

Jaumotte, F., S. Lall and C. Papageorgiou (2013). Rising Income Inequality: Technology, or Trade and Financial Globalization? *IMF Economic Review*, 61, 271–309.

Jimenez, R. (2017). Barriers to Electrification in Latin America: Income, Location, and Economic Development. *Energy Strategy Reviews*, 15, 9–18.

Johar, M., P. Soewondo, R. Pujisubekti, H. K. Satrio, A. Adji (2019). In Data We Trust? An Analysis of Indonesian Socio-Economic Survey Data. *Bulletin of Indonesian Economic Studies*, 55(1), 61–82.

Kaldor, N. (1955). Alternative Theories of Distribution, *The Review of Economic Studies*, 23(2), 83–100.

Kanbur, R. (2000). Income Distribution and Development. In: A.B. Atkinson and F. Bourguignon (eds), *Handbook of Income Distribution*, Vol. 1. North Holland: Elsevier, pp. 791–841.

Kataoka, M. (2018). Inequality Convergence in Inefficiency and Interprovincial Income Inequality in Indonesia for 1990–2010. *Asia-Pacific Journal of Regional Science*, 2, 297–313.

Kenkel, D.S. (1994). The Demand for Preventive Medical Care. *Applied Economics*, 26(4), 313–325.

Knack, S. and P. Keefer (1995). Institutions and Economic Performance: Cross Country Tests Using Alternative Institutional Measures. *Economics and Politics*, 7, 207–227.

Knight, J.B. and R.H. Sabot (1983). Educational Expansion and the Kuznets Effect. *American Economic Review*, 73(5), 1132–1136.

Kuznets, S. (1955). Economic Growth and Income Inequality. *American Economic Review*, 45(1), 1–28.

Lahoti, R., A. Jayadev and S. Reddy (2014). *The Global Consumption and Income Project (GCIP): An Overview*. Working Paper No. 1402. New York: New School for Social Research, Department of Economics.

Lankhuizen, M.B.M., H.L.F. de Groot and G.-J.M. Linders (2011). The Trade-Off between Foreign Direct Investments and Exports: The Role of Multiple Dimensions of Distance. *World Economy*, 34(8), 1395–1416.

Leigh, A. and P. van der Eng (2009). Inequality in Indonesia: What Can We Learn from Top Incomes? *Journal of Public Economics*, 93(1–2), 209–212.

Lewis, W.A. (1954). Economic Development with Unlimited Supplies of Labour. *The Manchester School*, 22(2), 139–191.

Li, H. and H. Zou (1998). Income Inequality is not Harmful for Growth: Theory and Evidence. *Review of Development Economics*, 2(3), 318–334.

Lundberg, M. and L. Squire (2003). The Simultaneous Evolution of Growth and Inequality. *Economic Journal*, 113(487), 326–344.

Mankiw, N.G., D. Romer and D.N. Weil (1992). A Contribution to the Empirics of Economic Growth. *Quarterly Journal of Economics*, 107(2), 407–437.

Manning, C. and D.S. Pratomo (2013). Do Migrants Get Stuck in the Informal Sector? Findings from a Household Survey in Four Indonesian Cities. *Bulletin of Indonesian Economic Studies*, 49(2), 167–192.

Manning, W., J. Newhouse, N. Duan, E. Keeler and A. Leibowitz (1987). Health Insurance and the Demand for Medical Care: Evidence from a Randomized Experiment. *American Economic Review*, 77(3), 251–277.

Mboi, N., I.M. Surbakti, I. Trihandini, I. Elyazar, K.H. Smith, P.B. Ali, S. Kosen, K. Flemons, S.E. Ray, J. Cao, S.D. Glenn, M.K. Miller-Petrie, M.D. Mooney, J.L. Ried, D.N.A. Ningrum, F. Idris, K.N. Siregar, P. Harimurti, R.S. Bernstein, T. Pangestu, Y. Sidharta, M. Naghavi, C.J.L. Murray and S.I. Hay (2018). On the Road to Universal Healthcare in Indonesia, 1990–2016: A Systematic Analysis for the Global Burden of Disease Study 2016. *The Lancet*, 392(10147), 581–591.

McCawley, P. (1971). *The Indonesian Electric Supply Industry*. Ph.D. thesis, Australian National University.

Milanovic, B. (2016). Global Inequality: A New Approach for the Age of Globalization. Cambridge, MA: The Belknap Press of Harvard University Press.

Ministry of Energy and Mineral Resources (2020). *Electricity Statistic 2019*. Jakarta: Sekretariat Jenderal Ketenagalistrikan.

Ministry of Health (2014). *Peraturan Menteri Kesehatan No 41 Tahun 2014 tentang Pedoman Gizi Seimbang*. Jakarta: Kementrian Kesehatan RI.

Ministry of Home Affair (2019). *Materi Direktorat Toponimi dan Batas Daerah*. Jakarta: Direktorat Jenderal Bina Administrasi Kewilayahan.

Miranti, R., Y. Vidyattama, E. Hansnata, R. Cassells and A. Duncan (2013). *Trends in Poverty and Inequality in Decentralising Indonesia*. OECD Social, Employment and Migration Working Paper No. 148. Paris: OECD Publishing.

Mishra, S.C. (2009). *Economic Inequality in Indonesia: Trends, Causes and Policy Response*. Colombo: UNDP Report for Strategic Asia.

Mulder, P. and J. Tembe (2008). Rural Electrification in an Imperfect World: A Case-Study from Mozambique. *Energy Policy*, 36, 2785–2794.

Murphy, K.M., A. Shleifer and R.W. Vishny (1993). Why Is Rent-Seeking So Costly to Growth? *American Economic Review*, 83(2), 409–414.

Nawaz, S., N. Iqbal and M. Khan (2014). The Impact of Institutional Quality on Economic Growth: Panel Evidence. *Pakistan Development Review*, 53(1), 15–31.

Nel, P. (2008). *The Politics of Economic Inequality in Developing Countries*. London: Palgrave Macmillan.

Nord, S. (1980). Income Inequality and City Size: An Examination of Alternative Hypotheses for Large and Small Cities. *Review of Economics and Statistics*, 62(4), 502–508.

Pagán, J.A., A. Puig and B.J. Soldo (2007). Health Insurance Coverage and the Use of Preventive Services. *Mexican Adults. Health Economics*, 16(12), 1359–1369.

Palma, J.G., (2011). Homogeneous Middles vs. Heterogeneous Tails, and the End of the ‘Inverted-U’: It's All About the Share of the Rich, *Development and Change*, 42(1), 87–153.

Pauly, M. (1968). The Economics of Moral Hazard: Comment. *American Economic Review*, 58(3), 531–537.

Perotti, R. (1996). Growth, Income Distribution and Democracy. *Journal of Economic Growth*, 1, 149–187.

Persson, T. and G. Tabellini (1994). Is Inequality Harmful for Growth? *American Economic Review*, 84(3), 600–621.

Piketty, T. (2014). *Capital in the Twenty-First Century*. Cambridge MA: Belknap Press.

Piketty, T. and E. Saez (2003). Income Inequality in the United States, 1913–1998. *Quarterly Journal of Economics*, 118(1), 1–39.

PT PLN (Persero) (1995). *50 Tahun Pengabdian PLN*. Jakarta, Indonesia: PT PLN (Persero).

PT PLN (Persero) (2017). *Rencana Investasi*. Jakarta, Indonesia: PT PLN (Persero).

PT PLN (Persero) (2019). *Electricity Statistics*. Jakarta, Indonesia: PT PLN (Persero).

Rahman, M., J.V. Paatero, A. Poudyal and R. Lahdelma (2013). Driving and Hindering Factors for Rural Electrification in Developing Countries: Lessons from Bangladesh. *Energy Policy*, 61, 840–851.

Rashad, K. and S. Markowitz (2009). Incentives in Obesity and Health Insurance. *Inquiry*, 46(4), 418–432.

Ravallion, M. and S. Chen (2003). Measuring Pro-poor Growth. *Economics Letters*, 78(1), 93–99.

Resosudarmo, B. and Y. Vidyattama (2006). Regional Income Disparity in Indonesia: A Panel Data Analysis. *ASEAN Economic Bulletin*, 23(1), 31–44.

Rodrik, D., A. Subramanian and F. Trebbi (2004). Institutions Rule: The Primacy of Institutions over Geography and Integration in Economic Development. *Journal of Economic Growth*, 9(2), 131–165.

Rukmana, D. (2015). The Change and Transformation of Indonesian Spatial Planning after Suharto's New Order Regime: The Case of the Jakarta Metropolitan Area. *International Planning Studies*, 20(4), 350–370.

Rukmana, D., F.Z. Fahmi and T. Firman (2018). Suburbanization in Asia: A Focus on Jakarta. In: B. Hanlon and T.J. Vicino (eds), *The Routledge Companion to the Suburbs*. London: Routledge, pp. 110–120.

Sagala, P., T. Akita and A.A. Yusuf (2014). Urbanization and Expenditure Inequality in Indonesia: Testing the Kuznets Hypothesis with Provincial Panel Data. *Letters in Spatial and Resource Sciences*, 7, 133–147.

Sakamoto, H. (2007). *The Dynamics of Inter-Provincial Income Distribution in Indonesia*. The International Center for the Study of East Asian Development Working Paper Series, No. 2007(25). Kitakyushu: The International Centre for the Study of East Asian Development (ICSEAD).

Sambodho, J.P. (2019). *From Clients to Citizens? Democratization and Everyday Citizenship in a West Javanese Village*. Ph.D. thesis, Universiteit van Amsterdam.

Sambodo, M.T. (2016). *From Darkness to Light: Energy Security Assessment in Indonesia's Power Sector*. Singapore: ISEAS, pp. 72-93.

Sinaga, R., Prastowo, B.C.H. Simangunsong, A. Liebman and A.H. Tambunan (2019). Analysis of Barriers in Supplying Electricity Using Interpretative Structural Modeling. *Energy Strategy Reviews*, 25, 11–17.

Stanciole, A. (2008). Health Insurance and Lifestyle Choices: Identifying Ex Ante Moral Hazard in the US Market. *The Geneva Papers*, 33, 627–644.

Stoneman, P. (2002). *The Economics of Technological Diffusion*. Oxford: Blackwell Publisher.

Sundrum, R.M. (1992). *Income Distribution in Less Developed Countries*. Oxon: Routledge.

Surroop, D., P. Raghoob , F. Wolfc , K.U. Shahd , P. Jeetaha (2018). Energy Access in Small Island Developing States: Status, Barriers and Policy Measures. *Environmental Development*, 27, 58–69.

Tadjoeddin, M.Z. (2013). Miracle That Never Was: Disaggregated Level of Inequality in Indonesia. *International Journal of Development Issues*, 12(1), 22–35.

Tadjoeddin, M.Z., W.I. Suharyo and S. Mishra (2001). Regional Disparity and Vertical Conflict in Indonesia. *Journal of the Asia Pacific Economy*, 6(3), 283–304.

Timmer, C.P. (2004). The Road to Pro-Poor Growth: The Indonesian Experience in Regional Perspective. *Bulletin of Indonesian Economic Studies*, 40(2), 177–207.

Todaro, M. (1976). Urban Job Creation, Induced Migration and Rising Unemployment: A Formula and Simplified Empirical Test for LDCs. *Journal of Development Economics*, 3, 211–226.

United Nations, Department of Economic and Social Affairs, Population Division (2018). *World Urbanization Prospects: The 2018 Revision*. Washington: United Nations.

Valeriani, E. and S. Peluso (2011). The Impact of Institutional Quality on Economic Growth and Development: An Empirical Study. *Journal of Knowledge Management, Economics and Information Technology*, 1(6), 1–25.

Van der Weide, R. and B. Milanovic (2014). *Inequality Is Bad for Growth of the Poor (But Not for That of the Rich)*. World Bank Policy Research Paper No. 6963. Washington, DC: World Bank.

Van Leeuwen, B. and P. Földvári (2016). The Development of Inequality and Poverty in Indonesia, 1932–2008. *Bulletin of Indonesian Economic Studies*, 52(3), 379–402.

Van Zanden, J.L. and D. Marks (2012). *An Economic History of Indonesia 1800–2010*. London: Routledge.

Vidyattama, Y. (2013). Regional Convergence and the Role of the Neighborhood Effect in Decentralised Indonesia. *Bulletin of Indonesian Economic Studies*, 49(2), 193–211.

Wadho, W. and S. Hussain (2020). *Ethnic Diversity, Concentration of Political Power and the Curse of Natural Resources*. GLO Discussion Paper No. 611. Essen: Global Labor Organization (GLO).

Wombwell, E., M.T. Fangman, A.K. Yoder and D.L. Spero (2015). Religious Barriers to Measles Vaccination. *Journal of Community Health*, 40, 597–604.

World Bank (2006). *World Development Report 2006: Equity and Development*. Washington DC: World Bank.

World Bank (2019). *Indonesia Economic Quarterly*. Jakarta: World Bank.

World Bank (2020). *World Development Indicators*. Washington, DC: World Bank.

Xu, K., D.B. Evans, K. Kawabata, R. Zeramdini, J. Klavus and C.J. Murray (2003). Household Catastrophic Health Expenditure: A Multicountry Analysis. *The Lancet*, 362, 111–117.

Yusuf, A.A. and B.P. Resosudarmo (2008). Mitigating Distributional Impact of Fuel Pricing Reform: Indonesian Experience. *ASEAN Economic Bulletin*, 25, 32–47.

Yusuf, A.A., A. Summer and I.A. Rum (2014). Twenty Years of Expenditure Inequality in Indonesia, 1993–2013. *Bulletin of Indonesian Economic Studies*, 50(2), 243–254.

Zweifel, P. and F. Breyer (1997). *Health Economics*. Oxford: Oxford University Press.

Zweifel, P. and W.G. Manning (2000). Moral Hazard and Consumer Incentives in Health Care. In A.J. Culyer and J.P. Newhouse (eds), *Handbook of Health Economics*. North Holland: Elsevier, pp. 409–459.

Summary

This thesis aims to answer what is happening to inequality in Indonesia for the past years. This empirical research consists of two parts, first, how is the relationship between inequality and development in Indonesia along with various aggregation levels; second, whether the accessibility to basic public services is the implication of inequality.

The key messages of the first part of this thesis are as follows: over the last decades relatively high economic growth in Indonesia is associated with rapidly increasing income inequality. Regional convergence of inequality across islands and provinces is driven by the fact that incomes of the rich in poor regions grow faster than those of the poor in rich regions. In other words, the middle class and especially the top incomes seem to benefit most from the economic growth dynamics in Indonesia. In relation to urbanization, our examination shows that if the current trend of increasing average agglomeration size continues, we can expect that inequality will further go up. Further, high inequality in the past will increase the subsequent growth, this high growth is then associated with high inequality in the current period. If this current inequality is too high, then the change in inequality will be positive and large enough to reduce growth in the future. In other words, the impact of high inequality in the past is associated with the slowing down economic growth and worsening existing inequality. It seems that our results support the Piketty (2014) argument that inequality keeps increasing as societies accumulate wealth. With regards to the institutional quality, the type of institutional quality plays a significant role in shaping the association with economic growth in the future. This role looks more important if we interact with the initial inequality suggesting that a combination of a certain degree of inequality and institutional quality is required to boost the economic growth.

The key messages of the second part of this thesis are as follows: given the health reform promoting inclusiveness in health access and given the complexity of connecting all people including from all the relatively small islands with huge disparities in terms of income and geography, insured people behave differently than

uninsured people, and within the insured group, the subsidized people behave differently from the non-subsidized people. Ex ante moral hazard exists in insured and subsidized groups. The disparity in access to electricity and its supply is high across Indonesia. Despite the fact that electrification already started more than one hundred years ago, the electrification ratio remains low and the speed of technology diffusion is slow. Population density as an internal factor and power supply availability as an external factor contribute to increasing the electrification ratio at the province level. This external factor can be translated as island barrier. Hence, the heterogeneity at the province level contributes in shaping the diffusion patterns.